

DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater

PETER ALEXANDER TORPEY

Bachelor of the Arts in Media Arts
University of Arizona, 2003

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology
September 2009

© 2009 Massachusetts Institute of Technology. All rights reserved.

Author:

PETER ALEXANDER TORPEY

Program in Media Arts and Sciences
7 August 2009

Certified by:

TOD MACHOVER
Professor of Music and Media
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by:

DEB ROY
Chair, Academic Program in Media Arts and Sciences

DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater

PETER ALEXANDER TORPEY

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on 7 August 2009, in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology

ABSTRACT

Early in Tod Machover's opera *Death and the Powers*, the main character, Simon Powers, is subsumed into a technological environment of his own creation. The theatrical set comes alive in the form of robotic, visual, and sonic elements that allow the actor to extend his range and influence across the stage in unique and dynamic ways. The environment must compellingly assume the behavior and expression of the absent Simon.

This thesis presents a new approach called Disembodied Performance that adapts ideas from affective psychology, cognitive science, and the theatrical tradition to create a framework for thinking about the translation of stage presence. An implementation of a system informed by this methodology is demonstrated. In order to distill the essence of this character, we recover performance parameters in real-time from physiological sensors, voice, and vision systems. This system allows the offstage actor to express emotion and interact with others onstage. The Disembodied Performance approach takes a new direction in augmented performance by employing a non-representational abstraction of a human presence that fully translates a character into an environment. The technique and theory presented also have broad-reaching applications outside of theater for personal expression, telepresence, and storytelling.

Thesis Supervisor: TOD MACHOVER

Professor of Music and Media
Program in Media Arts and Sciences
Massachusetts Institute of Technology

DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater

PETER ALEXANDER TORPEY

The following person served as a reader for this thesis:

Thesis Reader

DAVID SMALL

Associate Professor of Media Arts and Sciences

Program in Media Arts and Sciences

Massachusetts Institute of Technology

DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater

PETER ALEXANDER TORPEY

The following person served as a reader for this thesis:

Thesis Reader

WHITMAN A. RICHARDS

Professor

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology

DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater

PETER ALEXANDER TORPEY

The following person served as a reader for this thesis:

Thesis Reader

ALEX McDOWELL, RDI

Production Designer
DreamWorks Animation SKG, Inc.

ACKNOWLEDGMENTS

- TOD MACHOVER As my advisor for the past two years, you have provided me with numerous inspirational ideas, opportunities, and experiences for which I am greatly indebted. I very much look forward to our continued collaboration and learning from you.
- DAVID SMALL I wish to thank my thesis readers for their support. Your inputs and experience have been most valuable while formulating the approaches and ideas presented in this thesis. I have benefitted from your wisdom and surely will continue to do so.
- WHITMAN RICHARDS
- ALEX McDOWELL
- DIANE PAULUS Thank you for taking the time to participate in this research or for offering essential guidance. I look forward to working with each of you in the future, be it for *Powers* or other projects.
- JAMES MADDALENA
- CYNTHIA BREAZEAL
- ROSALIND PICARD
- MARK & NAOMI TORPEY Throughout my life, you've always been there to support me in innumerable ways. From you, I've learned the most important lessons in life, especially that I should never settle for anything less than my dreams.
- ELEANOR OLVEY Thank you, my dear longtime friend. You've always been there when I needed you and have always managed to encourage me, in spite of myself.
- ELENA JESSOP To my friend, colleague, and collaborator, I offer my thanks for your ideas, your contributions, and your company.
- NOAH FEEHAN I am grateful to my Opera of the Future colleagues for their kind friendship, assistance, and for the way in which they have challenged me explicitly or by example to better my work and myself.
- ANITA LILLIE
- ANDY CAVATORTA
- NINA YOUNG
- WEI DONG
- ADAM BOULANGER
- CRAIG LEWISTON

BOB HSIUNG
DEBBIE GIFFORD
ARIANE MARTINS

TAYA LEARY
KIRSTEN BROOKS
KRISTIN GALLAS
JIHYON KIM

BEN BLOOMBERG
MICHAEL MILLER
SIMONE OVSEY
JASON KU

OTHER INTREPID UROPS

DEBORAH FREY
WAYNE MONTIETH
JEFF SIMPSON
JUSTINE WEAVER
ROBERT BLACK
SUZANNE BRAYER
STUART REGES
JEFF IMIG
JANICE DEWEY

To those who have worked so tirelessly producing *Death and the Powers* and to the Opera of the Future administrative assistants, you have made much possible and life a little easier. You have my gratitude.

To date, 30 undergraduates have worked for Opera of the Future on *Death and the Powers*. Their intellect, expertise, humor, and labor have been invaluable to the opera and several have contributed directly to the work presented in this thesis, so I heartily salute my Intrepid UROPs.

I believe that who I am is the product of the experiences and people I've met throughout my life. With that in mind, I recognize these educators and staff from elementary school through college who have had a significant impact on the way I think and view the world in a way that is essential to the research I pursue in this thesis and beyond.

This thesis is dedicated to

my grandfather

MEYER LEVIN
1911–2008

I have learned much from Meyer. He taught me
to take nothing for granted,
to build and create,
and to always use the right tool for the job...
even if I have to invent it myself.

CONTENTS

Acknowledgments	11
Contents	15
List of Figures.....	17
1 Introduction.....	21
1.1 Motivation.....	22
1.1.1 <i>Death and the Powers</i>	22
1.1.2 Disembodiment	23
1.2 Framework and Objectives.....	24
1.3 A Caveat or Two	25
1.4 A Note on Terminology	26
1.5 Thesis Structure.....	26
2 Background.....	29
2.1 Augmented Performance.....	29
2.1.1 Media on the Stage.....	30
2.1.2 Dance and Visualization.....	32
2.1.3 Hyperinstruments	34
2.1.4 Color Music	36
2.1.5 Tele-Installation	37
2.2 Synaesthesia.....	38
2.3 Modeling Affect.....	41
2.4 Inference.....	43
2.4.1 Representation Mapping through Reified Inference.....	46
2.4.2 An Example of Reified Inference in Application	48
2.5 Mapping.....	50
2.6 Distinction from Related Work.....	52
3 Death and the Powers.....	55
3.1 Genesis.....	56
3.2 Synopsis.....	57
3.3 Production Design.....	60
3.3.1 Sound in Space.....	63

3.3.2	The Chandelier	64
3.3.3	Operabots	66
3.3.4	Walls	69
3.3.5	Conceptualizing The System	79
3.4	Control Systems	80
3.4.1	Operabot Control	84
3.4.2	Wall and Chandelier Control	86
4	The Disembodied Performance System	89
4.1	Pay No Attention to that Man Behind the Curtain	90
4.2	System Overview	93
4.3	Data Categories	93
4.4	Feedback	96
4.5	Performance Capture	97
4.5.1	Wearable Sensors	97
4.5.2	Audio Analysis	99
4.5.3	Computer Vision	100
4.5.4	Other Sensing Modalities	101
4.6	Character Modeling	102
4.7	Output Representation	104
4.8	Software Implementation	105
4.8.1	System Operation	106
4.8.2	Input Devices	107
4.8.3	Mappings	108
4.8.4	Cues	114
4.8.5	Output Renderers	115
4.8.6	Integrating with Other Systems	117
5	Disembodied Performances	119
5.1	Proof of Concept	119
5.2	Test Performance Capture Sessions	121
5.3	Recorded Data	124
5.4	Output Representations	129
6	Conclusions and Discussion	135
6.1	The Disembodiment Problem	136
6.2	This Side of the Uncanny Valley	138
6.3	Next Steps	139
6.4	Future Directions	141
6.4.1	Beyond <i>Death and the Powers</i>	141
6.4.2	Novel Applications	144
6.4.3	Representation Mapping	146
6.4.4	Other Applications of Representation Mapping	148
	Bibliography	151

LIST OF FIGURES

Figure 1: Josef Svoboda's <i>Polyecran</i>	31
Figure 2: Loïe Fuller	32
Figure 3: Ying Quartet performs with Soundsieve Live	34
Figure 4: Yo-Yo Ma playing the hypercello.....	35
Figure 5: Image from <i>VIDEOPLACE</i>	37
Figure 6: Bubelle	41
Figure 7: Sentograms of essentic forms	42
Figure 8: Spectrum of affective time courses	43
Figure 9: Kismet in a happy state.....	44
Figure 10: A joyous face.....	44
Figure 11: Typical mapping process	46
Figure 12: Reified inference.....	46
Figure 13: Screenshot of example inference application	49
Figure 14: Output color sets and chords	50
Figure 15: May 2008 production retreat	56
Figure 16: Conceptual sketches	60
Figure 17: <i>MATRIX II</i> , Erwin Redl	62
Figure 18: Matrix-of-light.....	62
Figure 19: Actuated hexagonal floor tiles	62
Figure 20: <i>Bird</i> , Constantin Brancusi.....	64
Figure 21: <i>Linear Construction No. 2</i> , Naum Gabo.....	64
Figure 22: Quarter-scale Chandelier model.....	65
Figure 23: Early Chandelier model	65
Figure 24: <i>Twenty-Five Spaces</i> , Rachel Whiteread	66
Figure 25: <i>Embankment</i> , Rachel Whiteread	66
Figure 26: Operabots at the start of the Prologue.....	67
Figure 27: Early cube-shaped Operabots.....	67
Figure 28: Cube Operabot prototypes	68
Figure 29: Operabot conceptual renderings	68
Figure 30: Operabot sketch	69
Figure 31: Final Operabot prototype	69
Figure 32: Image from <i>Deathly Still</i> , Dirk Reinartz	70

Figure 33: <i>Nameless Library</i> , Rachel Whiteread.....	70
Figure 34: Freestanding walls.....	70
Figure 35: Walls flown from a circular truss.....	71
Figure 36: Walls as periaktoi.....	71
Figure 37: Wall scale study	72
Figure 38: <i>Cold Dark Matter: An Exploded View</i> , Cornelia Parker.....	72
Figure 39: Book actuation	73
Figure 40: Blinking books	74
Figure 41: Pixel-per-book imagery.....	74
Figure 42: Wall actuation studies.....	75
Figure 43: Wall projection prototype.....	76
Figure 44: Wall projection test.....	76
Figure 45: Wall speaker cluster	77
Figure 46: Wall sonic prototype control software	78
Figure 47: Wall sonic prototype being played	78
Figure 48: HAL 9000 in <i>2001: A Space Odyssey</i>	79
Figure 49: Operabot control view	86
Figure 50: Conceptual rendering of set.....	87
Figure 51: Pit box animation.....	91
Figure 52: Human silhouette.....	91
Figure 53: Simon Powers in a pit box	91
Figure 54: <i>Wooden Mirror</i> , Daniel Rozin.....	92
Figure 55: Disembodied Performance System overview diagram.....	93
Figure 56: Disembodied Performance System schematic.....	94
Figure 57: Arm gesture sensor assembly.....	98
Figure 58: Breath sensor band	98
Figure 59: Shoe sensor assembly	99
Figure 60: Affect space	103
Figure 61: Control panel	106
Figure 62: Houdini VOPs	109
Figure 63: Mapping Designer view	110
Figure 64: Data Streams view	111
Figure 65: Trajectory in affect space.....	113
Figure 66: Parameter view	113
Figure 67: Performing with sensors.....	120
Figure 68: James Maddalena wearing sensors.....	122
Figure 69: James Maddalena performs with sensors	123
Figure 70: Proof of concept data and imagery.....	124
Figure 71: Data from first test recording.....	126
Figure 72: Comparison of “joy” and “anger” motion	129
Figure 73: Hues in the Arousal-Valence plane.....	130
Figure 74: Particle renderer.....	131
Figure 75: Fluid renderer.....	131
Figure 76: Lumigraph renderer.....	131

Figure 77: Book renderer 132
Figure 78: Stage renderer 132

Figure 79: VAMP 137
Figure 80: Uncanny valley 139
Figure 81: Personal Opera 148

All figures were created by the author unless otherwise noted in captions.

I INTRODUCTION

*A wonderful harmony arises from joining together
the seemingly unconnected.*

—Heraclitus, c. 500 BC

For millennia, we have been attending the theater to immerse ourselves in a story. We spend a few hours, lost in the crises and joys of fictional lives that become somehow meaningful to our own, watching them play out before us in a stylized form. An actor enters, hits his mark in the limelight, and delivers a line. We hear the words, but recitation or song is not the whole of acting. His lit form onstage communicates along with the quality of his voice. We understand his pain or elation or indifference. We empathize and understand what the character is going through. This is the hallmark of a good performance. How does this happen? What is communicated from the actor's body to our own that makes us understand beyond the dialogue or the libretto? What are we looking at to perceive and internalize this affect? If these questions aren't challenging alone, consider what the answers might be if the character we see onstage is the stage itself.

This thesis presents a novel approach toward thinking about representing a human theatrical performance called Disembodied Performance. This approach considers the salient qualities of a remote actor's performance to generate a compelling and provocative presence onstage in a completely new and abstract form. This is a challenging task and I will draw on work from fields such as affective psychology and cognitive science to define a theoretical framework in order to realize the goal of distilling the essence of a character being portrayed by an actor and reinterpreting it through arbitrary modalities. Also presented as the product of this research is a real-time implementation of the Disembodied Performance System that has been designed to facilitate the application of this technique in stage productions, and for the upcoming production of the opera *Death and the Powers* in particular.

In addition to presenting a theoretical conception and implementation, in this thesis I attempt to document process, both conceptual and technical. Granted, a document such as this cannot begin to provide a detailed account of all the creative decisions and concerns addressed from day to day in my two years working on *Death and the Powers*, let alone the project's twelve-year history. However, process is often as important as product and there are lessons to be learned, not the least of which is why things are the way they are. To this end, I will summarize along the way my contributions to the production and the ideas, decisions, and ambitions that influenced them.

1.1 Motivation

Several factors have contributed to the form this research has taken. The most important of these are the requirements of the production of *Death and the Powers*. Of course, I bring to this work my previous experience, expertise, and research interests drawing on my background in technical theater, video and film production, visual art, music composition, and computer software. In some way, all of these have contributed to the process of creating the Disembodied Performance System, if not the final result. I also have a tendency to see disparate parts of the world as interrelated. Consequently, I take an interdisciplinary approach to achieving this artistic goal.

Additionally, for some time, I've been interested in new human-computer interface paradigms that abandon the ubiquitous stale metaphors of computing in favor of interfaces that break down the bottleneck of interaction between the human and the computer. As part of these explorations, I have considered gestural interfaces, or what I would call *intangible interfaces*. As this thesis unfolds, traces of this idea will surface, tailored to the application at hand. A discussion of future directions will reveal how Disembodied Performance lays the groundwork for furthering new types of interfaces.

1.1.1 *Death and the Powers*

The primary impetus for the development of the Disembodied Performance System is Tod Machover's upcoming opera, *Death and the Powers*. In the story, the main character, Simon Powers has invented The System. Upon his death, he enters The System and transcends his human form and becomes embodied in the environment, interacting with the other characters in both omnipotent and familiar ways. This transformation is both technological and metaphysical, invoking death as well as immortality. As part of the production, the theatrical set must come alive as the main character, giving a performance that is both human and something much greater.

I assumed the task of realizing this unique performance onstage. I began by considering how an expressive performance could be represented using theatrical technologies while taking into account the intent and constraints offered by the opera's evolving production design. It became increasingly apparent to the production team and me that simple traditional control of theatrical elements would be inadequate to drive an expressive and evocative performance. Thus, the Disembodied Performance System was born as a way to have an actual human performance directly control the onstage representation of the character.

The challenge central to the research presented in this thesis is to capture a dynamic performance and reinterpret it in the environment. It is important that the audience have a profound sense of Simon's presence after his transference into The System. Simon in this new form must be capable of emoting, communicating, and interacting with the other players onstage.

While past technologies have enabled actors and dancers in the theater to extend their influence by manipulating sound and visual elements, Disembodied Performance takes a new approach. The story of *Death and the Powers* poses a unity between Simon Powers and his new form and representation in the theatrical set. Pre-recorded animation, artificial intelligence techniques, puppetry, and interactive projections do not purely translate a character from one form into another in the manner required by *Death and the Powers*. Disembodied Performance does provide a framework for this direct mapping of an offstage performer into an onstage representation.

1.1.2 Disembodiment

What I attempt to achieve through Disembodied Performance is an abstraction of communicative expression from the medium in which it is communicated, thereby liberating a performance from its form. By distinguishing the performance from its representation, it allows the performance to be re-embodied in any manner deemed necessary by the production. Disembodiment here, then, is not to be without form, as the literal interpretation would suggest. What then would we see or hear? The goal is to be able to, in essence, give an actor a new body, one capable of communicating and interacting through sound, music, the movement of the theatrical set, and visual projections in ways that the human form onstage cannot. The new body (or bodies) may be non-anthropomorphic or, more precisely, any substance or parameter that is perceptible in any modality. All the while, it is imperative that these new representations appear as an authentic character to the audience.

One way to consider how this is possible is to look at other evocative forms of expression. Artistic representations like poetry and music have the ability to resonate deeply within us. However, after centuries of study, there

remains no sufficient explanation as to how emotion is encoded in music and how a composer, through it, elicits an intended response. Stringing together words, the poet can invoke a similar form of magic that distinguishes good poetry as having heightened expressive power when compared to prose. While we can't put our finger on what it is about these media that can touch us, we can still feel it when it happens. Composers, poets, and painters have the cunning to speak in this language that we all comprehend, but do not understand. We cannot even fully identify the words and the syntax.

I rely on the actor's mastery of his craft to compose the necessary expressive behaviors that constitute a performance. With present technology, a computer could not synthesize a performance as convincingly rich and nuanced as a professional. The actor performs the same poetics of movement that we would see onstage, but I instead translate that expressive depth into a free multisensory experience.

1.2 Framework and Objectives

I have defined several goals and guidelines that the Disembodied Performance System must meet in order to be an effective tool for realizing the abstraction of a performance.

- The system must be capable of interpreting an actor's live performance and generating alternate representations in real-time.
- The output of the system must be expressive. It must retain the immediacy of the actor's behavior to provide a compelling presence onstage. The output will not rely on the image of the performer or the performer's presence before the audience to be expressive.
- The output will allow for non-anthropomorphic rendering of presence across many media including graphics, sound, and robotic motion. It will attempt to facilitate a poetics of form whereby the quality of the output representations is communicative without the need for explicit referents.
- The system's operation will be transparent to the performer. The actor will act as usual, only offstage, without needing to learn about or be skilled in deliberately manipulating the system.

- Just as with the actor himself, the system is capable of taking direction. The director can modulate the reinterpreted performance to achieve a desired appearance onstage. The system must be configurable with real-time feedback to facilitate the creative process in a traditional rehearsal context.
- The system should be user-friendly. Simple changes and standard usage should not require a specialized technician once introduced. The user should have a range of creative freedom without needing to concern herself with writing code or technical implementation details during the design process.
- The system must be robust and modular to accommodate the rigors of theatrical performance and the constant changes that occur in the development of a production.

1.3 A Caveat or Two

The work described in this thesis attempts to glean the emotional and cognitive state of a character from an actor's behavior. Disembodied Performance presents an approach to accomplish this task. However, the work described herein, was completed with an artistic purpose in mind, in the context of a theatrical production. The results are, by design, an abstractly qualitative interpretation of the actor's behaviors. This thesis does not present an absolute approach for inferring the affective state or emotional intent of a person for use as part of rigorous study, but rather for the artistic and stylized communication of these complex and intangible qualities. The mappings presented are designed and subjective. This approach may ultimately be useful in more scientific pursuits. From these results, formal techniques for generating mappings that are unambiguously grounded in cognitive theory may be attempted at a later time. For now, I will concentrate merely on the creation of an aesthetic or perceptual impression and the techniques that are sufficient to recover the expressive impact of a performance.

Additionally, it should be noted that Disembodied Performance does not create a character. Much research has been done on generating virtual characters for games, interactive experiences, and the stage. Disembodied Performance does not constitute a virtual character. There is neither artificial intelligence nor a deterministic behavior tree to independently represent an agent. What I am exploring is a method for translating the representation of a character, not the synthesis of the character's behavior or responses at the level of intent. In this case, the human actor does that.

However, the representation techniques could theoretically be applied to virtual character systems just as it is applied to a human portrayal of a character in this work.

1.4 A Note on Terminology

Throughout this thesis, I will use four similar terms: model, mode, modal, and modality. Of course, these terms are etymologically related and often can be used interchangeably or within a single context. However, in quite a few specific domains, such as computer science or physiology, these terms have more nuanced definitions and their usage may quickly become confusing or ambiguous. For clarity in this document, I will use the terms with the following particular definitions.

I will use *model* to refer to a structure that semantically represents or approximates an observed phenomenon. This usage is consistent with the statistical notion of modeling data. In discussion of the technique of reified inference that I propose, I will use the terms *model* and *intermediate representation* interchangeably. By the adjective *modal*, I shall mean something that is indicative of a property or unique feature in a particular data set or system. This invokes the concept of *modes* of a system found in physics or significant peaks in statistical distribution. I will typically reserve the adjectival form for such references, for example the term *modal regularity*. For the computer science and user interface definition of exclusive states of operation, I will use terms such as *state* or *configuration*. The noun *modality* or the adjectives *multimodal* and *crossmodal* will refer to types of sensation in the physiological sense or sensory media, as commonly referred to in the context of human-computer interaction.

1.5 Thesis Structure

This thesis comprises six chapters, including this introduction.

Chapter 2 provides an overview of related work in three relevant fields. Some examples of the use of technology and imagery to augment performance in theatrical contexts are given. Many of these examples have served as inspiration for the work at hand while others are presented to contrast with Disembodied Performance as I outline in this thesis. The contrasts are useful in demonstrating the advantages and novelty of the selected approach. I will also explain some work related to modeling affect and perception that I adapt to form the theoretical approach that is the underpinning of Disembodied Performance.

In Chapter 3, I discuss the opera *Death and the Powers*. I give a synopsis of the story, as well as a reflection on the theatrical production design. While the ideas and implementations presented in this thesis have relevance beyond *Death and the Powers*, the design of the Disembodied Performance System and the underlying approach were heavily influenced by the story and technical requirements of this production. I cover my involvement with the design and engineering of the production in preparation for a discussion of how Disembodied Performance is incorporated in it.

Chapter 4 begins by illustrating the theoretical methodology I have laid out as Disembodied Performance. I then proceed to document the Disembodied Performance System, an implementation of this approach for *Death and the Powers*.

Chapter 5 reviews the results of using the implementation described in Chapter 4 during preparations for *Death and the Powers*. I show performance data that was collected using the system and discuss the observed features. Example mappings and output representations using these data are given.

The final chapter discusses some of the theoretical implications of the approaches taken and enumerates possible future applications of Disembodied Performance, not just as a component of *Death and the Powers*. Example applications and directions for continued research are given that extend beyond the context of theatrical performance, along with the implications of the proposed theoretical approach.

2 BACKGROUND

We do on stage things that are supposed to happen off. Which is a kind of integrity, if you look on every exit as being an entrance somewhere else.

—Tom Stoppard, *Rosencrantz and Guildenstern are Dead*

In this chapter, I will present a brief survey of related work in the contexts of theater and performance. Along the way, I will elucidate what Disembodied Performance is by comparing and contrasting my objectives with existing practices, before moving on to a detailed explanation in subsequent chapters. I will also describe work from fields such as psychology and cognitive science that have influenced my approach and provide the theoretical basis on which Disembodied Performance relies.

2.1 Augmented Performance

Disembodied Performance falls under the broad category of *augmented performance*. Augmented performance encompasses the use of technologies to mediate some aspect of what is being performed onstage by interpreting the movement of performers. The term bears with it the connotation of *augmented reality*, suggesting that virtual or generated layers of imagery, sound, or other experience are layered on top of or incorporated within the real-world physical space.

Augmented performance has its origins with the first wearable, occasionally self-contained, electronics in stage productions as early as the 1880s. Use of electrical signals and keyboard-like controllers to affect stage lighting and theatrical effects predates this. Physiological and gestural sensors for performance would follow by the 1910s [80]. Today, technology has allowed for the elaboration of such techniques, though in many cases, the functionality is strikingly similar, as in Andrew Schneider's TwitchSet,

elementary gestural sensors that can be used to trigger media or lighting cues [76] or the performance company Troika Ranch [21].

2.1.1 Media on the Stage

Only about 400 years old, opera as an art form is relatively young compared to the 4500-year legacy of theater in general. As such, its role and substance may be considered formative and flexible [5]. It is a great nexus for many creative disciplines. Opera productions have become known to be some of the grandest and most technologically innovative forms of theater, spawning the developments of new scales of staging venues over the past three hundred years and some of the most lavish and elaborate scenic designs.

Although the use of video and projection is a technicality of the implementation of output representations for *Death and the Powers*, it is worth noting a few examples of the ever-growing body of work incorporating video elements and interactivity into stage productions. The use of video and film in theater has been, at times controversial. Critics argue that such a melding of media is a self-conscious attempt to compete with Hollywood blockbusters and television, undermining the very spontaneity and ephemerality at the heart of theater's essence. These views portending the end of theater as we know it, as it was meant to be, fail to recognize that in the over hundred-year history of media onstage, its application is quite varied and generally very different from cinema [30].

In theater, typically most uses of video, graphics, or interactive technology in general fall into one of four general categories: setting, abstract ambience, visualization, and mirroring. Often, static projections of images are used to represent a setting. An image of an environment or shape projected onto a surface adds detail to the environment, placing the action into a space or world rather explicitly, while being inexpensive and easily mutable. Projection can also serve as an extension of lighting, as in the work of designer Gilles Papain [61]. Other times, abstract graphics or artistic video is used to supplement the mood. Such content may be pre-recorded or generated through numerous means in real time. It is not uncommon for live relay from cameras located in the space or even handled by the actors to be used to show a different perspective or field of view of the action onstage.

The remarkable work of renowned scenographer Josef Svoboda has paved the way for incorporating dynamic projected imagery and motion pictures into set designs [86]. Svoboda's use of light and form is often striking. To this mix, he often adds unique visual elements, elevating projection from an intermedial role or one of setting to an active embodied element of stage design. Svoboda has used film, slides, and lasers to create the imagery in his work. He also pioneered the incorporation of projection on to moving

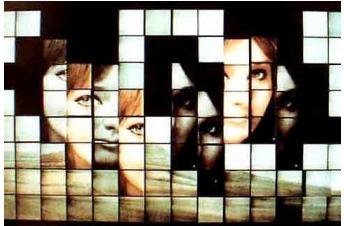


Figure 1: Josef Svoboda's *Polyecran*. The *Polyecran* rear-projected images onto the face of a wall of moving cubes. (Photo by Josef Svoboda)

screens and from inside objects or moving surfaces, as in his installation-like *Polyecran*, which consisted of a wall of cubes that moved in and out [12]. Inside each cube, two slide projectors present an image on the front face of the cube. The movement of the cubes and the slide projectors were synchronized electronically to a program recorded on a filmstrip.

Peter Sellars's 2005 production of Richard Wagner's opera *Tristan and Isolde* prominently featured projected imagery by famed video artist Bill Viola [33]. In this production, opera singers perform alongside their mute video counterparts. The imagery of the characters is typical of Viola's recent work: juxtaposing fire and water, brilliantly executed lighting, in-camera effects, and slow or reverse motion. The projections tower above the stage on tall screens, at times giving an impression of action and otherwise posing something more metaphorical.

Julie Taymor's direction of Elliot Goldenthal's opera *Grendel* became well publicized for its intricate mechanical set [32]. However, Taymor made extensive use of video projection on a variety of surfaces to underscore mood and to define abstract settings. While projection of large static photographic images has long been used to establish settings, particularly in opera productions, the imagery for *Grendel* in this role has a particularly unique graphic quality.

The 2006 operatic adaptation of Vladimir Nabokov's *Lolita* by Joshua Fineberg uses video projection and audio processing to set the entire piece within the delusional internal dialog of Humbert Humbert [44]. The actor reading the dialog is situated between the audience and the orchestra facing the stage. Live video of his face is projected onto a scrim in the center of the stage. Flanking screens show video images that metaphorically or loosely correspond to the action described in the dialog. At times, dancers can be seen through the projection scrims, appearing as if inside the character's head [8]. Audio processing allows the single actor's voice to take on singing-like qualities to signify the other characters he describes. However, the processing is not used by the actor to enhance his own performance. It is merely a layer on top of the performance over which he has little or no control.

The English National Opera production of Olga Neuwirth's adaptation of David Lynch's *Lost Highway*, directed by Diane Paulus, made extensive use of video projection to define an installation-like space in the round in which the action unfolds [45]. Much like *Lolita*, the imagery itself often served the role of symbol or metaphor, featuring loops of night driving on a road or images intended to represent a character's disturbed mental state [25]. At other times, live relay video from a character moving about the set with a camcorder is shown. This provides the audience with a close-up and immersive view of the action unfolding. Although evocative, the use of

video is decidedly objective, placing the audience voyeuristically into the story, providing otherwise impossible views, but still keeping the audience at a certain emotional distance from the action. It represents the mood surrounding the actors onstage more than it does a character.

Most recently, Robert Lepage's new production of Hector Berlioz's *La Damnation de Faust* has received much attention for its use of projection and technology, specifically of reactive imagery, which grew out of Lepage's work with Cirque du Soleil on *KÀ* [92]. In one scene, dancers in front of a projection screen are wearing sensors that manipulate the image of rippling fabric projected behind them. While technologically intriguing, such an approach begs the question of why actual naturally-rippling fabric was not used. Other imagery in *Faust* serves the common role of using metaphor to heighten the emotional content, such as flames that respond to the intensity of a singer's voice. In other scenes, the projection provides a more traditional, though at times dynamic, imagistic setting for the action, among other effects.

2.1.2 Dance and Visualization

A great many of the technological developments for the stage have their origins in dance performance. Also, many types of dance concern themselves with interpretation and representation as an artistic expression through movement. It is no surprise then that many researchers and artists are exploring techniques to incorporate more media and visualization approaches within their dance work.

In 1900, Loïe Fuller changed careers from actress to dancer and gained recognition for her abstract expressionistic routines. Fuller sought to take advantage of the possibilities electrical stage lighting could afford and made many significant advances in the technology and practice of theatrical lighting and the use of color [84]. Her performance work was equally innovative, developing new methods of dance and interpreting the colors of her lighting while wearing large, flowing silk dresses that could capture the full effect of the illumination.

Choreography is traditionally set to music as a physical expression of the rhythm and form of the music. However, artists have sought to reverse this relationship, as well. As early as 1965, Merce Cunningham and John Cage's *Variations V* incorporated photoelectric sensors and antennae to mark the positions of dancers [50]. The data gathered by these sensors and antennae then triggered and controlled electronic musical devices.

Today, the use of color and expressive lighting is essential to dance performance onstage. Combine that with performance artists' and choreographers' desire to extend the expressive range into control of lighting, visuals, and music using motion and gesture sensing and it is clear



Figure 2: Loïe Fuller
Fuller wore large, flowing, white silk gowns in her dance performances to enhance the effect of her colored stage lighting. (Photo by Frederick Glasier, 1902)

to see that the field is continually at the forefront of theatrical technology. Many recent dance works use wearable sensors or camera-based motion capture to drive interactive visualizations or the generation of sound.

For example, the 2006 work *Lucidity* used sixteen cameras and infrared reflective targets worn by three dancers to capture motion [36]. In each of the piece's three movements, the captured motion data was interpreted and represented differently. The movement of the dancers was analyzed for correlation in location and similarity of movement. These metrics were then used to control sound synthesis and 3D lines and surface sweeps that were projected onto a scrim in front of the dancers.

Mapping of output visualizations can be driven by proportional gestural movements, essentially a visual abstraction of the performer's form or movement, as is the case in Flavia Sparacino's *DanceSpace* [82]. A single dancer's movements are analyzed by computer vision and features and trajectories are displayed on a large projection screen at one side of the acting area. The visualization is a simple abstraction of the dancer's form. Once again, the captured movement is also used to create sound by attaching virtual musical instruments to the performer's limbs. Feedback from the visual representation and generated music to the dancer are encouraged.

More elaborate mappings and visualizations typically call on artificial intelligence approaches, leading to virtual character representations. Virtual characters are interactive computer-generated projections of a character into the space and are a subject of research in both theatrical performance and dance. The virtual characters need not have human- or animal-like form. Inspired by performance movements in Italian Futurism and Bauhaus theater, Sparacino introduces Media Actors, a type of interactive virtual character that can take the form of images, sounds, and displayed or spoken words.

Marc Downie, as a reaction to mapping-based approaches, theorized extensively on the role of artificial intelligence agents in the visualization and augmentation of dance theater [22]. Downie's works in dance and music performance use motion capture data, be it gestural or musical, to prompt agents. These agents, having goals and motivations of their own, then go about acting in accordance with the input stimuli. What we see are representations of the agents' performances, which may be something paralleling the human form or a complete visual abstraction.

The types of visual representations I've noted so far fall at two extremes of a continuum. In one type, the augmented space mimics, mirrors, or represents the motions of the performers purely for aesthetic reasons, having a sense of responsiveness as their only salient connection to the

performance. At the other end, new autonomous elements are projected into the space to generate additional behaviors and performances in conjunction with the live human performers. Distinct from these goals, another class of visualization is intended to explain the complexity of what they are seeing, having a form that conveys or replicates information from the performance, so as to heighten the appreciation of the experience through an enhanced understanding.

Reynolds, et al. developed a system to visualize a juggling performance by the Flying Karamazov Brothers [69]. Sonar was used to determine the location and gesture of the jugglers over time and the plan view of the stage and performers was projected on the upstage wall. A segmentation of the stage floor was also used to produce music. As another means of visualizing the complex patterns of juggling, special clubs were designed that would illuminate in different colors to reflect the function of that club and its path of travel from one performer to the next.

In 2008, I created a visualization system for real-time musical performance with colleague Anita Lillie for a performance by the talented Ying Quartet at the University of Iowa [90]. The system was dubbed Soundsieve Live due to a resemblance to Lillie's Soundsieve music visualization application, which inspired this project [43]. The Yings performed Ravel's String Quartet in F Major while real-time audio analysis was used to track pitch and performance characteristics of each part. The parts were then visually represented using shape, motion, and color and projected behind the performers in different configurations for each movement to illustrate pitch contour and articulation.

2.1.3 Hyperinstruments

Since the early days of electronic music, new types of music controllers have been the subject of much fascination. Gestural controllers are of particular interest for the performance of electronic music, but generally present a challenge for intuitive mappings that confound audience intuition and expectations that are firmly entrenched in the long tradition of the subtle and expressive performance of acoustic instruments [49]. We watch intently as Itzak Perlman feels his way through a piece of music, making the slightest of variations in motion to produce a sound that is as moving as the performance. The body language speaks to us. Indeed, there is a spectacle to such a concert. There is an apparent skill and talent and a performance that is almost as much about the instrumentalist onstage as it is about the music. The performance of electronic music struggles to find ways to capture that. For much of electronic and computer music performance, there isn't much to look at. A performer pushing buttons or turning dials or, worse yet, occasionally hitting a key on a laptop keyboard, offers little in addition to sound to the viewing audience. Eschewing old metaphors and instruments in favor of completely new methods of controlling the same

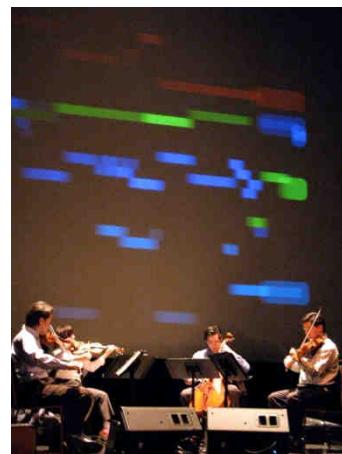


Figure 3: Ying Quartet performs with Soundsieve Live
(Photo by Tod Machover)

parameters is an appealing alternative pursued by thousands. Many of these efforts take the form of gestural interface that, unfortunately, end up providing more by way of spectacle of movement, more akin to dance, than they offer musically. The problem generally lies in the mappings from gesture or control to music, tending to control complexity or parameters in overly simple or overly intricate ways that are not particularly musical and lacking in expressive subtlety and control [37,49].



Figure 4:Yo-Yo Ma playing the

hypercello

(Photo provided by Tod
Machover)

For Tod Machover's opera *VALIS*, he wanted to have only a small number of instrumentalists be able to perform complex and rich music live during the production. The systems developed for the opera to control electronic sound sources, while preserving the musician's intention and the live feel of the performance, evolved into what has become known as hyperinstruments [46]. Early hyperinstruments, such as the hypercello constructed for Yo-Yo Ma, introduced unintrusive sensors on the instrument and, in some cases, on the performer. The information gathered from the performance was used to control digital signal processing of the instrument's sound and for the synthesis of additional layers of sound.

Hyperinstruments don't follow the troublesome path of many instruments and controllers for electronic music. Instead of bringing new instruments to electronic music, classic hyperinstruments bring electronic techniques to traditional instruments for virtuosic musicians with the specific intent of understanding the performer's expressive movement and extending their range [47].

The same technologies can be used to provide access to musical creation for anyone, not just trained musicians. A new generation of hyperinstruments allows anyone of any age, regardless of musical training, to perform and compose music expressively. These ideas have been at the core of our research group's work at the MIT Media Lab since the early 1990s.

The parallels to musical expression are, however, fitting and informative to this work. A Disembodied Performance will have a musical quality. This is more than appropriate in the operatic context, though it entails more than analogy. Some abstract notion of musical form is relevant. The expressive output will represent emotional content without explicit referents in much the way music does. In some sense, the Disembodied Performance System is a hyperinstrument not unlike those developed previously in our research group. Although not treated quite like a musical instrument, given the requirement I set forth that the process of Disembodied Performance is transparent to the performer, the objective of the proposed practice and its implementation share much in common with hyperinstruments. It is an interface intended to expand and enhance expression, leveraging the skills and attributes already mastered by the performer.

2.1.4 Color Music

Although British painter A. Wallace Rimington gave us the terms *color music* and *color organ* in 1893, the idea of a synaesthetic medium for music-like performance of color extends back to Ancient Greece [57]. The first recorded implementation is the Clavecin Oculaire by Louis-Bertrand Castel in 1734 [62]. Many subsequent devices for color music performance had a similar form, augmenting an existing keyboard instrument to control color filters over some source of illumination.

