Clemson University

TigerPrints

All Theses

Theses

5-2014

LONG-TERM MULTI-DIMENSIONAL INTERACTIVE TIME-LAPSE PHOTOGRAPHY USING MICROSOFT KINECT

Daniel McCoy
Clemson University, apexanalog@gmail.com

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

Part of the Communication Commons, Computer Sciences Commons, and the Fine Arts Commons

LONG-TERM MULTI-DIMENSIONAL INTERACTIVE TIME-LAPSE
PHOTOGRAPHY USING MICROSOFT KINECT

A Thesis
Presented to
the Graduate School of
Clemson University

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

by
Daniel Dean McCoy
May 2014

Accepted by:
Donald House, Committee Chair
Bruce Whisler
Anderson Wrangle

ABSTRACT

In this thesis, a method is presented for the capture and interactive presentation of long-term multi-dimensional time-lapse photography. Time-lapse capture is commonly used for the observation based study of relatively long term phenomena such as plant growth and weather patterns. In terms of filmic devices, the visual time compression effect is complementary to slow motion and is nearly as prevalent. In this project, commonly available camera and computer equipment is used to capture images autonomously with minimal system supervision. A set of images is established, using long term, short interval continuous capture at a fixed position. Results are presented demonstrating dynamic movement within this set using the Microsoft Kinect sensor for Xbox 360 to evaluate participant gestures in real-time. Viewers' tracked movements and positions motivate specific frame selection and playback order, allowing independent navigation through the time-lapse, independently by time of day, time of season, and in any order participants can obtain with their own movement and performance.

DEDICATION

  I dedicate this thesis to my family for their support, encouragement, and love. I am especially grateful to my grandfathers, Roger "Max" Fry and Howard "Mac" McCoy. Time spent with them taught me that anything is possible, and their passion for work, love of life, and dedication to family continue to influence everything I do.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION AND BACKGROUND

1.1 Rhythmic Interplay

A musician all my life, I appreciate the layered and syncopated nature of the world. Everything that circles the sun follows its downbeat, and the interplay of rhythms is infinite and beautiful. Our continuous movement through space and time is marked by milestones, singular events, frozen and static. *Fragments of Time* is my expression of the way these moments describe a path, draw a picture, and tell a story.

1.2 Project Origin and Influences

As a student of several disciplines, it seemed impossible to choose only one as the focus of my thesis. Instead, I set out to create a work that embodied the common ground in my multidisciplinary curriculum. I've always had an affinity for drawing, but never significantly developed those skills until I took up figure drawing and learned to storyboard for animation. I've enjoyed photography for as long as I can remember, but studying the way light propagation is modeled in the computer has changed the way I see. This new visual learning, in combination with my musical background in piano performance and experience in audio production, allowed my appreciation for temporal art forms and live performance to encompass interactive art and non-linear narratives. Together with these elements, learning to program has fundamentally changed my understanding of the computer from a utilitarian tool to an expressive instrument. I sought a project that would emphasize this shift in perspective and motivate others to

consider the coherence of art and technology through direct, real-time interaction with both.

To visualize my search for a project that encompassed these influences, I created a drawing in the style of the color mixing Venn diagram, as shown in Figure 1.1.
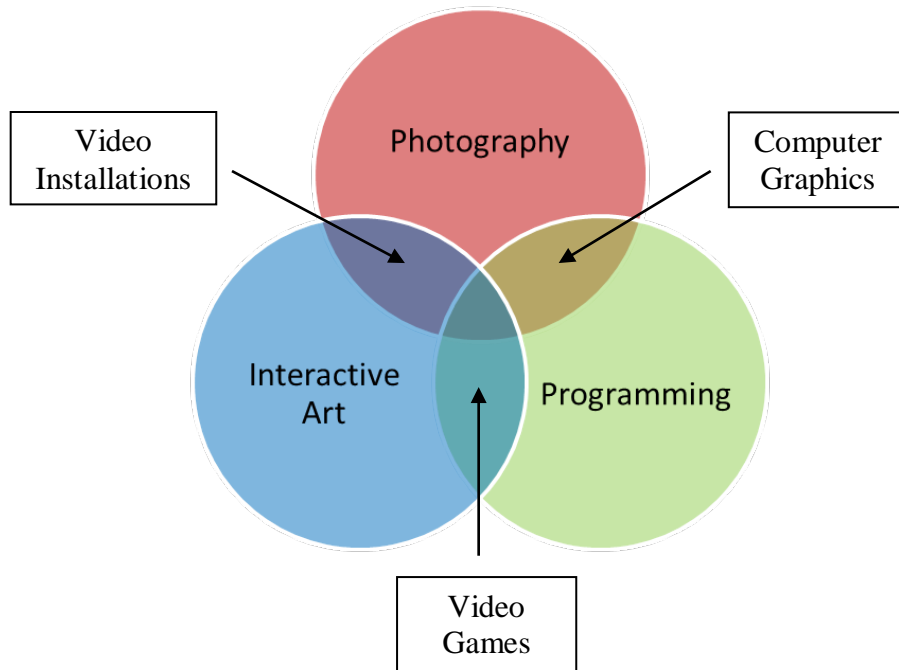
Figure 1.1: Project Genesis