The Italian Futurists, Bauhaus, and the resonating idea of Wagner's *Gesamtkunstwerk*, which he proposed in *The Art-work of the Future* in 1849, inspired many works of synaesthetic expression during the 1910s [84]. Painter Wassily Kandinsky had been experimenting with the visual representation of musical form [95]. In the year prior to his publication of his *On the Spiritual Art*, he completed the first of his "color-tone dramas." *The Yellow Sound* was an opera without dialogue or plot, relying heavily on the *mise-en-scène* and expressive lighting. In 1915, Alexander Scriabin held the New York premiere of his, *Prometheus: Poem of Fire*, Op. 60, which was scored for orchestra and tastiera per luce, a color organ called the Chromola based on Rimington's design.

Later visual music devices produced not simply colored light, but abstract images of light. Thomas Wilfred coined the term for these projections: *lumia*. Wilfred's instrument, the Clavilux, was introduced in 1922 and used a keyboard of sliders to control internal prisms and light sources. Wilfred did not believe that there was a correspondence between visual and music, but his Clavilux performances nevertheless had musical qualities in terms of timing and structure [62].

Nearly thirty years later, animator Oskar Fischinger's Lumigraph bears some similarity to the Clavilux, in that it uses a pure abstraction of light and color in live performance [57]. Unlike Wilfred, Fischinger did sense the potential for a proper correspondence between image and music, often performing the Lumigraph to musical accompaniment. The device consisted of a frame containing the lighting elements and various color filters. The light was emitted in a thin plane a short distance from a white latex screen. The performer could use his hands or any other object to deform the screen into the plane of light to produce imagery. Unlike the other examples mentioned, Fischinger's instrument was played gesturally and was capable of extreme degrees of subtle and virtuosic expression.

In the mid 1970s, Laurie Spiegel developed the Video and Music Playing Interactive Realtime Experiment (VAMPIRE) system at Bell Telephone Laboratories [83]. The system added real-time generative graphics to Max Matthew's and Dick Moore's GROOVE music system. Numerous input devices to the computers—from an organ keyboard to a joystick—

controlled the many parameters available to the user. The result was a system capable of generating music and color music by procedural means in conjunction with gestural input.

Instruments for color music are devices that address the desire to create expressive real-time imagery. As I'll discuss below, the links among color, music, and expression are important for the work presented in this thesis. However, Disembodied Performance does not employ such a direct manipulation of imagery by the performer as we've seen in all of these examples. Instead, the proposed system takes a more structured approach to defining the mappings from intent to output by using an abstracted intermediate representation.

2.1.5 Tele-Installation

Perhaps the most influential work that has informed the architecture and the spirit of the Disembodied Performance System is that of Myron Krueger. His research expanded on the groundwork laid by Ivan Sutherland and others in the 1960s that led to what would become known as virtual reality. He observed that traditional human-computer interactions were symbolic, not perceptual [38].

Krueger was strongly motivated to liberate interaction with the virtual realm from the accoutrements of human-computer interaction devices and displays [70]. The result is what Krueger termed a "responsive environment" that could be an extension of the user. Creating spaces that were sensing and aware allowed for notable installations as *METAPLAY* and *VIDEOPLACE*. In these pieces, interactions among different spaces were made possible by transmitting the abstracted camera images and visual output (video projection) from one space to the other, a form of remote presence. Initially, out of view of the participants, Krueger would remotely manipulate the computer imagery generated and displayed. Users began to find affordances for communicating with him, though they were not aware of his role, through the system. Krueger was acting in the role of an intelligent and responsive environment. Later, these spaces were imbued with enough computational power to facilitate interactions by participants with the systems and each other, but their true expressive nature was an emergent property of the interaction, of the human experience.



Figure 5: Image from *VIDEOPLACE*. Myron Krueger's *VIDEOPLACE* used computer image processing and logic to mediate remote human interactions. (Image from [38])

Much like Krueger's works, for the Disembodied Performance System, I employ sensing technologies that provide ample information about gesture and interaction without requiring explicit physical interaction from the performers. The system responds and transmits the presence from one space (the actor offstage) to another (the walls and environment onstage). While the emphasis here is not one creating responsive or sentient environments, as was the case for Krueger's explorations, the ideas he explored are still of immediate relevance. Such reactive uses of video are increasingly

commonplace in installations like those of Zachary Booth Simpson, which caught media attention in the early 2000s. Though influenced by Krueger's work, these sorts of interactive experiences become more about the reaction of the computer to the participant than participants interacting with each other in a virtual or augmented space. The Disembodied Performance System is not a reactive system in the sense that its intent is to process the captured performance by a set of rules, but rather to provide a means for Simon Powers to be expressively realized as the environment and thereby interact with the actors onstage and, albeit more passively, the audience. The experience is directed outward, not reflected back toward the participant/actor.

2.2 Synaesthesia

Synaesthesia provides an important lens through which we can attempt to understand the role of representation in communication, particularly in the communication of abstract ideas or ideas that are difficult to quantify, articulate, or that are inherently intangible. Clinically, synaesthesia is considered a pathology, or at least an abnormality, of cognitive and neurological processing where a stimulus presented to the sensory pathway for one modality activates a second sensory pathway for a different modality [64]. The pairings of modalities are not necessarily reciprocal. We may not all be clinical synaesthetes, however, it is clear that synaesthetic experiences are common to all of us at some level of reasoning. Our cultural artifacts, languages, and means of expression are replete with multisensory forms and crossmodal representations that have the capacity for broadly understood meanings. This 'loose' view of synaesthesia refers to supposedly voluntary mappings and might be considered an intellectual synaesthesia rather than a true condition [84]. However, the use of such multimodal representations to communicate ideas and their often 'intuitive' application, as in design, may suggest some as yet not fully-understood cognitive basis.

Synaesthetic representations occur in many forms. Many types of linguistic metaphor may constitute synaesthesia and, if the usage of *metaphor* is allowed to extend beyond discreet symbolic representations, the two terms may be considered in some ways synonymous. Inter-sensory catachresis is very commonly used in evocative descriptions: a bright child, a frigid glance, a soft shade of green, a biting wind, an abrasive personality, to have the blues. None of these words means, by definition, the qualities that they convey in such instances. Yet, we have a clear understanding of the intended meaning—perhaps a clearer understanding. Their usage is not without some intelligible structure [39]. We use such constructions all of the time to create shades of meaning that don't exist in the palette of language itself. Language and symbolic systems in general are mappings of an abstract signifier to some concept, the signified. This mapping can itself

be considered as a parallel for synaesthesia. Correspondingly, it has been shown that the angular gyrus is the center for both metaphor and most types of true synaesthesia in the brain, providing a neurological basis for these experiences, leading to a plausible explanation for the origin of language [66].

Poetry often uses metaphor to great effect in creating an emotion in the reader. Music acts similarly in its ability to convey, very exactly, feelings. However, in the case of music, unlike language, the substance with which a composer scores his meaning has no ascribed referents at all [41]. Visual artists have long honed the techniques of generating meaningful swaths of color, shape, and texture that can evoke in the viewer a poignant response in images that are both representational, in that they depict actual objects or scenes, and those that are not. Designers use the same techniques to explicitly create a desired response in the viewer within a variety of media. There are heuristics and theories about the use of colors and typefaces or qualities of lighting, but the knowledge of the response that should be created and the execution of the design that succeeds in creating it are often processes that cannot be well-defined or articulated. An interpretive dancer knows how to move to elicit a response, but does not necessarily know how he knew which movements to perform. In its genesis, the movement is spontaneous.

For created movement, image, or sound to be communicative, it must evoke a response—particularly an emotional response—in the perceiver. A synaesthetic view of processing the stimuli suggests that this is possible. Perhaps there is indeed a cognitive and neurological basis for synaesthetic expression. Most forms of synaesthesia have been shown to be sensory rather than cognitive phenomena, though they can be influenced by higher-level cognitive reasoning [66]. In order for these representations in these media to communicate—for the creator of such representations to translate his message into an abstract medium and have that message received by the perceiver of his work—these sensory pathways must be linked with the emotions they are capable of communicating, and perhaps each other, at a very low level in the brain. The aesthetic experience relates perceptual stimuli to emotional responses. Support for the importance of synaesthesia in generating emotional responses was given by Richard Cytowic, who indicates that synaesthesia emerges from the subcortical limbic system, also responsible for emotion [19]. Thus, the connections from perception to emotion may not be a conscious phenomenon.

In the case of movement, mirror neurons explain a high-level neurological response to observed actions. The percept of an action invokes in the perceiver activation patterns identical to those that would occur if the perceiver were performing the action herself. At a cognitive level, this sort

of internalization of perceived concepts is accomplished by proxy agents in the model of anigrafs that Whitman Richards details [71].

Mirror neurons alone cannot explain the communication of emotion in modalities other than movement, however. Perhaps the phenomenon holds for representations that have some quality in common with movement, but are not themselves movement. Consider perhaps a drawn line or a melodic or rhythmic contour in a piece of music. Manfred Clynes proposes an answer. I'll elaborate on essentic forms in the next section, contours that are apparent in human behaviors that are unique for different emotional states. It is worth mentioning here, however, that Clynes has identified in visual art and music appropriate uses of the same essentic forms, suggesting an innate biological basis for the gestural communication of emotion [17,16].

What of communicative properties such as color? It is widely known that color can induce physiological responses and it is also perceived as being closely related to emotion [91]. Recent research by Mark Changizi suggests that color vision evolved in humans and other primates for the purposes of discerning subtle changes in the skin tone of another individual due to variations in blood oxygenation that specifically indicate changes in emotion [14].

In a recent experiment, subjects were asked to complete two tasks involving pairing colors with sketches of faces intended to represent six basic emotions. In the first task, only one color was associated with each face, whereas a set of three colors were requested in the second task. The results show a high consensus in both tasks for which colors and qualities of sets of colors are associated with the six basic emotions [20].

It is quite apparent that the conflation of color, music, and emotion is of significance. This particular trio has been the subject of artistic exploration for quite some time resulting in forms such as color music and lumia, as described in the previous section. Johann Wolfgang von Goethe introduced the relationship between color and music in his *Theory of Colours*:

Before we proceed to the moral associations of colour, and the æsthetic influences arising from them, we have here to say a few words on its relation to melody. That a certain relation exists between the two, has been always felt; this is proved by the frequent comparisons we meet with, sometimes as passing allusions, sometimes as circumstantial parallels. The error which writers have fallen into in trying to establish this analogy we would thus define:

Colour and sound do not admit of being directly compared together in any way, but both are referable to a higher formula, both are derivable, although each for itself, from this higher law. They are like

two rivers which have their source in one and the same mountain, but subsequently pursue their way under totally different conditions in two totally different regions, so that throughout the whole course of both no two points can be compared. Both are general, elementary effects acting according to the general law of separation and tendency to union, of undulation and oscillation, yet acting thus in wholly different provinces, in different modes, on different elementary mediums, for different senses.

In this writing, we see Goethe allude to a synaesthetic interpretation of the two media. Additionally, though he is referring mostly to the physical production of the phenomena, his explanation of a common origin parallels the premise I shall posit in Section 2.4 below.



Figure 6: Bubelle
(Photographs from [77])

The work cited throughout this section suggests that we've evolved to communicate our emotional state without aid. However, as I do with Disembodied Performance, technologists and artists alike see the potential of being able to leverage this congruency for expressive purposes to further enhance that communication. Dynamic expression of a person's emotion has evolved with recent technologies from the mood ring to pieces such as Philips Design's Bubelle, part of the larger SKIN probe project to find applications for emotional sensing. The Bubelle, or "blushing dress," changes color and patterns in response to the sensed emotion of the wearer [77].

2.3 Modeling Affect

As an actor moves about and gesticulates onstage, such behaviors give the audience clues as to what the character that actor is portraying is thinking and feeling. The quality of the motion, the voice, the interaction with other characters onstage all sign the affective state of the character from which the observer must infer the other's affective state [64]. This sort of sentic modulation is involuntary in normal interactions and is to some large extent involuntary, or at least intuitive to a degree, in acting. The performer may not always be aware of the reasons for making a particular movement or aware that such movement has been made, even if he is consciously aware of the character's personality, emotional state, and objective. It is this emotional expression that I seek to preserve in translating a performance from an offstage actor to an onstage representation.

The study of the expressive capacity of gesture took form in the late 1800s. François Delsarte catalogued numerous gestures that repeated with consistency in everyday life. From his observations, he asserted that the gestures are generated in accordance with the psychological state [82]. Delsarte's notion went on to influence schools of dance and theater.

Manfred Clynes also suggested that sentic modulation is essential to communicate emotion with others through the phenomenon of sentic equivalence [16]. He describes the Equivalence Principle as:

The sentic state may be expressed by a variety of motor modes: gestures, tone of voice, facial expression, a dance step, musical phrase, etc. In each mode the emotional character is expressed by a specific, subtle modulation of the motor action involved which corresponds precisely to the demands of the sentic state.

Disembodied Performance relies on this concept, but does not specifically suggest that the elements of communication are Clynes's essentic forms.

Some means of quantifying emotional expression is required in order to examine and record affective state. Philosophers, psychologists, and neurologists have long theorized models for representing and classifying emotion. One class of models hypothesizes some number of basic emotions that are atomic, discreet states. Complex variations in emotion are experienced by simultaneous basic states being blended by physiological, cognitive, and social factors [64]. Although compound states exist, contradictory states cannot be simultaneously expressed.

Clynes revealed that certain forms of emotional expression have well-defined universal qualities for a set of basic emotion states [16]. He developed an instrument called a sentograph to measure the quality of button presses by subjects in response to spoken prompts that were the names of seven basic emotions. Repeated trials from subjects from a broad range of cultures and backgrounds demonstrated commonalities in the responses to each emotional prompt. The transient shapes of the button presses, normalized for time, reveal unique contours, or essentic forms, for each affective concept (Figure 7).

Several psychologists posit that emotions should be viewed not as discreet states, but rather as occupying some portion of a continuous space and defined in terms of basis dimensions. In 1980, James A. Russell proposed a circumplex model in two dimensions on orthogonal axes [74]. He placed common affective concepts along the unit circle in the plane. Russell was primarily concerned with the relative position of each concept, measured by an angle about the origin, derived from experimental data. The orthogonal bases for this space were labeled at unit extents *pleasure* and *misery*, from positive to negative on one axis, and *arousal* and *sleepiness*, from positive to negative on the other. Today, the circumplex model is a common representation and we typically refer to the former axis as *valence* and the latter axis as *arousal*. The advantage of using such a metric space to define emotion is that it accounts for variability in magnitude of expression and

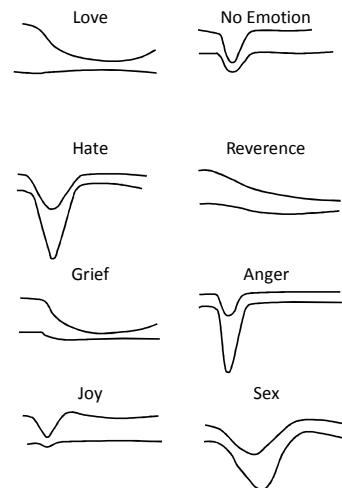


Figure 7: Sentograms of essentic forms

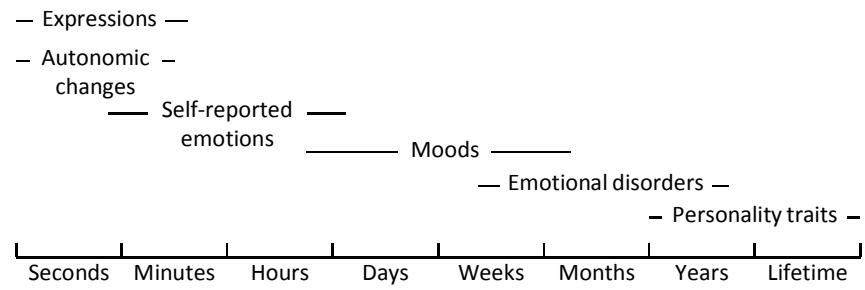
Manfred Clynes observed these universal essentic forms with a sentograph in responses to the emotion prompts. For each, the top contour is the vertical displacement of the button and the lower contour is the horizontal displacement. Each trace is about 2 seconds in duration. (Redrawn from [16])

can suggest expressive features that different emotional states may have in common.

The dimensionality of emotion spaces is generally considered to be between two and five dimensions, most commonly two or three [81]. In her work to create sociable robots that can understand human emotion and themselves emote in a communicative way, Cynthia Breazeal adapted Russell's circumplex model, adding an additional orthogonal dimension of *stance*. Stance is defined as the extent to which the individual approaches or is engaged with a stimulus. Breazeal used this affective model as part of a sophisticated emotion system for the humanoid robot Kismet to model the robot's response during the course of human interactions. The model was used to drive a range of parameters for facial expression, allowing Kismet to emote as part of its interactions [11].

Another aspect of affective modeling to consider is the time scales of different phenomena. Oatley, Keltner, and Jenkins define several of these [59]. We are interested in two of these: *expressions* and *moods*. The term emotion is commonly associated with short-term affective perception and is manifest through expressions lasting from fractions of a second to several minutes. This is the phenomenon that we wish to model for Disembodied Performance. At a longer time scale are moods. Moods may have the same qualities as some emotions, but persist from hours to months and can modulate the short-term response to stimuli. As I'll explain in Chapter 4, moods will not be gleaned directly from the performer in the way that short time scale responses are. The representation of moods will be designed to fit the emotional tenor of a scene or part of a scene. It is also worth noting that *personality traits*, the characteristics that persist over the longest period of time are expected to be emergent in the short-term behaviors and dialogue, particularly apparent in the comparison of the human performance onstage to the disembodied performance that is afforded in *Death and the Powers*.

Figure 8: Spectrum of affective time courses
(Redrawn from [59])



2.4 Inference

Our capacities for synesthesia give us the ability to understand affective expression. However, the challenge faced in Disembodied Performance is to translate from the expressive stimuli in one medium to another. I believe

the answer to bridging this divide lies in research on perception. Indeed, the problem Disembodied Performance proposes is one of perception: how can we perceive from arbitrary forms and modalities of representation an expressive performance? We're interested in both sides of the equation, though. What do we normally perceive from a performer and how can we preserve its underlying meaning when observed in other forms?

Considering the image of the face in Figure 9, I think it is safe to say that most people would assume that this person is joyous. How do we know this? In part, we learn that certain shapes of facial features (actually states of the facial muscles) are used to express an affective state that the expresser is experiencing. However, we also have low-level cognitive mechanisms for recognizing these emotive representations [18]. We can infer from our perception of the face how it feels. This is not tied to the human face. If we look at a humanoid robot (Figure 10) we are capable of making the same inference. Again, it could be a recognition of the physical shape of salient features, which the robot face preserves, but it transcends anything that looks human in style of features and materials. Looking at these representations, though, it is difficult to describe exactly what the configuration of these features are that unambiguously allow us to infer an affective state. They are not ritualistic gestures, *per se*, as would be a genuflection or applause, which are symbols to which we have ascribed a certain meaning. We could diligently describe the contour of the lips or the distance of the eyebrows with respect to the eyes, but any level of verbal description would not convey the affective state, let alone the subtle combinations of affective states that are possible. The description of the features and even the features themselves don't reveal what we know them to mean. In the case of artistic abstract representations, no analogy between the facial features or the painting can be drawn, yet we understand the affective content of both. In the case of art, the features cannot possibly be learned, for every piece of art, as a whole, is unique.

The power of emotional and literal representation in music is familiar. Music can be marvelously emotionally evocative. It also has the capability to communicate imagery. Leitmotif is a musical phrase or passage that is used in association with a character, generally found in longer works, such as opera and more recently film. The practice of leitmotif defines a musical symbol for the character that we learn and then interpret over the course of the work. The evocation of imagery though is not limited to explicitly defined symbols. We can understand a story or scene in a musical work without having to first establish an operating lexicon. Works such as the tone poems of Franz Liszt, Edvard Grieg's numerous pieces of incidental program music, and Modest Mussorgsky's *Pictures at an Exhibition* paint vivid mental pictures. Listeners can interpret the story or come up with their own images from the music alone. Ludwig van Beethoven's infamous Piano Sonata No. 14 in C-sharp minor, Op. 27, No. 2, was not written as a



Figure 9: A joyous face
(Photograph from [68])

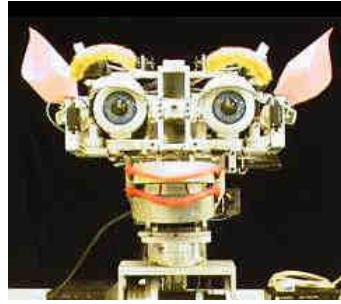


Figure 10: Kismet in a happy state
(Photograph from [11])

tone poem, but as funerary music. It became dubbed “Moonlight Sonata” after German music critic Heinrich Friedrich Ludwig Rellstab described the piece’s delicate quality in terms of the moonlight reflecting off of Lake Lucerne. Yet, this is a quality and an image we all see in the music, hence the moniker has stuck for over 200 years.

My argument in these examples is that evocative representations have intrinsic meanings that reflect the represented. There are complex perceptual and cognitive processes that account for our understanding of the underlying intent, though they are not necessarily conscious processes. Despite these hidden levels of processing, we can understand something relatively unambiguous without knowing how we understand it or the manner in which the features of a medium convey the message that we receive. We know what these representations mean without knowing how we know or what about them exactly tells us.

In visual perception, our brain tries to make judgments about the image of the world that our eyes see. Extensive research has been conducted on how we can translate from images to concepts and enable us to describe what we see. However, Witkin and Tenenbaum made an important contribution to research in perception by recasting perception from a problem of description to one of explanation [94]. They claim that the goal of perception is to explain the features and properties of an image in concepts, not merely identifying what constitutes the image, but why the image and its features are arranged the way they are. The consequence of this approach is that it assumes that the properties we notice in an image are the deliberate, causal result of some generating process that explains how the image was formed. The image is a representation of the generating process.

In such representations, key features are those that are meaningful in the sense that they allow an accurate perceptual inference to be made. The world is not random. The natural and man-made worlds are both highly structured and, in that structure, the properties of the world cluster accordingly [72]. These special non-accidental cases occur in greater probability than chance and thus form modes of the distribution of possible configurations of properties [73]. In representations of the world, these modal regularities occur in the key features that allow us to make an inference about the generating process.

A representation, particularly a cognitive representation of an idea or a physical stimulus or state, does not inherently have the properties of the original. Donald Norman suggests that good representations present only the features of importance and ignore unnecessary details in such a manner as to facilitate the ability to perceive regularities and reason about the represented world [58].

The premise here is that we can extract information from a representation. We want to leverage the brain's ability to find and interpret regularities in affective representations. Aside from the text of the play that reveals its story, it is reasonable to assert that emotion is the currency of performance. What we will see onstage, for example the projections of color shapes, is a representation of the character and the regularities that the audience sees in those representations can contain the perceptual information. We experience this when we watch the physical actor onstage and, as we always do, are able to infer the character's emotional state [64]. At some level, there needs to be a correlation in the output representation that corresponds to the input representation.

In the case of Disembodied Performance, we adapt these inference techniques from the realm of visual perception to that of emotional perception. The typical approach to this task would be to consider the actor's performance as the generating process for which we create output representations. As a step in this process, we'll need to model the input from the actor, so that we can map it to multiple output representations. However, if we consider the behavior that the audience can see when the actor is onstage to be a representation of the character's emotional and cognitive state, then the physical manifestations are themselves representing, not what is to be represented. The character's cognitive state is the generating process for the gestural representation. Once we frame the problem of Disembodied Performance in these terms, we can see how inference may be applied. We have an input representation of the generating process and we want to model the generating process using an intermediate representation, so that we can synthesize new representations from that process. This provides a new, structured approach to interpreting and mapping performance using the method described below.

2.4.1 Representation Mapping through Reified Inference

To map one representation into another, we need to know something about the salient properties the original representation conveys. We assume that the meaning of the original representation is perceived through regularities in its key features. That is to say that the parameters of the representation space have a high codimension where important features exist. We want to preserve the perception of such regularities in the newly generated representations. Thus, we can consider the original representation as providing the input parameters for an overdetermined system that can be modeled by a lower-dimensional intermediate representation. The values of this intermediate representation can then be used to generate additional output representations (Figure 11).

Another way of looking at this method of representation mapping is that of trying to recover a model of the generating process of all possible representations, including the input representation itself (Figure 12). If we

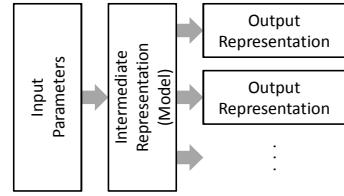


Figure 11:Typical mapping process
In a typical mapping process, input parameters directly drives output parameters via some intermediate representation that performs a transform.

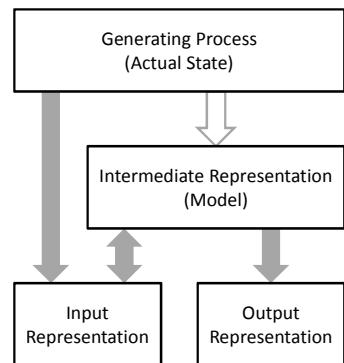


Figure 12: Reified inference
Reified inference treats the input representation similarly to the output representation: as a representation of the generating process. An intermediate representation serves as an approximate model of the generating process.

assume that the non-accidental properties of the input representation are the result of a generating function that created the input representation, then we use that to infer an intermediate representation that is informed by (or could itself have been generated by) the generating process. The goal is that the output representations may be derived from the intermediate representation to have the same perceptual effect as if they had been created directly by the generating process. I will call this abductive process *reified inference*. The mappings to and from the representational space to the model space should preserve the perception of the model. The perceiver should be able to infer the model being represented from the new representation just as he would be able to from the original representation.

One method of verifying an intermediate representation given an input representation in a suitable medium would be to recover the input representation from the intermediate representation. Given the difference in dimensionality of the input representation and the intermediate representation, some variability may be introduced in the output representations, just as additional variation and features may be present in the input representation, though not essential to the perceived effect and therefore not modeled.

Creating the intermediate representation model from the input representation is a task of dimensionality reduction. The choice of the model in reified inference is necessarily of lower dimension than the input representation and the model must retain sufficient dimensionality to capture the information we want to know and communicate from the generating process.

For analogy, let's consider Schenkerian analysis in music, which seeks to arrive at the *ursatz*, or fundamental structure, of a piece of tonal music. The piece itself is a prolongation of the *ursatz* or a series of elaborations on it. The dimensionality of the piece of music is analyzed and reduced to its fundamental form. However, since the *ursatz* is prescribed and not parametric, the intermediate representation of intent in reified inference more likely corresponds to a slightly higher level in the Schenkerian analysis. The idea here is to find the lower-order form of meaning and structure from an input representation in as few dimensions as possible and use that to model the generating process.

The mapping from input representation to intermediate representation extracts the key features of the input representation, from which the model state is inferred. We want to strip the high-dimensional representation of its elaboration and noise in order to determine the model state. It is akin to looking at the low-frequency component of a signal or image to get a sense of its essential form without unnecessary detail. The model is an abstraction and generalization of the input and output representations.

Expressive gesture demonstrates essentic forms [15]. Thus, for the purposes of Disembodied Performance, these essentic forms present possible regularities to be used in the inference process.

Like many methods of dimensionality reduction, a consequence of both the general theory of reified inference and the lower-dimensional model means that the transfer of representation is not invertible, i.e. the representation process is not lossless. For example, in principal components analysis, a linear transformation (a change of bases) is applied to a dataset in order to maximize the variance of a reduced number of uncorrelated dimensions. Analysis is then done on data in the reduced space. However, ignoring the trivial, the values in the reduced-dimensionality space are a projection from a higher-dimension space and the original values cannot be determined. As Donald Norman suggests, though, if the representation is ideal, the missing detail is not relevant to what is being represented [58].

In reified inference, the true dimensionality of the generating process is unknown. We choose a model with a known low number of dimensions that approximates the salient characteristics of the generating process. In between, we capture a high number of dimensions of the input representation in an attempt to overconstrain the model. In modeling the input representation alone, it is clear that we are losing some information. This information ideally, if the input mappings are optimal, is not relevant to the model, but it is a part of the output representations. Consequently, for all but the most trivial output representations, the output will be of a higher dimension than the model. How do we recover the additional dimensions? For one, the model values will contribute to most dimensions of an expressive output representation, just as they contribute to many dimensions of the input. Still, additional information will be required. Introducing random values to any of the output dimensions, even stochastic noise that is modulated by the model values, intuitively feels like a problematic choice and John Maeda would admonish us for doing so in an interactive representation [48]. However, many of the parameters of the input representation likely contain a fair amount of noise so, on second thought, it shouldn't be too surprising that some noise that should have been filtered out by the input mappings would need to be reintroduced. Executed properly, typical pseudo-random generators or more elaborate stochastic methods could be injected judiciously into the output to create rich representations that preserve the intended meaning. Adding detail in this way is like the texture of a brush stroke when one is trying to paint a flat shape or the expressive improvisational elaborations in the upper voices of a motet.

2.4.2 An Example of Reified Inference in Application

In order to demonstrate some of the possibilities of the proposed method of remapping a representation using reified inference, I created a simple

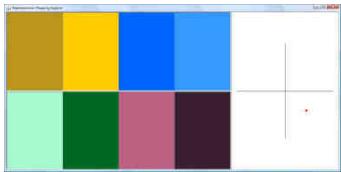


Figure 13: Screenshot of example inference application

This program demonstrates an application of reified inference, using a two-dimensional model to map sets of four input colors to a different set of four colors and a musical chord.

computer application to generate two novel representations based on an input set of four colors. The user is presented with a window consisting of two rows of four rectangles on the left and a two-dimensional plot on the right (Figure 13).

The user specifies the input set of colors by selecting a color for each of the top four rectangles. As colors are selected, the program generates a new set of four colors in the bottom row. Simultaneously, the user hears a one-second playback of a polyphonic chord using MIDI synthesis. The output set of four colors are intended to have a similar perceptual effect or impact as the input color scheme, though the color palette may be entirely different. While the color output is a new representation having identical form to the input, the chord is a representation of the same distinctive perceptual characteristics of the input representation in a different modality.

To accomplish this, a two-dimensional model was chosen that is comprised of two linear metric orthogonal axes representing the perceptual properties to be preserved. These parameters were selected to represent intuitive qualities that do not have standard well-defined semantics in either of the color theory or musical harmony domains, but do represent potential correlations in the input. I call the horizontal and vertical axes of this mode “variation” and “boldness”, respectively. The value of the model for the current input of four colors is displayed as a red dot in the plot at the right-hand side of the interface. The two-dimensional model is used to generate the two output representations that exhibit the same quality. The four input colors are each defined in a three-dimensional color space, yielding a twelve-dimensional input (or this may be viewed as four three-dimensional samples from the distribution of possible representations having that model state, depending upon the model and inference technique chosen).

Understandably, the color output representation has the same form. The chord output representation produces a triad with each of the three notes having a single parameter: pitch. Properties of the input vectors having high codimension when viewed in the context of the model parameters are used to derive the state of the intermediate representation. The function in the example program uses a nonlinear dimensionality reduction method based on the variance of the means and the means of the variances of each component vector, with the assumption of modal correlations consistent with the semantics of the model.

The output representations are computed using the values from the model state, the intermediate representation. The location of an initial value for one color component is made at random, since we’re trying to reproduce a quality of the set of four colors in the input representation, not the original representation. Subsequent values for the remaining components are selected from the uniform distribution defined by the variances and this location. The semantics of the model mean that the notion of hue is not

preserved while saturation is particularly well preserved in the “boldness” dimension, as it is influenced by the variance of the color components. Whereas the color representations had four elements with each of three parameters, the mapping needs to be different for the chord representation as it has only three to five elements each with one parameter. As with the color representation, the location (the root of the chord) is chosen at random from the octave range including middle C. The subsequent intervals are chosen from the diatonic scale with the “variation” value dictating the degree to which subsequent intervals vary chromatically and with the “boldness” parameter defining the range over which the entire chord may span. The use of these model parameters in generating the chord intervals is analogous to the perception of distance in the color swatches as captured by the model.

The output representations from the program are anecdotally consistent with the perceptual concepts being modeled given the input color swatches. Groups of four similar colors produce outputs demonstrating closely related colors. The chords generated for such output are also reasonably consonant, spanning a short range. The model parameters in these cases fall into the lower right corner of the plot, indicating low “variation.” If the colors are similar and muted, the “boldness” parameter is also low. Introducing one or two colors in the input set that vary dramatically in hue, saturation, or brightness from the others in the set produces an output that contains greater variation and a chord that sounds more dissonant. When all four colors vary greatly from each other in hue, saturation, and brightness, the model parameters move toward the upper right, indicating high “variation” and high “boldness”. Qualifiers such as “reasonable” and “similar” are used when describing these behaviors because the probabilistic nature of the remapping is non-deterministic and introduces some variation per trial and with variations in the input. However, the overall perceptual effects are consistent. It is also interesting to note that the two-dimensional model space exhibits properties that roughly classify well-known color harmonies, thus suggesting some validity in the intermediate representation as defined with respect to perceptual color theory (Figure 14).

2.5 Mapping

For Disembodied Performance, the process of transforming input representations to the intermediate representation and then to a multiplicity of output representations is accomplished through *mapping*. Mapping is a general term for many types of transformation, but it does presuppose that the process from input to output is deterministic. In our case, the overall effect is not deterministic because our input, the actor’s behavior, will vary from one performance to the next. Disembodied

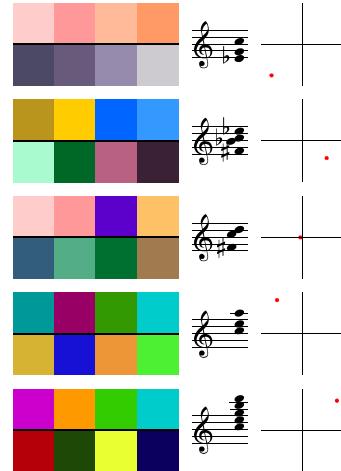


Figure 14: Output color sets and chords

The vertical axis is “boldness” and the horizontal axis is “variation”. The hues for the sets of colors for inputs and outputs with respect to the quadrants in which the locus of their model states fall reveal known color harmonies. Analogous color schemes fall in the lower left quadrant (low “variation” and low “boldness”; top). Complimentary color schemes fall in the lower right quadrant (high “variation” and low “boldness”; second), while split complementary lie near the center of the model space (third) moving toward roughly triadic color schemes near the top left quadrant (low “variation” and high “boldness”; fourth). The most bold and discordant color schemes, as termed by the model, are tetradic (bottom).

Performance does not intend to model the process, simply translate the regularities from one representational medium to another.

In [22], Marc Downie makes an impassioned critical argument against the use of the term mapping and related approaches. He argues that mapping-based approaches are naïve and inexpressive, citing examples that use very linear interactions between input and output. While many performance works do use banal approaches to map input and output, this should not be a criticism of the functional approach of mapping or the tools designed to facilitate the connections of inputs with outputs. It is a criticism of the types of mappings that are typically created by artists. I fully endorse the idea of mapping, in its broad definition of transforming input into output and suggest that artists that bring to their work a critical eye and a theoretical framework can create powerful and meaningful mappings.

I believe that Disembodied Performance escapes many of the pitfalls that Downie attributes to mapping approaches. Reified inference brings to the process a level of abstraction that distills the semantic qualities that are of importance to the artist and the audience. This abstraction preserves intension while discouraging the more direct and meaningless mappings that Downie decries. The mapping is merely the necessary means to interface with the intermediate representation. Additionally, in the Disembodied Performance System that I will describe in Chapter 4, the interface deliberately deemphasizes the properties of the input and output devices. The mappings are indeed made up of the composition of basic functions, but it is my intention that they facilitate the development of powerful modes of expression, not the details of connection.

In opposition to mapping, Downie offers an elaborate agent-based approach. While an artificial intelligence is compelling for some applications, it does not suit the type of application required for *Death and the Powers*. In agent-based works, it is often unclear from where any emotional resonance comes, if it is perceived. Whereas, to visualize the character of Simon Powers, the representation must clearly embody the emotional quality of the actor's performance. It is insufficient to allow autonomous agents to generate a performance with intentions potentially distinct from those of the character of Simon Powers. The representation is the character. Thus, the computational distance between the actor portraying the character and the onstage representation must be as short as possible. The intermediate representation of the character model does not constitute artificial intelligence. It is merely a consistent means for parameterizing the input and output mappings for the system. The system is *essentially* mapping-based. Not only do I believe that this approach is crucial for making the immediacy of gesture and emotion intelligible, but that it also affords the ability for the system to take direction and be tuned for the desired performance.

2.6 Distinction from Related Work

The Disembodied Performance System for *Death and the Powers* that I will present in Chapter 4 is notable for its departure from the theatrical and performance works reviewed above. I propose and demonstrate this new approach relying on reified inference to distill the actor's performance and translate it to the stage in a non-representational way. The actor is offstage. We do not see the actor's human form or the suggestion of the human body in the representation. We don't see his video image. What we see does represent the essence of his performance interpreted in whatever form is most appropriate for the production.

Disembodied Performance differs significantly from the above examples of related work involving media in theater in several ways. For *Death and the Powers*, projection is used as a visual representation of a character, not merely as the setting or to non-diegetically augment the mood. The visual languages for output representations need not make use of representational images, as in all of the above examples. Like Lepage's *Faust*, I will be manipulating imagery in real time, creating reactive visuals, but this process is transparent. It is merely a means of interpreting or transferring the actor's performance from one to another. It is not, as it is in *Faust*, a reaction to or commentary on the performance. It is also not representational, as imagery occasionally is in many of the stage examples, particularly in *Lolita*.

Certainly, sound and music will be intimately tied into the generated visuals and will even be generated by the Disembodied Performance. Disembodied Performance, though, seeks not a language that unifies visuals and sound, but a language that uses visuals and sound and other parameters to represent the essence of a character.

I have introduced reified inference, a method of inference and modeling that focuses on the output. Roughly stated, it doesn't quite require us to know *why* or *how* something is perceived and understood, merely *that* it is understood and *what* about it we want to understand. This approach is clearly not suitable for all modeling applications and academic endeavors, but it is a reasonable course of action when translating the representation from one form to another when that which is being represented is too complex to be modeled realistically or for artistic pursuits, as in its application for Disembodied Performance.

The approaches and considerations outlined above arm us with the tools to create Disembodied Performance. We can model affective expression through reified inference. Then we can communicate the affective intention palpably onstage by leveraging synaesthetic representations. The entirety of this approach is a new application for sensing and display in the theatrical context.

Disembodied Performance doesn't make very many additional demands on the performers, in contrast to other intermedial applications of technology in theater [30]. The timing and behavior for both the disembodied performer, the actors onstage, and even the role of the director remains constant, with the only difference being that the director and onstage actors must treat and address the set as the displaced character.

None of the past techniques presented in this chapter accomplishes what my system for *Death and the Powers* must strive to do: be a character. Much like the suspension of disbelief required for an audience to appreciate a marionette performance, The System must appear as the character itself, not as a controlled device or augmentation of the human actor. The proposed system is not representational: photographic imagery will not be presented as the character. To show images or video of the character or the actor offstage portraying the character belies an important plot point: Simon Powers has transcended human and material form. What will be seen and experienced in the theatrical environment is not passive. It is not merely a display, nor is it mindlessly reactive in the manner of many video-based interactive installations. The character as the set is aware of and active in its environment, to the point of both omniscience and omnipotence. The walls respond to and interact with the other characters onstage.

3 DEATH AND THE POWERS

*Every creative act involves... a new innocence of perception,
liberated from the cataract of accepted belief.*

—Arthur Koestler

Death and the Powers: A Robot Pageant is a new opera by Composer and Creative Director, Tod Machover, being produced at the MIT Media Laboratory. Drawing from its urtexts *Oedipus at Colonus* and *King Lear* [23], *Death and the Powers* explores what it means to die and questions the legacy one can leave behind. The libretto by Robert Pinsky, former U.S. Poet Laureate Consultant in Poetry to the Library of Congress, tells the story of Simon Powers and his family. Powerful Simon is not obsessed so much with living forever as he is with leaving something of himself behind in the world. To this end, he develops The System into which he can upload his essence upon the moment of his death. Simon enters The System at the end of Scene I and for the remainder of the opera, we see how he has retained agency and awareness in his new form. His family—and the world—struggle to understand and cope with this new way of being.

Under the direction of Diane Paulus—renowned artistic director of the American Repertory Theater and director of numerous opera and theater productions, including the Tony-Award-winning Broadway revival of *Hair*—and with production design by famed Hollywood designer Alex McDowell (*Minority Report*, *Fight Club*, *Watchmen*), *Death and the Powers* will premiere at the Salle Garnier of the Opéra de Monte-Carlo, Monaco in September 2010 under the haut patronage of His Serene Highness Albert II, Sovereign Prince of Monaco with a contribution from Opera Futurum, Ltd. Following that, the production will have a U.S. premiere at the American Repertory Theater in Cambridge, Massachusetts and is expected to tour throughout the United States and worldwide. The production will feature a host of new concepts and technologies for the opera genre and

theater in general, from new forms of hyperinstruments to redefining the interplay between actors and theatrical sets.

During the past two years, I have had the good fortune of being involved in many aspects of the production including discussions that ultimately suggested modifications to the story, experiments in sound design, as well as countless meetings about production design, engineering, and the task of realizing the opera onstage. It has been a remarkable collaborative process in which to take part. In general, I am primarily responsible for creating the visual elements of the opera that help tell the story and integrate into the overall production design, be they prerecorded motion graphics and video or live computer-generated imagery. I also am presently in charge of the development of new show control systems and consult on the theatrical feasibility of proposed scenic designs and staging.

The story, production design, and overarching concepts for *Death and the Powers* were all important motivations for the development of Disembodied Performance and the control system. In this chapter, I will discuss how the design of the production and the needs of the story have evolved, touching on some of the ways I have been involved with these aspects.

3.1 Genesis

The commission that spawned *Death and the Powers* began in 1997 when the newly appointed president of the Association des Amis de l'Opéra Monte-Carlo, Kawther Al-Abood, approached composer and MIT Professor of Music and Media, Tod Machover, to create a new opera. Al-Abood wanted a piece that would revolutionize the opera community, particularly in Monte Carlo, attracting younger audiences. She sought something fresh and innovative, not just in terms of musical style, but also in staging and production technology. Early ideas Al-Abood considered included a performance that took place on the Monaco Opera stage, which would then open up revealing action on the Mediterranean Sea beyond.

Creative discussions were underway for the project by 1999. One of the early collaborators Machover brought to the project was renowned juggler Michael Moschen. Together, Machover and Moschen conceived of a production that was based on the motion of objects and abstract forms onstage. Moschen also brought to the table the beginnings of a story that dealt with issues of trust within a family, between a young girl and her mother. While the dramaturgical aesthetic was generally abstract and gestural, one key element of the production was that the young girl's toys would come alive, presumably through the use of robotics. While the story has changed since this time, several of these core ideas have persisted in some form.



Figure 15: May 2008 production retreat

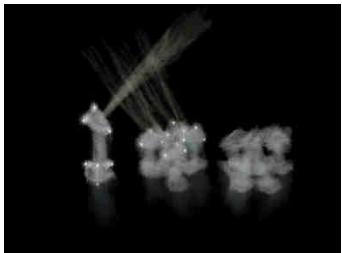
From these early stages, the overarching theme that Machover continued to embrace is the notion of a choreographic language of physical objects that would be intimately related to the music and the story. Machover wanted to extend scenographic technique beyond the commonplace uses of static scenery and flat video projections [10].

Around this time, Machover sought out the verbal talents of poet Robert Pinsky to further develop the story and pen the libretto. MIT Media Laboratory robotics professor, Cynthia Breazeal, also joined the project to help realize the ambitious mechanical scenery and props that were envisioned. Moschen's contributions to the production and the story were influential, but did not ideally fit the longer-form medium of opera as the project developed. The production would continue to evolve without him.

The next additions to the creative team, in 2003, were Randy Weiner and Diane Paulus. Weiner worked closely with Pinsky and the rest of the creative team to flesh out the story that would be known as *Death and the Powers*. From this, Pinsky drafted the full libretto. Paulus would serve as the opera's director. The first preview of the production occurred in Monte Carlo in November of 2005, featuring the first scene of the opera set to Machover's score with James Maddalena assuming his lead role as Simon Powers. Just prior to this, the final member of the core creative team would sign on: Production Designer, Alex McDowell.

3.2 Synopsis

PROLOGUE & MEMORY DOWNLOAD

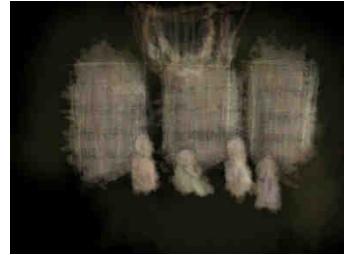


As the audience enters the theater, they are immersed in a strange environment. The stage and mood is set before the show begins. A geometric assembly of objects onstage appears to be a set. However, as the opera commences, these objects come to life. They are a community of intelligent robots set some time in the future that have awakened to perform a ritual reenactment of the story of their creator, Simon Powers.

As the robots take their places for the reenactment, a sequence known as Memory Download begins. During this process, the audience observes the robots accessing the story of the characters that will be portrayed. Images of Simon Powers, his daughter from his first marriage, Miranda, his third wife, Evvy, and his research assistant and protégé, Nicholas, are drawn from space and illustrate the back-story of these main characters. The images coalesce, transforming four of the robots into these characters. The remainder of the robots carry on setting the stage and acting somewhat like a Greek chorus or portraying earlier versions of themselves throughout the remainder of the show. The action of the inner play commences.

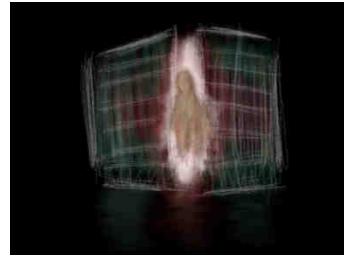
SCENE I: SIMON AND THE SYSTEM

Simon Powers—an eccentric inventor, business mogul, and wealthy entrepreneur—energetic in spirit, but physically withering, is about to enter The System he has created throughout his home to preserve his essence and agency after his imminent death. He is unfazed by the prospect of living on through his technological creation and Nicholas, who has benefitted himself from Simon's ingenuity and experimentation, is eager to initiate the transition that Simon must undergo. Evvy and Miranda, on the other hand, are rather frightened at the prospect of losing Simon and the whole procedure. Simon pauses to reflect on his family and the idea underlying his creation of The System: one's life and essence is not one's body or one's possessions, but the spirit, the intangible movement and meaning. Miranda and Evvy are uncertain as to what this will mean. At last, Nicholas completes the preparations and Simon vanishes into The System.



SCENE II: INSIDE THE SYSTEM

Once inside, The System comes online and Simon's consciousness begins to experience the result of his life's work for the first time. In an aside, we see him struggle to reassemble his thoughts and make coherent sense of this new way of being. He searches for memories, a trace of his identity free of his body, and eventually finds his footing, discovering that what he has become is, in truth, no different from his mortal self.



SCENE III: GETTING TO KNOW YOU

Meanwhile, outside of The System, Nicholas, Miranda, and Evvy observe the machinery functioning, but search for some sign that Simon still exists. Nicholas tinkers and checks to make certain that The System is operating correctly. Evvy is torn between the as yet unconfirmed hope that Simon will somehow return and mourning the loss of her beloved husband. Young, naïve, and sheltered Miranda is more skeptical and afraid that she has lost her father, the only true family she has. She too searches for some sign of life from The System. Sure enough, we begin to feel a presence in the house. The walls and the furniture come alive, at first with a sign of intelligence, and then with behaviors that resemble Simon himself. Soon, the house starts exhibiting the same playful and energetic qualities we observed in Simon before his death.

SCENE IV: EVVY'S TOUCH

There is a lull in activity in the Powers home. Nicholas retreats to his workshop and Miranda, still uncertain about her father's transformation, retires. We see Evvy alone. She is desperate to reconnect with her husband and talks to the house as if he were there. She reminisces about their past together and Simon in The System responds. He presently inhabits the chandelier, which begins to move and sound as it descends to envelope Evvy. Together, they learn how to touch and interact across this new divide.



Evvy can feel that The System is in fact Simon, as they share an intimate opportunity to get to know each other once more.

SCENE V: NICHOLAS AND THE ROBOTS

In his laboratory, Nicholas celebrates what he believes to be the success of Simon's transference into The System. Years of toil realizing Simon's dream have paid off. Nicholas, in many ways, was a guinea pig for technologies that would be incorporated into The System. Simon benevolently rescued Nicholas at a young age and raised him as the son he never had. The young Nicholas was considerably disabled and missing limbs, including his arm. Simon was able to create remarkable prosthetics for Nicholas, not just to restore his normal movement, but also enhancing his capabilities. Nicholas views these additions as an improvement on the human form and as steps toward becoming part of The System himself. As he expresses his joy, he dances about with several of the robots who are not only utilitarian assistants, but to Nicholas are his companions and even his kindred.

SCENE VI: THE WORLD REACTS

Some time has passed. The System has grown in complexity and scale, yet it itself is fading from materiality. Evvy, Miranda, and Nicholas have become more accustomed to Simon in his new form. As when he was alive in material form, Simon continues to transact business dealings and trading in international markets. However, his renunciation of the material world has only been affirmed by his time within The System. His actions have shown a blatant disregard for the well-being of world economies, industry, and the communities that depend on these institutions. Evvy wanders about in a daze, in constant communication with Simon, hearing words and speaking words only the two of them can hear.

Miranda announces the arrival of a delegation of world leaders who have come to seek an audience with Simon and plead for aid and support, a reversal of the economic turmoil he has caused. When she presents them to Simon in The System, they grow indignant that they must address the house itself, thinking Simon's omnipotence and antics to be a trick at their expense. Simon, unsympathetic to their cause, taunts and humiliates the Delegates. Miranda is now torn. She tries to defend her father as the Delegates impugn his motives and very existence. On the other hand, she is appalled at her father's indifference, thinking that if it were truly her father in The System, he'd not be so callous, and asks for his understanding.

SCENE VII: INTO THE SYSTEM

At this point, Nicholas has begun shedding his biological and mechanical body. His conviction that a better and truer life awaits in The System has been bolstered by recent events and he is prepared to join Simon inside. Evvy, who has been in contact with Simon for some time now, understands Simon's experience and she too is eager to reunite completely with her

husband in the realm free of matter and the body. Miranda finds it difficult to accept what her family is doing as she watches Evvy and Nicholas vanish into The System.

SCENE VIII: MISERIES, MEMORY, AND MIRANDA

Miranda is desperate to save Nicholas, Evvy, and Simon from their self-absorbed mindset and abandonment of humanity. She wants to remind them of the virtues and needs of the physical world and implores, particularly Simon, to re-engage and be sympathetic to the needs of the world's people and her own need of a father and companionship as well. She has been left with only the memories of some semblance of a normal life and her loved ones inside The System for which she cares. To persuade Simon, she summons the world's miseries, oppressed and downtrodden masses, as an example to her father of what his lack of compassion has wrought. The miseries, however, do not have the intended effect and she is again left alone. Dejected, Miranda is astonished to see Simon reappear to her in his human form. In this final confrontation, he entreats her to shed her mortal and material life and join him and the others inside The System.

EPILOGUE

The reenactment has concluded. Though the robots have performed this ritual pageant many times before, they still fail to grasp the notion of death and the significance of the story the "human creators" have left as their legacy to be retold ad infinitum.

3.3 Production Design

The production of *Death and the Powers* seeks to be highly innovative in the world of opera in terms of the quality of its staging and the technology it employs. Production designer Alex McDowell, in his first departure into theater from film, in collaboration with the rest of the creative team has developed a look and an environment that serve the unique story and provides many opportunities for creating a theatrical experience that redefines the genre of opera. The production design has evolved over the years since McDowell formalized initial concepts in 2005 and continues to do so as production approaches. McDowell brought to the table a wealth of artistic inspiration and a clear vision for the character of the story world, from the architectural style of elements to their materials. I have worked closely with McDowell and other members of the team toward realizing his designs within the constraints of theater and the technology available.

A publicity document describes the aim of the overall production design:

The production of *Death and the Powers* will be at once spectacularly innovative and practically simple. The stage will represent Simon

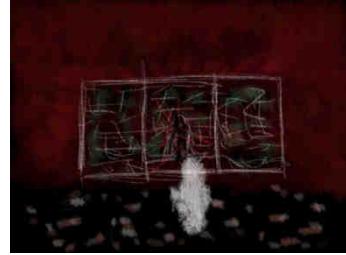


Figure 16: Conceptual sketches
The sketches accompanying this synopsis were created by the author to explore aspects of the staging and use of color for *Death and the Powers*.

Powers's house, but this room will gradually reveal itself to be a vast, interconnected, intelligent system that will allow the room itself to change its shape, undulating, vibrating, pulsating, or pounding. [...] Embedded in The System will be a number of image display surfaces and sound-producing elements, capable of showing the disparate, fleeting thoughts and memories from Simon's inner world. Also essential to the production of *Death and the Powers* is an overall technical design that will allow the opera to be installed, rehearsed and performed in extremely diverse performance spaces. [55]

It has always been expected that the production would have considerable longevity, like many past Machover projects. After the premiere, the production is expected to tour worldwide in a variety of venues. The originally scheduled premiere venue was the impressive Salle des Princes at the Grimaldi Forum in Monaco. This massive theatrical venue contrasts greatly with some of the smaller theaters to which the production may be brought. Ultimately, the premiere location was moved to the smaller Salle Garnier, a traditional opera house. Original designs for the set were designed for the larger scale, but would have difficulty in smaller venues. Consequently, many of the designs were reduced in size and alternative staging. Modular expansions were also considered to be a solution to this problem, but proved too expensive for current production budgets. Nevertheless, in engineering the set for *Death and the Powers*, modular designs play a critical role in approaching the goal of a "practically simple" implementation, allowing parts to be easily replaced or repaired. An important moment in the scaling-down of the show's design occurred when a partnership was struck with the American Repertory Theater. The A.R.T.'s Loeb main stage at Harvard University in Cambridge, Massachusetts was one of the smaller theatrical venues considered. It was decided that the United States premiere of the opera would take place there, which meant that the production had to fit superbly into that space. A reconception of the scale and the design led to a more intimate and immersive production. Adding to this new take was the ability to use a thrust configuration in the Loeb. Although the changes to the show's engineering at this point were not extensive to accommodate the thrust space, it changed the way the production team thought about the design in all venues.

Traditional theatrical venues are not the only target spaces for staging *Death and the Powers*. Parts of the set, such as the walls and the Chandelier, were envisioned to appear in other contexts ranging from demonstrations to interactive museum installations. An ancillary project is Personal Opera, in which individuals could create their own version of The System, a multimedia legacy of their own. Personal Opera would span multiple platforms from mobile devices to the Web. An installation version was also proposed where a wall or periaktos from the stage production would be

situated in a space and visitors could design their Personal Opera using the massive interactive wall display. The Personal Opera concept could also be incorporated into stage productions of *Death and the Powers*. Users across the globe could contribute material or influence aspects of the performance through an online interface in a manner similar to Machover's *Brain Opera*, which, in its second act, incorporated material generated by participants in the Mind Forest, a collection of hyperinstruments that were explored in the first act, and by users of the production's website [93].

The present design for the production features two large set pieces that are central to the representation of Simon in The System: the Chandelier, a large Hyperinstrument that flies in and out above the stage and plays a critical role in Simon's interaction with his wife, and three large periaktoi that resemble bookshelves and move and rotate ballerically across the stage to form the setting for each scene. Other elements that may be incorporated as representations of Simon in The System include robotic furniture that walks or moves as Simon Powers, an elaborate futuristic wheelchair that may serve as Simon's portal into The System, and a small mechanical bird that is both companion and toy of the young Miranda. A third major component of the design is the Operabots, which may exist both in The System and in the material world. An effect called the matrix-of-light, a grid of numerous floating white points of "intelligent" light that fills the stage and possibly the entire house of the theater, represents The System itself at the very end of the opera. The matrix-of-light is similar to the 2005 installation *MATRIX II* by Erwin Redl. Several elements present in original designs have been abandoned for practical reasons and due to lack of importance to tell the story, such as an animatronic floor with large agonal tiles that could elevate some distance and form complex surfaces, flying pianos, a portrait of a young Simon Powers that comes alive, a mischievous rug, and a large sculpture of a bird.

Machover's conception of an "opera of visuals" has been a mainstay of the production design from its earliest inception. The physicality of moving objects survives in many of the elements of the set design, such as the Operabots and in the books of the walls, which I will discuss in the sections below. Promotional literature for *Death and the Powers* describes the importance of these ideas as a departure from conventional and contemporary practices:

Death and the Powers proposes a totally new approach. Rather than sterile projection screens, objects on stage—from huge to tiny—will move magically to delight the eye and imagination, while performers will be naturally extended by set and robots. [54]

This spatiality and physicality is seen throughout the design of the production, from the earliest concepts and renderings of the set assembled



Figure 17: *MATRIX II*, Erwin Redl
(Photograph provided by Alex McDowell)



Figure 18: Matrix-of-light
The matrix-of-light effect represents the inside of The System and fills the stage, extending into the house.
(Rendering by Arjuna Imel)

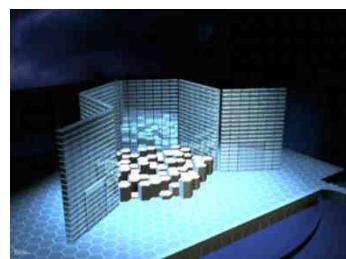


Figure 19: Actuated hexagonal floor tiles
The early concept of an actuated floor created opportunities for vertical elevation of the action and Simon in the System to express himself in another part of the environment. (Rendering by Arjuna Imel)

by McDowell and the creative team, to the current incarnation. It pervades every aspect from the number and behavior of the stage elements to the blocking and even sound reproduction and sound design. Unfortunately, realizing such a level of physical and mechanical complexity, with reliable autonomy of robotic elements, is an extraordinary engineering, logistical, and financial challenge. Budget, time, and personnel constraints have weeded out all but the most essential of these elements and compromises have been made and will continue to be made. However, I find the vision that Machover touts of a choreography of objects that break free from traditional constraints and representations especially appealing. In my contributions to the project, I have tried wherever possible to preserve these ideas in developing more budget-friendly and practical implementations.

The opera is in one act with no intermission and no curtain. While these deviations from traditional theatrical vernacular are increasingly common, they pose certain technical challenges for the scenography. All entrances and exits of scenery must occur before the audience's eyes, while concealing any theatrical tricks that would be necessary to avoid breaking the audience's suspension of disbelief.

3.3.1 Sound in Space

In any opera, the quality of the sound of music and singing is of utmost importance and concern. Although often controversial in the opera world, this production necessitates the use of sound reinforcement. One reason for this is so that the singers' voices blend well with the electronic sounds common in Machover's compositions, as well as with amplified instruments and hyperinstruments. Additionally, effects will be applied to performers' voices, especially in the case of Nicholas, who has a prosthetic arm that allows him to manipulate his voice, and Simon Powers, who is omnipresent and free to move any aspect of himself around once inside The System.

To complement the concept of a choreography of visuals, spatializing sound is an important aspect of the production. Machover's original conception was that each of the various objects moving about the stage would also be a sound source. The promotional website for the production describes this idea [53]:

Although much of the music of the opera will be "electrified," special sound projection techniques will be used so that everything will have a lovely, shapely, three-dimensional quality, capable both of filling the entire theater with viscous, enveloping waves and also of whispering ever-so-delicately into the ear of every audience member.

This also includes robotic musical instruments, coined *sonitronics* as portmanteau of "sonic animatronics", such as the Chandelier [55]. Making all objects onstage practical sound sources introduces a considerable

amount of complexity to the logistics and infrastructure of the show. Consequently, alternative approaches have been explored that not only achieve the creation of a dynamic sonic landscape on the stage, but extend sound into the house and afford many new opportunities for sound design in theater that are amply capable of realizing the evocative description above.

Early experiments for *Death and the Powers* using ambisonic spatialization techniques led to their incorporation to Machover's 2008 premiere of another opera, *Skellig*, to great effect. The third-order ambisonic sound format encodes audio sources in six dimensions, using their position and velocity, as 16 channels related to the spherical harmonic decomposition of the sound pressure of the source in space. The sound can then be decoded to an arbitrary number of speakers at any scale for a full periphonic experience. Given the success of ambisonics in *Skellig*, further advancements are planned for the premiere of *Death and the Powers* as well as investigations into other techniques including wave field synthesis.

3.3.2 The Chandelier

The Chandelier is a large architectural string musical instrument. It hangs above the central focus of the play area onstage. During Scene I, it reads much like a piece of sculpture or an illumination chandelier. However, once Simon Powers enters The System, it is one of the main elements of the environment that comes alive. It can move and subtly gesture. The music it produces can be tied to Simon's voice. We can hear Simon through the Chandelier. During Scene IV, the Chandelier has its virtuosic moment as it descends to the stage. Its wings open and close to envelope Simon's wife, Evvy. Together they have an intimate and musical encounter through the instrument. Original blocking for this scene had Evvy being drawn up into the air by the Chandelier or had her dancing with the structure in space, high above the stage. One version of the libretto called for Evvy's entrance into The System by way of being flown out inside of the Chandelier. Due to safety and insurance concerns, Evvy's aerial acrobatics and flying inside of the Chandelier have been cut, leaving her to be enclosed in the instrument at the stage floor for Scene IV.

The design of the Chandelier evolved quite a bit over the years. Original conceptual drawings and renderings show a graceful wire form reminiscent of the illusory curved shapes Michael Moschen used in some of his routines. The design of the Chandelier began to reference the mathematical stringed surfaces of Naum Gabo and the sculptural forms of Constantin Brâncuși. At one point in the history of the production, a large sculpture of a bird was to be onstage as an object that Simon in The system would occasionally inhabit. This sculpture was dropped from the production and its bird-like qualities migrated to the form of the Chandelier. The Chandelier's wings would be the bird's wings and tail and a blob-like body



Figure 20: *Bird*, Constantin Brancusi
(Photograph provided by Alex McDowell)

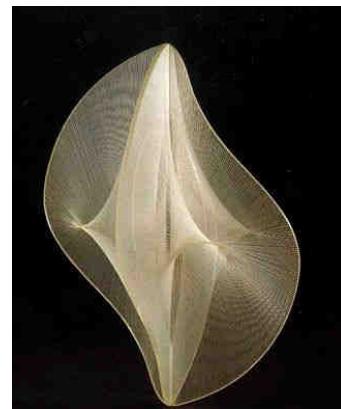


Figure 21: *Linear Construction No. 2*, Naum Gabo
(Photograph provided by Alex McDowell)



Figure 22: Early Chandelier model
In this model, the wings open upward and a bird-like form is at the center of the structure.
(Model and photograph by Steve Pliam)



Figure 23: Quarter-scale Chandelier model
This more recent model has the wings oriented downward so that they can enclose Evvy.

was in the interior in some designs. The wings were later inverted so that the instrument could envelop Evvy. The current design of the Chandelier invokes not only the bird image, with a beak-like structure at the top, but that of an egg, a heart, and a womb.

I refer the reader to [65] for a detailed documentation of the Chandelier's design and to [27] for a discussion of its musicality. Both have evolved and changed since the time of those writings, particularly the latter. The Chandelier's core will likely become larger and more solid, as it will need to house the electronics for aggregating signals from its strings and amplifiers and drivers for sound reproduction.

Musically, the Chandelier has become a pitched instrument, relying only on electromagnets for actuation. The electromagnets are driven by signals at specific frequencies, causing the strings to resonate when the frequencies are at the fundamental for the string length or any of its overtones. Guitar pickups at the opposite end of each string convert the string's vibration back into an audio signal. Strings then can be played at any of these frequencies, not simply at their fundamental as in traditional instruments. Additionally, a tactile player can touch an actuated string, damping it, changing its length, or touching a node to produce a harmonic. Of course, strings may be plucked or hit to produce a sound, as well. This presents interesting possibilities, such as a duet on a single instrument with one player physically touching the strings while another player controls the signals sent to the electromagnets.

Since the harmonics and overtones can be played as well as the fundamental, a contemporary 12-tone equal temperament tuning is insufficient, as the overtones of different strings do not line up well. Instead, the Chandelier will be tuned in 31-tone equal temperament, the next tuning where both overtones and fundamentals line up perceptibly well with each other and the traditional 12-tone tuning that other instruments in the orchestra will use. Since the tuning is unusual, multiple notes can be played on one string, and there are 48 actuated strings, of 96 total strings, capable of producing many frequencies and timbres each, new musical controllers and keyboards are being developed to allow an instrumentalist to perform the Chandelier.

As a Chandelier, the structure will provide a motivated source of illumination for the interior of the Powers's home. Additionally, the truss structure and strings of the instrument will be lit in expressive and architecturally compelling ways. One of the issues common to remote instruments like the Chandelier and other uses of technology throughout the show is that of demonstrating that the effect is real. Just as the question of where is the actor portraying Simon Powers when the character is in The System exposes a disembodiment problem, so does the playing of the

Chandelier. It is an actual instrument that will be creating sound during the opera. If the audience is not close enough to the hanging device to see and feel that it is actually doing so, how can they know that what they are hearing is actually coming from the instrument? Possible resolutions have included tying decorative elements or crystals to the strings or adding a laser parallel to each actuated string. In both cases, the idea is to make the string vibrations as visible as possible from a distance. Neither of these approaches is practical due to physical interference of objects connected to strings and precision alignment difficulties with lasers. However, current designs include addressable LED lighting at either end of each string that will react to and highlight the string's actuation, possibly using stroboscopic effects as in Jeff Lieberman's *Slink* [42].

3.3.3 Operabots

The Operabots are unique among the technological set pieces being developed for *Death and the Powers* and very much reflect the notion of a choreography of objects. Unlike the Chandelier or the walls, the Operabots don't represent Simon in The System. They are a collective of independent characters that frame the story of Simon Powers and his family, enacting a pageant for reasons they fail to comprehend at some distant time in the future.

The Operabots are conceived of as the physical remnants of The System or the robots that are Nicholas's assistants (or their descendants) from Simon Power's time. They are the first things we see alone onstage when the audience enters the theater. At first, they are inanimate, but as the pageant begins, they come alive. Early conceptual renderings depict the Operabots arranged in a regular grid in their static state at the start and end of the opera, reminiscent of artist Rachel Whiteread's sculptural works *Twenty-Five Spaces* (1995) and *Embankment* (2006). This arrangement mimics the grid of the matrix-of-light that represents The System, of which they may be a part or to which they can freely enter. Recent discussions have the stasis patterns resembling a cellular structure, a mass from which the individual robots separate, though the final robot design is not well suited for this.

As the Operabots activate, they exhibit social structure and behavior. They swarm in hive-like ways and demonstrate some sense of hierarchy. The Operabots are communicating and have dialog in the libretto. Though we hear only their strange language set to music, the supertitles for the production will provide a translation. One important design goal for the Operabots was to ensure the capability of giving expressive performances. Cynthia Breazeal, with her experience in sociable and expressive robotics, consulted with the creative team and Diane Paulus is eager to direct the Operabots onstage as she would human actors.

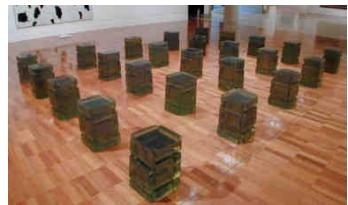


Figure 24: *Twenty-Five Spaces*,
Rachel Whiteread

The grid arrangement of cube-like objects influenced early designs for the Operabots in the Prologue. (Photograph provided by Alex McDowell)



Figure 25: *Embankment*, Rachel
Whiteread

Original Operabot designs resemble Whiteread's white cubes. (Photograph provided by Alex McDowell)

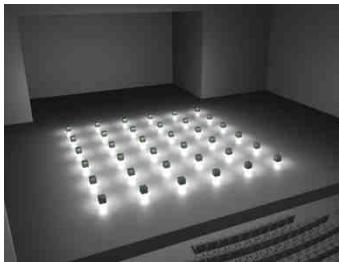


Figure 26: Operabots at the start of the Prologue

An early rendering depicts Operabots arranged as a grid at the start of the opera. As the Prologue begins, the robots activate and begin moving. (Rendering by Arjuna Imel)

As Memory Download begins, four of the Operabots transform into the four main characters. Physically, each Operabot elevates to the exact height of the human character it is about to become prior to the transformation. The suggestion here is that the Operabots can change their material form and become human in appearance. Numerous approaches have been proposed for achieving the transformation from Operabot to human onstage ranging from lighting and projection tricks to careful blocking and indirection. These four Operabots are the actors reenacting the tale of Simon Powers entering The System. The four leaders of the robot society would play the main characters, which leads to an identification of status with each of the human characters. In the Epilogue, it would be revealed that the leader of the robots was not Simon, as may be expected, but Miranda. In recent discussions, this pairing of specific robots with characters and the corresponding implications have been eliminated.

Throughout the inner play, the Operabots serve a number of different roles. With respect to the metanarrative, they act as a Greek chorus, observing as an audience, directing attention or reacting in commentary to the action. Some serve as set pieces for the reenactment; as furniture, for example. They also participate in the action, namely as Nicholas's assistants and companions. They have the ability to provide illumination to the scene and some may have display surfaces (or the impression of such) for use as Nicholas's instrumentation. As the story unfolds, the Operabots have a more active role. In Scene IV, some versions of the blocking have them choreographed in Busby Berkeley fashion, heightening the intensity of Evvy's interaction with Simon as the Chandelier. In Scene V, Nicholas dances with the Operabots. In Scene VI, they may mirror or taunt the Delegates.

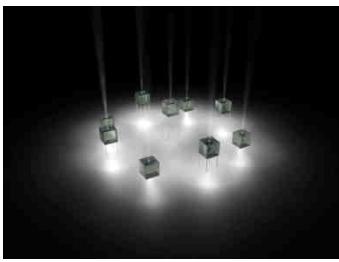


Figure 27: Early cube-shaped Operabots

Early Operabot designs were transparent cubes that could elevate on four thin legs. The Operabots had a down light to enhance the impression of floating and a strong vertical beam of light. (Rendering by Arjuna Imel)

The design of the Operabots has changed more during the history of the production than perhaps any other set elements. A few properties are common to all designs, though. All versions could elevate some height above the stage, at least enough to assume the height of the human characters. All designs also incorporated internal illumination and a spotlight at their top that would extend the vertical impression of the Operabots.

An original design had the Operabots as transparent two-foot cubes. The cubes had four thin telescoping legs that would allow the Operabot to hover just above the stage floor and then extend up over six feet, all the while being able to translate and rotate simultaneously at high speeds. To enhance the effect of the upward-pointing spotlight, the Operabots were illustrated with the ability to release fog. This light and fog effect was one possible method for concealing Simon's disappearance into The System. A downward pointing light enhanced the Operabots' appearance of floating. Although McDowell expected that some machinery would be visible inside

the transparent cube, the cubes would later become a translucent white to reveal some of the interior mechanics, enhance the visibility of the robot's internal illumination, and reflect a new material aesthetic. McDowell chose the translucent white material to match the quality of rapid prototyping resins. He envisioned that the Operabots' ability to transform matter was a form of rapid prototyping that they used to create everything from the set and props for their pageant to the human characters' costumes and themselves.

A first generation of Operabots was developed by Cynthia Breazeal's Personal Robots Group at the MIT Media Laboratory. Four identical prototypes were constructed as scaled-down translucent cubes. They had internal illumination, but no spotlight, and no mechanism for elevation. Instead of legs, the prototypes had holonomic omnidrive systems at their bases. This implementation was intended primarily as a proof of concept of a unique control system that would use choreographed animations of virtual representations created in Autodesk Maya as control data for the robots. A Vicon vision system tracked fiduciary infrared reflective markers on the top of each robot to provide wireless real-time updates to the robots, ensuring that they remained on their choreographed path with respect to each other. The system was extended with multi-robot obstacle avoidance, so that Operabots could maintain their choreography without colliding with each other or actors [63].

Later designs would change the shape of the Operabots to have triangular components. One version had a triangular base and a triangular head, with a spotlight, that was actuated by three triple-jointed arms at the vertices. This design could be covered with a translucent white stretchable fabric skin and provide very organic expressive animations; complex deformations of a triangular prism. This shape reflected the change in the configuration of the walls to periaktoi and the idea of a cellular, fractal-like structure to the system. In this way, the robots could be read as smaller versions of the walls. This evolved into a subsequent design where each Operabot was composed of two tetrahedra connected in the middle at a single vertex atop a triangular base. The head plate was individually supported and could move with the top face of the uppermost tetrahedron or independently, emerging from the volume of the upper tetrahedron. McDowell imagined that each tetrahedron edge would be a piston with hydraulic actuation allowing for complex movements of the structure. The tetrahedron design reflected not only the triangular nature of the periaktoi, but also the tetrahedra subtly embedded into the design of the Chandelier.

The final design of the Operabots was arrived at during the summer of 2008. Each Operabot has a triangular base, two-feet on a side, containing batteries, control, and drive systems. Nine transparent rods that can be illuminated independently at each end rise from the base to the triangular



Figure 28: Cube Operabot prototypes

The initial prototypes of the Operabots had a cube shape. The prototypes' movement and glowing illumination was controlled by an animation and had real-time collision avoidance during playback. (Video frame from [63])

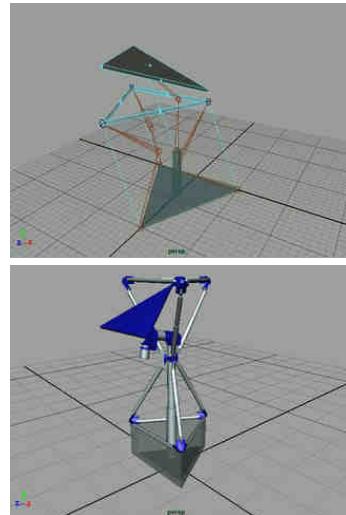


Figure 29: Operabot conceptual renderings

These frames from animations demonstrate two possible Operabot designs based on a triangular form. Both are highly articulated for expressive animations. (Renderings by Arjuna Imel)

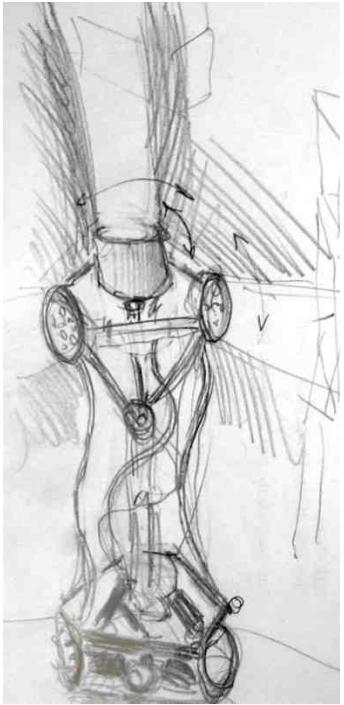


Figure 30: Operabot sketch

This May 2008 sketch would evolve into the final Operabot design. (Sketch by Alex McDowell)

head. The head contains a spotlight and itself can be illuminated. A bright point light is located at the three vertices of the head and three top vertices of the base. These point lights provide options for revealing the Operabots triangular-prism form and references the matrix-of-light representation of The System. The Operabots can move forward or backward rather swiftly and turn in place. The head elevates from approximately 4.5' to 7' feet. McDowell wanted the normal size of the Operabots to be slightly greater than half-human height and feel subservient while still being able to extend to the heights of the actors and beyond for an imposing stance. The head can also tilt forward approximately 90°. This final design was determined by a careful interplay between engineering constraints and creative requirements. Several iterations of a prototype of this design have been constructed by the Opera of the Future research group at this time.

The possibility of incorporating acoustic sound sources on the Operabots to serve as their voice is being investigated. An instrument, such as a daxophone, could provide a mechanically-actuated expressive voice-like quality. This sound may be controlled by instrumentalists in the orchestra. In fact, any aspect of the Operabots' expressive articulation, though not translational movement, is performable by musicians in the orchestra pit using the control system described in 3.4 below.

Concepts for the cube and tetrahedral Operabots called for 36 robots to be onstage. In a grid configuration, this would be a 6 by 6 matrix of robots (Figure 26). In discussions over the past year, the number of Operabots has been reduced to a number between 9 and 16. Budget restrictions lead to the reduction as well as considering the scale of the production for smaller venues. The triangular Operabot forms don't lend themselves to a rectangular grid, so other options for their initial sculptural appearance onstage are being considered.

Concerns for safety and battery life of each unit led to plans that would reduce the amount of time the robots would be onstage and approaches such as limiting the number of fully articulated robots to four hero Operabots while the remainder would be less sophisticated, though identical, versions or props that would be moved across the stage using common theatrical devices. Consummate engineering, however, has afforded control, drive, and power systems that are expected to be able to operate within considerable tolerances for the duration of the entire production, allowing the Operabots to perform onstage as much as desired.