Just as red, green, and blue wavelengths are components of pure white light, I hoped to find the convergence of photography, computer programming, and interactive art. The pairwise combinations of these three core influences are video-based art installations, videogames, and traditional computer graphics. My project is at the center of these disciplines; an interactive time-lapse photography exhibit that uses the principles of

object-oriented programming to create a dynamic, live, interactive visual experience using affordable consumer hardware.

1.3 Fragments of Time

The purpose of the interactive portion of this project, titled *Fragments of Time*, is to explore the idea of an artistic and technological continuum interactively through a combination of current digital imaging technologies. This provides perspective on the transitional state of these technologies by allowing participants to manipulate them in real time using only their own body, thereby effectively creating their own experience.

Artists have long sought to immerse their audience in their works. *Fragments* accomplishes this by relinquishing artistic control of time. Like a painting, the borders of the canvas and individual compositional forms are fixed. In a film, these things are dynamic, but movement through time is relatively constant if not literally linear. A fixed window of time is the essence of a photograph, each one a detailed record describing the travel of a specific amount of light. *Fragments* is made from photographs and has the mechanics of film, but time is the paint and the audience is the brush.

1.4 Basic Project Layout

The project is presented as an interactive video installation designed to be projected on a large flat surface such as a screen or wall. The space in front of the screen is open, and the movement of viewers standing in this space affects the content of the projection via hardware and software that tracks viewers' positions in real time. The

projected imagery consists of still frames from a set of images captured at regular intervals from a fixed viewpoint. This is an otherwise typical time lapse photo shoot with an exceptionally long duration, ideally an entire year.

The location and orientation of viewers within the presentation space controls the playback order of frames, using predetermined control schemes; the playback is not randomized. Playback must have a high enough frame rate to produce continuous motion for sequential frames, and frame blending is used to smooth the transition between non-sequential frames.

It is my intention that viewers will transition from passive observers to active participants when they realize they can control the visualization. Once this ability is established, viewers can create their own version of the visualization as they see fit.

1.5 About Time Lapse Photography

Time lapse photography has parallel origins with the development of motion pictures. It was first used by special effects pioneer George Melies to compress an amount of time into a short shot or sequence, the technique's most direct use. [Liu12] The time-lapse effect is created by replaying sequenced still frames at a higher rate than their capture rate. [Denv90] In this way, time-lapse photography is similar to slow-motion photography, which is produced when the capture rate is significantly greater than the playback rate. Just as slow-motion photography reveals high-frequency temporal detail, time-lapse photography reveals fluidity in low-frequency motion. The technique is commonly used to capture and temporally compress natural phenomena such as organic

growth and celestial motion. This type of movement is imperceptibly slow with direct observation. Time-lapse photography, like slow-motion or even photography in general, allows us to use technology to augment our perception to better understand the natural world.

I chose time-lapse photography for this project because the aesthetics of time-lapse dovetail with the technology used for both image capture and interactive presentation. For viewers to have a sense of significant control of the final presentation, the captured frames must span several months, so that seasonal changes are recorded. Continuous video capture for this length would produce an unmanageable amount of data, but the occasional, regular nature of time lapse photography allows a continuous visual experience covering a large amount of time, creating smooth motion from low frequency events. For example, the incremental position of the sun as it breaks the horizon at sunrise describes a path modulated by Earth's motion. Every type of weather and sky is represented in colorful motion. The growth of new leaves on all the trees is transformed from something we intrinsically know into a story we can watch. When presented interactively, these visual elements empower viewers, allowing them to feel as if they can control celestial movement, affect the weather, and even create or destroy life. Awakening this sense of artistic and technological empowerment is the purpose of this project.

1.6 Ubiquity of Digital Imaging Technology

The increasing availability of high-resolution, low cost digital image capture equipment, along with the decrease in the price of storage media, enables individuals to produce time-lapse photography sequences and films that were once the domain of large film studios or professional artists. As shown in the Figure 1.2, the cost of film media alone for this project would be prohibitive, and cameras and associated control equipment would also be very expensive. The affordability of professional quality digital cameras substantially diminishes these costs and also makes it possible for anyone to produce works that would have previously required a team of professionals and a large budget. This is especially true for time lapse projects, where using the computer to control the shot interval and to store captured images allows capture to proceed autonomously once properly set up.

Project Statistcs

- 310,789 images
- 20 second shot interval
- **<$10,000 for entire project**

Using 35mm Film Camera

- 8634 rolls of film
- **$43,166** at $5/roll

Using 35mm Movie Camera

- 15,840 frames/1000' reel
- 20 1000 ft. reels of film
- **$12,000** @ $600/roll
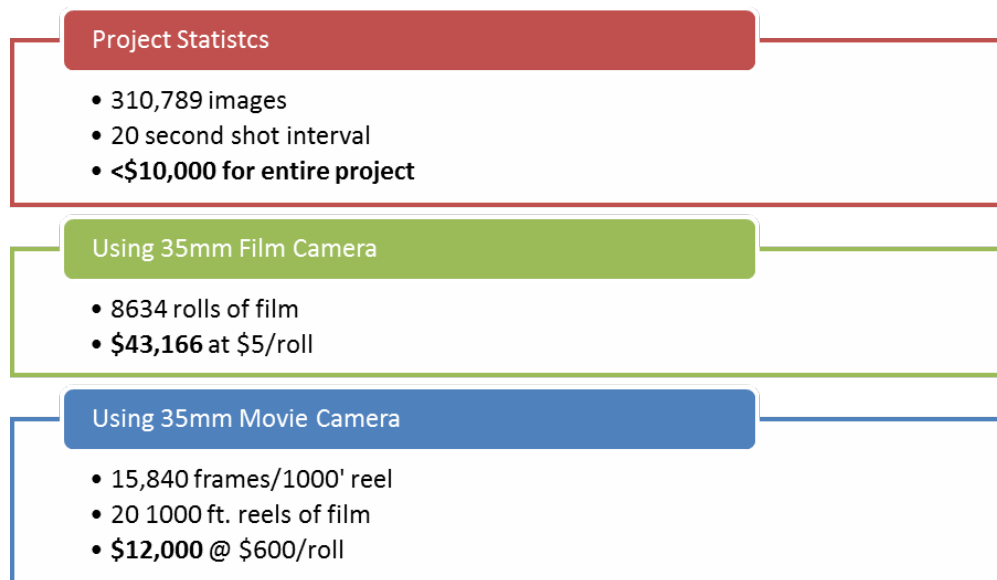
Figure 1.2: Cost Comparison

This ubiquity of digital imaging technology has also fueled technological innovation in the interactive entertainment industry. For example, the Microsoft Kinect is a consumer-grade multi-sensor bar containing a 2D image sensor similar to those found in webcams, a 3D infrared depth camera system, and a microphone array. Originally designed as a video game peripheral, the Kinect is available for less than $100, and serves as the sole audience-interactive controller in the presentation of this project.

CHAPTER TWO

RELATED WORK

2.1 Baraka

Ron Fricke's 1992 film Baraka [Fric08] is the work that drew my attention to time-lapse photography. It is a non-narrative film, shot in 70mm format around the globe. The film has the visual look of a documentary, but has no voiceover or explicit subject. Stunning visuals and the accompanying soundtrack pull you in by your senses, immersing you in an experience in place of a narrative. Nevertheless, the film's intense sequences of frantic human movement, traffic, and industrial elements are bookended by serene natural elements and beautiful time-lapse sequences created with precisely controlled camera movement, revealing the motion of Earth itself. Baraka uses a sensual experience to pose questions about what is constant and what is most important.

2.2 Roman Signer

I encountered the work of visual artist Roman Signer in a photography class. His film, Signer's Koffer [Sign14] features his installations, which at first viewing can seem absurd. Chairs being launched from hotel windows simultaneously, streamers tied to rockets launched into Stromboli, remote control helicopters creating dramatic tension; Signer's work often has a time based element with a visual punchline that makes me laugh out loud. Underneath that entertainment element, however, all his work seems to say, "Why not?" He uses anything and everything to create his installations, often constructing elaborate contraptions to accomplish the result, which is captured on film. I

came to appreciate his limitless approach. Signer caused me to think differently about art and to see it everywhere. He supplies narrative elements, asking viewers to pick them up and assemble the meaning. Film is Signer's medium, but he is a performance artist.

2.3 John Ott

Author of *My Ivory Cellar* [Ott58], John Ott was a pioneer in time lapse photography and built some of the first motion controlled movie cameras. Using time-lapse for the scientific study of plant development, it didn't take Ott long to figure out how to turn real, captured imagery into choreographed movement in the style of animation of the period. In the 1960's, Ott produced a film demonstrating his methods titled *Exploring the Spectrum*. [Ott75]

2.4 Alan Kim Boling, Jr.

Professional photographer Alan Boling [Boli14] produces high-dynamic-range photographs that capture the essence of a location. I've often taken a photo in a beautiful place, sometimes with my phone, only to find later that it was a poor replacement for standing there taking in the scene. Boling's photos, however, can take you to places you've never been. I especially appreciate how he can seamlessly combine several timespaces in one photograph. Your eye moves around the composition to areas that seem like they couldn't exist together, yet the motion is smooth and the transition is imperceptible.