3.3.4 Walls

The walls of the Powers's home are the primary set pieces in which Simon's presence will be manifest and are central to the design as a whole. All of the walls are vast bookcases filled with numerous books. Their conception, as indicated in the libretto and notes on the production design, is that the



Figure 31: Final Operabot prototype

books represent Simon's memory. Originally, the repeated forms of the books were to be animatronically actuated to create dynamic variations in the surface, thus lending themselves to a choreography of objects that exhibit expressive and musical properties (Figure 42). I intend to retain a similar notion, though due to cost and safety measures, I have changed the medium as described in the following section.

The visual design of the walls began with a photo McDowell selected from Dirk Reinartz's 1995 series *Deathly Still* of the prisoner files in the record room of the Theresienstadt Nazi German concentration camp. Other sources of inspiration included the cast sculptures of Rachel Whiteread, particularly *Nameless Library* (2000). The resulting designs all incorporated monolithic bookshelves lined with numerous over-scale books and some earlier plans show large doors similar to those in both Reinartz's photograph and Whiteread's sculpture.

The materials McDowell envisioned for the walls match those of the Operabots. In early renderings of the walls, the shelves and books were both transparent. In some concepts, one side of each wall appears to have a mirrored or partially mirrored surface. As with the Operabots, the transparency of these designs gave way to a translucent white plastic finish to suggest their story-world rapid prototyping resin construction. The transparent and translucent versions of the books have always been seen as incorporating some form of illumination. This idea later evolved into the walls becoming large display surfaces, as discussed below.

The configuration of the walls onstage has changed considerably during the past several years. Original renderings depict nine large, nearly 20' square, freestanding walls that can move about the stage and join together at one edge to form groups of three walls at 120° angles. These cell-like clusters could then create a hexagonal interior space, mimicking the hexagonal tiles of the actuated floor envisioned at the time, and rotate open to create the various settings for each scene.

The method by which these walls would stay upright and move around was not, to my knowledge, addressed until the subsequent design, which featured a large 45' diameter circular truss that hung above the stage. The walls were suspended from the straight and circular tracks on the truss with an individual linear truss segment for each wall to translate and rotate, as well as slide and orbit around the truss. An open circular area at the center of the truss allowed the Chandelier to fly in and out. The entire structure would be flown in and out to reveal and hide the walls during the prologue and epilogue. Animations by McDowell demonstrate the abundance of settings that could be created in this manner and beautiful choreography of the walls with respect to the actors. For example, a scene change occurs as a wall sweeps around on the circular track with a stationary actor passing



Figure 32: Image from *Deathly Still*,
Dirk Reinartz

(Photograph provided by
Alex McDowell)



Figure 33: *Nameless Library*, Rachel
Whiteread

(Photograph provided by
Alex McDowell)

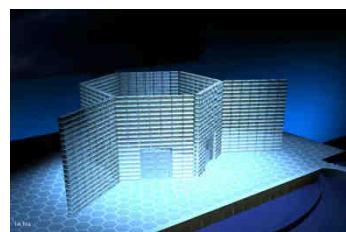


Figure 34: Freestanding walls

This early rendering shows eight large wall sections forming a hexagonal cell that would open to reveal the interior of the Powers's home. Several of the wall sections show paneled doors like those in the artworks above. (Rendering by Arjuna Imel)

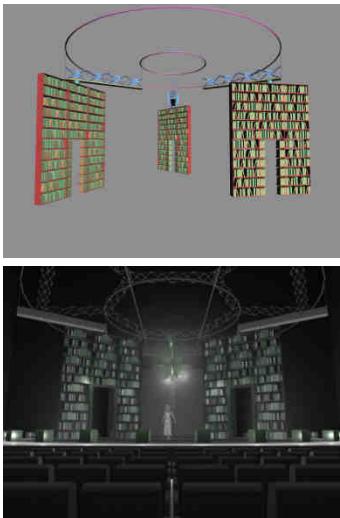


Figure 35: Walls flown from a circular truss

Walls flown on the large circular truss could orbit, rotate, and translate along a shorter attached truss segment. Note the Operabot cubes on the shelves, as well. (Renderings by Arjuna Imel)



Figure 36: Walls as periaktoi
(Rendering by Arjuna Imel)

through the moving doorway. Eventually, the truss was abandoned for many reasons. It was too large to suit smaller venues. It was technically complex giving the walls' motion many degrees of freedom that would be difficult to coordinate and hanging walls might not be able to resist properly the recoil of actuated books (which I'll discuss shortly). The structure also interfered with other stage rigging for lighting. Despite all of this, the two primary reasons given for the truss' demise were that it was prohibitively expensive and kept the Chandelier too far upstage. Director Paulus wanted most of the action, particularly Evvy's encounter with Simon as the Chandelier in Scene IV, to remain in the downstage third.

At the time I joined the production, a third wall design had been established. Like the original design, this approach had three groups of three walls joined at the ends and resting on the stage floor. However, now the walls were configured into triangular periaktoi with a constant internal volume. Each wall face would be 15' wide by 20' tall. The traditional-looking paneled double hinge doors were replaced with slightly narrower openings. The opening would be revealed by retracting split bookcases, inspired by a design by Rikiya Fukuda, or simply by having a return wall set back from the opening. Some renderings show the return wall flush with the front plane of the wall and moving inwards to allow entrance and egress. The periaktoi had no interior floor to allow stagehands concealed in the interior to move the walls about the stage from one blocking to the next. Also, the blocking for the production at the time had actors and Operabots making their entrances and exits by way of the periaktoi, instead of the wings, in the manner of a shell game, as the walls moved about the stage.

The scale of the periaktoi has been chosen by McDowell specifically in relation to the proportions of the human body and the size of the stage and proscenium at the Salle des Princes. However, given the visual weight of the periaktoi compared to earlier designs and considerations of their radii to avoid collisions when moving among each other and with respect to the Chandelier, it was determined that they were too large for many of the other planned venues. I worked with Technical Consultant Peter Colao to design a modular version of the periaktoi (Figure 37). The shelf lengths were visually subdivided so that, for larger venues, two additional narrow columns of shelves and one additional row could be incorporated during the assembly of the set. This yielded slightly increased wall face dimensions for larger stages and smaller dimensions for smaller venues. Ultimately, the modular expansion approach was not used and the overall size of the walls was reduced for all venues, in accordance with making the staging have a more intimate and present feel. The final wall dimensions are 12' wide by 18' tall. A consequence of the reduced size was that it would be unlikely that the Operabots could enter the walls and that the already-cramped interior space for actors and stagehands would be even smaller.

Many mechanisms for actuating the individual books on the shelves were considered. For some time, the books had been conceived of as being able to move in and out perpendicular to the plane of the wall, though concept renderings of even more elaborate versions of books with as many as seven degrees of freedom per book had been considered (Figure 39). Until engineering began on the walls, it was assumed that the mechanism would be pneumatic or hydraulic. It was determined that both of these methods presented many problems, failed to generate the subtle movements envisioned, and were far too noisy. Electromagnetic linear actuators were inefficient and the power requirements of driving thousands of books posed numerous issues in terms of power distribution, heat dissipation, and concerns for the safety of personnel and actors inside the periaktoi in close proximity to power equipment and exposed mechanisms. Drive systems were then developed that would gang books within a shelf, preventing individual addressability and some of the suggested gestures of the walls' surface, while still maintaining the ability to create high- and low-frequency patterns within each shelf. Another problem was the force of actuation of one face required several tons of counterweight inside the structure to prevent it from tipping. The periaktoi would need to have their own drive mechanisms to avoid having stagehands inside the dangerous interior. The power required to actuate and illuminate the books with internal lighting in each necessitated a tether out the top of each periaktos, which was deemed unsightly and could present shadow problems when lighting the stage. The rotational motion and lack of a base meant that the tether could not come from the bottom, though constraining the walls to a track system in a custom show floor was briefly entertained. Eventually, the decision was made to remove book actuation altogether, the implications of which for the display surface are discussed in detail below. Doing so reduced the power necessary for each periaktos and they could be made battery operated and wireless, free of their tether. The doorways were also completely removed.

The libretto presents another requirement of the walls. When Miranda summons the Miseries, they swarm the stage and destroy the walls. Early designs with the walls as freestanding faces show them breaking into pieces and being suspended in mid air above the stage, as if debris of an explosion has been stopped in time. The inspirations for this effect are the sculptural works of Cornelia Parker. During this process, the Miseries are scripted to climb over the walls, loot the memory objects on the shelves (in versions where they exist), and tear apart books producing a "controlled vortex of paper" practical effect. The theatrical destruction of the walls is a complex effect and even more so in the case of the periaktoi. Because of that, the walls no longer break apart. Only two of the faces of each periaktos are actuated and display surfaces. The third surface, which is visually identical under front lighting, is never presented full on until Scene VIII, when the Miseries storm the stage and climb the walls. This side is loaded with

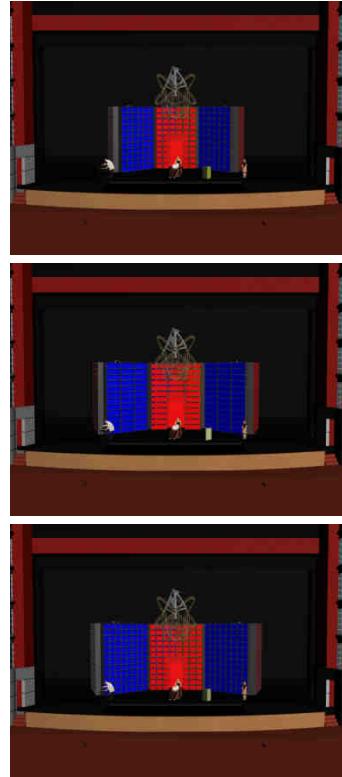


Figure 37: Wall scale study
The middle rendering depicts the planned wall size with respect to the proscenium of the Salle des Princes. The top and bottom figures illustrate the sizes of the periaktoi with the proposed modular approach.

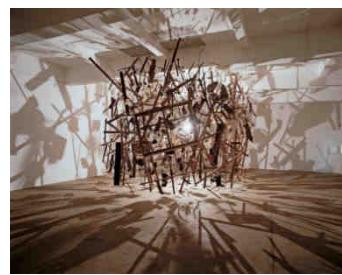


Figure 38: *Cold Dark Matter: An Exploded View*, Cornelia Parker (Photograph provided by Alex McDowell)

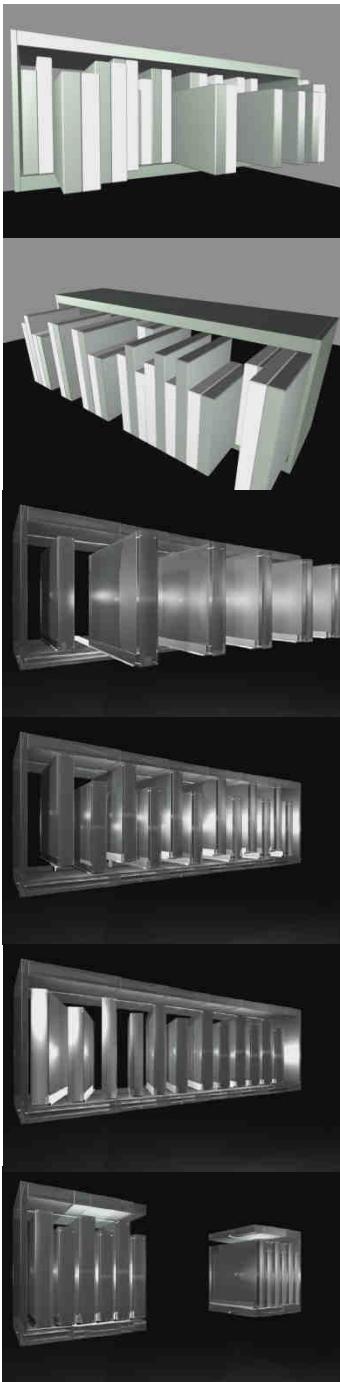


Figure 39: Book actuation

The top two images show the effect of actuating individual books moving in and out. The bottom four images illustrate a more complex actuation scheme that gave each book seven degrees of freedom. (Renderings by Arjuna Imel)

removable gimmick books containing reams of paper and other practical effects that can give the appearance of the destruction of the walls.

Continued concerns about the production's budget may lead to the elimination of the periaktoi form of the walls. I am presently exploring alternatives that would still provide an expressive means for representing The System and that would serve as a Disembodied Performance representation. Current ideas include simplifying the walls to single faces that are flown on battens and have only a few degrees of freedom, such as rotation and traveling on track in line with the batten, instead of being able to translate freely about the stage. I am also investigating methods of volumetric projection that can be used at a theatrical scale. Volumetric projection techniques would provide a completely new palette for representing sets and Simon in The System that could extend beyond or replace the library and book metaphor that so far has been integral to the production of *Death and the Powers*. A technique such as wire mapping can be difficult to implement in such large spaces as a stage, requiring precise placement of strings throughout the space and careful registration of video projectors. However, such an approach would allow for the creation of unique effects that would tie in with the matrix-of-light representation of The System, bear a visual similarity to the Chandelier, and allow for the spatial presentation of surfaces that harkens back to the undulating actuated floor, even if not tangible.

Though the design of this crucial element to the opera and to the output of the methods described in this document remains undetermined, I will continue to assume that the walls are three periaktoi with internal rear projection on the books.

3.3.4.1 DISPLAY SURFACE

As the primary representation of Simon in The System, the walls need to be capable of producing identifiable, nuanced, and expressive visual effect communicating Simon's continued presence to the audience. They must assume the role of the main character onstage. The actuation of the individual books to extend and retract has always been considered the medium in which Simon in The System would express himself. Early designs of transparent books blinking in controlled patterns over the surface of the walls added an illumination component to the physical movement (Figure 40). By the time the walls had assumed the periaktos form, they had become an elaborate display surface.

The display system was planned for the walls, incorporating custom high-resolution color LED matrices manufactured by Barco into a curved transparent resin spine on each book. When assembled side-by-side, the walls would form an image 45' across by 20' high. It was anticipated that

visually rich cinematic sequences for Memory Download would be shown on the walls. Simon would display his thoughts and reactions using a rich photographic language. While this would be an impressive canvas for the video work I would have the opportunity to work on for the production, I felt that it was a regression to the sorts of screens and technologies Machover wanted to avoid. Story-wise, I didn't feel it was appropriate to use the walls for the Memory Download sequence and I didn't think it appropriate that Simon in The System, who has renounced materiality, would express himself in photographic images. What is more, the sheer expense of producing an enormous LED display with individual elements that moved with the book actuation was far too expensive for the production's budget.

I proposed a solution that I called a *pixel-per-book* display. Instead of building an enviable high-resolution display, I suggested that we embrace the regular forms of the books, which are 2" wide by 13" to 17" tall, and provide a single color LED strip, of the sort manufactured by Philips Color Kinetics, in the spine of each. The resulting display on the smaller-scale periaktoi would be 162 pixels wide by 9 pixels high. This simplified control requirements for the display and the construction of each actuated book, as well as the cost of building such a system. However, in the end, it too proved to be too expensive when illuminating the 2,916 books of all display faces, two per periaktos, in the smaller design.

I felt that the pixel-per-book approach was unique, in contrast to high-resolution displays and projection that are increasingly common in stage productions, dance events, concerts, and urban environments. Giving a single book a color reinforced the concept of the book and the real physical nature of the display surface. To present the idea to the creative team, I prepared examples of what pixel-per-book graphics would look like at wall resolution, including photographic imagery. Even at such low resolutions, photographs and video can be presented very intelligibly, yet with a unique and unexpected quality. Nevertheless, I staunchly advocate the avoidance of photographic imagery on the walls, if possible. I believe that photographic representations can undermine the intent of The System in the context of the story, unless it is imagery that Simon in The System is deliberately conjuring to show those observing from outside of The System. Although it has been considered, I am especially opposed to showing video imagery of Simon's human form, the actor, when he is inside The System. Instead, I am in favor of a more abstract and visceral expression of Simon Powers's thoughts and emotions, as will be the visual output of the Disembodied Performance System that I will cover in the following chapters. Fortunately, even though such imagery can be shown well, the pixel-per-book approach is more of a deterrent than a high-resolution display. Though, in recent discussions with McDowell, some high-resolution imagery may be used



Figure 40: Blinking books
(Rendering by Arjuna Imel)

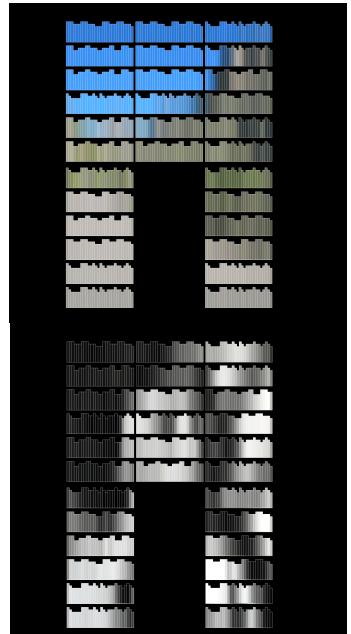


Figure 41: Pixel-per-book imagery
These tests demonstrate the legibility of photographic imagery rendered at pixel-per-book resolution, where each book is a single solid color.

toward the end of the opera to give the suggestion that the materiality of the walls is breaking down.

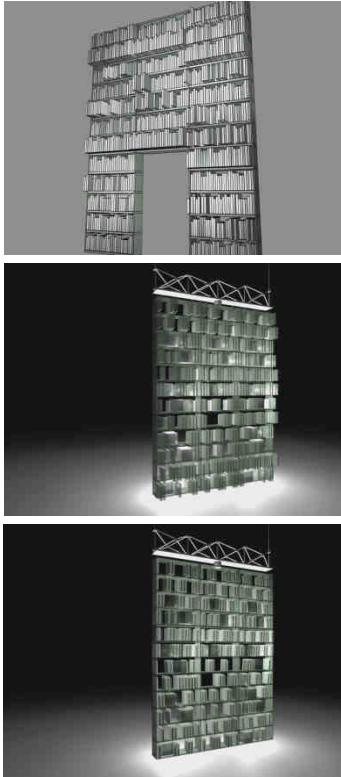


Figure 42: Wall actuation studies
Actuated books could produce sharp noise displacements (top two renderings) or subtle undulations that make the structure appear to deform (bottom). (Renderings by Arjuna Imel)

I also had a concern that the 13" throw of the actuated books may not be readily visible from typical audience distances of 40' to 120' from the walls onstage. All of the 3D animations at that time had been made from much closer distances where the motion was apparent. I conducted several studies with 3D computer models and found that the book motion was not particularly perceptible. McDowell suggested adding a second colored illumination source to the sides of the books to make their extension more visible. While this probably would have had the desired effect, it would have doubled the cost of an already expensive proposition.

A third option for book display remained: rear projection. This approach would mean the demise of actuated books, which was already being considered for the reasons mentioned above. Projection would open the door for numerous advantages, such as a cost-effective solution for a display surface, the capacity for high-resolution display if need be, and sufficiently low power requirements to achieve truly wireless unmanned periaktoi.

Using projection, especially on two faces with three projectors per face, requires that the interior volume of each periaktos be mostly empty. Actors, stagehands, and robots could not enter the walls, so the doors were removed. Consequently, the movement of the periaktoi certainly required self-locomotion. The drive systems, now based on a robotic palette lift, air casters were moved into the base of the wall. Since the walls no longer needed to be flush with the stage floor for entrances and exits as well as human locomotion, an 8" base was planned to add precious space for the drive systems and numerous batteries that would be required to power control computers, projectors, and the drive. The batteries would also provide ample ballast to lower the center of gravity of the structure for stability. These components could extend upward slightly into a skewed pyramidal volume that would not interfere with projection. To prevent projector stacks from getting in the way and to gain a bit more depth to cover each face, "bumpers" were added to the periaktos design. These curved portions filled in the 120° region bounded by the sides of adjoining walls. The bumpers could be externally illuminated and the base of the walls extended to add stability and prevent collisions with other periaktoi or objects in the concavities.

Even though I was proposing a projection-based solution, I was adamant about retaining the physical quality of the books. I did not want the shelves to appear very different from their original actuated counterparts. The spines would remain curved and the books would have depth that would be visible, especially under front lighting. To accommodate the rear projection, I redesigned the shelves so that adjacent books were contiguous without

internal divisions and reduced the book depth to 4" from 14". The area surrounding the books would be painted a flat black to become a visual void and enhance the apparent depth. The depth reduction and elimination of dividers was essential to minimize self-shadowing. Since the image being projected onto the back surface of the spines is not incident to the surface at its normal, some areas of shadow are inevitable. Carefully arranging the patterns of book heights can minimize the shadows due to variations in the books.

Video rear projection does bring with it some disadvantages aside from shadowing. The only suitably-sized projectors for this application are not extremely bright, and fall slightly short of the intensity of emissive LED strips in each book. The choice of book spine material would be critical to color reproduction, balancing light transmission, maintaining decent black levels under front light, and the structural integrity required for the spines to retain their form. We have conducted experiments with several materials including sandblasted vacuum-formed PETG, rear-projection materials, and gel diffusion. An important constraint is that the projector must be securely mounted and precisely registered with the book projections for accurate pixel-per-book imagery and to prevent the image from moving independently of the projection surface, spoiling the impression of glowing books.

To convince the creative team of the projection approach, a two-shelf mockup was constructed and projected on. The tests were very successful. Both pixel-per-book and high-resolution imagery was displayed on the book in no light and front-lit conditions. I also demonstrated that the vocabulary of the books moving in and out could be maintained using a visual illusion in the projected pixel-per-book imagery, even if they were not physically actuated. In fact, a side-by-side comparison was made between projected book movement and physical LED-illuminated books being extended and retracted by a distance of 13". The projected book movement was readily perceptible at great distances while the physical movement was difficult to discern from an orthogonal viewpoint as close as 15'.

Most illusions of depth are dependent on viewing angle. This is a challenge given that audience sightlines are incident with the walls from a fairly large angular spread. On top of that, the walls move and rotate, so the theoretical viewing angle must change in response to wall motion with respect to the audience and as the wall moves through the lighting in the space. There are some less physically accurate illusions of depth that can be created that are reasonably invariant under viewing angle. One of these techniques involves changing only the scale of a book proportionately without shadowing. This results in the appearance of glowing books moving inward instead of



Figure 43: Wall projection prototype

The projection prototype used rear-projection screen for the spine of each book. In this photograph, the 4" depth of the books can be seen.

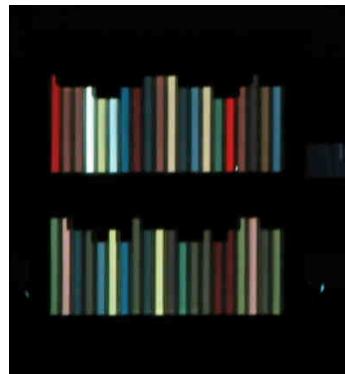


Figure 44: Wall projection test
Rear-projection of 'book colors' onto the curved spine surfaces proved successful.

extending past the frame of the bookcase and was used in the projection tests.

The successful demonstration of the effectiveness and possibilities of rear projecting the book display in the periaktoi led to the current doorless, wireless, unactuated design.

3.3.4.2 SOUND SPATIALIZATION

In keeping with the unity of sound and visual envisioned for the production, the walls would need to not only move expressively, but make sound as well. Given the importance of the actuated and illuminated books in representing Simon in The System, Machover felt that the books should also be musical. He envisioned that each book would have a unique sound that corresponded immediately with its movement and that this per-book linkage would be apparent to the audience. Since acoustic sounds are generally easier to localize, he imagined that some sort of plectrum or percussive instrument would be integrated into the actuation mechanism for the moving books. Short of this, he considered that each book could be outfit with a tiny speaker inside, though this would impact the use of books as a display surface.

In January of 2008, tests of localizing sound on the walls were conducted. Working with representatives from Bowers & Wilkins, a British manufacturer of loudspeaker technologies, an approach was devised that required as few drivers and minimal complexity as possible. A reduced-scale 10' × 15' wall face was erected and covered with an acoustically transparent projection screen. On variable-height platforms behind this screen were arranged several configurations of midrange and tweeter drivers. Low range frequencies are not as localizable, so a single subwoofer and two bass drivers were placed in fixed locations behind the screen. Due to the actuation of books and their use as a display, in the final implementation, the speaker drivers would be embedded in the supporting struts of the shelves or hidden in the empty spaces above the books within a shelf.



Figure 45: Wall speaker cluster
These mid-range and tweeter clusters were positioned behind the acoustically transparent screen for the sound spatialization tests of the wall prototype.

I developed an experimental control system that would allow books on the wall to be played. Each book was assigned a color and a sound sample. Additionally, the user specified the locations of the speakers relative to the wall using a graphical interface or with numerical measures. The wall could then be played using a MIDI keyboard connected to the control system, by clicking or drawing patterns of books on the graphical representation with a Wacom Cintiq display and stylus, or by playing back a MIDI file. A visual representation of the books was projected onto the wall mockup, physically registered in accordance with the computer model.

Reflecting the idea of a pixel-per-book display, as a book is played, the book is highlighted in its color both in the graphical interface and the projection onto the wall. Playing a book generates a MIDI event on a particular channel. Up to 16 books can be played simultaneously in this implementation, one on each channel at a time. The MIDI messages are sent from the control computer to a second computer via a MIDI-over-IP protocol. Native Instruments' software sampler, Kontakt, runs on the second computer that generates a sound at the system's optical audio interface on a channel corresponding to the MIDI channel. The control system generates this by computing a Delaunay triangulation of the speaker arrangement and uses barycentric coordinates to determine the contribution of speakers in reproducing the localized sound. This contribution value for each speaker is sent via proprietary MIDI messages to a Yamaha LS9 audio mixing desk, which uses matrix routing to set the volume level of each output channel, corresponding to the speaker arrangement on the wall, for each audio input based on the location of the book represented. The output signals of the LS9 pass through the second computer, once again, for DSP and crossover before being output to amplifiers driving the speakers.

The results of these experiments clearly demonstrated that it was in fact possible to localize sound on the surface of a wall, particularly without relying on physical mechanisms of actuation. What is more, at a scale similar to the final set piece and at distances proportional to where an audience would be in the premiere venue, adequate vertical and horizontal resolution of sound could be achieved with as few as nine midrange/tweeter assemblies arranged in a rectangular grid. Books did not need to move to produce sound, so additional actuation mechanisms were needed. The sound palette could be limitless by using live audio or samples as sources. Individual books did not require their own small speakers. The projected books were 1.5" and roughly 13" in height with the sound sources computed at the center of each book. Of the twelve rows of books displayed, the row on which a sound sample was to originate could be identified and the horizontal location could be perceived with a resolution of a few books. It was found that perception of spatialization was enhanced with the listener's eyes closed. A sense of localization was greatly improved by highlighting a book simultaneously with its sounding, even in the case of multiple books being played at one time, as opposed to an absence of visual feedback. While these tests proved highly successful at creating the desired effect with an economy of audio infrastructure, it may be the case that a dedicated sound reproduction system is not needed with the installation of an ambisonic or wave field synthesis system for the entire production.

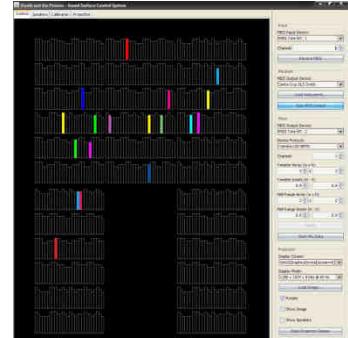


Figure 46:Wall sonic prototype control software

This screenshot shows the control software while the wall is being played. The software computes the sound source and signal routing levels, which are sent to a mixing board and another computer to produce the audio.

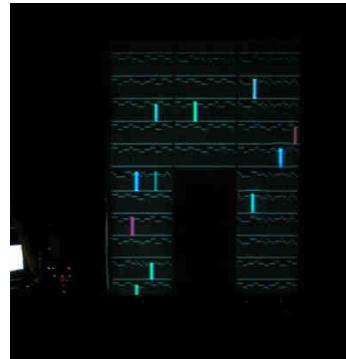


Figure 47:Wall sonic prototype being played

In this photograph, the 10' × 15' wall is being played. Each illuminated book corresponds to a sound localized to that book by an array of speakers behind the wall.

3.3.5 Conceptualizing The System

Using walls lined with bookshelves as the setting for the interior of the Powers household clearly invokes the notion of a library. Tying this to the primary representation of The System, the books represent Simon Powers's memory, his knowledge, and his legacy. Some stage directions in the libretto and production notes have Miranda as the keeper and custodian of the household and particularly the walls. Some early designs have shown not only books on the shelves, but the cube-shaped Operabots and "memory objects" that are supposed to have special significance to Simon and to which Miranda is attached after he has gone. The books and items on the shelves are precious because they become all that physically remains of Simon. The walls are the machinery and memory banks of The System. Their regular vertical patterns are similar to the design of the HAL 9000 in Stanley Kubrick's *2001: A Space Odyssey* (1968) and other computer representations in television and film. They resemble a computer mainframe and exhibit a visual quality of blinking lights common to the stereotype.



Figure 48: HAL 9000 in *2001: A*

Space Odyssey

(Image from the film)

Considering this dramaturgical perspective has been an important influence on my conception of the representation of Simon in The System. A strong proponent of pixel-per-book and non-photographic imagery, I wanted to develop a rationale for the imagery I would create to represent Simon's presence. In the world of the story, how do the walls work and what is the meaning of what we see displayed on them? Are the walls something like phantom limbs that amputees experience or the tremendously pliant visual canvas that the blind report [75]? Can we see Simon's thought processes? Is there some discernable Von Neumann architecture with unique areas for input, output, memory, and processing spread across many set pieces or individual walls? Do the patterns of illumination reveal complex data structures or patterns of neurological activation? In July of 2008, I wrote this preliminary description of how The System works in the world of the story in order to inform my subsequent designs:

The System: A Technical Explanation

Functional chunks of memory are stored in each book. Thus, a book may store some level of abstraction over a class of ideas: Evvy, research projects, foods. Whole thoughts are comprised by activations among all books containing concepts for that thought. The Evvy book and the food book might activate to recall a memorable meal with Evvy. Most concepts would entail many more activations. Activations may vary in intensity with the most salient concepts of a memory or thought. Activation patterns may change as specific details are recalled once a specific memory is brought forth. Books storing concepts may be arranged in the bookcases by some ordering of similarity or conceptual relation, revealing a hierarchy or

structured pattern to activations. Large, intimately familiar concepts, such as Evvy or Miranda or Nicholas, may have their representations spread over a larger area of several books to contain the necessary breadth of stored information.

Books also are a tangible and emotive skin for The System. These responses are typically independent of memory and thought activation, though may induce in Simon a memory or thought. These responses look noticeably different, as there is an apparent geographical response. These behaviors are often not deliberate. They are gestural or like facial expressions: unintentionally communicative. Responses to interaction are autonomic, like jerking away from a hot surface or having goose bumps when hairs are lightly brushed.

Despite having a distinct mechanism for storing memory, The System does not have a von Neumann architecture. There are specialized components and pathways, but processing takes place in the same medium as memory through patterns of activation, much like the brain. Additionally, the physical structure of this memory is also part of the I/O mechanism.

There is an initial learning phase and a secondary learning phase. While entering The System (S2), The System learns about the individual as the individual's essence or prolonged consciousness learns about sensing and behaving inside The System. During the secondary phase, there is still a central element of consciousness that is almost embodied in a single place, but its presence is subtle. At some point, The System sublimates. There is no longer the need for physical storage of memory or processing. The individual again exists, but purely in the Matrix of Light, not in The System proper.

With some grounding for the types of visuals on the walls, I will later discuss in the next chapter the manner in which they are generated and how they actually convey a performance.

3.4 Control Systems

Beginning in the summer of 2008, I took on the task of designing a unified show control system for the opera. At the time, the system would be responsible for a new generation of Operabots, the movement and musicality of the Chandelier, and the movement, actuation, and pixel-per-book illumination of the periaktoi. It was expected that other elements, such as furniture or moving set pieces, would be added to the technologies and components being implemented at the Media Laboratory, so the

system would need to adapt easily to any such additions and changes in plans. It was conceivable that such a show control system might integrate the theatrical lighting or, at least, be able to communicate with the lighting console that would ultimately be selected. The advantages of such a unified system would be great. Light and motion could be designed holistically, with elements and gestures easily moving throughout all aspects of the set in a seamless manner. The appeal is especially significant considering the representation of Simon in *The System* is to be omnipotent and omnipresent. This section sketches, without getting into a detailed explanation, how the control system was conceived, as it has influenced the design of the Disembodied Performance System.

One of the goals for the unified system was to incorporate high levels of autonomy and interactive behavior. It was to be responsible for dynamic movements reflecting Simon in *The System*, such as complex patterns of book actuation in the walls. Reproducing the choreographic functionality of the first-generation Operabot controls was a requirement and automatic collision avoidance, as had been beautifully achieved in the original system, would ideally be incorporated. The original system was not robust enough in implementation and hardware to be incorporated directly and its creators were no longer working on the project.

The unified control system needed to be reliable enough to safely operate numerous robotic elements of varying scales. High degrees of precision in spatial movement and timing would be required to keep swarms of robots from colliding with each other, other set pieces, or actors. Redundancies and many levels of safety features would be essential to keep robots or the periaktoi from running into moving actors, rolling off the edge of the stage, or failing while obstructing the fire curtain at the theater's proscenium. Critical to this aim is the capacity for high-accuracy dead-reckoning. The location of everything onstage, including actors, would need to be known at all times. For some of the planned choreography of the robots and the sets, this would require centimeter tolerances. Unfortunately, solutions for accomplishing this were deemed prohibitively expensive and many addressed only planar tracking. At the scale of a theatrical stage with elements that vary in elevation, perspective parallax would greatly reduce accuracy. Full three-dimensional tracking would be needed. Additionally, most solutions were optical, including the Vicon system that was used for the first generation Operabots and similar systems by other manufacturers. Aside from the expense, I had concerns about using optical systems, particularly with reflective markers rather than active coded markers, under theatrical lighting. Requirements for line-of-sight between tracking targets and cameras could not be guaranteed with the large walls and robots that elevate and remain close to each other. Constraints on rigging would also be an issue. Relative positioning would have to be used, relying on motor encoders and the control system to make best guesses about the position of

elements. Until the accuracy of the control system and robotics could be determined, it was decided that it would be necessary to keep a human in the loop, especially for all potentially hazardous motion, such as translation onstage.

The system would be responsible for many elements. Managing the large number of robotic and illumination output devices would be a challenge. At this time, we were planning to have 36 Operabots with as many as four or five degrees of freedom in movement and articulation and several channels of lighting. Three periaktoi with all three sides actuated (as the design stood at the time with one degree of freedom per book) would have had 9,072 individual drives. All told, it was estimated that there would be 93,742 control channels required to run the production.

With this in mind, I began to design a system architecture that would allow all of these disparate robotic and illumination devices to coexist and exchange information. The outputs would potentially be controlled by a number of live inputs from what would become Disembodied Performance and for puppeteering of Operabots, not merely preprogrammed animations. The basis of the architecture is that both input and output devices are duals of each other and would be treated similarly by the system. Both inputs and outputs are *devices* that have one or more *axes*. For example, an input device may be a joystick with a forward-backward axis and a left-right axis. Correspondingly, an output device may be a robot that has two degrees of freedom or two axes: the ability to move forward and backward and the ability to turn left or right. The axes of input devices are connected or *mapped* to the axes of output devices. In our joystick and robot example, this mapping is rather straightforward: the forward-backward axis of the joystick is linked to the drive of the robot and the left-right joystick axis is mapped to the rotation of the robot.

Of course, merely passing the value of an input axis to an output axis is not likely to yield expected results because these values may have different ranges and interpretations. A joystick may produce an integer value with 10 bits of resolution while a motor controller may expect a value having only 7 bits of resolution. This communication was simplified by providing a common interface for devices and a lingua franca for allowing components to interact. Consequently, the control system wraps each device in an abstraction layer that is specific to the device hardware. This layer knows how to communicate directly with the device and the ranges and types of values each axis produces or consumes. All mappings are accomplished by communicating with a device's interface. The lingua franca for the mappings is to represent all values as normalized floating-point numbers. The value of 0 is an axis' home or off value. A value of magnitude 1 is the maximum value of that axis. For some axes, the range of values may be $[-1,1]$ instead of $[0,1]$. Axes may be declared *absolute* or *relative*. Absolute

axes have a value that corresponds to an exact position, such as a slider on an input or a linear actuator or light intensity on an output. Relative axes specify a value with respect to the current position for *unbounded* axes, such as the position of a robot onstage. Relative values express a rate of change of the unbounded value. In the case of a robot's drive system, this value is proportional to the robot's velocity and may be negative to allow backward motion.

Given the theatrical context for this show control system, we originally planned to implement all control using the traditional DMX and MIDI protocols with off-the-shelf components. Wireless DMX and DMX-over-IP solutions, such as ArtNet, were considered to easily distribute the numerous control signals the show would require. As I began to design the control system, it became clear that a newer IP-based protocol would greatly simplify the system architecture and the amount of hardware needed on robotic set pieces. Wi-Fi antennae and routers provide a lower-cost solution than wireless DMX transmitters and receivers and DMX USB or DMX Ethernet interfaces. I settled on using the Open Sound Control (OSC) protocol, developed at CNMAT, over UDP as the protocol for all output devices. Each output device would be addressed using its IP address and OSC addresses would target sub-devices, device axes, and other methods for querying or configuring devices.

The implementation of this system became known as *Core*. I provided several of my undergraduate workers with a clear description of the system architecture and worked with them to implement Core in Java 5.

Core was heavily inspired by the Architecture for Control Networks (ACN) protocol suite for theatrical control being developed by the Entertainment Services and Technology Association. ACN provides a number of very useful features and is a standard intended for controlling technologies within the theater, so it would have been an ideal choice, particularly for a production that is expected to tour and be run by theater staff [35]. However, the ACN suite is rather complex. With an approaching premiere, then scheduled for September of 2009, I did not feel that my team or myself could afford to take the time to implement ACN. An open source implementation of the protocol suite, OpenACN, is being developed, but we deemed it too incomplete and unstable to use at this time [60]. Nevertheless, I have attempted to preserve many of the desirable features ACN affords in the implementation of Core, such as IP-based communications, device description files and managed configurations, and automatic discovery of devices on the network [26].

In Core, each device has an associated XML *device description* file that enumerates its axes and the properties of each axis so that the system knows how to interpret and write values to them. Additionally, XML files are used

to manage *shows*, which are assemblies of devices with their addresses, *sequences*, and *cues* for a specific production. Each show begins in a default cue. The transition to other cues is managed by an input device trigger. Cues maintain the mappings of input device axes to output device axes. Sequences are fragments of keyframe or procedural animations that can also be mapped to output axes. Keyframes specify the value of an axis at a given point in time relative to the duration of the sequence. Keyframes can be interpolated in various ways, much like their implementation in most computer animation software. Procedural animation sequences produce a value for an axis based on a mathematical expression that is a function of time or the value of some other axis, such as the position of a robot. In this way, behaviors such as mirroring, mimicking (slaving), and look-at constraints can be achieved. The current implementation is not presently robust enough to produce flocking or collision-avoidance behaviors.

As part of Core, we implemented a non-linear animation system similar to those found in 3D modeling and animation software. Non-linear animation allows multiple animations to be mixed and blended in different ways. The output axis value of a non-linear animation is a logical operation on the value at that time given by the *tracks* of keyframe sequences, procedural animation, or live input mapped to that output. Axes are controlled independently, so animations can be carried out on lighting channels while drive systems are mapped to live joystick input. However, animation sequences can also be blended with the joystick data. A puppeteer may drive a robot in a straight trajectory, but that input can be modulated by an oscillation that would cause the robot to take a serpentine path. Altogether, complex animations can be built up by layering simple motifs that are expressive or goal-directed. Individual robots can have recurring motifs that give all of that specific robot's behaviors a unique characterization. Although much of this functionality has been implemented in Core, it is not yet exposed through an authoring environment interface, which will be the subject of future development on Core.

3.4.1 Operabot Control

I also worked with the current Operabot engineering team to specify and design the control systems onboard each Operabot. The final Operabot design has a two-wheel drive system controlled by a Roboteq AX2850 motor controller and is capable of moving forward, backward, and turning in place. The vertical elevation and head articulation are handled by a smaller Roboteq AX1500 motor controller. All drive trains are fitted with optical encoders and the two actuation systems have limit switches at their travel extents. The 11 channels of lighting on each robot are powered using LuxDrive 3023 BuckPuck constant current drivers controlled by 8-bit PWM signals from an Atmel AVR ATmega1280-based Arduino Mega microcontroller. Both motor controllers and the Arduino Mega are

controlled via serial connections over USB from a One Laptop per Child XO laptop. After exploring several options, I decided on the XO laptop because it provides several unique advantages: it is capable of running a full Linux operating system, is extremely power efficient, is extremely tolerant to supply power without the need for special batteries and voltage regulation circuitry, natively supports wireless communication and mesh networking for robust communications (though this feature is not currently being used), and has a compact form factor for our space-limited application.

Wireless communication is always a concern in theatrical venues. The physical properties and materials used in the construction of such spaces often pose problems for signal propagation. For an internationally touring show, such as *Death and the Powers*, few assumptions can be made about the availability of radio spectrum. Analog wireless microphones can occupy 470 MHz to 854 MHz while digital microphones can occupy 900 MHz or 2.4 GHz, varying by country and regional spectrum allocation. The XBee modules used for wireless props and sensors in the production communicate on 2.4 GHz (see 4.5.1). Thus, to avoid crowding these bands, which would negatively impact the reliability of critical time-sensitive communications with robotics equipment, wireless communication with the Operabots and walls is handled using 802.11n at 5.6 GHz. Since the XO's internal wireless antenna is for 802.11g communication (at 2.4 GHz), an external USB 802.11n wireless antenna was added.

I replaced the default operating system shipped with the XO laptop with an OLPC-modified Debian Linux distribution and the recommended RedHat kernel. I then wrote the on-board control software, called XOBot, in Java 6. XOBot manages the dispatching of commands received wirelessly from the Core system, translating the OSC messages into serial commands to the motor controllers and lighting controllers. The XOBot program also implements additional safety features that ensure that the robot enters a secured state in the event of communications loss or device failure. For example, if the Operabot fails to receive updates from the Core system in a certain short period of time (approximately one second), it will stop translating, which could be potentially dangerous when not controlled. Behaviors that may not cause injury to the robot or humans, such as animation of lighting, are not stopped in the event of a communications loss. Therefore, if it is possible to do so, the robot does not appear to die completely in the middle of a performance. XOBot reports the status of the robot's systems—battery power, encoder data, motor controller failure responses, communications signal strength—back to the Core system for monitoring by operators. This is accomplished using Java's native logging mechanism, for which I implemented a log handler to transmit log

messages as OSC packets back to the Core host computer, affording a variable level of feedback granularity for testing and performance scenarios.

As noted above, for safety reasons, Operabot translation is almost always under the control of a human. Several offstage operators will have joysticks in order to control the movements of a group of robots at any given point. The joystick also can be configured to allow the other parameters of head tilt, elevation, and illumination to be puppeteered or to have animation sequences modulated by the operator. Since the number of operators is fewer than the planned number of Operabots, only as many robots as there are operators will translate at a given point. Unless the Operabots demonstrate sufficiently accurate and reproducible ability to follow pre-programmed choreography in future tests with multiple Operabots, rarely if ever will Operabots translate autonomously. Operators will have the ability to trigger cue changes in the Core system that will remap their controls to a different Operabot. When not being puppeteered, Operabots can execute programmed animation sequences on non-critical axes so the entire population of Operabots always appears to be alive and engaged. Procedural effects, such as a look-at constraint, can allow non-puppeteered Operabots to rotate in response to Operabots that are being controlled.

Even though Operabot translation is puppeteered, the system will still need to know the location of each robot onstage. Encoder feedback will provide a mechanism to accomplish this, though it is not entirely reliable. When Operabots drift from the location that the Core system thinks they occupy, human intervention is needed. A camera located above the stage will capture a live view of all of the elements in the play area. This view is provided on a large display to Operabot operators. Overlaid on this display is an iconic representation of all of the Operabots tracked by the Core system. Visual indicators provide feedback about the current state of the Operabots' controls and on-board systems to operators. When an operator observes a discrepancy between the tracked location of an Operabot and its physical location, he or she can use a mouse or stylus to correct the system's location estimate by dragging the icon for that Operabot into alignment with its image.

3.4.2 Wall and Chandelier Control

When the decisions were made to not actuate the books on the walls and to use video projection instead of LED strips to illuminate each book, the system controlling the wall visuals had new requirements that were more readily achieved by means other than the current architecture of the unified control system. Rendering imagery to a raster that could be directly displayed by a computer connected to a video projector inside a periaktos was more natural than controlling thousands of axes. Also around the time of these decisions, the current conception of Disembodied Performance began to take shape. The mapping requirements for the Disembodied

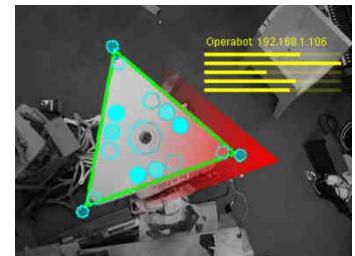


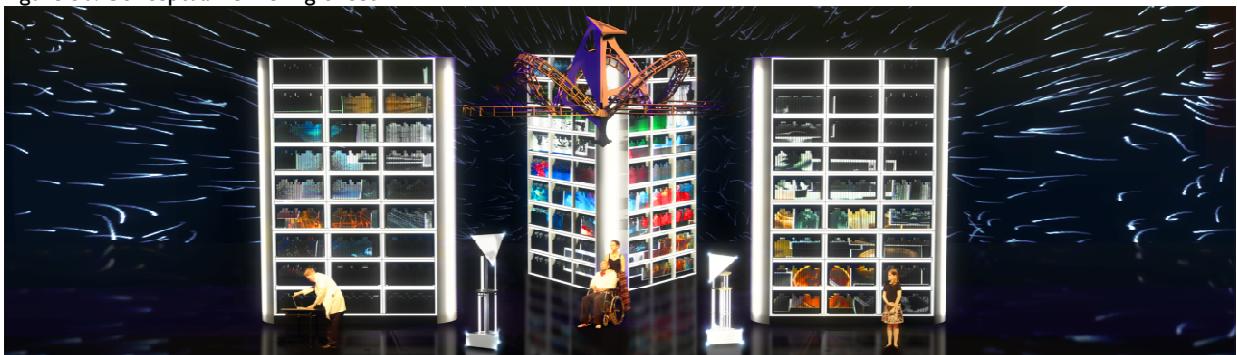
Figure 49: Operabot control view
On an overhead image of the stage, the Core system overlays the status information and position icon for each Operabot. This view allows operators to monitor the robots and correct position and orientation discrepancies.

Performance System were also different from—a superset of—those in Core. As a result, my work on the systems for the walls began and the architecture diverged from that of Core. Many of the concepts were maintained in the Disembodied Performance System and were independently reimplemented, as described in Section 4.8. Perhaps, as both systems mature, they will be reintegrated, providing a single robust and powerful show control system.

Although Core is capable of driving the Chandelier, my colleague Andy Cavatorta, who is now the lead developer of the Chandelier, has begun to independently develop a different control system for the instrument and for musical robotics, in general. Cavatorta's system is less general and follows a similar architecture and functionality to Core, but enforces some musical semantics on the exchange of data and provides a convenient web-based interface for managing robotic installations. Significant implementation differences make it unlikely that elements of Cavatorta's system will be integrated with Core, though the two systems can communicate with each other fairly easily. At present, with separate systems for rendering Disembodied Performance and playing the Chandelier, only the Operabots remain running on Core.

The challenge remains to capture a dynamic performance and reinterpret it in the environment. It is important that the audience have a profound sense of Simon's presence after his entry into The System. Simon in this new form must be capable of emoting, communicating, and interacting with the other players onstage. My solution to this task, The Disembodied Performance System, is presented in the next chapter.

Figure 50: Conceptual rendering of set



4 THE DISEMBODIED PERFORMANCE SYSTEM

*I looked for the movement, the vibration,
Not the matter, the system!*

—Simon Powers, *Death and the Powers*

The Disembodied Performance System is a hardware and software implementation of a show control system for *Death and the Powers*. Unlike the unified control system described in the previous chapter, the Disembodied Performance System specifically handles the translation of an offstage performance into its onstage representation. It coexists with other control systems and has the capability of integrating with them, as required.

The intent of the Disembodied Performance System is to create a framework and environment to facilitate the creation of a remote expressive performance that can be mapped into an arbitrary form. Its design and structure reflect the theory of Disembodied Performance, allowing feature representations to easily be recovered and translated across modalities.

In this chapter, I will provide an overview of the Disembodied Performance System, indicating how it reflects the notions of performance abstraction and inference of a character model. Along the way, I will point out architectural decisions that were made to satisfy the production requirements of *Death and the Powers*. I will then continue by explaining the sensing hardware developed specifically for the character of Simon Powers in The System, as well as a description of the software implementation of this system and its role in a theatrical production.

4.1 Pay No Attention to that Man Behind the Curtain

One question that has plagued the design of the production until recently is: Where is the actor playing Simon Powers once Simon enters The System? The title of this section, a quote from 1939 motion picture *The Wizard of Oz*, is as good an admonition as any. Some members of the creative team posited that the actor needed to be present somewhere in view of the audience. Not only would the actor still be visible to communicate his performance to the audience in a traditional manner, having him physically there would demonstrate that he is still in fact performing. It is feared that having the actor out of view might lead the audience to believe that he is off somewhere and that what they see and hear onstage is a recording. Showing live video of the actor offstage, perhaps on the display surfaces of the walls, would be a familiar approach. Other suggestions included: keeping the actor onstage, but such that the other characters no longer see him; having the actor suspended at the top of one of the walls onstage, again unable to be seen by the other characters; placing the actor in a box in the theater house. One idea that gained a fair amount of traction was to have the actor enter into the orchestra pit and perform, perhaps illuminated, alongside the musicians. The orchestra pit, just outside of the diegesis, would become The System: the metaphysical realm that Simon enters. The play on layers of reality led so far as the idea of having the pit elevator raised as Simon is about to enter, bringing the instrumentalists and conductor to the level of the stage, before it descended with Simon on board.

In September of 2007, at a preview of *Death and the Powers* held at a benefit for the California State Long Beach University Art Museum, James Maddalena performed new music from *Death and the Powers*, namely the end of Scene I and Scene II, in which Simon enters The System. The presentation was simple, as previews typically are, with Maddalena singing at a podium with some video and visuals projected on a screen behind him. The performance was well-received and, at least to my surprise, some audience members noted that having the actor there in front of them singing after the character has died was eerily compelling. This response gave further merit to the notion of keeping the actor visible to the audience even though the character is deceased and has gone into The System.

Around this time, I began working on the production. Despite all of the arguments in favor of keeping the actor visible, I found this to be incommensurate with the story. Simon's argument for entering The System is one of abandoning corporeality. We see, throughout the opera, the struggle between the material and the realm of energy or formless presence. To my mind, seeing the actor's human body after he enters The System—whether that is having the actor onstage or in the pit or shown in video

projections, even in extreme close-up—undermines the very dichotomy the story probes.



Figure 51: Human silhouette
(Photograph provided by Alex McDowell)



Figure 52: Pit box animation
This animation by the author demonstrates how the silhouette of the actor can change shape and how the actor's gesture can be extended across the stage.



Figure 53: Simon Powers in a pit box
This conceptual rendering of the set shows an early idea for a pit box in which the actor playing Simon Powers would be backlit from inside a frosted Plexiglas construction protruding from the stage floor.
(Rendering by Arjuna Imel)

In thinking about where the actor should be located, as a compromise I sought to find a way to maintain the actor's presence to the audience while obscuring the troublesome literal quality of having the actor somewhere onstage. A photograph in Alex McDowell's image bank for the production inspired my answer. One of the first conceptual animations I created for the production showed a frosted Plexiglas screen embedded into the stage. Inside would be the actor back lit with diffuse lighting that could change color appropriately. I later discovered in archived production notes that a similar idea had been considered at one point by the creative team. Both treatments were informally called *pit boxes*. In the original design, the box protruded outward and upward from the stage, whereas in my design, a single translucent sheet was cut into the stage on an angle such that the projected shadow on its surface would be visible from any sightline. In both cases, custom decking over the orchestra pit would likely be required. In my concept animation, I demonstrated how the actor could move, as part of his performance, toward or away from the translucent Plexiglas to change the quality of his silhouette from an amorphous blob to a decidedly human form. I also tied the color of the backlight to the color of the wall displays, creating a clear visual link between the undeniably live presence of the silhouetted actor and the entirety of The System.

Simon in The System's influence was described as pervasive throughout the Powers home and in the very environment. By this point, several aspects of the original production design that would physically allow Simon's omnipresence to be demonstrated, such as an animatronic floor, had already been cut from the production. I wanted to extend Simon's influence beyond the confines of the pit box in a more direct way than the playback of animation and imagery on the walls that was envisioned at the time. The concept animation of the pit box began to get at this idea of extending the actor's performance throughout the space in a very simple and straightforward manner. In it, I demonstrated that a video camera could be trained on the illuminated pit box to capture the expressive motion of the silhouette. Then, with some basic image processing and distortion, the captured image could be projected onto the whole set, from stage floor to props. A shadowy breathing blob would imbue the whole set with an organic pulse. A flail of an arm would extend outward across the stage. It is in this germ of an idea for directly extending the actor's visual performance in an intelligible way that Disembodied Performance was born.

Subsequent ideas attempted to use the captured image of the actor's performance in the pit box to drive other parts of the set, not merely projection. For example, computer vision techniques could provide a morphological analysis of the shape of the actor's silhouette that could then

be used to actuate the books in a particular pattern. Previous to this, I had considered embedding sensors, such as proximity sensors, into the walls at regular intervals so that they would automatically respond to the action of other characters, modulating pre-recorded or procedural graphics animations on their display surfaces. These approaches are reminiscent of numerous interactive display, such as Daniel Rozin's *Wooden Mirror* (1999), J. Meejin Yoon's *Low Rez / Hi Fi* (2005), or the work of Zachary Booth Simpson, which caught media attention in the early 2000s. A small test screen was constructed that combined both computer vision and simple sensing techniques. While the wall representation did appear responsive, there was no sense of presence, only of states and elementary reactions.

Another early experiment in creating a visual language for the wall displays revealed that basic semantic representations of Simon's cognitive state—for example, directly assigning referents and concepts to individual books as signifiers (see 3.3.5)—are neither very compelling nor readily intelligible. The text of the libretto and stage directions were used to generate the lexicon for these mappings. A cued playback system would synchronize the wall visuals with the singing and action. Unfortunately, these purely semantic mappings alone resembled random patterns, if the viewer had not been informed of the semantic grounding or observed the behavior for long enough to learn patterns that related to concepts in the text.

Apart from lacking in visual expressivity, in both the reactive sensing and semantic mapping cases, it did not make sense that the computer system controlling the walls would make the simple reactions as a representation of the main character. In order to properly tell the story and to be a convincing representation of the main character that could sustain a 90-minute production, the representations of Simon in The System have to be driven by the actor. The walls themselves or the show control systems that run them are not the intelligence that can portray the character of Simon Powers. The actor is. The capture methods and mappings used have to be rich enough to support the expressive equivalent of the performer onstage and the output representations need to be equally as dynamic and have ample degrees of freedom of representation. Indeed, research on facial expressions and body language, color symbolism, and contour perception must greatly influence the mappings that will be used. With a system capable of this, I believe the immediacy of the actor's performance and the compelling power of the onstage representation no longer requires the actor to be onstage to prove that the performance was genuine or to make up for the lack of expressivity elsewhere. This is the idea of Disembodied Performance. This is the concept for the system I will now explain; a system that is architected to be capable of faithfully translating the performance of an actor into any onstage representation. Perhaps the epigraph for Chapter 2 regarding the integrity of a performance should read for our purposes, "We do off stage things that are supposed to happen on."



Figure 54: *Wooden Mirror*, Daniel Rozin

(Photograph provided by Alex McDowell)

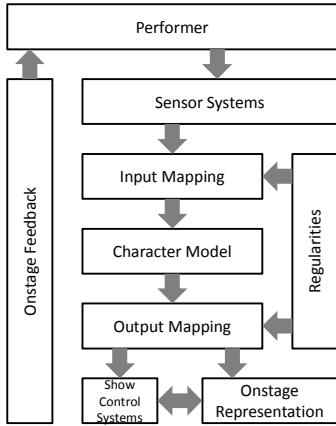


Figure 55: Disembodied Performance System overview diagram

4.2 System Overview

The Disembodied Performance System expands on the concept of reified inference that I introduced in Chapter 2. Its architecture can be segmented into four parts. The first part, *performance capture*, uses several sensing modalities to acquire the performance of the offstage actor as an input representation. The next phase, *character modeling*, uses mappings to translate the input representation into a semantic model of the character's cognitive and affective state. Here, the character model serves as an intermediate representation of the generating process, which is the supposed psyche of the character as portrayed by the actor. Subsequent mappings from the model lead to *output representations* that use the model data to generate new perceptible representations of the character. A final component provides *feedback* to the actor about the action onstage, so that his performance is not made in isolation of essential context and cues. In the sections that follow, I will explain each of these segments in detail and remark on their specific implementations for *Death and the Powers*.

This system architecture can be viewed as an instantiation of the ubiquitous model-view-controller design pattern used in computer science. The intermediate representation appropriately provides the model component. The performance capture and actor himself are the controller and the output representations are the multiplicity of views of the model. This effectively separates the content from the representation in accordance with the goal of abstracting performance from its form. Considering the system in this terminology indicates that there is an inherent modular design. Each of the components can take many possible forms and, as we'll see below, all of those forms can interact through a common interface.

In the detailed system diagram (Figure 56), it can be seen that the model-view-controller design does play a role in segmenting the system for practical implementation, not merely conceptually. Because of this, wireless connections provide the interfaces for inputs and outputs to the modeling process. All processing and modeling occurs within a single computing system that is responsible for the mapping to and from the intermediate representation and ensuring that regularities in the inputs and outputs are preserved. It is also important to note the sources of metadata, such as show-wide timing information and the state of other related show control systems, that can control global parameters of the mapping process and output representations.

4.3 Data Categories

From the system diagram, four general categories of data stream can be identified in the Disembodied Performance System: modeled data,

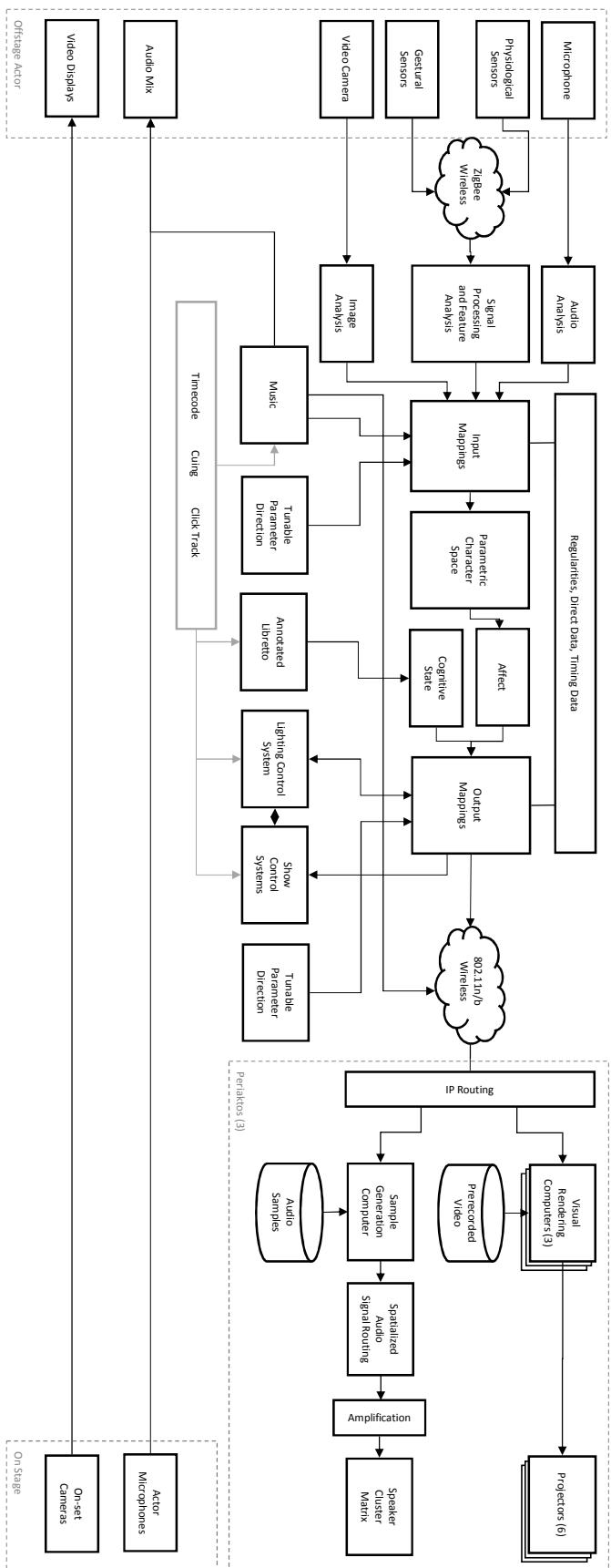


Figure 56: Disembodied Performance System schematic

This diagram illustrates the components of the Disembodied Performance System implementation for *Death and the Powers*, including the distributed audio and video rendering system for each periakos.

direct data, timing data, and metadata. These four categories are distinguished by the type of data, the pathways through the system that the data takes, the source of the data, and how it is processed. I believe that all four categories are necessary for any non-trivial Disembodied Performance.

The first category I will call *modeled data*. These data streams originate from the performer's sensed gesture and vocalization, as described in 4.5 below. These are the streams for which we are looking at the quality of the movement. Modeled data streams typically undergo some preprocessing to extract statistical derivatives and features and are then mapped into the character model. Modeled data forms the core of the Disembodied Performance input.

Direct data refers to data that is captured from the performer, but does not contribute to the character model. These streams are not heavily pre-processed and their mappings translate from one coordinate frame to another, but the semantics of the parameter are preserved. For example, in this implementation, I have defined a two-dimensional parameter called *focus*, which represents a geographic location onstage. The offstage performer's hand location can be used to point out this focus, such as to gesture to another actor who is onstage, and that relative location is translated to the appropriate location on the stage. The location semantics are preserved with only the scale (and possibly covariance or aspect ratio) altered for the corresponding output.

Although the Disembodied Performance System and the sensing methods in the implementation described below is intended to be as transparent to the performer as possible, leveraging his natural performance gesture without the need for consideration of the disembodyment process, direct data streams do allow for the system to respond to deliberate gestures, if desired. In this way, the system can function as an instrument, of sorts. Specific actions, rather than simply qualities of action, can control or trigger the output in an intentional manner. When outfitted with these kinds of direct mappings, the system functions much like typical controllers, be the modality gestural or otherwise. Events, such as the playback of specific sounds, can be achieved by the performer.

Timing data is something of a hybrid between the modeled and direct categories of data. While it is typically derived from physiological recordings or gesture, its pathway through the system is more like direct data. Though the stream may also contribute to the model as modeled data, as timing data it is mapped to the output representation more directly. However, it does not have defined or relational semantics as do direct data streams. In the current implementation, one example of timing data is the performer's respiration. The relatively long time-varying quality of a breath provides a clear indication of the natural phrasing and pace of the music

and action. Phrase onsets can trigger subtle events or allow changes in the model to have an effect on the output representation in a manner that closely links the timing of the change with the performer's speech or song. Additionally, breathing is a behavior that is very closely associated with the perception of something possessing life. Respiration data can be used to subtly modulate aspects of the output representation in order to impart an appearance of the representation being alive.

The category of *metadata* includes input to the system that does not originate with the performer. Metadata streams can come from a variety of sources—offstage personnel (stage manager, conductor), props, other show control systems (lighting console), and show timecode—and may include cues, timed triggers, color or musical information, position information, or libretto text. Metadata allows for long-timescale events to be triggered, such as a cue that defines the look of the output representation for a specific scene, or changes in mappings. This capability is critical for the Disembodied Performance System to integrate seamlessly with the theatrical design. Color palettes may be associated with cues, the color and intensity of stage lighting, or other production-wide dynamic parameters may be synchronized precisely with the Disembodied Performance representations.

4.4 Feedback

Feedback is an essential concept in any system and Disembodied Performance is no exception. Since Simon in The System must interact with and respond to other characters onstage, the Disembodied Performance System provides a feedback loop to the offstage actor. This feedback may be as simple as an array of video monitors in front of the actor connected by closed circuit to cameras embedded in the set onstage. For *Death and the Powers*, this provides the actor with Simon's omnipotent point of view, allowing him to perceive the world that Simon can perceive from inside The System. The actor will then be able to gesture and react to the action onstage. The actor will also have an audio monitor feed to hear a mix of other vocalists, selected audio effects, and music.

Typically, when an actor is onstage, he is aware of his surroundings, his relationship to the set and other actors, as well as their reactions. In theater, however, the actor's awareness is not limited to the diegesis. He is also aware of the location of the audience and the subtle cues that the audience issues in response to the action onstage. These cues can desirably affect how an actor proceeds. In Disembodied Performance scenarios, particularly when the represented actor is playing offstage, it is likely essential that the ability to be aware of the audience and their reaction be provided. This can

be achieved in several ways, such as incorporating signals from microphones in the theater's house into the mix that the actor hears.

It is important to note that the feedback views provide a point of view for the actor and do not specifically show (and perhaps should not show) the actor his onstage representation. For example, in *Death and the Powers*, the primary output representation of the performance is on the set walls. Thus, the feedback views the actor sees should be from the walls, but not of the walls. The reason for this is to reinforce the transparent nature of the system. The actor's performance is intended to be natural. If the actor sees a representation of his performance onstage, then he may alter his performance in order to attempt to achieve certain responses from the system. The actor begins playing the system, rather than performing as a character. Due to the nondeterministic nature of the input mappings to the character model, however, deliberately trying to manipulate the system will prove difficult, likely with unpredictable results, and would surely be a distraction from the performer's natural methods.

4.5 Performance Capture

Several approaches are used to capture the expressive performance of the actor. In the current implementation of the Disembodied Performance System for *Death and the Powers*, we use audio analysis of the actor's singing, wearable gestural and physiological sensors, and some basic computer vision techniques to translate characteristics that would normally be perceptible by the audience looking at an onstage actor into the system.

In the case of an offstage performer, we have the luxury of being able to outfit the performer and the room he is in with the necessary sensing technologies without concern as to their appearance. Nevertheless, for the sake of the actor and authenticity of the performance, these sensors and setups have been designed to be as unobtrusive and unencumbering as possible.

4.5.1 Wearable Sensors

In order to record the expressive gesture of the actor, we outfit the actor with several wearable sensors. These sensors were constructed by my Opera of the Future colleague, Elena Jessop, with the intent of being minimally intrusive with the potential for integrating them into a costume. Our intent is to make the wearable sensors unnoticed by the actor during performance and thus their presence does not detract from the actor, make him self-conscious, or intimidate him.

Four $\pm 3g$ Analog Devices ADXL330 three-axis accelerometers are placed on the actor, one at each forearm near the elbow and one on the back of each

hand. A fingerless glove on each hand secures the accelerometers in place with a known orientation. The forearm accelerometers are fastened to a muslin band that is secured on the arm with Velcro. A flexible ribbon cable from each accelerometer runs to a second muslin band worn on the upper arm and also secured to the actor with a Velcro strap. A pocket on this second band contains a microcontroller, battery, and radio device. Other work has used variations on this approach, such as replacing the accelerometers with inertial measurement units in [29], giving position-accurate data from which the gesture qualities can be derived. Such precise information is extraneous for the application at hand.

An additional wearable sensor, a breath sensor, is used. Breath sensors typically measure thoracic or diaphragmatic respiration [64]. The latter is often more difficult to measure. The choice of which is also dependent upon the person. In this application, we are sensing the respiration patterns of an opera singer and from where a singer breathes is a matter of vocal technique. Our singer has been trained to breathe from the chest, so the sensor used is designed to record thoracic respiration.

The breath sensor consists of two muslin straps that secure in the front with Velcro. At the back, the two straps are joined with a stretchable spandex tube inside of which is an Images Scientific Instruments STRX-04 stretch sensor. The stretch sensor is a conductive rubber cord that increases resistance in proportional response to extension. As with the other wearable sensors, a pouch contains a microcontroller, battery, and radio device. The entire assembly is worn firmly about the rib cage and is sensitive to the change in volume of the chest during inhalation and exhalation. The stretch sensor does exhibit a slight decay in resistance when returning to an unstrained state, but this non-linear response on exhale does not pose any adverse consequences in the application at hand.

The final set of wearable sensors constructed measure foot pressure. Elastic straps fit over the actor's shoes and secure two independent pressure sensors on the sole of each foot: one at the front and one at the rear. The pressure sensors are made from conductive foam that increases in conductance when compressed. Conductive adhesive foil forms the electrodes on both the top and bottom side of each foam pad to which the electrical leads are soldered. As before, a microcontroller, battery, and radio unit accompany each sensor assembly and are attached to the straps on the back of the shoe above the heel.

For each of the wearable sensors described, the sensor units are wired to the 10-bit analog-to-digital converter pins on the microcontroller boards. The microcontrollers used are Funnel I/O boards, which are based on the LilyPad Arduino using Atmel AVR ATmega168V microcontroller running with an 8 MHz clock. The Funnel I/O boards were chosen because they

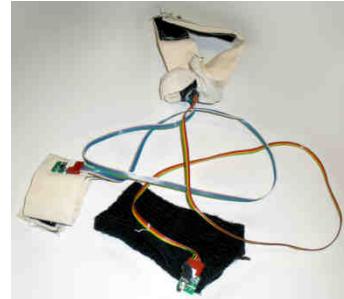


Figure 57: Arm gesture sensor assembly

The arm assembly consists of two three-axis accelerometers, one on the hand and one on the forearm.



Figure 58: Breath sensor band
The band is worn about the chest. A stretch sensor inside the band records the expansion and contraction of the chest cavity.



Figure 59: Shoe sensor assembly

Two pressure sensors are attached to the sole of each shoe with elastic straps.

possess a number of desirable features—such as a recharging circuit, voltage regulator, and connector for a 3.7v lithium-ion polymer battery and wired headers for ZigBee modules—in a small form factor. The batteries used provide ample power for over three hours, covering all performance needs. Digi XBee (a vendor specific implementation of the ZigBee specification on top of the IEEE 802.15.4 physical layer and medium access standard) modules with chip antennae provide wireless communication from the Funnel microcontroller board to a ZigBee coordinator connected to the control system computer. Using XBee for wireless serial data communication provides a robust solution for our needs, supporting ranges up to 133 feet indoors (or greater, if need be, with other models and antenna configurations) on the 2.4 GHz band, which avoids other trafficked bands in our production. The Funnel I/O boards run a modified version of the standard Firmata firmware for convenience and extensibility. The analog-to-digital converters are configured to sample at 30 Hz and transmit the values to the control system at 60 Hz over a 19200 baud serial data link.

4.5.2 Audio Analysis

Although unusual for opera, *Death and the Powers* makes extensive use of reinforced sound. One reason for this is so that the performers' voices can better blend with some of the electronic and amplified acoustic sounds of the orchestra. Additionally, the voices of two of the characters, Simon and Nicholas, are treated with effects and are spatialized at different moments of the show. Particularly, in the case of the actor portraying Simon, a microphone is required to present the voice of the offstage actor to the audience in conjunction with his onstage representation. I take this opportunity to use the actor's voice as another source of expressive input for the Disembodied Performance System.

An audio signal from the actor's microphone, upstream of any effects or processing, is sent to the Disembodied Performance System computer where it is analyzed. I have chosen to extract three parameters from the voice signal: amplitude, instantaneous frequency, and a parameter that I call consonance. Processed by the Disembodied Performance System, the parameters are sampled and computed at the system's data sample rate (see 4.8 below) from a 44.1kHz audio stream. The amplitude is the instantaneous sound intensity.

In an opera context, pitched vocalization is prescribed by the score and not sufficiently expressive for our purposes. Thus, determining pitch is of little significance and we use frequency primarily as a relative parameter that may encode something of the actor's vocal expression or, over longer time scales, an expressive parameter of the score. Without the need for the accuracy of autocorrelation or harmonic product spectrum pitch detection algorithms, I use a simple and computationally efficient method to compute

instantaneous frequency. The incoming audio stream (at 44.1 kHz) is buffered with a window size of 1024 samples. A fast Fourier transform (FFT) is computed for each window and the median frequency of the maximal bucket is reported.

Timbre is perhaps one of the most expressive attributes of the singing voice. Indeed, Tod Machover's score for *Death and the Powers* calls for the variations in vocal timbre. The third parameter, which I have called *consonance*, is intended to capture something of this timbral quality. It is a measure of spectral purity and can be used to continuously classify pitched versus unpitched vocalizations. Using the FFT of the sample window already computed for frequency, the consonance c is computed as

$$c = \left| 2 \left| \frac{f_1}{f_0} - \left\lfloor \frac{f_1}{f_0} \right\rfloor - 0.5 \right| - 2 \left| \frac{f_2}{f_0} - \left\lfloor \frac{f_2}{f_0} \right\rfloor - 0.5 \right| \right|$$

where f_0 , f_1 , and f_2 are the ordered median frequencies of the three maximal buckets. Essentially, this is a measure of how close the second two maxima are from being harmonics of the first. This yields a consonance parameter that varies continuously between 0 and 1 indicating unpitched (substantial inharmonic partials, as in obstruent sounds or speech) and pitched sound (substantial harmonics or nonexistent partials, as in a sung “ooh”), respectively. As formulated, the value is sufficient for our purposes and inexpensive to compute in real-time. The accuracy of the metric can be improved by increasing the order, the number of maximal buckets examined, or, as with the frequency parameter, by replacing f_0 with a more accurate frequency estimate derived from autocorrelation or other robust methods. The value is set to zero when the amplitude of the maximal bucket is less than some threshold of noise.

4.5.3 Computer Vision

While the accelerometers capture the quality of motion, it is non-trivial to integrate accurately the accelerometer data twice to recover position information. In order to recover some precise spatial quality of certain gestures, such as a gesture toward another character or part of the set, computer vision techniques are used to accomplish basic motion tracking. Four tracking points are placed on the actor's body: one on the back of each hand and one on each shoulder. The tracking points are bare LEDs that mount into sockets on the arm accelerometer assemblies described above and are powered by the same battery. A suitable LED color may be chosen, most likely infrared, to be tracked by a camera. The camera image defines a coordinate frame that can be mapped relative to the proscenium or onstage display surfaces providing a means for recovering the focus parameters discussed in 4.3. These coordinate frames, or defined regions of the camera's view, also are registered with the feedback views noted in 4.4 so that the actor's gesture on the stage corresponds spatially to what he is

seeing of the stage. This absolute mapping allows the actor to precisely define an effect at a specific location onstage. The shoulder points also provide a relative reference for the hand points and can be used in determining posture. While the hand points provide a focus parameter, the recovered posture would likely feed into the model.

A video or web camera connected to the control system can be used for the motion capture. An input module was written for the Disembodied Performance System, but the necessary image processing to recover the tracking points from the image is sufficiently computationally expensive to slow the system to an unacceptable rate. In production setups, a separate computer system would be responsible for the computer vision, which then passes the point locations or computed focus location to the Disembodied Performance System.

4.5.4 Other Sensing Modalities

The choice of which gestures to capture and the types of sensors with which to outfit our actor was made with several considerations in mind. A multimodal approach, such as the one described, not only ensures that a range of expression to span the affective space is captured, but provides enough redundancy to minimize the effects of aberrant outliers and to reinforce consensus [79].

Actors are trained to communicate to audiences their character's thoughts and emotion through gesture. In some schools of acting, these gestures are intended to communicate directly to other characters and the audience's comprehension of them is seen as a desirable byproduct [6]. These are exactly the sorts of signals and cues that we want to capture and reproduce in Disembodied Performance. Thus, the arm accelerometers and shoe pressure sensors are ideal for recording these sorts of gestures that are within the actor's control and movements that are naturally made when acting onstage.

The theatrical context poses several constraints on the types of sensing modalities employed. In everyday interpersonal interactions, gestures tend to be more subtle and less histrionic than those of an actor onstage. Such fine movements, such as minute differences in facial expression, are difficult to read from the audience in typical theater settings. While research has been done on automatically inferring cognitive states using computer vision to analyze facial features, as in [24], this approach would not work in our context. Additionally, the models used to recognize facial expressions would need to be retrained for the types of exaggerated expressions an actor might produce and, particularly in the case of opera, the models would likely not be accurate while the performer is singing, as an open mouth distorts the proportional relationships of facial features.