CHAPTER THREE

IMPLEMENTATION

3.1 System Summary

The three primary components of the system, as shown in Figure 3.1, are a capture and storage system, a software environment for image processing, and an interactive output system for final display and presentation. Readily available, consumer-grade hardware and software comprises system components.
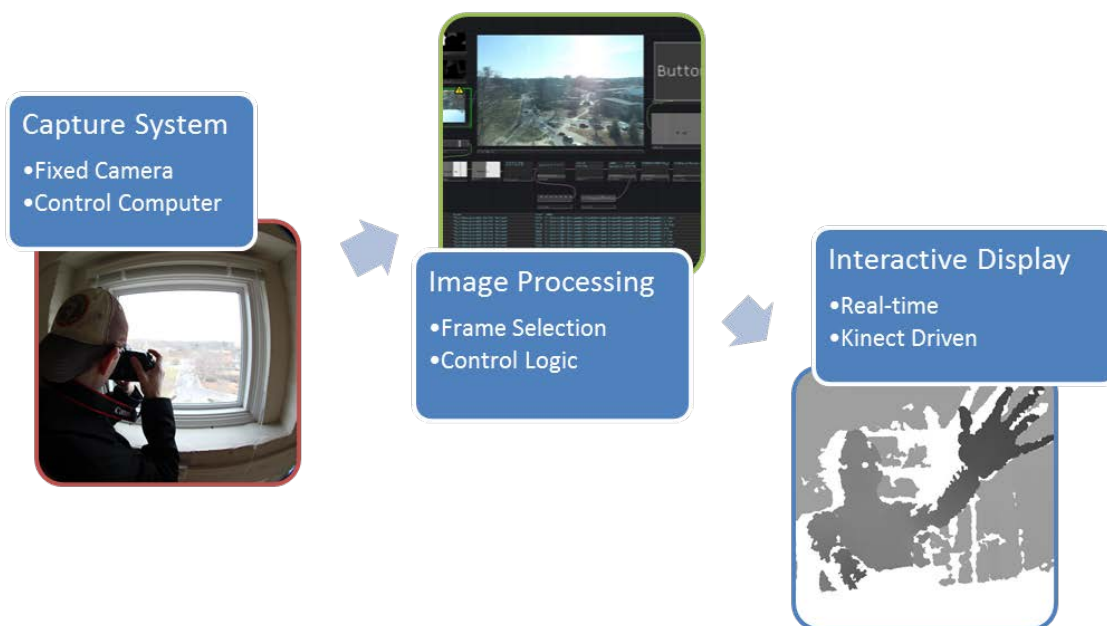

Figure 3.1: System Workflow

3.2 Capture System

The capture system is comprised of a fixed, tripod-mounted Canon EOS 5D mkII digital SLR camera with a Canon 24-70mm f/2.8L lens and a dedicated control computer running Canon's EOS Utility. This software functions both as an intervalometer, accurately timing and controlling exposures, as well as a storage management system. This configuration allows complete camera control and image management without physically touching the camera, which would compromise the final result. Since it is critical that the camera is not moved throughout the shot, an AC adaptor provides continuous camera power, and captured images are immediately stored on twin external hard drives which are regularly swapped for collection and backup. Images are saved in jpeg format at a resolution of 2784x1856. Each shot's file name contains a specific capture time, ISO, and unique shot number. Photographic parameters are also stored as metadata in each image file.

3.3 Shot Location

As shown in figure 3.2, the capture system is installed in a sixth story bell tower window in Tillman Hall on the Clemson University campus, overlooking the plaza in front of the building. This secure location is chosen for its view of a busy part of campus featuring pedestrian and automobile traffic, plant life, and a good view of the sky.

Figure 3.2: Shot Location

3.4 Camera Parameters

   Camera settings are set remotely via Canon EOS Utility software. The shot interval is 20 seconds with a fixed exposure time of one tenth of a second, allowing some blur in fast moving elements such as automobile traffic while capturing slower moving elements, such as clouds, as if they were static. Without this blur, time-lapse playback of fast movement can feel choppy and abrupt, like a digital video on fast forward. The exposure time must also be shorter than the shooting interval. This is trivial in most time-lapse shoots where the interval is often much greater than the exposure time, which varies to accommodate dynamic lighting conditions. For this project, the shooting interval was too short to use this method. Instead, aperture and ISO settings are allowed to vary with lighting conditions. Variable aperture produces smooth lighting changes across daylight sequences and dynamic ISO, which is impossible with film photography, produces relatively clear night shots without increasing the exposure time, at the expense of increased digital noise. This noise only appears in dark shots, is analogous to film grain, and is negligible when frames are played back as video. Specific settings are determined by experimentation after the camera is installed at the location because the settings depend on local lighting conditions. In general, variable aperture and ISO with a fixed shutter speed is a suitable setup for any time-lapse sequence created with a DSLR camera, as it allows for widely varying lighting conditions. All other automatic camera settings, including autofocus, are disabled to prevent any shooting delay or in-camera image processing, which would appear upon playback as timing jitter or flicker.