Of the sensing modalities described above, only one sensor, the breath sensor, records a physiological signal. While a plethora of physiological signals can be recorded—temperature, blood volume pressure, galvanic skin response, electromyogram—they pose a particular problem in the case of stage performance. Physiological signals are usually involuntary and, while we are interested in natural expressive responses rather than conscious gestures, these signals are indicative of the emotional state of the actor. The experience of the actor is not what we're looking for, but rather *the experience of the character*. They are distinct [64,49]. By definition, the actor alters his outward appearance, the behaviors that are observable without instrument, to represent the character, not his own experience. The actor may be exerting himself or having a bad day or apprehensive about opening night or be uncomfortable under theatrical lighting. All the while, the character he is portraying may be in a pleasant, relaxed, carefree mood. Physiological signals would more likely correspond with the former qualities rather than the latter, producing an inappropriate representation when processed by the Disembodied Performance System. Thus, for the current application, the sensors used capture behaviors and features that would be apparent to the audience if the actor were onstage.

With these considerations in mind, it should be noted that the types of sensors discussed and the features being sensed are not specific to the Disembodied Performance approach. An existing body of extensive research into gesture and wearable sensing exists and, though many of these approaches do not capture the qualities of expression as required for Disembodied Performance, they may suggest alternative methods or refinements. The implementations given here are merely the initial sensor selections for the particular production of *Death and the Powers*, though I believe that they are representative of the general types of sensing that are necessary and sufficient in performance contexts. Additional modalities and gestures could be sensed in a variety of ways and could be incorporated into the system.

4.6 Character Modeling

The crux of the Disembodied Performance technique is to infer the state of a model as an intermediate representation of the generating process and use that to drive output representations. In this case, we're modeling the character's affective state. Using reified inference, we want to capture the state of a low-dimensional model that contains enough salient features to generate an output representation that can convey to an observer a similar sense as the input representation would give.

While our goal is to reduce the dimensionality of the input data, we don't want to select a model with too low of a dimension to be sufficiently

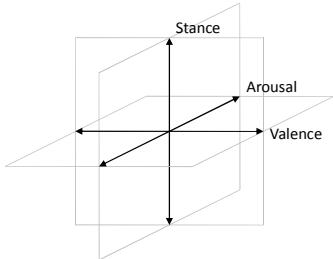


Figure 60:Affect space

The affect space model has three orthogonal dimensions: arousal, valence, and stance.

capable of expressing the richness of the actor's performance. Additionally, the nature of the inputs and the parameters of the output representation that will be varied over time by the model suggest the use of a continuous space. To model the character's affect, I have chosen to use the same three-dimensional metric space with orthogonal axes representing the normalized signed affective bases of stance, valence, and arousal that was used by Cynthia Breazeal in her work on sociable robots [11]. The 3D space allows for a broader range of captured expression than the traditional 2D circumplex model, providing more unique loci for the generation of output mappings from model parameters. Nevertheless, it has a much lower dimension than the set of all signals recovered from the performer, so we apply a set of mappings to reduce the dimensionality of the data and project it at a given time to a point in the affect space. As the actor performs, he will carve out a continuous path through this *affect space*.

The system employs a combined inference approach to determining the model state. In combined inference, all of the inputs contribute to a single model, as opposed to hierarchical inference where features of each input would generate a model state and then all models are arbitrated to produce a result [79]. The lack of semantic features in the particular types of input captured from the performance necessitates this approach. The affective state cannot clearly be determined from any one input axis in isolation or in small groups. The input mappings seek regularities across inputs, the correlatives for which can vary from one acton to the next.

As noted in the discussion of reified inference in Section 2.5, the output representations will have a higher dimension than the intermediate representation. To generate the values for all dimensions of the output representation without trivially spanning the representation space, some sort of additional information needs to be added. Incorporating some sort of randomness in the output mappings would accomplish this. However, a better approach in many cases would be to use original input data as direct data and timing data to fill out the necessary dimensions of the output representation. These data will contain useful noise and be correlated with the rest of the performance.

The quality of the mappings from the input performance to the onstage expression is the key to implementing this system successfully. Unlike systems that attempt to synthesize emotional representations on their own, the mappings in Disembodied Performance do not need to account explicitly for sigmoidal activation responses, thresholds, or decay, as these properties are inherent in the actor's expression. Each of the parameters measured (such as the velocity of the performer's hand or the timbre of the performer's voice) does not directly represent some aspect of the character's affective state. However, regularities in the change and variation of these parameters are expected to be consistent with the portrayed emotional state.

To ensure that the audience can immediately comprehend an emotion or a thought presented on the bookshelves, the modal regularities in the human actor's performance must be preserved in the output representations. Just like any other actor, the performance system is designed so that it can also take direction. As the director will coach and guide the performance of the actor portraying Simon, so too will the system have parameters that may be tuned to generate the desired vision.

4.7 Output Representation

So far, I've explained how I capture the performance and how it is placed into a parametric space defining the character. The final step is to take that intermediate abstract representation and transmute it into the sights and sounds of a sentient environment. The simple low-dimensional model state that results from the mapping of input data, along with a few control instructions, provides an efficient mechanism for transmitting data to distributed components responsible for generating the output representations. The output representations may be multimodal and take any form. The power of this technique is to have output representations of arbitrary scale and in any medium, creating a completely non-anthropomorphic presence onstage. In some productions or applications of Disembodied Performance, it may be desirable to have the output representation be an anthropomorphic character, but for *Death and the Powers*, this is not the case. Numerous approaches to anthropomorphic character animation, indeed stretching back for the entire history of puppetry, have been studied at great length and present alternatives to Disembodied Performance. This approach seeks to create a wholly different sort of expressive experience.

For the most part, in *Death and the Powers*, the remapped performance parameters will be rendered as a video rear-projection onto the surface of the books. The design of the books and the rationale given in Section 3.3.5 will inform the way expression is rendered on the books. Sonic cues and events will also be generated and spatialized over the active wall faces. Furthermore, the system will have the potential to interact with other show control systems, such as lighting, to extend the influence of Simon Powers beyond the periaktoi and into the entire surround as light, robotic motion, and sound.

In the periaktos design, two faces of each column will be capable of such projection, the third reserved for physical effects during the final scene. A distributed approach to generating the visuals and sound is used. Each of the "active" faces will be illuminated by three projectors stacked along the height of the wall. Thus, each column will house six projectors for a total of eighteen in all. Three computer systems, each with two DVI video ports,

will drive the projectors and run the output renderer software, which I will introduce in Section 4.8.5 below. Both computers, along with other on-board computer systems, will connect to an internal switched IP network that will communicate with offstage systems over a wireless 802.11n gateway that will allow for untethered remote control and transmission of the performance data to be visualized on the walls. Any pre-recorded imagery or sampled media required will be duplicated on each on-board system to avoid transmitting high-bandwidth real-time streams of audio and video. Sonic cues and events will also be generated and spatialized over the active wall faces or through generalized spatialization systems.

For the purposes of this thesis and a theatrical context such as *Death and the Powers*, the output representations are explicitly designed. They are not automatically generated from some universal rule set. In the general case of the theory of Disembodied Performance and reified inference, I believe that structured approaches to generating output representations are possible. For example, with proper semantic grounding of an intermediate representation, specific applications in information visualization could automatically generate meaningful output representations using a well-defined vernacular such as the one expounded by Jacques Bertin in his *Semiology of Graphics*[7]. For more expressive and abstract representations, the research of Heider and Simmel shows that simple forms can be imparted with movement that suggests purpose and emotion [34]. More recent work explores some of the formal properties of animacy and the perception of intent, including the movement geometric primitives derived from human motion, and finds that a social context greatly contributes to the inclination to perceive an animation as sentient [51]. Present understanding is insufficient to suggest an automated formal approach to producing effective and evocative non-anthropomorphic output representations of human performance, though future research will shed additional light on the matter, perhaps aided by the techniques presented here.

4.8 Software Implementation

The software implementation of the Disembodied Performance System manages the communication of input devices with the distributed system of output renderers. Running on a central computer, the software handles the mapping of input data, analysis of data, and mapping data to and from the intermediate model. The system then broadcasts the computed values to the output devices. The Disembodied Performance System also provides a user interface for the creation and management of show configurations, mappings, cuing, and provides status feedback about the performance and remote devices. The implementation strives to facilitate the application of

Disembodied Performance in an actual production, such as *Death and the Powers*.

The software for this project was implemented in Java 6 and borrows a great deal from the original concept for the unified control system outlined in Section 3.4. Due to their similarity, it is reasonable to expect that these independent, but parallel, implementations may be reintegrated at some point. Parts of the implementation have been deliberately written in a general fashion so that the system can serve not only in multiple productions, but components can be reused in further explorations of Disembodied Performance and in applications of the underlying inference-model-mapping approach.

4.8.1 System Operation

As in the Core show control system, a *show* is the set of information that defines the configuration of the Disembodied Performance System for a particular production, equivalent to a “project” or “document” in many sorts of computer applications. The show configuration can be saved and reloaded as an XML file. The XML file contains sections that enumerate the input devices used by the show and their calibration settings, a list of cues and their properties, and the mapping for each cue. The devices and mappings sections can also be saved independently of a show so that new shows can reuse configurations that have already been developed. Input devices are specified by their fully qualified Java class name. From this name, the `InputDevice` wrapper implementation is dynamically loaded by the system so that new devices can be added in a plug-in fashion without needing to rebuild the executable code for the whole system.

In the system as a whole, there are essentially three different rates at which parts of the system are updated. Data from sensors is sampled at a rate specific to each sensor assembly and transmitted to the system where the current value is buffered. Output renderers may also determine an update rate at which they compute new rendered “images” in their medium. At the core of the system, a third update rate determines how often the mappings are processed and the model state is adjusted. This value is specified with the configuration of a show.

When a show is loaded, any of its aspects can be modified, including: adding and removing devices; adding, removing, resorting, and altering cues; modifying mappings. On the right-hand side of the interface, is a panel presenting most all of the global controls for managing shows, input devices, output renderers, and creating and navigating cues. A panel also displays a log of system events for monitoring the success or failure of operations and the status of remote components. Above the main part of the interface are five tabs that allow the user to switch among different task-oriented views. The Mapping Designer view is where mapping associated

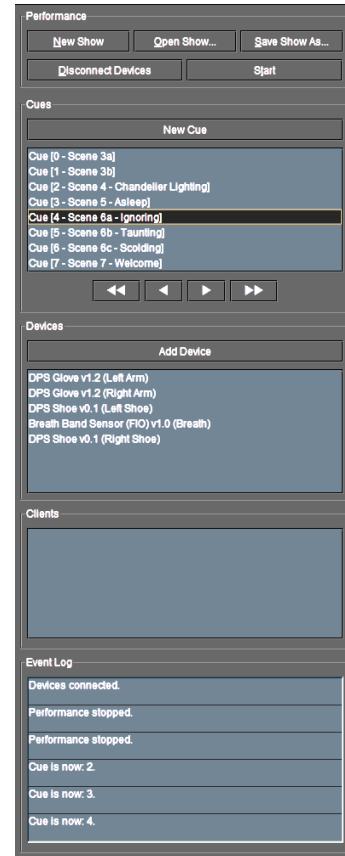


Figure 61: Control panel

The system control panel provides access to show, input device, and cue management as well as system status feedback.

with the current cue is edited. The Parameter Director view displays the tunable parameters for the current mapping. The Input Streams tab provides a live time series view of each input device axis. The Model Viewer shows a real-time representation of the inferred model state. Finally, the Render Preview tab allows certain visual output renderers to be loaded for design-time reference when independent output renderers are not set up.

A show need not have input or output devices connected to the system when it is loaded. This allows modification of shows without requiring the entire system to be set up in place. Before running a show, the user must tell the system to open the devices. If input devices are not connected at this time, the system will produce an error indicating that one or more devices are missing. Output renderers need to be registered to run a show. Once the system has established a connection with the input devices, the show can be run. This starts the system update timer, which causes mappings to be evaluated and output broadcast. Time-dependent cues rely on the show to be in a running state to compute the absolute or relative elapsed time for cue events.

4.8.2 Input Devices

Input devices constitute any hardware or algorithmic process that provides modeled, direct, or timing data to the Disembodied Performance System. Each device has one or more *axes* that have a value at any given point in time. These devices are connected to the system and their axes appear in the interface. The values of these axes are the inputs to the mapping and character model. For example, the wearable sensors connect to the performance system computer by a wireless serial connection. The ZigBee coordinator that receives the wireless serial data from the remote hardware is connected to the system via USB port. Each transmitting microcontroller appears as an input device to the system. Similarly, audio input from a microphone is connected through an audio interface and a camera for computer vision may also be connected via USB or FireWire. In the case of audio and video, however, the hardware is not itself the device, but rather the code module that performs the audio analysis or computer vision.

The hardware and analysis algorithms are presented to the Disembodied Performance System by implementing the `InputDevice` interface. This interface exposes methods for enumerating axes and retrieving the current value of an axis, as well as metadata about the device, such as device name. For serial devices, the implementation of `InputDevice` is simply a wrapper around the serial connection, storing the axis values locally as the device samples them. For analysis algorithms, the `InputDevice` implementation may manage the retrieval of the audio samples from hardware or images from the camera and compute the analyses, storing the results for each axis.

Local values for device axes are updated asynchronously to avoid blocking reads from the devices that would cause stuttering in the transmission of messages to renderers. Many of the implemented input devices utilize the RXTX serial implementation compatible with Sun's abandoned Java Communications API. This API, as well as several others used in `InputDevice` implementations, provides event-based notification for asynchronous updates from the hardware. Furthermore, all devices are updated in their own thread. The device values are accessed by the system's data processing thread through synchronized methods. If a device read does block, the data processing thread will still be able to read the currently stored value for an axis. Repeated values over a small number of intervals have virtually no noticeable effect. When the broadcast rate is the same as the device update rates, nominally a value is never repeated (aliased) for more than two time steps.

The values for device axes exposed by the `InputDevice` interface are normalized into the range [0,1] or [-1,1], depending upon the axis' semantics. The normalization is carried out with respect to the calibration of the device, including the raw hardware value extents and "home" or mean value. This practice unifies all of the disparate inputs to the system and is an important design decision in terms of creating mappings, as described below.

Some useful properties of the incoming data are desirable for most axes in most mapping scenarios. The properties are generally statistical and can indicate something about the quality of the input signal. For this reason, each device axis has associated with it a `DataStream` object. Each `DataStream` maintains a buffer of the incoming data over some window of time, usually on the order of one to five seconds, but this varies according to the needs of the output representations. For each system update interval, the `DataStream` object computes the following values over the window size: normalized value, mean, maximum, minimum, instantaneous derivative, integration, and rugosity. In order to maintain real-time responsiveness, a cyclic array buffer stores the input values for the window duration and the derived metrics can then be computed in constant time, with the exception of maximum and minimum (upon which the computation of the normalized value depends), which require linear time. The normalized value is set to zero when the variance is small. All of the properties of each `DataStream` are visualized live in the Input Streams tab of the interface when a show is running.

4.8.3 Mappings

The system relies on mappings to connect the values of input device axes to output renderers. This is a critical process in the Disembodied Performance approach, as the mappings to and from the model determine the success of the model and output representation in maintaining the meaning of the

input. Consequently, this is one of the most elaborate components of the software implementation. Although the process of deriving informed mappings is the subject of future work, I wanted the Disembodied Performance System to provide an environment where a user could craft mappings that preserve semantics. In early versions of this system and other mapping-based systems I have created, I would write mappings in the code and need to rebuild the software and reintroduce the inputs to evaluate the mapping. This iterative process was slow and tedious. The implementation of the mapping system and user interface in the Disembodied Performance System is my solution.

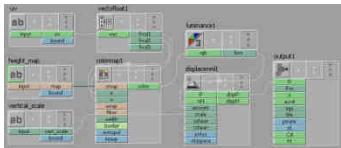


Figure 62: Houdini VOPs

Houdini’s graphical view of its VEX programming language is a node-based flow that inspired the mapping interface for the Disembodied Performance System. (Screenshot from Side Effects Software Houdini 7.0.99)

While I almost exclusively prefer writing text-based code to visual programming languages, as I find the process more flexible and efficient, the process of creating mappings seemed like an ideal application of a graphical approach. Many multimedia development environments, from audio synthesizers to visual effects compositing packages, employ some sort of visual programming language or node-based flow. I looked to several of these for inspiration before implementing the flow-based mapping described in this section. The primary source of inspiration for the interface for this implementation was the VOPs (vector expression operators) component of the 3D modeling and animation software Houdini by Side Effects Software. In Houdini, VOPs provide a visual programming language analog for the C-like VEX language Houdini uses internally.

The node-based dataflow programming interface for mapping in the Disembodied Performance System closely reflects the internal structure; the user interface and the internal representation are tightly coupled. The mapping consists of a set of nodes with varying numbers of input ports and output ports. The nodes can be of several types that each carry out a specific operation or otherwise tie the node to some other part of the system. The view of the nodes can be freely organized and named to be relevant to the user. In order to facilitate the creation of mappings with minimal regard to the technical details of the implementation, I set several design goals to ensure simplicity of use. These goals are met by this system and distinguish it in marked ways from existing applications, including Houdini, Max/MSP, vvvv, and Secret Systems [85].

- All values in the mapping are of the same data type and cardinality, so there is only one type of link and any port on any node can be connected to any other port on any other node. The only restrictions on connections are due to a few requirements of the graph topology, as described below.

- Parameters of the mapping should have some semantic importance. They are called out to a separate interface for tuning the mapping once its structure is in place.
- The user should not be concerned with the particulars of a device, its connectivity, or its configuration. In a theater setting, people from a broad range of backgrounds may wish to use the system. Those responsible for constructing mappings or tuning parameters need not concern themselves with the technical particulars of the input and output devices during the creative process.
- Mappings can be modified in real-time while the show is running so the effects of a change can be seen immediately.
- With only one data type, complexity can be reduced by providing a relatively small set of simple and intuitive operators.

All abstraction of hardware is done before and after mapping, unlike some existing systems. As with the unified control system described above, all values are mapped into a lingua franca of normalized floating-point numbers. Only the developers of the `InputDevice` wrappers and output renderers need concern themselves with the various data types and ranges the input and output hardware supports. Even though the values generated as inputs to the mapping never have a magnitude greater than one, the mapping implementation allows values to exceed this magnitude as the result of computations, as such values are occasionally useful in output representations. Values can be clipped by the user at any point in the computation, if desired.

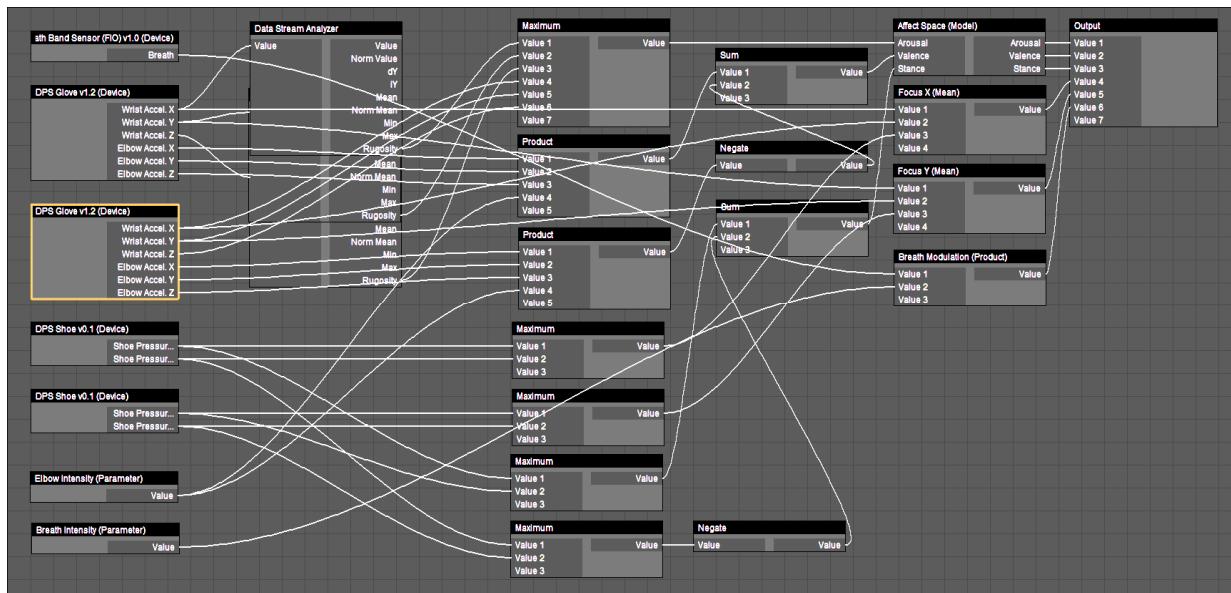


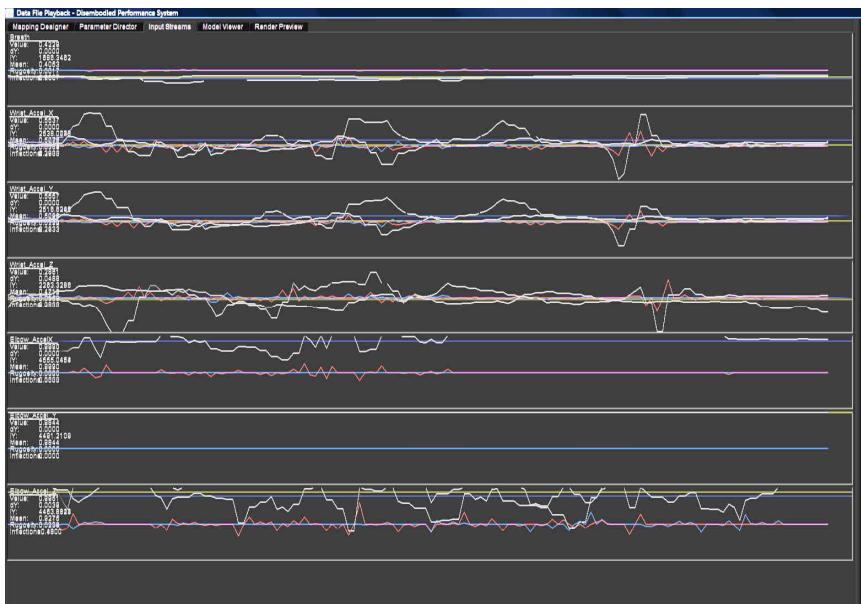
Figure 63: Mapping Designer view
The mapping view displays the connections from input devices to the model and output. The user can manipulate, add, remove intermediate nodes and adjust the connections in real time.

Mappings are acyclic directed graphs. Each node in the graph implements an abstract class `Node` and has any number of input `Ports` and output `Ports`. Some subclasses of the `Node` class can declare themselves as variadic if a variable number of inputs is appropriate for the operation computed by the `Node`. Variadic `Nodes` always have zero or more fixed input `Ports` followed by the variable arguments. One unused variable argument input `Port` is always presented. When this `Port` is connected, a new `Port` is added. When the last used variable argument input `Port` on a `Node` is disconnected, the last `Port` is removed. An output `Port` may be connected to any number of input `Ports` of other `Nodes`. However, an input `Port` can only connect to at most one output `Port`. A `Node` cannot connect to itself or to another `Node` in such a way that a cycle is created. The mapping is computed at each update interval in a recursive depth-first manner starting with the root, the `OutputNode`. All nodes in the connected component containing the `OutputNode` are visited. For each `Node`, if the values of its input `Ports` are not up-to-date, the `Nodes` connected to the input `Ports` are updated. When all of the input `Ports` of a `Node` are up-to-date, the values of the output `Ports` are computed. All input `Ports` have default values that they assume when not connected to the output `Port` of another `Node`.

At the minimum, a mapping contains a `DeviceNode` for each input device in the show and a single `OutputNode`. No links among these nodes are provided by default. A mapping can only contain one `OutputNode`, at present. In the future, different `OutputNodes` may be provided for different types of renderers. For now, all renderers receive the same output parameters on update. Additional nodes can be added from a context menu in the mapping view. All nodes can be moved around within the view and given a meaningful label. Added nodes can be deleted, along with all incoming and outgoing connections.

Figure 64: Data Streams view

This tab presents a plot of all incoming data from input device axes in real time. Each section identifies the data source and provides a numerical readout of the current value. All outputs from the `DataStream` associated with each axis are displayed along with the actual value. This view provides a mapping designer with an understanding of the incoming data and can be used during performances to monitor devices.



Currently twenty-two node types are available:

Node Category	Node Type	Variadic
Statistical	Maximum	■
	Minimum	■
	Mean	■
Arithmetic	Sum	■
	Product	■
	Negate	
	Invert	
	Clip	
	Scale	
	To Signed	
	To Unsigned	
Generation	Random	
	Noise	
	Threshold	
	Impulse	
	Switch	■
Data	Span	■
	Data Stream	
	Parameter	
	Device	
	Model	
	Output	

Nodes are loaded dynamically by class name from show files and in the context menu by a configuration file. Because of this, the types of nodes available can easily be extended. Many of the default node types' operations are self-explanatory. I will briefly describe the nodes that have a special function in the system.

DeviceNodes have an output Port for each device axis exposing the live values of the device. **DataStreamBreakoutNodes** can be added to access the descriptive statistics already computed for device axes. If the input of to the **DataStream**'s node is not a device node's axis, then the node maintains a new **DataStream**, introducing new buffers to compute statistics after some point processing in the flow.

The **OutputNode** is a special type of **Node**. Only one **OutputNode** can exist in a mapping. It is a variadic **Node** with an arbitrary number of inputs. The values of the inputs are those that are transmitted to registered output renderers. The values are combined into a floating-point array that is the payload of the broadcast OSC message. By convention, the output values are ordered with the model values first, followed by direct data and timing data. The close coupling to the float array used in the OSC message was

implemented for reasons of efficiency. Future versions may allow multiple `OutputNodes` in a mapping so that different types of renderers receive only the values of interest to them by OSC method address with all messages transmitted at once in an OSC bundle at each time step.

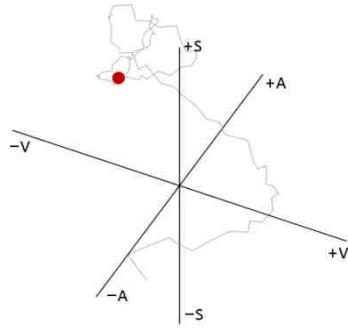


Figure 65: Trajectory in affect space

The trajectory shown moves from a boredom-like state to anger, which is sustained in the high arousal, high stance, negative valence octant.

Another special type of `Node` represents the multidimensional model. In the case of Disembodied Performance, this model is the three-dimensional affect space. A mapping may only have one model node. The model node performs no operation and the values of its input `Ports` are copied unchanged to its output `Ports`. The value of the model node, however, is used to illustrate the state of the model in the Model Viewer tab of the interface. This view shows the dimensional space with a point for the current value of the model and a history trajectory of the values through the space over recent time steps.

Once a mapping has been developed for a cue to capture the salient features of the performance for modeling and output, the mapping will likely remain untouched. One of the goals of the Disembodied Performance system is to allow the performance to be directable by a production's director and creative personnel. During performances, actors take direction. The director may instruct the actor to behave in a certain manner or give emotional and motivational cues. The disembodied performer would respond to such direction as if he were acting onstage in a traditional manner. The influence of these changes will be reflected in the onstage output representation automatically. However, there may be cases where the director wishes to elicit a specific type of movement from the actor and instruct him to move in that manner. These movements are more for visual and stylistic effect and would not be interpreted in the same manner by the performance system. When working with Disembodied Performance, I want to afford the director the same ability to control aspects of the output representation. My solution is to incorporate tunable parameters into mappings. While the structure and process of the mappings remain unchanged for a cue, parameters of the mapping that could affect variances or amplitudes can be exposed to the director for manipulation during rehearsals. The director may not be involved in the creation of the mappings, but is in the tuning of them. The parameters are not intended for dynamic control during performance. Once the mappings are created



Figure 66: Parameter view

Parameters of the current mapping can be adjusted using sliders in the Parameters tab of the interface. Tunable parameters allow the director to obtain the desired performance without altering mappings.

and the parameters tuned, the system requires no direct input other than that of the actor to realize the performance.

Tunable parameters are added to mappings by means of `ParameterNodes`. Parameter nodes have a single output `Port` presenting a signed or unsigned normalized floating-point value. It is configurable whether or not a `ParameterNode` is signed. Like all other nodes, `ParameterNodes` can be given a meaningful descriptive name. In the Parameter Director tab of the system's user interface, sliders for all parameters in the current mapping are displayed with appropriate bounds and their descriptive name. This interface is deliberately simple and familiar, resembling the faders of a lighting or sound control console. From this view, the value of any parameter can be adjusted and the results observed in the output representation in real-time.

The flow representation of the mapping assists the user in creating meaningful mappings. The visual overview makes it readily apparent which data sources contribute to which model parameters and how they are combined. Mappings upstream and downstream of the model can be seen in a single view, which is useful when attempting to preserve correlations in input parameters with those of output parameters, though the correspondences are not necessarily one-to-one. Different pathways of data from input to output are also apparent. Modeled data flows through the `ModelNode`. Direct data and timing data bypass much of the dimensionality reducing operators and flow fairly directly to the `OutputNode`.

It is important to note that this flow-based approach to mapping reflects a key principle of Disembodied Performance. The system has no autonomous behaviors. All data being output is computed from the data coming from the input devices capturing the actor's expression. There are no intervening routines, since the output of the system is intended to be driven by the offstage performance in order to represent it entirely.

4.8.4 Cues

Like most theatrical show control systems, the concept of a cue is at the heart of the control behavior in the Disembodied Performance System. In traditional theatrical practice, a cue is a verbal command, an electronic signal, or occasionally a light that represents a sudden change that synchronizes one or more events—actors, technicians, orchestra, recorded sound, lighting, scenery movement—with the action onstage. Cues also typically define the look of a scene or part of a scene that persist over a relatively long period of time. So too is the case for cues in the Disembodied Performance System.

All shows have at least one cue. A default cue containing a default mapping is provided when a new show is created. A cue stores a particular mapping

and other configuration parameters that may affect the output representation. Cues are ordered and the position of a cue in the list can be changed. At a given point, one of the cues in the list is the current cue for which, when the show is running, its mapping will define how the output values are calculated. An external trigger or controls in the user interface can advance the current cue to the next cue in the list. It is also possible to have cues last for a fixed duration. The cue then advances after the specified amount of time elapses from the moment the cue became active. Cues typically advance to the next cue in the ordered list, but a cue can be made to jump to a specific cue in the list, enabling cycles for repeating sequences of configurations.

Since each cue has a mapping associated with it, cues provide a mechanism for defining properties and qualities of the output representations over a relatively large timescale. Unlike the actual output values of a mapping, which vary continuously, the parameters for a cue are discreet. Cues are important in theatrical design, setting the mood for a scene, for example. The choice of colors and lighting that designers make for a scene do communicate something about mood in the affective sense. If the Disembodied Performance mappings produce affective values that represent emotions, which have a short time span and result from the actor's performance, then cues are responsible for the longer-term affective concept of mood that can be the purview of the production's designers.

4.8.5 Output Renderers

Output renderers are responsible for taking the modeled, direct, and timing data, as well as metadata, and producing the visuals, sounds, and motion that will be seen and heard onstage. Renderers interpret the mapped performance data as needed to control hardware and software implementations and are designed to be a truly distributed rendering system. Renderers could be anything that understands the particular OSC protocol implementation and the semantics of the output parameters of the mapping. This includes other show control systems, media servers, music synthesizers, and graphics generators with layers to interpret the appropriate messages. Renderers are somewhat unlike the output devices in the Core control system that I explained in Section 3.4 in that the input values to the renderer do not map directly to a physical property. Renderers are not quite the dual of the input devices used for performance capture.

In order to accommodate the numerous mobile set pieces in *Death and the Powers* that represent The System and will be controlled using Disembodied Performance, I chose to implement renderers as separate computer programs or devices receiving commands via OSC over UDP. This allows IP-based networks, both wired and wireless, to create a distributed system of output representations throughout the theatrical set. The individual renderers register themselves with the central system software when they

come online. The software maintains a list of all of the renderer clients to which it sends subsequent data. The system can also use IP broadcasting to deliver messages to renderers on the same IP subnet without requiring them to register. This process reduces the amount of time it takes to deploy and launch the system and affords considerable flexibility. Using this observer design pattern, none of the details of IP addressing or application-level addressing need to be considered in actual usage scenarios.

After a renderer registers itself with the performance system, it will receive an acknowledgment from the system that includes some global parameters that it may need to do its work. To improve robustness, current renderer implementations will try to register several times with the control system until they receive an acknowledgment. Once the renderer has successfully been registered, it will begin to receive two types of messages from the performance system. When a show is running, the output data from the current mapping will be sent to each renderer for every time step (see 4.8.3 above). Additionally, metadata is broadcast to renderers as needed. This metadata includes cue changes and parameters associated with the current cue so that the renderers can modify their functionality accordingly. A cue change for a scene, for example, may install a new color palette in a visual renderer or a different sample library in an audio renderer. In the same manner as for the Core system, renderers may also send messages to the performance system so that feedback about their status can be monitored remotely from the main system offstage.

A limitation of the current architecture requires that renderers understand the semantics of the OSC messages that they receive containing the mapping output data. Since the output of a mapping is specified merely by convention, the renderers need to know something about the mappings being used for each cue. A future implementation may use an ACN-like device description framework allowing a renderer to semantically declare the properties it is interested in receiving. OSC bundles may then be used to synchronize messages that have meaningful OSC addresses to renderers.

Renderers have two options in handling the timing of their output. They may synchronize themselves to the messages arriving from the Disembodied Performance System or they may asynchronously compute their output responding to changes in mapping values as they arrive. Individual renderers, such as robotic elements, will likely use the latter method, keeping efficient control loops running as quickly as possible. The current values from the performance system would be buffered and used for potentially multiple time steps of the device's loop to ensure smooth and reliable operation. In the case of robotics, certain watchdog safeguards are required to handle anomalous loss of communication. Visual renderers that constitute an array of contiguous screens may use the former approach, so that a single image remains synchronized over multiple devices. Thus, for

distributed rendering of deterministic graphics, the update messages from the performance system provide a trigger to draw a new frame. At sufficiently high frame rates, depending upon network latency and burstiness, this method can produce reasonable results. If the update messages arrive at renderers with varying delay or at a nonlinear rate, the renderers may produce undesirable results.

To date, several renderers have been implemented to test the system and experiment with outputs. All renderers so far have been visual renderers with the exception of two. A Max/MSP patch implemented a renderer to manipulate live sound in response to performance data. Another renderer was implemented to synchronize the playback of a QuickTime video with performance data (described in 5.1).

The existing visual renderers have been implemented in Java 6 using Java 2D and the JOGL Java Bindings for OpenGL abstraction layer for hardware accelerated graphics. Many of these renderers rely on vertex and fragment shaders to offload processing onto the GPU to achieve sufficiently high performance. In order to facilitate the development of visual renderers, an abstract renderer class was written to take care of the registration and OSC handling as well as parsing configuration files, entering full-screen exclusive mode for graphics contexts, synchronization and threading, and the display of debugging information. The renderers then implement this abstract class needing only to override `config()` and `render()` methods to set up global state and draw each frame, respectively. The abstract renderer also has the functionality to allow visual renderers to specify a portion of a larger canvas for which they are responsible. Consequently, many machines supporting one or more graphics contexts can be assembled into a large multi-screen display, as is required for the rear projection of the walls for *Death and the Powers*. This approach is similar to Daniel Shiffman's Most Pixels Ever library for Processing [78]. Each renderer knows the bounds it must draw for its part of the whole image and culling drawing by these bounds can optimize redraws.

4.8.6 Integrating with Other Systems

Since the Disembodied Performance System uses OSC for communicating with output devices and has a structured approach to gathering input, many other systems can be made to interoperate with it. The OSC protocol is becoming increasingly standard in many types of equipment that would be used in the theater and elsewhere. It is also quite straightforward to develop new software using this protocol. As the ACN protocol mentioned in Section 3.4 gains traction, this system could be reimplemented to use it instead of or in conjunction with OSC.

As part of a unified theatrical design, the output of the system can extend the influence of Simon in The System to beyond the walls and other output

renderers to numerous parameters in other systems and scenery to present the omnipotence of The System throughout the environment. For example, the representation of Simon in The System would likely require control over lighting effects and color in response to mood or gesture. Since lighting will be controlled by off-the-shelf systems, it is trivial to translate the OSC messages emitted by the system into a form that a lighting console or other common show control system can understand. Similarly, the Disembodied Performance system can be used to control sound generation, synthesizers, and trigger events. All of these systems constitute output renderers.

It is also possible for the performance system to receive data from other show control systems. Just as it is likely that lighting will be controlled by the system at times, output representations may require feedback about the stage lighting in order to render certain illusions and effects. For this purpose, a generic OSC `InputDevice` implementation was written. The device can allow for the development of new modeled data sources using OSC or, more likely, to provide direct data, timing data, and metadata from other show control systems or interfaces. OSC musical interfaces or other devices can be used to trigger cues, for example.

Altogether, a broad range of possibilities for integration into a theatrical environment exists. The Disembodied Performance System can readily take its place in the realm of technical stage production as a new tool for creating expressive theatrical experiences.

5 DISEMBODIED PERFORMANCES

If you know your character's thoughts, the proper vocal and bodily expressions will naturally follow.

—Konstantin Sergeyevich Stanislavski

Utilizing the Disembodied Performance System detailed in the previous chapter, three sets of data were collected during the summer of 2009 for analysis, testing, and offline design of output representations. These data will serve as an invaluable resource as the design of *Death and the Powers* is finalized prior to rehearsals scheduled for the summer of 2010.

Due to budget and schedule changes, equipment was not purchased and the actual output devices and set pieces were not constructed.

Consequently, this work has yet to be evaluated extensively at the envisioned theatrical scale. Nevertheless, the evaluations of the system and approach that are described in this chapter present promising preliminary results showing the efficacy of maintaining an actor-driven presence through alternate expressive representations.