3.5 Filtering Artifacts

A circular polarizing filter was carefully fit to the camera rig during the time-lapse. This was done to capture more vibrant colors in the skies and trees, but it also created some undesirable artifacts from point source lights such as street lights and automobile headlights. As shown in Figure 3.3, the artifacts are more noticeable in night shots. Their effect on daytime shots is negligible.

Figure 3.3: Filtering Artifacts

3.6 Image Processing Software Environment

Initially, all software was created in a Linux environment using the C++ programming language. As described in Figure 3.4, a Python script was used to batch-rename all image files with frame number, timing data, and ISO separated by periods for simple parsing.
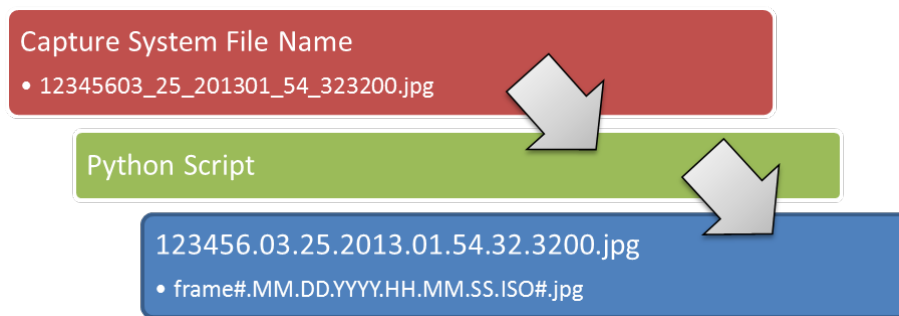
Figure 3.4: Original file names rewritten for parsing

Custom code was combined with open-source material from the OpenKinect project [OpenKinect] to access the depth data from the Kinect sensor and process it as a 2D grayscale image. In order to create iterative prototypes more quickly and to access Kinect skeletal tracking data directly, the project was migrated to the node-based TouchDesigner visual programming environment [Deri14] on the Windows 7 platform. The node-based approach used in TouchDesigner, as shown in Figure 3.5, is increasingly the standard in visual production software as it allows artists to create interactive visual environments using the object-oriented approach common in conventional program design.

15
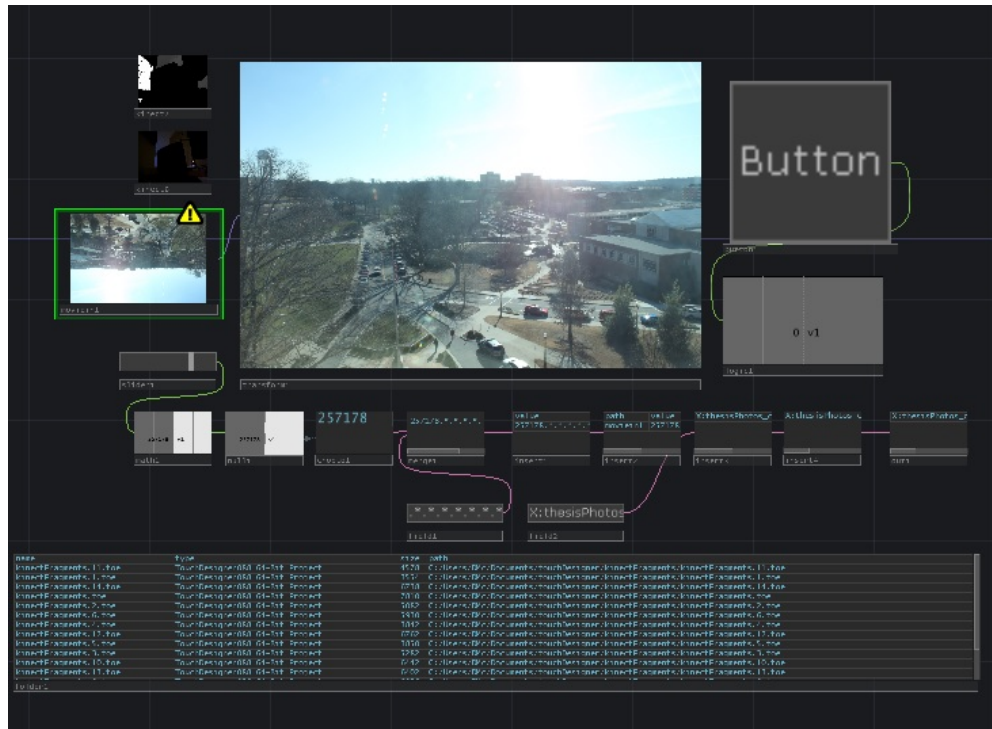
Figure 3.5: TouchDesigner Software Environment