5.1 Proof of Concept

Three capture sessions were conducted with the Disembodied Performance System. A performer was outfitted with various wearable sensor prototypes. In each case, the `DataRecorder` output renderer was used to record raw data values directly from the input devices. The data was sampled and recorded at 60Hz. The `DataRecorder` renderer writes the input values for each axis to a plaintext file, including some metadata about the input axes calibration settings, cues, and time code. The recorded files can then be replayed through the Disembodied Performance System using the `DataPlayer` input device. In this way, the real-time input can be recreated from a recorded

performance, allowing mappings and designs to be explored without requiring a performer to be present and connected to the system.

For each of these capture sessions, a high-definition video of the performer was also recorded. The video provides a record of any comments made during the session and a visual reference for the gestures when additional clarification is needed. The audio of the performer singing is also recorded if further experimental analysis is desired. The video was synchronized with the captured sensor data using an audible tone mixed with the microphone input on the camcorder at the start of recording by the Disembodied Performance System.

The first capture session was conducted for the purposes of demonstrating how the system works. My colleague and designer of the wearable sensor assemblies, Elena Jessop, wore the breath sensor and a first prototype of the arm gesture sensors on her right arm. For this session, no other sensors were used, nor was audio analysis incorporated. Like its successors, the arm assembly featured two three-axis accelerometers, but was constructed as a full-arm glove with straps that would cinch the glove in place on the arm. On a large screen, a video trailer that had been created to convey an impression of the story of *Death and the Powers* was shown along with music from Tod Machover's score. Jessop has experience in theater and interpretive dance. I asked her to gesture expressively and sing or vocalize along with the video playback. She rehearsed this choreography several times and three recordings of her performance data were made subsequently.

For the later demonstration of the Disembodied Performance System, one recording of Jessop's captured performance was played back through the system, synchronized with a second computer showing the trailer video. A third computer was set up with an output renderer depicting a 3D computer rendering of the stage setting for the opera. The output renderer was written to interactively control the color and intensity of a simple five-point lighting setup on the virtual set. Additionally, the display surfaces of the virtual walls could show visual renderings, simulating the look and function that the actual walls for the set will possess. In a manner similar to the system's usage in the actual production, both the lighting and visuals displayed on the walls are controlled by mapping the modeled output from the Disembodied Performance System.

An initial look at the recorded data reveals several notable properties. There were, however, several errors with the sensors due to a poor connector used in the wiring. Much of the time, the elbow accelerometer axes are clipped at the maximum by the analog-to-digital converter, particularly the x-axis, which was aligned to the gravity vector. Furthermore, the x- and y-axes of the wrist accelerometer were coupled. Despite this, distinct sections of



Figure 67: Performing with sensors

For the proof of concept data collection, Elena Jessop reacted expressively to a trailer video for *Death and the Powers* while wearing one arm gesture sensor assembly and a breath band sensor.

different qualities can be apprehended on visual inspection. The breath signal clearly defines phrases, which include the onsets of the sections. Comparing the data with the video that prompted the performance, the sections correspond precisely with sequences in the video. The quality of each section is given, for the most part, by the amplitude of the accelerometer data.

Mappings from the input data to the affect space model were kept relatively simple, due to the problems in the accelerometer data. Meaningful correlations across axes were difficult to achieve with clipped and coupled values, so normalized amplitude was the most important factor. To compute statistical values in this demonstration, a one-second window was used for each axes' `DataStream`. The one-second value was chosen as it yielded changes in the output representation that appeared to synchronize most closely with the average shot duration in the trailer video during simultaneous playback. Correlation in acceleration rugosity contributed to the valence model parameter while amplitude was the primary component influencing the arousal parameter. The stance parameter was computed as a function of the maximum in the sample window and respiration amplitude.

Color targets within the model space were chosen to correspond with the various palettes in the video with consideration for the emotional connotation in the affect space. This worked well, since affective design decisions had been made as a normal part of the video production and color grading. For the output representation, arousal was mapped to both lighting intensity and density of the visuals on the walls. Valence controlled the orderliness of visual elements on the walls and breath information served as a direct data stream influencing the rate of motion in wall visuals. A threshold of the derivative of the respiration signal also triggered changes in the representation, synchronizing the phrasing of the movement with sequence changes in the video. Overall, this test demonstrated that an expressive intent, prompted by the video and interpreted in the performer's gesture, could be translated into a non-anthropomorphic representation through the Disembodied Performance System.

5.2 Test Performance Capture Sessions

Buoyed by the results of this initial session, two recording sessions were scheduled with the actor and opera singer who will be portraying Simon Powers, James Maddalena. In this section, I will describe the procedure followed for these two sessions. A review of the data collected will be given in the next section.

In preparation for the next session, the design and construction of the arm assembly was refined to the form described in 4.5.1, so that it would be

easier to don and more comfortable to wear. The electrical connectors were replaced in order to eliminate the erroneous sensor values that were observed in the first session. With the revised design, a second arm assembly was built. Some minor refinements were made to the Disembodied Performance System software to streamline data recording for test scenarios, ensure regularity in sampling intervals, and add additional planned functionality to the user interface.

Maddalena wore both arm assemblies and the breath sensor. As before, the performance was recorded on synchronized video for reference and the sensor data and audio analysis (amplitude, frequency, and consonance) were recorded through the Disembodied Performance System.

I asked Maddalena to perform three sorts of activities, for which data was collected. The first activity recorded some baseline sensor readings of the actor simply wearing the sensors without making any deliberate gesture. Also, for reference, data was recorded of the performer conversing calmly with others in the room and then singing a few selections of his choice, without any direction beyond, “Feel free to gesture naturally as you sing.” No explanation of the sensors, the types of data being recorded, or how the data would be mapped was given before the session, so as not to encourage deliberate control of sensor parameters.

In the second activity, I provided spoken prompts to which Maddalena would respond by singing a scale while acting with the appropriate emotion, pausing for a period of time between each response and the following prompt. The actor was allowed to consider the prompt before beginning to sing and gesture. This consideration period was intended to be a substitute for the rehearsal that would be a part of a normal production scenario. An actor’s performance onstage is generally not a spontaneous response to unexpected stimuli, but a carefully considered behavior. As such, the responses to my prompts were allowed to be somewhat planned.

In order to develop input mappings for an affect space model, the prompts given were single words describing basic emotions. The emotion words were derived from two sources: the eight emotions Manfred Clynes used as prompts in his sentograph experiments [16] and words for given locations in the affect space used by Cynthia Breazeal [11]. Not all words from each set were given as prompts and a couple of words overlapped. The prompts from Clynes were: no emotion, joy, grief, sexual desire, anger. The prompts from Clynes represent one set of basic emotions and the data for these prompts could be examined for the known sentic forms. The prompts from Breazeal were: calm (neutral/no emotion), joy, sorrow, disgust, fear, surprise, acceptance, tired, anger. Breazeal gives the locations of that set of prompts in the affect space that span the space, which provided known loci



Figure 68: James Maddalena wearing sensors
Opera singer James Maddalena performed during test capture sessions wearing two arm sensor assemblies, breath sensor, and shoe pressure sensors.

when developing input mappings, allowing detected features to be pinned to known locations.

The third activity conducted by Maddalena was to make deliberate instrument-like gestures with his arms and hands, the kinds of features that would constitute direct data. While the Disembodied Performance System is not intended to be used in an instrument-like fashion, it has this capability. *Death and the Powers* composer, Tod Machover, who was present at this test capture session, wanted to explore the possibility of having Maddalena trigger sonic events and effects with such deliberate hand and arm gestures, including pointing in space.

During this session, Machover, Jessop, and I observed several additional behaviors Maddalena exhibited while performing the activities. Most notably, Maddalena shifts his weight up and back between left and right feet in an expressive manner. Also, there seemed to be a fair amount of expression conveyed in the orientation and posture of his shoulders.



Figure 69: James Maddalena performs with sensors

Maddalena sings and gestures “grief”. Note the LED tracking markers in addition to the sensors.

With this in mind, the foot pressure sensors were added for the third test session. Also for that session, LEDs were added to the hands, connected to the arm assembly, and on each shoulder. The relationship between the two shoulders and each shoulder with its corresponding hand would provide a modeled data stream while the hand locations could supply direct data streams. White LEDs were used as they showed well on the camcorder used to record the session (especially when the image was slightly defocused) and could easily be tracked. Unusually, this particular high-definition camcorder was not sensitive to infrared light. A computer to run the motion-tracking algorithm was unavailable at the time of this session, so no computer vision data was included in the recording. The recorded video could be processed to recover the tracking information for analysis and incorporation into the Disembodied Performance tests, but this has not yet been done.

The procedure for the third session was similar to that of the second, with the addition of the new sensors. The audio analysis was not recorded during this session, but could be recovered from the synchronized video recording. Maddalena completed two activities. The first was again responding to verbal prompts of emotional labels in the same manner as described for the second session. The prompts this time included: calm (neutral/no-emotion), love, grief, hate, joy, frustration, bored, tired, happy, elated, surprise, fear, soothed, disgust, acceptance, and anger. Many of these prompts were repeated to verify the results from the previous session with over a month of intervening time.

In a second task, I asked the actor to sing and respond as he might onstage to four scenarios. The scenarios given were brief (due to limited time)

moments from *Death and the Powers*. Unlike an actual performance of these moments from the production, the actor did not yet know the music and lyrics and did not receive any direction. My hope with this activity was to record small samples of acted emotion varying over time, given the mindset of the character, not simply single-emotion responses without any context.

All three test recording sessions provided invaluable information and demonstrated the effectiveness of the system. In the next section, I'll review some of the recorded data and the implications for the use of the Disembodied Performance System in an actual production context. The actor was very amenable to wearing the sensors and found them to be comfortable and unintrusive, as we hoped. Having already become familiar with the procedures that will be used in the production of *Death and the Powers*, Maddalena will be more comfortable as the production develops. The reference responses to prompts will play an important role during rehearsals as the mappings for the production are created.

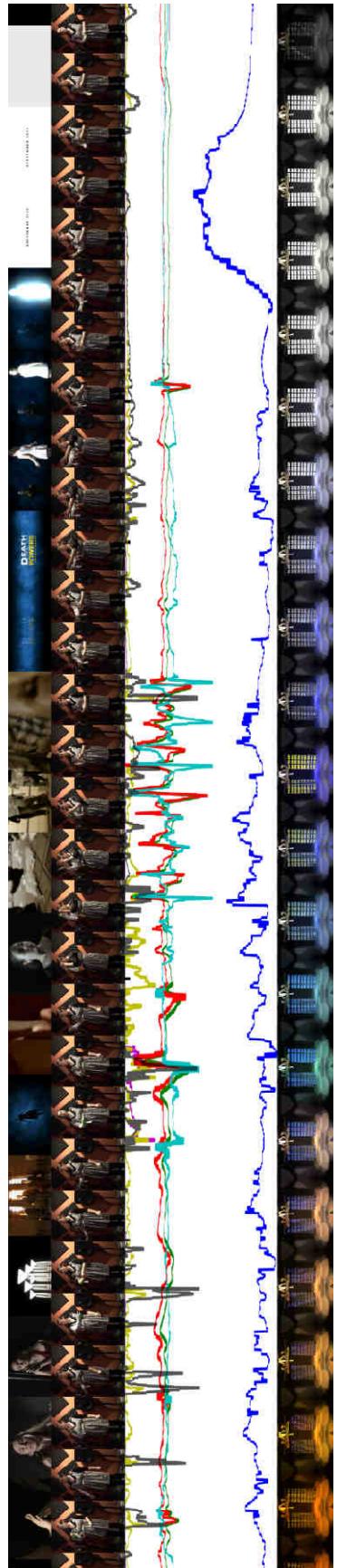
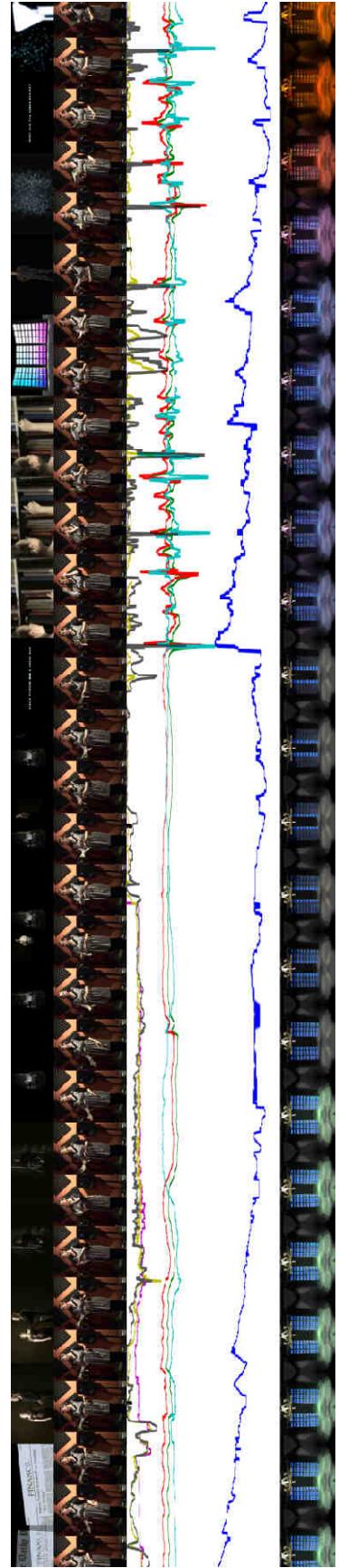
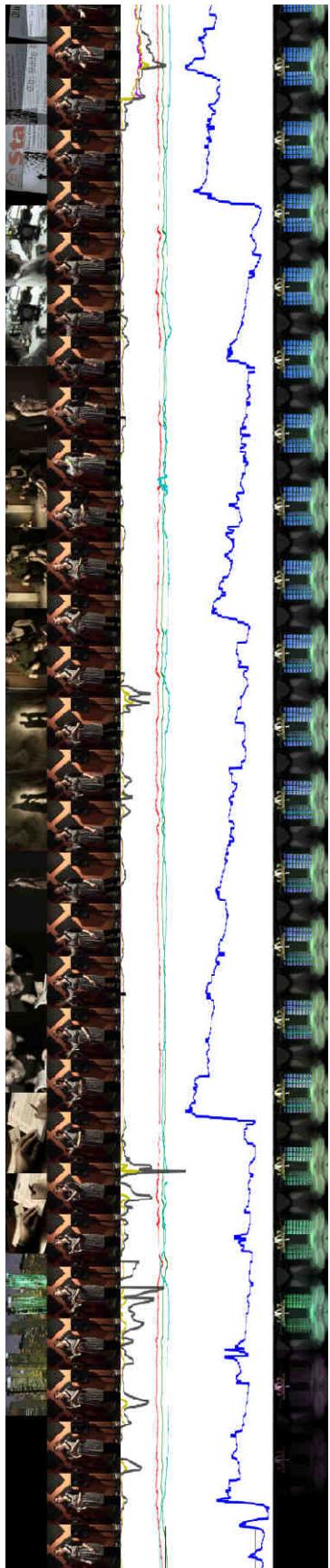
5.3 Recorded Data

In these tests, the Disembodied Performance System was configured to update at a rate of 60 Hz. This matches the rate at which sensor assemblies reported their values to the control system and also provided decent performance on the dual 3.2 MHz Pentium 4 computer on which the software was running. At what rate should the system ideally update itself? The answer to this question depends upon a number of factors including the types of sensors being used, the sensing modalities, and the needs of the output representations. In [49], a rate of 4 kHz is suggested, though this is due to the frequency of the physiological signals, notably the electromyography, that is being captured. Looking at the properties of actual physical movement, the rate of change over time is generally not very fast compared to these sampling rates. Clynes conjectures that expressive actions, atomic units of expressive motor movement, have a duration of about 0.2 seconds [16]. At that rate, 60 Hz is sufficient to recover the contour of actions. Indeed, empirically, I believe I have been able to capture sufficient gestural information at this rate, as can be seen in the plots of sensor data that follow.

Of greater concern is overall system latency from an actor's sudden gesture to a change in the output representation. Through profiling, I have estimated the total system latency to be in the range of 12 to 25 ms and it is expected that additional optimization can minimize these values. The visuals generated from the performance must appear in sync with the actor's singing and music, some of which may be acoustic and some of which is routed through independent systems for audio processing and reproduction, particularly in response to sudden onsets. For comparison,

Figure 70: Proof of concept data and imagery

This sequence (facing page) of four bands illustrates the setup described in Section 5.1. The sequence runs two-and-a-half minutes. At the top are frames from the trailer video. The second band of frames is from video of Elena Jessop wearing sensors and gesturing in response to the trailer. The traces in the third band show the recorded gesture and respiration (in blue) data. Note the issues with this data discussed above. The fourth band shows the stage renderer, described below, which alters the lighting and imagery on the walls of a computer-generated image of the set for *Death and the Powers* in response to the performance data.



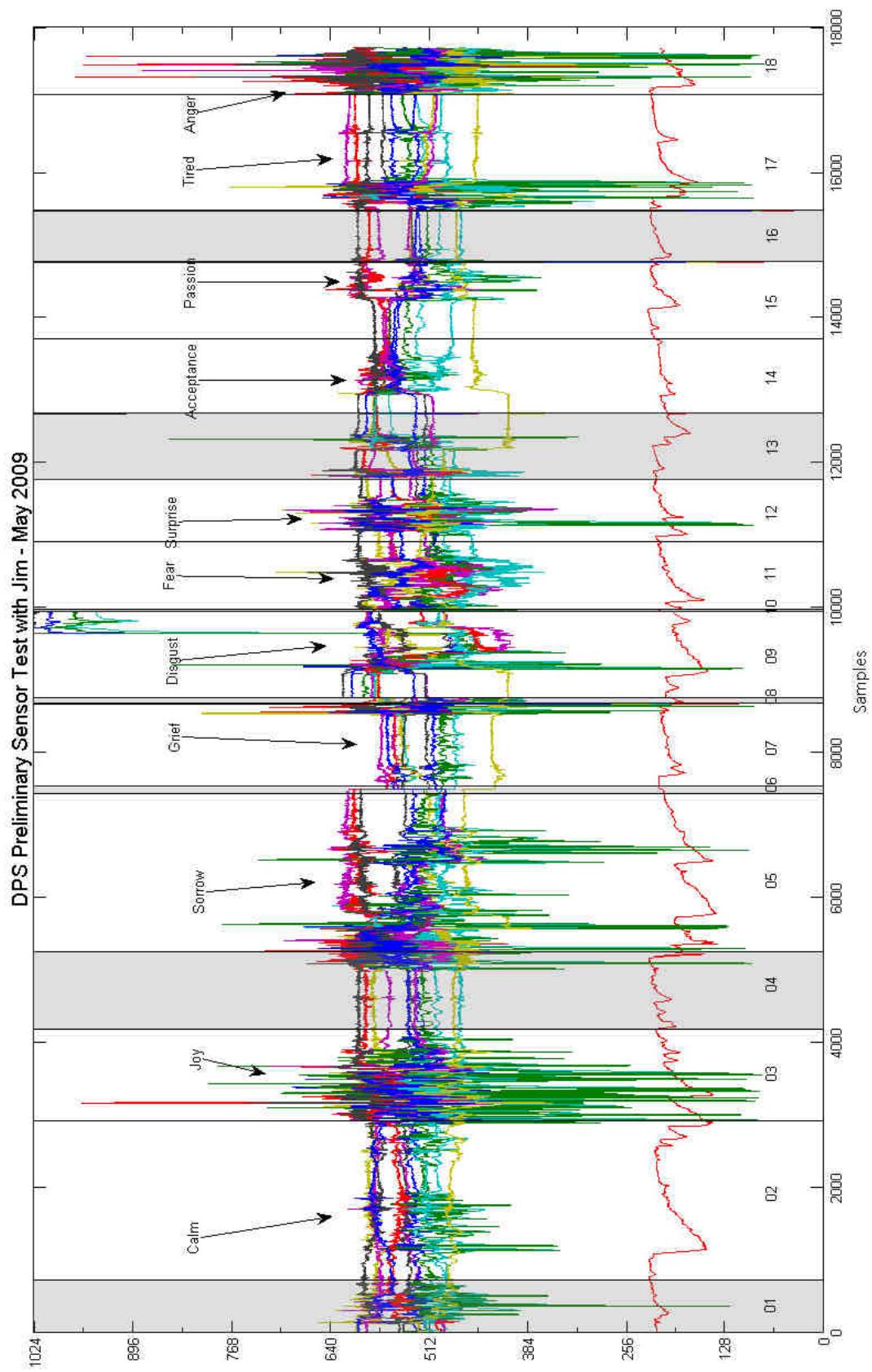
the studies by the Implementation Subgroup of the Advance Television Standards Committee have found that the audio and video programs of broadcast television have a 60 ms tolerance for synchronization [3]. For the audience, the current visual latency should not be a problem as long as other systems have comparable latency. Tolerance for latency can be as low as 2ms, for an individual generating the output that is presented to him. This is a familiar problem in user interface design, computer gaming, and electronic music. However, this is not a concern here since the Disembodied Performance System generally does not provide the performer with a view of the output, as discussed in 4.4.

Looking at the data from the proof of concept recordings, it was immediately apparent that certain qualities could be distinguished by eye. Subsequent numerical analysis showed that many of these properties would be easily captured by simple arithmetic operations during mapping. The mapping interface did prove to be a useful tool for exploring the data, though initial analyses were made in other applications. As noted above, distinct sections were apparent in the recordings of Jessop's performance that corresponded with the emotional tenor of the scenes in the trailer. The rugosity and amplitude of the axes accounted for the visual distinctions. Breath data also indicated accurate onsets of sections or subsections.

Distinct sections and qualities were also apparent in the data collected during the test sessions with Maddalena. Having an arm sensor assembly on each arm revealed that important features can be determined not only from the accelerometer axes at one location, but by distinguishing data from left and right sensors. Respiration data again clearly marked the onset of phrases as Maddalena sang as well as the start of each response to my prompts. It is well suited for timing data, as expected. Data was recorded in the periods between prompts, as well, which shows incidental action that is markedly different from the performance behaviors. Such actions included fidgeting, scratching, pointing during conversation, and adjusting clothing. On rare occasions, these actions occurred during a performance section. Given the informal context of these tests, I expect the actor will control such behaviors even more so in an actual performance context. These gestures clearly appear as outliers with respect to the trends of the performance section, starting and ending suddenly with abnormally high amplitude compared to the surrounding values. In future refinements to the system, it may be desirable or necessary to develop methods for filtering out these unwanted actions.

Looking at the data in Figure 71, we can see that simple trends in the data suggest possible mappings to the affect space model. Amplitude is unquestionably a factor in arousal level. High-arousal prompts such as "anger," "joy," "surprise", and "disgust" all have a generally high amplitude over the course of their performance. The peak locations in the course of

Figure 71: Data from first test recording
This plot (facing page) shows the accelerometer and respiration sensor data recorded from James Maddalena during the first test capture session with him. The numbered sections were segmented during the recording and the gray sections were disregarded, as they do not contain action in response to a prompt. The white sections do contain extraneous motion, but it is generally clear where the acted region begins and ends. The breath data also indicates the onset and conclusion of acted regions. The prompt for each region is indicated. The accelerometer data is clustered through the center of the plot, since acceleration is a signed quantity. The traces break into two groups. The more sparse upper group are the accelerometer axes under the influence of gravity. Raw 10-bit sensor values are shown. The respiration signal (red trace near the bottom of the plot) is inverted due to the nature of the sensor. The downward-pointing peaks are inhalations.



these emotions, however, are distinct and may be important in some mappings. As expected, “tired,” “acceptance,” “reverence,” and other similar affective states exhibit low amplitude.

Amplitude can also contribute to the value for stance, though its relationship is more relative to the contour of the expression than absolute. The variation in accelerometer amplitude for emotions such as “disgust,” “grief,” and “sorrow” is small. Emotions that would require more engagement with the stimulus have greater variance. Some information about stance can also be gleaned from the shoe sensors. Intuitively, for emotions in which the actor is more engaged with the stimulus, he leans forward applying pressure to the front of the shoe. Similarly, a distancing response is enacted when the actor applies pressure to the heels of his shoes. Most of the time, however, the actor stands fairly level and the sensors report little information of interest below their threshold of sensitivity. In these cases, the contribution of the shoe sensors to the stance value should be nil. In actual performance situations, depending upon the mobility of the offstage performer, it may be necessary to disregard footfalls recorded by the shoe sensors. It is expected, though, that the offstage actor will travel very little and the shoe data will play a role in the inference of the character’s affective state.

Low-pass filtering the data over a window of a few seconds results in smooth contours of the accelerometer data. The approximate frequency of the contour provides a clue to the valence of the emotional state. Slow-varying contours generally equate to negative valence expressions, such as “grief” and “fear.” Even though there may be a substantial amount of high-frequency content in the expression, the overall trend of these actions is a smooth dip or decline. Positive valence emotions seem to exhibit either more constant or more erratic trends. In the Disembodied Performance System’s mappings, low-pass filtered data can be accessed from the `DataStream` objects for each axis.

Valence also can be derived by notable differences between left and right sensors. For example, if we compare “joy” and “anger” reactions from the second test recording session with Maddalena, we see several differences (Figure 72). Both emotional states are high arousal and generally moderate to high in stance. However, “anger” is strongly negative in valence and “joy” is strongly positive in valence. First, we observe that the amplitudes of both expressions are relatively high, overall. If one were to compute an envelope for the amplitude, “joy” would be more consistent whereas “anger” has a lower-frequency contour, as noted above. The motion for “anger” is also much higher-frequency than for “joy,” though both are fairly high-frequency with respect to many of the other responses. The most significant difference, however, is that the left arm and hand are more expressive in the positive valence case, while the right side dominates in negative valence. It

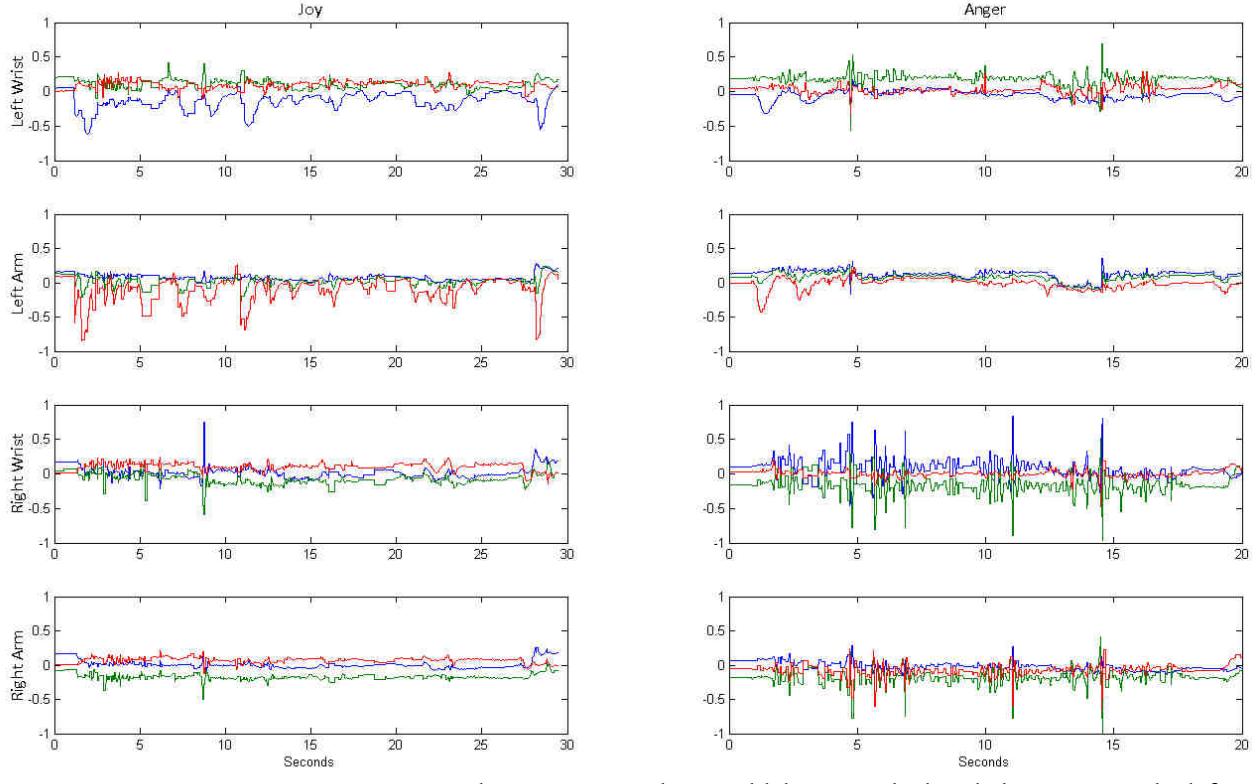


Figure 72: Comparison of “joy” and “anger” motion

Data shown was collected during the second test recording session with James Maddalena. The four accelerometers’ data are compared side-by-side for the “joy” and “anger” prompts, both high arousal but with opposite valence. Each accelerometer’s x-, y-, and z-axes are shown with blue, green, and red traces, respectively.

is worth pointing out that Maddalena is right-handed. Moreover, the left wrist accelerometer data shows most of the movement on its x-axis, which was oriented perpendicular to the axis of the body and parallel to the floor plane, suggesting that the motions for “joy” move laterally. By contrast, the movement of the left hand for “anger” is predominantly in the accelerometer’s y- and z-axes. The correspondence suggests that the hand was canted slightly and the movement is more vertical. The orientation of the movement can be refined by accounting for the influence of gravity and can be used to play a critical role in inferring valence.

Guided by these observations, I experimented with several preliminary mappings in the Disembodied Performance System. The mappings collapse the input data into the three-dimensional affect space (Figure 65). As rehearsals for *Death and the Powers* get underway, the final mappings for this production will be designed to meet the aesthetic and performance needs of each scene or cue in the opera.

5.4 Output Representations

Let us now look at a few examples of visual output representations. These representations were created as part of the construction of the Disembodied Performance System and to research possible approaches for *Death and the Powers*. However, due to the changes in the production schedule, the final output renderers for the opera have not been designed, though they may be

derived from those presented here. Future discussions with the creative team and additional design work are needed to arrive at final implementations that demonstrate the true expressive power of this approach within the constraints and particular aesthetic of *Death and the Powers*. The examples here illustrate only the basic capabilities and offer points for discussion of the process of designing an output representation of a performance.

While I will show sample frames from these output renderers, their expressive power lies with their dynamic response. In these examples, the input mappings to the affect space model abstract the actual input modalities, but they do not project time into the model. Since affect is encoded transiently in sentic modulation, it is expected that a time-varying output representation would contain more expressive information than a static one. Nevertheless, the images on these pages do preserve the qualities that are modeled and demonstrate the beginnings of what can be expected from Disembodied Performance.

Five example visual output renderers are shown. Each of them was implemented by extending the abstract renderer in the manner described in Section 4.8.5. OpenGL provides hardware acceleration and the ability to stylize the renderer with fragment shaders that add various types of glow or polygon smoothing. Since renderers share a common interface, some of the renderers can also be composited for cue transitions or to build compound renderers. All but one, the `ParticleRenderer`, includes the use of color chosen based on the affective model.

The output mapping for selecting a color value was similar in each case and is most easily described using a hue, saturation, and luminance color space, though the actual implementations may use other color models internally. Hue values were pegged along the unit circle of the valence-arousal plane (Figure 73). This approach is similar to the analogy of color space to emotional space described in [9]. Although the hue mapping used is neither monotonic nor linear, positive arousal contributes warmer hues, while negative arousal is cooler. Positive valence tends toward purples and yellow, while negative valence introduces greens. For every discreet instance of a color at a given time, the hue was sampled from a distribution centered around the hue defined by valence and arousal. A function of stance and arousal controlled the variance of that distribution.

In all renderers, there is a two-dimensional focus that is mapped from the x and y focus parameters of the direct data. The focus is used to generate localized phenomena, typically when the value for stance is high, to suggest Simon in The System's attention is on a specific character or area of action onstage. The focus serves as the particle emitter or the source of perturbations from which motion phenomena originate in the display. In

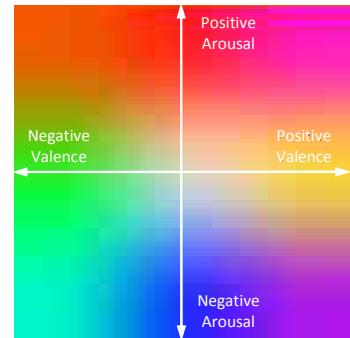


Figure 73: Hues in the Arousal-Valence plane

In output representations, hue was mapped onto the Arousal-Valence plane of the affect space model.



Figure 74: Particle renderer
Particles swarm to an image of a woman as the force gradient.

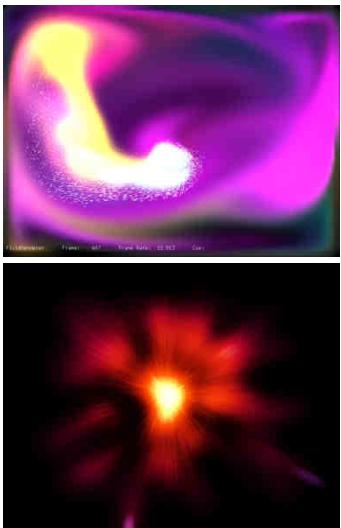


Figure 75: Fluid renderer
In response to the “passion” prompt [A = 0.6, V = 0.8, S = 0.9] (top) and the start of the “anger” prompt [A = 0.7, V = -0.8, S = 0.8] (bottom).

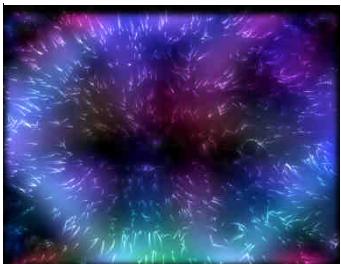


Figure 76: Lumigraph renderer
The Lumigraph output renderer is inspired by Oskar Fischinger’s color instrument. A virtual surface is displaced so that it intersects a plane of light.

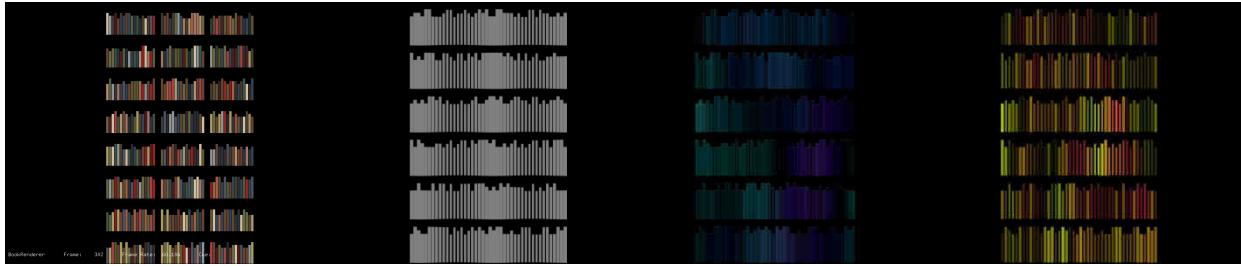
some renderers, the size of the focus is governed by the value for stance, with a larger focus area corresponding to low stance.

The first output renderer for the Disembodied Performance System was an adaptation of code created as a concept for the Memory Download scene in *Death and the Powers*. In this **ParticleRenderer**, a 2D particle system responds to a force gradient. For Memory Download, the gradient would be photographic images so that the particles would be attracted to areas of highlight in the image. The force attenuates the particle’s velocity, but particles can achieve a velocity to overcome the force, so they do not become static. The result is a dynamic monochrome rendering of an image where the brightness of a point in the image corresponds to the probability of a particle crossing that point. The images can be changed on the fly or video can be used as the gradient to create fluid transitions as the particles continually adapt.

As an output representation for the performance system, the force gradient is an image of the focus with decay. The decay and particle velocity can be controlled in response to modeled parameters. **ParticleRenderer** can also be used as a composite renderer with one of the other renderers described below providing the force gradient for complex behaviors. Controlling particle velocity with the breath timing data produces an interesting effect and sense of presence. However, the **ParticleRenderer** overall is not especially expressive. Additionally, the imagery it generates is not well suited for projection onto the books for *Death and the Powers*.

The **FluidRenderer** uses a fluid dynamics simulation to generate glowing imagery. Perturbations in the fluid produce a visual trail that combines with the existing gradient of the simulation to produce very dynamic results. The direct data that drives the focus can be somewhat expressive in this renderer, since the effects of a motion in the simulation compound over time before decaying. The most important parameter for this renderer is the viscosity of the fluid, which can be modulated by arousal and stance to produce a variety of gestures. Valence is indicated by the color injected into the fluid at the focus and by the quality of the perturbations at the focus. Negative valence increases the amount of noise in the focus location.

Inspired by Oskar Fischinger’s Lumigraph, I created the **LumigraphRenderer** that operates on the same principle as Fischinger’s instrument. A triangle patch surface covers the viewing area to simulate the rubberized screen of the Lumigraph. The vertices have a 3D location that is computed with a damped spring dynamics simulation along the edges between them. Impulses are applied to displace the surface as a function of modeled data and the current focus value. As the surface is displaced, it intersects a narrow region illuminated by three light sources, which rotate around the axis perpendicular to the surface at a rate controlled by the



arousal level. The color of the light sources is drawn from the hue distribution based on the modeled data as described above.

The `BookRenderer` is designed specifically for display on the walls for the set of *Death and the Powers*. It implements a pixel-per-book display, with each book having a single color and displacement value at any given moment, yielding seven degrees of freedom: (t, x, y, h, s, l, d) where t is time; x and y are spatial coordinates of the book in the array of books; h, s, l are the parameters of some three-dimensional color space to describe the color of the block; and d is the displacement or distance the book is extended outward from a resting position. The displacement is a holdover from designs for the set featuring actuated books (see Section 3.3.4) and is implemented as a visual effect where the size of the book is increased along its width and upward to simulate the book moving forward perpendicular to the plane of the wall. This approach was used in the projection tests and effectively gives the impression of a moving book when rear-projected onto the curved spines. Although expressive on their own, the `FluidRenderer` and `LumigraphRenderer` can be downsampled to pixel-per-book resolutions and used to control the color and displacement of books in the `BookRenderer` to achieve complex expressive results.

The final renderer I'll mention is the `StageRenderer` that we've already seen in the discussion of the proof of concept demonstration. This renderer uses pre-rendered images of the theatrical set to control interactive lighting of the set. The hues of the virtual light sources are chosen in a similar manner as in the `LumigraphRenderer`, with the modification that the most complementary hue is assigned to the backlight. The intensity and saturation of the illumination is influenced by arousal and stance values respectively. Direct data and valence also determine whether lighting comes from only one direction or multiple directions. The `StageRenderer` is also a composite renderer. The output of another renderer can be texture-mapped onto the books of the walls.