3.7 TouchDesigner Nomenclature

TouchDesigner nodes are called operators, arranged in families based on the data

on which they operate. The Python programming language is used for operator

interconnections. This project uses DAT operators to access files and to create and

modify tables, texture operators – denoted TOP – for 2D image manipulation, and

CHOP's, or "channel operators", for most mathematical and logical operations. COMP

operators are premade control components such as buttons and sliders. Any network of

operators can be saved as a COMP and used in other operator networks, allowing

complex, nested project designs. The complete project is stored as a TOE file, short for

TouchDesigner Operator Environment.

3.8 TouchDesigner Project Structure

The TouchDesigner project for *Fragments* is built using a similar structure as the C++ prototype. A Folder DAT and a MovieIn TOP point to the image files, which are stored on a local solid state drive for fast access, and it also creates a table of their access paths. A Sort DAT rearranges the table created by the Folder DAT in capture order and assignes an index to each frame. This index is mapped to a MovieIn TOP, which cues the video at the specified location and begins caching upcoming frames. Manipulating the index via control structures comprised of CHOP operators allows the video to be cued at any point at any time.

Kinect skeletal data is acquired from the sensor with a Kinect CHOP, and is accessible by downstream nodes as channel information at thirty frames per second. The 3D location of each joint is tracked as three linear channels aligned with the cardinal axes. The sensor can track two full skeletons and the hip position of an additional four people, but this project uses only the horizontal and vertical position of the head and hands to calculate control gestures. Joint locations and stored and delays are used to calculate movement vectors. This data is used to specify the cue index, allowing users to control playback interactively.

3.9 Run-Time Hardware

The interactive system used to present the time-lapse imagery as described consists of a computer, two digital displays, and a Kinect for Xbox 360. This project was implemented and presented on a custom built computer running Ubuntu 12.04 LTS and Windows 7 Ultimate. Most computers should have no trouble running the TouchDesigner software, but it is critical to eliminate bottlenecks in image file access in order to maintain high frame rates. For this reason, my implementation uses a Samsung 840 EVO 1TB solid state drive. Lower resolution proxies of the original images were created using Image Magick, a powerful open source API. [Imag14] These proxies are used for fast real-time playback, and high resolution, non real-time renders can be created by replaying recorded movement data and using an alternate MovieIn TOP.

CHAPTER FOUR

RESULTS

4.1 Diversity of Results

The project has two components that produce results. The capture system produces digital photographs describing the scene. These stills are used as the input for the interactive part of the project, designed as an immersive art installation where the result is a unique video experience created in real time by the audience.

4.2 In-Place Prototyping

Off-site experimental shots were produced to test photographic parameters and computer control. Once the capture system was installed and the final shot composition was established, the camera could not be moved. Day-long, in-place test shots were created and evaluated, and camera control parameters were adjusted accordingly. The camera system was in direct sunlight each morning and in near complete darkness each night. These experimental test shots established settings that would produce desirable images across a range of lighting conditions. Settings were fixed at the beginning of the long-term shot, and images from the experimental shots were not included in the final image set used for display by the interactive system.

4.3 The Image Set

The capture system used traditional linear time-lapse photography techniques. Time lapse projects of this length are common, although the interval between captures is typically much longer. The resulting image set contained 310,789 high-resolution jpeg

photographs. The set describes the scene at any time of day or night over an 88 day period. As shown in Figure 4.1, the scene is represented at every time of day and also from winter to spring. The seasonal growth of tree leaves is the slowest motion element, captured with an approximate time resolution of twenty seconds. The combination of a very long term shot with a relatively high frequency produced a highly diverse and vibrant image set. Every type of sky is included, providing a moving visual reference for sunrises and sunsets, numerous cloud types and color palettes, and varying weather conditions. Barren trees seen at the beginning of the interval are flush with new growth in later images. On playback, the short interval between shots produces smooth movement of seasonal growth, shadows from sunlight, moving weather systems, and even the moon.
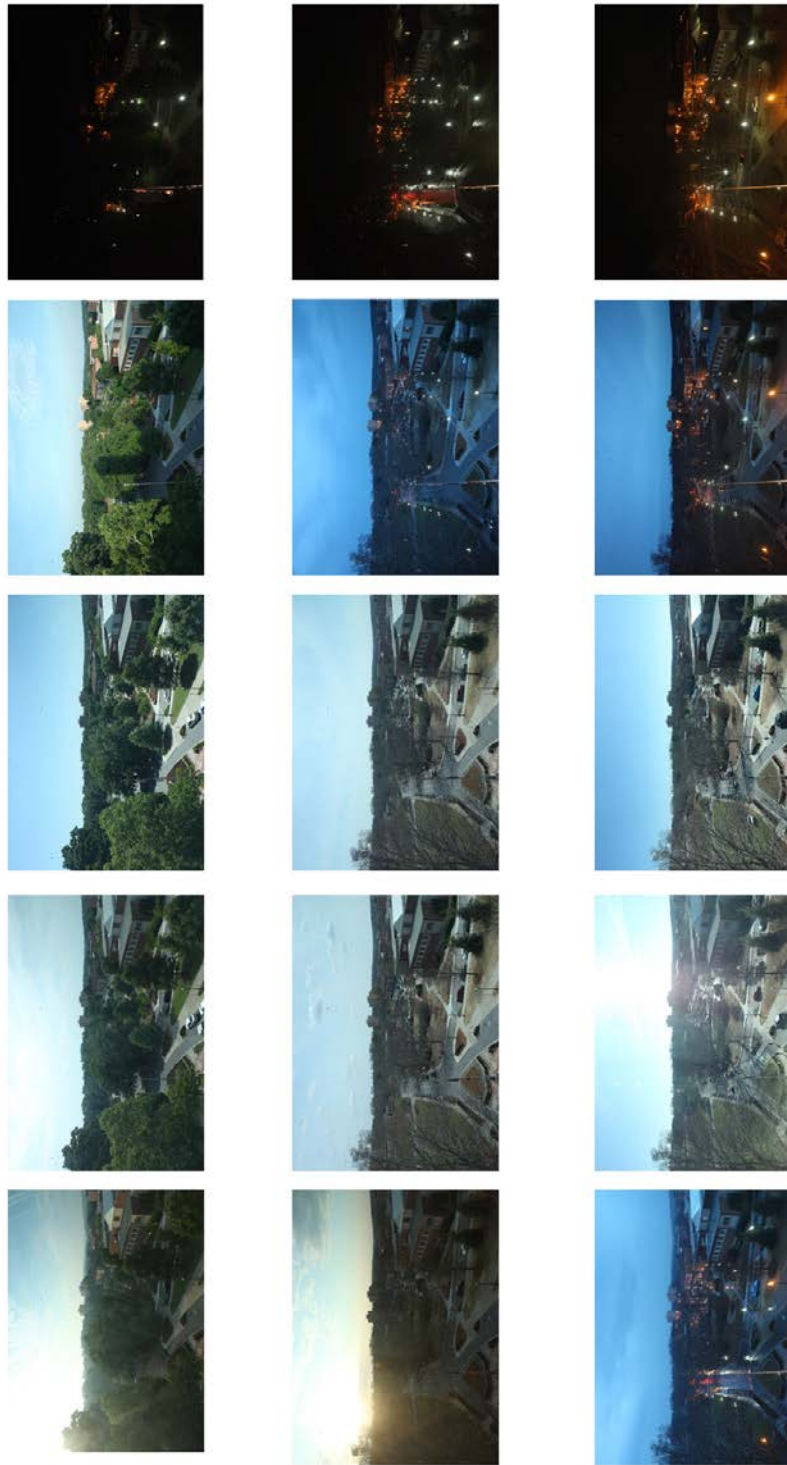
Figure 4.1: Sample frames arranged by days and by seasons

4.4 Continuity Lapse

There were some discontinuities in the image set due to camera shutter mechanical failures and computer image file writing errors, mostly in the earlier part of the shot interval. It was discovered that the computer's file system was not ideal for storing large numbers of files with long filenames in one directory, so the external hard drives were reformatted as a workaround. Timing jitter was created by these hard drive swaps and down time, but the problems were corrected without significantly interrupting the time-lapse shot, as the camera was never touched or moved.

4.5 Interactive Time Lapse Experience

Results produced by the second stage of the project are presented as an interactive art installation. Viewers' movements and gestures, recorded by the Kinect, affect which images the system selects and retrieves. Sequences are played back quickly to produce smooth motion, allowing participants immediate control. Sequences can also be cached and used to render video files.

The way the system translates viewers' physical movement into image selection parameters is controlled using a graphical user interface out of audience sight. This GUI allows the presenter to control frame selection, playback speed, and any other system parameters. The presentation display and the GUI are both part of the TouchDesigner project file, so the specific control parameters available can be adjusted and customized for a particular installation or purpose. This technological fluency is both a subject of the

installation and its core methodology. It is meant to provide the viewer a perspective about what they may be able to create with the ubiquitous technology that surrounds them. The following discussion explores some perspectives on this work.

CHAPTER FIVE

DISCUSSION

5.1 Humanization of Technology

One of the most powerful features of this project is its human, organic feel. Viewers can attempt to retrace a path, but producing the same result twice is effectively impossible. This is not a technical limitation, but the inherent result of linking the playback to human movement. Even in a world where anything imaginable can be rendered in full photorealism with enough mouse clicks, artists still spend considerable time extracting that realism. Like any CG rendering, this project is just a computer program under the hood, a digital expression comprised of millions of colored squares. In that way it is flawed and false by design, like theatre or videogames. We only get out what we are willing to invest; it requires our input and the resulting visual experience represents our participation.

5.2 Ideal Implementation and Variations

This project is a prototype, a test case designed to explore an idea and to refine it. Ideally, the time-lapse would span an entire year, creating a set of frames representing a complete annual cycle just as smaller spans like days and weeks are represented in this project. This way, the playback would have no effective start or end point. As long as there were no major changes, such as building construction, within the shot over the course of the year, the changes from the last frame to the first would be no different than changes from day to day.

Also, the relatively narrow, fixed perspective could be expanded using advanced photographic equipment and control systems. Ron Fricke's work [Fric08] uses computer controlled camera movements to produce traditional smooth pans over several hours, and that kind of implementation is no longer prohibitively expensive. If a stereo camera set was used, user input could control parallax, and multiple synchronized cameras could be used to produce results that create surreal visual scenes built from real, captured imagery.

Although it was a success, the project's implementation could be improved. Holes in the sequence due to technical problems can create abrupt changes on playback and should be minimized. Captured images could be stored in an uncompressed or raw format, allowing a greater range of artistic control during presentation. Any photographic parameter that can be adjusted in post-production becomes another real-time variable in the live presentation of captured frames.

Fragments would also benefit from increased optimization. This project used independent capture, image manipulation, and presentation systems, but using one system to accomplish all these stages would allow for more seamless control of the material.

5.3 Digital Economies of Scale

As the cost of capturing and storing a single digital image approaches zero, the possibilities are endless for artists. Paint and brushes can still be expensive, but the idea of not being able to paint for want of the technology is comical. Large record companies are essentially obsolete, movie studios will follow, and the technical capacity for capturing, storing, and manipulating high resolution, high frame rate imagery is as

common as a cell phone. Anyone with the imagination and inclination can create the

kinds of imagery described by this thesis.

CHAPTER SIX

CONCLUSION

Fragments of Time was developed as an experimental confluence of three disciplines in the Digital Production Arts MFA program: computer science, fine art, and performance. It was challenging in these respects, and in its intention to promote consideration of the continuous and intertwined nature of art and technology through real-time interactivity, it is successful.

Time lapse photography forms the visual vocabulary for this project, and a visual narrative is created by the audience, expressed via their performance. This interaction requires no training and there are no physical controls to learn; results are seen immediately, presented in a visually immersive projection. Digital imaging and motion tracking technologies are combined such that the technology is transparent, and the narrative experience is emphasized. It is the same in our everyday lives; the immediate availability of technology empowers individual expression and simultaneously brings us together. Technology has a profound and direct effect on our lives, but when it is operating as designed, it is invisible; we only see each other.

Critical thinking and artistic expression are the essence of human nature. My hope is that this thesis brings to light the way widely available technology and contemporary art forms continuously converge to produce infinite forms of human expression.

The only constant is change.

REFERENCES

[Liu12]   Liu, Wei, and Li Hongyun. (2012) "Time-Lapse photography applied to educational videos." *Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on*. 3669-3672.

[Kamd11]   Kamdar, D.G. and Vithalani, C. H. (2011) "A novel approach to design and implement Differential Time Lapse Video in real time application," *Emerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on* ,pp.731,735.

[Denv90]   Denvir, B. (1990). Photography. *The Thames and Hudson encyclopaedia of impressionism* (). London: Thames and Hudson.

[Dima11]   DiMare, P. C. (2011). *Movies in American history an encyclopedia*. Santa Barbara, Calif.: ABC-CLIO.

[Ott58]   Ott, J. N. (1958). *My ivory cellar; [the story of time-lapse photography].*. Chicago: Twentieth Century Press.

[Ott75]   Ott, J. N. (Director). (1975). *Exploring the spectrum* : International Film Bureau.

[Fric08]   Fricke, R. (Director). (2008). *Baraka* : MPI Home Video.

[Liec95]   Liechti, P. (Director). (1995). *Signers Koffer* : LOOK NOW.

[Sign14]   Signer, R. (n.d.). Roman SignerKunstmuseum St.Gallen/CH. *Roman Signer / Start*. Retrieved , from http://www.romansigner.ch/en

[Boli14]   Alan Kim Boling, Jr.. (n.d.). *Alan Kim Boling, Jr.*. Retrieved, from http://alanboling.com/

[Deri14]   Derivative TouchDesigner. (n.d.). *Derivative TouchDesigner*. Retrieved , from https://www.derivative.ca/

[Imag14]   Convert, Edit, and Compose Images. (n.d.). *ImageMagick: Convert, Edit, Or Compose Bitmap Images*. Retrieved , from http://www.imagemagick.org/