The overall display is modulated by the breath timing data, which provides the impression of life in the output rendering. For the `FluidRenderer` and `LumigraphRenderer`, the breath influences the viscosity and spring tension, respectively. In the case of the `BookRenderer`, the books undulate subtly in and out in response to the actor's breath. This gives a clear impression of breathing or presence. At low values of arousal, when changes in book

Figure 77: Book renderer
For low values of arousal and stance, the current implementation transitions to 'book colors' (far left). Modifying the size of the book produces a displacement effect when rear-projected onto curved spines, as if the books were moving in and out. Here, ripples in book displacement are shown (center left). The two images on the right show the `BookRenderer` displaying output of the `LumigraphRenderer` at pixel-per-book resolution.



Figure 78: Stage renderer
Rendering of wall projection and dynamic lighting on a virtual stage from the proof-of-concept demonstration.

depth and other parameters are attenuated, the periodic effect of breath is quite perceptible. As arousal increases, other factors dominate the expression. Sergei Eisenstein believed that editing film at approximately the pace of human heart rate made it particularly compelling [84]. In our case, breath is an ample substitute and can time transitions in color to breath onsets providing a visceral connection to the actor.

The possibilities for output representations are endless. In the examples above, most of the generated imagery has a fluid organic form that is subtle and controllable. The amorphous `LumigraphRenderer` and `FluidRenderer` outputs are well suited to full resolution display on the walls or to downsampling to pixel-per-book resolutions, while retaining their expressivity. In other contexts, geometric imagery could be used, as it engages other aesthetics and has a broad range of parameters. Geometric imagery was not explored for *Death and the Powers* because I felt it would clash with the rigid grid structure of the books and bookshelves.

6 CONCLUSIONS AND DISCUSSION

As for the future, your task is not to foresee it, but to enable it.

— Antoine de Saint-Exupéry

In the course of this work, I have defined a new type of expressive human performance that ventures outside the bounds of traditional theatrical representations of actors onstage. Combining ideas from cognitive science and perception, I posed the method of reified inference, the beginnings of an approach to modeling representations and synthesizing new ones while maintaining the perceptually salient properties of the original. Armed with this theoretical technique and an awareness of the traditions and innovations in theater technology and dramaturgical context, I present a new way to think about and implement augmented performance systems called Disembodied Performance. This thesis also documents the history and process of design on the particular production of *Death and the Powers*, which played an important role in defining the requirements of Disembodied Performance and how the technique should be applied in practice. I employed reified inference to distill the essence of a character from parameters recovered from an actor and allow the performance to be extended out into the environment. This mapping-based implementation abstracts away the body in a meaningful way, allowing a larger-than-life character to become anything, to exist anywhere, even in non-anthropomorphic manifestations, offering the potential of greater ranges of evocative, intelligible, and compelling expression. This Disembodied Performance System addresses the critical design goals, laid out at the beginning of this document, and has been flexibly designed to coexist alongside other designed elements and theatrical control systems and technologies. Three experimental sessions were held to capture data from performers that show the efficacy of the capture methods in ascertaining the character's affect. These data were then used to demonstrate the expressive possibility of alternate representations of performance in preliminary representation designs.

It is hoped that the cross-disciplinary research demonstrated in this thesis is just the beginning of this new way of considering performance representation and that Disembodied Performance will find a welcomed role in the portrayal of appropriate characters in theater, film, and new media. Certainly, the work of refining the Disembodied Performance System and crafting the visual and multimodal aesthetic of Simon Powers in The System for *Death and the Powers* will continue leading up to the opera's 2010 premiere. I also intend to develop further these ideas in my continuing research in the areas of crossmodal representation and mapping and apply them to numerous applications in performance, telepresence, storytelling, and interaction design.

I conclude this thesis with a discussion of some of the issues Disembodied Performance addresses. Following this, I lay out the next milestones with the current implementation, in preparation for *Death and the Powers*. I then step back and take a broad look at the potential applications for the theory of Disembodied Performance and a discussion of open areas of exploration.

6.1 The Disembodiment Problem

The Disembodiment Problem, as defined by Teresa Marrin Nakra, tends to arise whenever technology is introduced in performance contexts. Simply stated, the problem is that technology often adds a level of indirection to a performance that obfuscates what is responsible for the perceived results [49]. This is a common problem with electronic instruments, and the Chandelier in *Death and the Powers* is no exception. The instrument functions, but is controlled often indirectly by a human performer. Since the instrument is not acoustic, its sound must be generated by physical, as with the Chandelier, or electronic means and reproduced by speaker systems that are often not physically associated with the instrument. How does a listener then know that the instrument is producing the sound? How can the listener tell that a performer at a MIDI keyboard, or even a computer keyboard, is performing the instrument? With these two levels of indirection, the expressive and visceral quality of a performance is difficult to maintain. The disembodied problem also makes gestural instruments and interfaces weak, as their mappings are often opaque and don't seem to relate to the output.

In contrast to much work with using sensor based and gestural instruments, I cite two examples that overcome this problem in different ways. The long-revered theremin is an electronic instrument patented in 1921 that allows a player to move her hands to create sound. Each hand controls one dimension of the output: frequency and amplitude. This simple mapping makes the correlation between the performer's gesture and the sound produced intuitively apparent to an observer. While the controls may be



Figure 79:VAMP

The Vocal Augmentation and Manipulation Prosthesis uses a gestural vocabulary that physicalizes the sound of a singer's voice. Here, Elena Jessop is grabbing the note she is singing in order to manipulate it.

simple, the instrument exhibits a subtlety common to many traditional acoustic instruments and entails a virtuosic mastery [31]. A second example is the Vocal Augmentation and Manipulation Prosthesis (VAMP), which uses a simple sensor glove with relatively complex mappings to allow gestural control over the performer's voice. While the mappings from sensor data to processed vocal output are less direct than the distance of a hand from an antenna controlling frequency and amplitude of a theremin, VAMP is designed specifically to virtually embody the voice. The gestural vocabulary was constructed to give the impression of physicality to the performer's voice, allowing for a sung note to be grabbed and held, for example [37]. The result is a gestural instrument that is especially compelling to an audience and does not suffer from the confusing or distancing indirection that typically plagues such new technological interfaces.

Disembodiment is generally an undesirable consequence of the applications of these technologies. However, in this thesis, I have actually cultivated the notion of disembodiment rather than gone to great pains to avoid it. In general, this is out of interest in devising techniques to overcome the problem and maintain a compelling experience despite technology. In the specific case of *Death and the Powers*, embracing disembodiment was a necessity of the story. In fact, the technology was essential to creating a disembodied experience that does remain compelling and expressive. I believe to have arrived at an implementation that satisfies this goal.

Having the performer offstage and invisible to the audience is both a blessing and a curse. As noted in Section 4.1, there was concern about where the actor playing Simon Powers would be once the character enters The System. A need was perceived to have the actor somehow visible to *prove* that a genuine performance is being given and to demonstrate that the behavior of the environment is indeed related to the actor's movement. Certainly, it is a difficult enough challenge to make the connection between a gestural musical instrument and the sound produced when you see the instrumentalist before you producing the sound. Surely, it would be even more of a problem with the performer offstage with no way to relate the output to the actor's input for the benefit of an audience. I would argue that this belief points out a problem not of the idea of disembodiment, but of the mappings we've come to expect from these instruments. If the output representation is required to have its own presence, then a different approach to mappings must be taken: one not of generating an effect, but of generating a complete transference of the prized qualities of intent and effort and immediacy. Disembodied Performance addresses this explicitly. Even though it is a mediated system, I have not conceived of it as a layer of indirection. It is not to be viewed as the output of a process. What the audience sees synthesized onstage, regardless of the form, is the character.

The analogy to puppetry is apt here. The system is nothing more than the strings of a marionette connecting the handheld controller to the body.

While having the actor driving the system offstage presents a challenge in terms of the quality of the mappings, it brings with it some leniency of not needing to make the connection between body and generated representation so explicit. For example, it's quite intelligible to have the vertical height of the actor's hand from the floor control the intensity of blue displayed on the walls, but it is not expressive. Without the need to address the direct connection between the cause and the effect, the mappings can focus on communicating the intent, rather than the mechanics. This is particularly the case for representations that are non-anthropomorphic. They can be expressive and communicate in ways that we, as audience members, aren't consciously aware. Within some instant of time, the brain can perceive and interpret more than we can take the time to reason through. This is likely the result of our tendency to apply the metaphor of a musical instrument to a performer onstage having control over a device. We have experience with and can understand the causality of a drummer striking a drum or a harpist plucking a string or a violist trembling his hand to create vibrato and can see that the sound stops when the bow leaves the string. If that causal link isn't apparent, as it often is not in mediated systems, then the audience feels dissociated, or worse yet, loses the suspension of disbelief that holds them in the world of the story. If the audience is not engaged with the story or the music, in the case of a concert, the significance of the experience is lost.

Avoiding the instrument paradigm also presents an advantage to the actor. The system is transparent to him and he needn't learn specific behaviors or be trained to play the Disembodied Performance System. Again, in the case of sensor-based electronic instruments, the mappings tend to be complex and not particularly intuitive, especially to performers other than the creator of the instrument. In most cases of new instruments, it is the inventor that performs with the instrument and has the intimate understanding of its functionality to produce deliberate results. On the other hand, these inventors are not necessarily trained musicians and so necessary musical ability may be absent from the performance or musicality may be missing from the design of the instrument as a whole. Since we're interested in the actor's natural movement, he doesn't need to learn how the system works or make complex and unnatural movements to create a desired effect.

6.2 This Side of the Uncanny Valley

With the freedom to choose any representation with the actor offstage, the representation avoids the expectations associated with the instrument

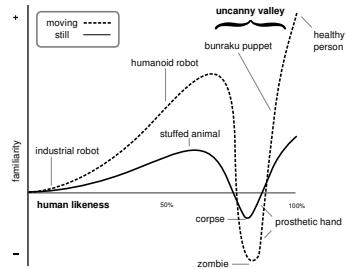


Figure 80: Uncanny valley

The uncanny valley refers to representations that fall just short of actual humans (or animals) and make those that interact with them uncomfortable. (Reprinted from [56])

paradigm. Furthermore, truly abstract and non-anthropomorphic representations not only have the potential to be expressive, they escape the uncanny valley. If the representation of the actor onstage were humanoid, as in much of the work with virtual actors and dance, anything short of a realistic appearance or the most expressive performance would be off-putting. What is seen onstage is not a human and not supposed to look human. It's also not an instrument. In both cases, the representation or combinations of representations that haven't been seen before, for which the audience has no frame of reference, and that don't exist in reality or tradition have no point of comparison in the audience's mind to fall short of.

Without relying on the instrument metaphor or the form of the body, we find other ways to make the presence compelling and to enable the audience to engage with the character. I believe the immediacy of response to the action onstage and to the music is a very important factor in creating a believable representation. The performance, in its abstract form, is itself genuine. It's the representation that is unique. If the changes feel right and have a natural cadence that fits within the whole of the production, it will feel alive and not computer-generated or pre-recorded. However it looks onstage, whatever properties convey the meaning and emotion of the character, it is the timing and the sense of variability and chance that will sell the performance. Like most of the qualities Disembodied Performance communicates, it is difficult to articulate what exactly it is, but we know it when it exists. I believe this is a significant distinguishing factor of this work from related work. Even in describing Disembodied Performance to Media Laboratory visitors and during presentations, without them actually seeing the performance, the most common comment I receive is spoken with surprise and awe: "So the performance will be different every time?" The answer is of course, "Yes, absolutely." We're watching an actor perform. The question isn't asked of a human onstage, but is something that is apparently incongruous with public conception of bringing technology to the stage.

6.3 Next Steps

As in any software or system, there are areas for improvement and further development. Some minor improvements to the current implementation of the Disembodied Performance System are planned for the near future, particularly to add additional flexibility to the output mappings and file representation. The user interface will be enhanced to provide access to more of the capabilities of the underlying architecture and provide improved cue management. For example, in addition to the panel controls and list in the current version, a new Cues tab will display a detailed list of cue properties and allow them to be adjusted. Currently, some of these

adjustments can only be made in the XML files for saved shows. Cue transitions need to be developed further, as well. At present, cues can only cut from one to the next, though additional cue types are defined to allow linear or non-linear blends from one state to the next over a defined period of time. This will require computing the values for multiple mappings simultaneously. Planned technical improvements will modify the threading model and event handling of the application to optimize performance.

I am very much looking forward to my continued work on *Death and the Powers* as scenery is constructed, working with actors, both James Maddalena and others that have to interact with the disembodied representation, and with music from Tod Machover's score. Leading up to the rehearsal periods scheduled for mid-summer 2010, the physical set pieces and infrastructure will be completed. Several questions about the show's production design remain to be resolved, including budgetary considerations, before this happens. Shortly, however, a final design of the walls and other set pieces that constitute The System will be decided. Also during this time, the look of the show will be refined and output representations will be designed and implemented to fit the aesthetic and story needs of each scene. The integration of the Disembodied Performance System with other show control systems will be addressed with appropriate production staff. For the visual renderers for the walls, coordination of color palettes and lighting moments is critical.

In the interim, I will continue the process of authoring the output renderers using the data already collected from the actor, Maddalena. Future test capture sessions may be required, though I believe at this time that the collected data is sufficient for my work until rehearsals begin. It is possible, though, that additional exploration of representations will suggest a need for additional sensors or modifications to the existing ones. If this is the case, the changes will be constructed and test capture sessions can verify their efficacy.

During rehearsals, we'll be able to see the responses of the output renderers in situ, incorporated at scale into the full theatrical set, and given the data of the actual performance of Maddalena as Simon in The System, complete with Machover's score. This period will be essential, not only in fine-tuning the performance of the system, but for truly evaluating the effectiveness of the system with input from these experts, other production staff, and those outside of the project. I'll have the opportunity to receive essential feedback from the creative team on the success of the Disembodied Performance System in practice. An effective onstage representation will deliver the character's presence. Feedback from the other highly skilled actors onstage will assess how well they feel they are interacting with a character inhabiting the environment. Most importantly, I will assist Director Diane Paulus in drawing out her desired performance from the Disembodied Performance

System. Treating this new form just like any other actor in rehearsal, if she can achieve the emotional resonance she envisions from this performance, the system will have been successful at providing a representation that can take direction. Early in the rehearsal process, additional changes to the output renderer implementations may be made. However, the bulk of the work will be the creative process of tuning the mappings in order to achieve Paulus's vision.

6.4 Future Directions

The work I've presented has far-reaching implications for future research and applications in the realm of theatrical performance. Furthermore, the methodology and many aspects of the software infrastructure can offer new perspectives in the domains of remote presence, personal archiving, and storytelling. All of these areas are ripe for exploration, using the ideas in this thesis as a springboard for future innovation.

6.4.1 Beyond *Death and the Powers*

The conceptual and technical contributions of the research presented in this thesis, although heavily motivated by the production needs and the story of *Death and the Powers*, are by no means tied to this opera. The ideas at the foundation of Disembodied Performance are immediately generalizable to other theatrical works and the implementation of the Disembodied Performance System presented can be applied to new productions without modification of the system software. At a high level of granularity, the system effectively separates content from representation, encapsulating the roles of performance capture and character modeling as independent from output representation, allowing them to be readily reused. Output renderers specific to the production design of another performance would need to be created and the type of performance may necessitate alternative sensing systems.

When bringing new technologies to the stage, it is of the utmost importance that the role of the technology is considered. The question is not one of what can be done with the technology, but rather that of what can only be accomplished using the technology. Disembodied Performance supplies numerous answers to that question, opening the door for completely new ideas for dramaturgy and scenography that were not previously possible with the quality and immediacy required.

The system is not just for replacing a presence onstage, but can be used to augment presence. Nothing about Disembodied Performance precludes the actor from being onstage alongside the representations generated by his performance. This did not suit *Death and the Powers* well, due to the nature

of the story, but could be useful staging in other plays and operas where it would be appropriate.

Additionally, though the system was designed for stage acting, with modifications to the model and to the methods of data capture, the system can be applied to other sorts of performances. For instance, a musician could play his instrument while generating a visualization or other representation of his experience of performing, rather than a visualization of the sound or music he produces. The performer's music could be accompanied solely by the visual—not unlike Simon Powers singing is heard in conjunction with his onstage representations—but the performer can be visible onstage as well, allowing the real-time novel representation to extend the range and perceptibility of his performance. The types of gesture or physiological features sensed would need to be changed so that the affective model would not be influenced by the necessary movements of playing a physical instrument.

Further study of the output representations will result in improved effectiveness of mappings. The present implementation relies on mappings being explicitly designed to fit the visual language of the production and the qualities of expression to be preserved. The ability to generate output renderers from a set of primitive behaviors for the output modality would ease the construction of renderers that preserve and communicate features of the model. Currently, there are little to no affordances in the interaction or output. This works within the Disembodied Performance context and has great advantages as discussed above. However, this isn't necessarily a good thing in all cases. Taking the notions of perceptually meaningful elements of the representation and re-rooting them in affordances or meta-affordances that take place at a level of abstraction may make the representations more accessible. There exist several possible approaches including incorporating metaphor, an adaptation of the 12 Principles of Animation [87] to non-anthropomorphic elements, or invoking perceived animacy [34].

As part of the design goal to make the Disembodied Performance System transparent to the actor, the methods of sensing the performance were designed to be as unobtrusive as possible. The wearable assemblies have been constructed so that they would be very easy to put on and remove. They also need to be lightweight and not constricting; barely felt by the actor. Most importantly, wearing the sensors cannot restrict or alter the actor's natural gesticulation, to which using wireless data communication contributes significantly. Since the actor is offstage while sensed, for *Death and the Powers*, I was not concerned with having invisible wearables, though this may be an issue in other performance contexts. Nevertheless, the ideal setup would be completely unencumbering, requiring no sensors on the body. A computer vision system would be the most apparent choice to take

this next step toward completely passive off-the-body sensing. However, the current state-of-the-art computer vision techniques cannot capture all of the gestures with the accuracy and detail of the wearable sensors. With only a camera, without visible tracking targets, it is extremely difficult to segment, identify, and reliably track human motion, such as the movement and orientation of limbs. Some promising solutions have been prototyped, such as the Prakash and Second Skin systems developed by Ramesh Raskar, et al. [52,67]. These systems use infrared coded light and inexpensive passive markers to record high-frequency, high-precision position, orientation, and incident illumination intended as a motion capture solution for film and video. Although this does require the actor to wear special electronics, the tags are sufficiently small to be comfortably and inconspicuously embedded in some types of clothing and costume.

Disembodied Performance has positive implications for the future of theater and storytelling media. Clearly, it is a technological approach for augmenting performance and is not intended to replace, nor can it replace, actors or their presence in a space. Lacking this potential is distinct from the ultimate aim of general research into virtual characters. There is a focus on live performance as a collaborative process in a traditional venue, though the system presents opportunities beyond that. It is also not intended to create a notion of a skilled disembodied performer, an actor who is specially trained and experienced with giving disembodied performances. The transparency of the process to the actor ensures this. The actor acts normally, as if onstage, and the system relies on that type of behavior.

Through Disembodied Performance, stories can be told onstage that cannot be told by human forms alone. Theater has utilized countless methods to do this in its long history, but only puppetry has come close, allowing a human to assume another form onstage without engaging a representational metaphor or indirection. Such non-anthropomorphic representations are historically accomplished with suggestion or stage trickery. Cinematographic technique, such as point of view and unseen characters, and visual effects and animation in television and film have further extended the range of performance by non-humanoid forms. However, in the cases from theatrical gimmicks to visual effects, these representations are not generated by an actor, but rather by technicians and animators. In some cases for animation, the representation is created in part by data from a physical performance, such as tracing or motion capture, which can preserve expressive qualities, but the reliance on the physical form of the actor constrains the representation and may not preserve expressive intent. For this purpose, Disembodied Performance can lend a freedom of form to expressive performance, not just on the stage, but in film effects. The liberated forms that visual effects for film can achieve can be enhanced, when viewed as output representations, offering new avenues for generating non-humanoid characters and expressive animacy.

Subjective assessments of representations have been and will continue to be made throughout the creative development for *Death and the Powers*. This model is generally sufficient for artistic projects, as it maintains creative control and relies on well-honed design intuitions. Evaluation of such a project is generally left to critical reviews once the design has been formalized or, more commonly in film and television, qualitative assessment by focus groups and other analytical market research methods.

A rigorous study would provide extremely useful results for numerous application domains as well as to contribute to the growing body of understanding of how we perceive affect. It should be noted, however, that most research until recently focused on examining perception of affect using primitives, such as color, or humanoid features, such as faces and voices, as stimuli [13]. Using the Disembodied Performance architecture, multisensory stimuli can be investigated in a non-anthropomorphic or representational form, providing a glimpse at how the brain processes properties of representations rather than whole representations. Such studies can evaluate mappings of affective message to representation through self-reported emotional interpretations of stimuli. The stimuli would be the output representations generated by inference of an emotional state from sensed parameters. It would also be very interesting to use affective sensing on study subjects to determine if a mirrored emotional response is created in the perceiver corresponding to that which generated the stimuli. A positive correlation would validate a representation's capacity as a communicative medium for affective content.

6.4.2 Novel Applications

In *Death and the Powers*, the offstage actor is expected to be performing in the wings or a room near the stage. However, given the distributed nature of the Disembodied Performance System, nothing precludes the performer from being anywhere in the world. This enables a new type of remote or distributed performance for the stage where performers could be giving their performance from any place and have that performance integrated with a stage production. The relationship between performance and presentation is not just many-to-one, but can be one-to-many or many-to-many. A Disembodied Performance can be distributed to many presentation venues, each of them being simply one or more output devices. Disembodied Performances can be simulcast from the theater to the Web or other output media. An interesting twist on having multiple disembodied performers is to allow them to interact in the model space. This would require modifications to the current Disembodied Performance System implementation and likely a model with somewhat different semantics.

Beyond remote and augmented performances, Disembodied Performance can readily bring affective and expressive representations to many

computing domains. For a detailed explanation of affective computing applications, to which the process of Disembodied Performance and the sensing methods developed in this thesis can be applied, I refer the reader to Rosalind Picard's book *Affective Computing* [64]. I believe that Disembodied Performance can be particularly useful in computer-mediated interpersonal communications, where the output representations can be viscerally understood. Informal online correspondence in instant messaging, e-mails, and social networks has demonstrated a need for affective augmentation of textual communication, as evidence by the introduction of modern-day emoticons in 1982 and their widespread adoption and continued use [2]. Disembodied Performance can fill such a niche by providing a simultaneous channel of affective communication to accompany text.

Extending from the augmentation of text communications, Disembodied Performance techniques can contribute to telepresence. Since one of the fundamental goals of Disembodied Performance is to convey a sense of presence, it can supply alternatives to or enhance video communications in several ways. In this case, the system is much as it is used in theatrical productions, only the disembodied participant is simply taking part in an interaction across distance without acting.

Presence in virtual worlds and games can also benefit from the ideas I've laid out. Presently, avatars take several forms, but are generally humanoid representations in appearance and often gesture. In many cases, a user has the opportunity to configure their virtual self as a means of expression. However, the movement of the avatars is generally functional and related to navigating space or occasionally the state of activity, such as walking, standing and waiting, gesturing while talking (typically as the user is typing text to "speak" to another individual in the virtual space). Rarely, if ever, can these avatars convey the affective content of the communicated words or the emotional state of the user. In some cases, a few basic commands are provided to change the behavior or appearance of the avatar, but the result is much like the use of emoticons. Applying Disembodied Performance approaches, the avatar can emote continuously, informed by the affective model. The intermediate representation can be efficiently transmitted in distributed environments. The modes of sensing into the model would be different from those for performing in an opera. In the case of console games, for example, the system can sense properties of how the game controller is being held as input data. At a higher level, how the avatar is being moved in space (erratically, slowly, quickly toward an object or other avatar) can contribute additional information to the model. Affective avatars could greatly enhance the immersive qualities of these virtual environments and interactions. The emotive humanoid form is only one possible output representation. Having expressive dynamic non-

anthropomorphic avatars greatly broadens the palette of representations of an individual's identity.

Disembodied Performance can play an innovative role in asynchronous interpersonal communications and memory-sharing. Technology is increasingly common in greeting cards and keepsakes to play music, record voices or, in the case of digital picture frames, to display changing images. Now, imagine an emotion-capturing keepsake that presented the recipient with a completely personalized and genuine experience of the giver's unique affective signature. A thank you card could capture some expression of just how grateful the sender was when she signed it. A holiday ornament could replay the awe and enthusiasm of your grandson on Christmas morning year after year. The simple model of affective state I have used can be readily stored in such devices and represented as sound and image.

6.4.3 Representation Mapping

An area of subsequent research to which I intend to contribute builds on the process of representation mapping using reified inference. Looking beyond modeling affect for performance contexts, I believe this approach can shed light on methods for translating representations in an informed way, preserving important characteristics while changing the presentation of information for artistic, scientific, and communicative applications. The ultimate goal of continued research would be to extend the concept of reified inference in order to formulate a general theory and practice of representation mapping.

The selection of models for intermediate representations and the mapping to and from input and output representations can benefit from rigorous methodology. Additional modeling methods would be one method of extending the technique. Non-linear and higher-order models can preserve more detail in the intermediate representation, lending the technique to applications where the information that must be reinterpreted and conveyed is more complex than the two- and three-dimensional linear models I've shown in this thesis.

Machine learning approaches and Bayesian methods can be used to compute the model state from input representations. Machine learning was specifically avoided for Disembodied Performance to simplify the extensibility of the method for multiple actors without the ability to gather a sufficient corpus of data for training a model, as well as to promote the artistic process of designing mappings. Typical machine learning approaches, like hidden Markov models that are becoming increasingly common in modeling patterns of gesture as state transitions, excel at discrete classification, whereas placement in a metric space is essential for Disembodied Performance. Human intervention in the mapping process does make use of the reference data acquired, so it is, in some sense, a

training process. However, I feel that the lack of an automatic ability for the system to find key features and generate, or assist in generating, informed mappings is a current drawback of the Disembodied Performance System. In the future, it is conceivable that unsupervised learning may be used and that machine learning may prove essential for the extension of representation mapping to domain-specific modalities and their parameters.

The purpose of representation mapping can be to illustrate things that we normally have difficulty expressing or to observe higher-level structures than what we usually can see; to find the emergent patterns in a system or to simply feel them. In order to produce meaningful and compelling output results in a variety of representation contexts, whether they are problems of information design or more artistic pursuits, as is Disembodied Performance, a greater understanding of the perceptual qualities of the input and output modalities will prove useful. I conjecture that information theoretic analysis of representation modalities and their dimensions can shed light on the optimal or possible encodings of meaning in a medium. How can color or motion be used to express information? What parameters of a multi-tier musical structure can and must vary to create a piece of music that conveys a certain quality of feeling? I believe that these questions and others can be answered in terms by representation mapping in terms of information theory.

One can imagine a design environment for visual, auditory, or symbolic representations that provided a live quantitative assessment of information content and density of a representation based on its parameters as the user is creating it. Coupling this with known information about human perception and response to stimuli can go a long way toward codifying the intuition and practice of information visualization and design, allowing for tools that understand these heuristics and can assist in the creation of meaningful representations that communicate quickly and effectively. Although there has been much work in the area of information presentation, though very significant, the results are generally anecdotal guidelines. For example, the well-known work of Jacques Bertin and Edward R. Tufte provides techniques for conveying information and addressing such notions of parameterization of representation and information density [7,89,88]. However, these approaches must be studied and learned by practitioners. A formal approach would allow for tools to assist designers in applying appropriate, rich, and concise representations made possible by abstracting the intent or salient features of the information to be represented. This would be invaluable in the visualization (or sonification) of large data sets of high-dimensional data, like those commonly found in economics, computational genomics, and bioinformatics. The work of researchers in information visualization, such as Tamara Munzner, has taken steps in this direction by creating new types of controls and applications for representing data, though the choice of

representation and an understanding of what must be discovered from the data is still the responsibility of the researcher using the data or the designer (ex. [40,1]). In many cases, researchers explore these datasets through visualization without knowing exactly what it is they are looking for and struggle without much grounding to find representations that will reveal the hidden regularities in the data. Powerful forms of representation can greatly assist in this process.

6.4.4 Other Applications of Representation Mapping

With the ability to freely map between arbitrary representations, new types of human-computer interactions become possible. I described some general applications of representation mapping, mainly in the field of information design, above. Another general implementation could be a computer user interface that is a proper superset of current interface technologies. Such an interface can present any information in one or more modalities, fully customizable to the user's needs, aesthetic preferences, and accessibility requirements. In this environment, a user can see, feel, or hear information as well as manipulate it in multiple ways. I believe that an interface capable of providing users and developers with this degree of configurability to choose the substance and manner of representation and change it at whim could revolutionize how we use computers.

Let us conclude with a brief look at how representation mapping can influence storytelling applications. In Chapter 3, I mentioned an ancillary project to *Death and the Powers* called Personal Opera. The goal of this project is to create a platform for individuals to tell their own story—to share their legacy—in a fun and intuitive way. Leveraging the power of music as a central narrative thread, Personal Opera will provide a simple interface to assemble images, video, sound, and text in a compelling manner, much like the innovative design of Hyperscore facilitates music composition by anyone using a language of gesture and color to represent high-level musical properties [28]. Personal Opera will leverage similar abstractions of shape and contour to control parameters of the story being told, the arcs of tension and resolution, and how media will be combined. Personal Opera may appear in a number of forms that allow for different types of creation and storytelling experiences. It will have an architecture well suited to a computer application, web application, or even a mobile application where it can draw on a user's personal media collection and additional sources of information. Personal Opera can also appear as a large-scale interactive installation piece.

A somewhat similar project called Atticus is being explored by Plymouth Rock Studios, a MIT Media Laboratory Center for Future Storytelling partner. Atticus is a massive architectural video display that would be the focal point of the film studio's campus and a regional attraction. The studio's concept for the display is that it will showcase a live visual



Figure 81: Personal Opera

This conceptual rendering depicts Personal Opera as large-scale installation based on a periaktos from *Death and the Powers*. Users would be able to create their legacy by manipulating content through expressive gesture.

aggregation of media from around the world. For special events, Atticus will be controlled by a notable personality, such as a film director currently working at the studio. The result is a multimedia performance by that person and it is envisioned that the method of control will be gestural, as the director would be conducting the assembly of the available media, facilitated by the system, for public display. Like Personal Opera, a gestural vocabulary for manipulating multimedia content would be required, along with technologies for sensing and interpreting the performer's gesture. Disembodied Performance and evolved methods of representation mapping would satisfy these needs in the creation of a compelling experience.

As this thesis takes its place in the ever-growing library of work that calls upon the many sciences in the furtherance of the arts, I am hopeful that the ideas and the implementation I have presented will be of use to others. Perhaps, this work will facilitate and inspire others, you, to tell a story that could not have been told before onstage, in other media, and in one's daily life. I look forward to the creations that will result from the research and theoretical foundations I've presented in this thesis.

Surely, there will be more research to come and I eagerly set out to continue my own part. I hope that the process of conceiving Disembodied Performance and implementing the Disembodied Performance System demonstrates the importance and power of a synergistic and multidisciplinary approach to reasoning about our world. With this understanding, we can embrace the tremendous potential for new modes of personal expression and interaction with each other.

BIBLIOGRAPHY

- [1] D. ARCHAMBAULT, T. MUNZNER, and D. AUBER. “TopoLayout: Multi-Level Graph Layout by Topological Features.” *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 2, 2007, pp. 305–317.
- [2] Associated Press. “Digital ‘smiley face’ turns 25: Web’s ubiquitous ‘colon-hyphen-parenthesis’ celebrates milestone.” MSNBC, September 2007.
- [3] ATSC Implementation Subcommittee. “Relative Timing of Sound and Vision for Broadcast Operations.” Advanced Television Systems Committee, Washington, D.C., Findings, IS-191, 2003.
- [4] C. BAUGH. *Theatre, Performance and Technology: The Development of Scenography in the Twentieth Century*. New York: Palgrave Macmillan, 2005.
- [5] BBC News Newsnight: Opera—but not as we know it. [Online]. <http://news.bbc.co.uk/2/hi/7366214.stm>, April 2008.
- [6] R.L. BENEDETTI. *The Actor at Work*. Englewood Cliffs, New Jersey: Prentice-Hall, 1976.
- [7] J. BERTIN. *Semiology of Graphics: Diagrams, Networks, Maps*, trans. W.J. Berg. Madison: University of Wisconsin Press, 1983.
- [8] Boston University. Research Magazine 2008: Inside the Mind of a Madman. [Online]. <http://www.bu.edu/research/spotlight/2008/arts/fineberg/index.shtml>
- [9] W. BOUCSEIN, F. SCHAFER, E.N. SOKOLOV, C. SCHRÖDER, and J.J. FUREDY. “The Color-Vision Approach to Emotional Space: Cortical Evoked Potential Data.” *Integrative Physiological and Behavioral Science*, vol. 36, no. 2, pp. 137–153, February 2001.

- [10] Bowers & Wilkins Society of Sound Fellows: Tod Machover. [Online].
<http://www.bowers-wilkins.com/display.aspx?infid=3517>
- [11] C. BREAZEAL. “Emotion and Sociable Humanoid Robots.” *International Journal of Human-Computer Studies*, no. 59, pp. 119–155, 2003.
- [12] J.M. BURIAN. “Josef Svoboda: Theatre Artist in an Age of Science.” *Educational Theatre Journal*, vol. 22, no. 2, pp. 123–145, May 1970.
- [13] G. CALVERT, C. SPENCE, and B.E. STEIN. *The Handbook of Multisensory Processes*. Cambridge, Massachusetts: MIT Press, 2004.
- [14] M.A. CHANGIZI, Z. QIONG, and S. SHINSUKE. “Bare skin, blood and the evolution of primate colour vision.” *Biology Letters*, vol. 2, no. 2, pp. 217–221, 2006.
- [15] M. CLYNES. “On Music and Healing” in *Music and Medicine*, eds. R. Spintke and R. Droh. Berlin: Springer Verlag, 1987, pp. 13–31.
- [16] M. CLYNES. *Sentics: The Touch of the Emotions*. Garden City, New York: Anchor Press/Doubleday, 1978.
- [17] M. CLYNES. “Time-forms, Nature’s Generators and Communicators of Emotion.” *Proceedings of the IEEE International Workshop on Robot and Human Communication*, Tokyo, 1992.
- [18] L. COSMIDES and J. TOOBY. “Origins of Domain Specificity: The Evolution of Functional Organization” in *Minds, Brains, and Computers: The Foundations of Cognitive Science*, eds. R. Cummins and D.D. Cummins. Malden, Massachusetts: Blackwell, 2000, ch. 35, pp. 523–543.
- [19] R.E. CYTOWIC. *Synaesthesia: A Union of the Senses*, 2nd ed. Cambridge, Massachusetts: MIT Press, 2002.
- [20] O. DA POS and P. GREEN-ARMYTAGE. “Facial Expressions, Colours and Basic Emotions.” *Colour: Design & Creativity*, no. 1, 2007.
- [21] S. DIXON. *Digital Performance: A History of New Media in Theater, Dance, Performance Art, and Installation*. Cambridge: MIT Press, 2007.
- [22] M. DOWNIE. Choreographing the Extended Agent: Performance Graphics for Dance Theater. 2005.
- [23] A. EGGERT. “The Role of the Dramaturg in the Creation of New Works.” *OPERA America*, ed. K. Rourke, pp. 24–27, Spring 2008.

- [24] R. EL KALIOUBY and P. ROBINSON. “Real-Time Inference of Complex Mental States from Facial Expressions and Head Gestures.” *IEEE International Workshop on Real Time Computer Vision for Human Computer Interaction*, 2004.
- [25] English National Opera. ENO Lost Highway: behind the scenes with Philip Bussman. [Online]. <http://blip.tv/file/772947>, January 2008.
- [26] “Entertainment Technology - Architecture for Control Networks.” American National Standards Institute, ANSI E1.17, 2006.
- [27] M. FABIO. “The Chandelier: An Exploration in Robotic Musical Instrument Design.” Massachusetts Institute of Technology, S.M. Thesis, 2007.
- [28] M.M. FARBOOD. “Hyperscore: A New Approach to Interactive Computer-Generated Music.” Massachusetts Institute of Technology, Cambridge, S.M. Thesis, 2001.
- [29] A. GALLAGHER, Y. MATSUOKA, and W.-T. ANG. “An Efficient Real-Time Human Posture Tracking Algorithm Using Low-Cost Inertial and Magnetic Sensors.” *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, 2004, pp. 2967–2972.
- [30] G. GIESEKAM. *Staging the Screen: The Use of Film and Video in Theatre*. New York: Palgrave Macmillan, 2007.
- [31] A. GLINKSY. *Theremin: Ether Music & Espionage*. Champaign, Illinois: University of Illinois Press, 2000.
- [32] E. GOLDENTHAL. *Grendel*, dir. J. Taymor. Los Angeles Opera. 2006.
- [33] M. GUREWITSCH. “Visualizing the Depths of ‘Tristan’.” *The New York Times*, April 2007.
- [34] F. HEIDER and M. SIMMEL. “An Experimental Study of Apparent Behavior.” *The American Journal of Psychology*, vol. 57, no. 2, pp. 243–259, April 1944.
- [35] J. HUNTINGTON. *Control Systems for Live Entertainment*, 3rd ed. Burlington, Massachusetts: Focal Press, 2007.
- [36] J. JAMES et al. “Movement-based Interactive Dance Performance.” *Proceedings of the 14th annual ACM international conference on Multimedia*, Santa Barbara, 2006, pp. 470–480.

- [37] E.N. JESSOP. "The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-Based Gestural Controller for Vocal Performance." *Proceedings of the 9th International Conference on New Interfaces for Musical Expression*, Pittsburgh, 2009.
- [38] M.W. KRUEGER. "VIDEOPLACE: A Report from the ARTIFICIAL REALITY Laboratory." *Leonardo*, vol. 18, no. 3, pp. 145–151, 1985.
- [39] G. LAKOFF and M. JOHNSON. *Metaphors We Live By*. Chicago, Illinois: University of Chicago Press, 1980.
- [40] H. LAM, R.A. RENSINK, and T. MUNZNER. "Effects of 2D Geometric Transformations on Visual Memory." *Proceedings of the Symposium on Applied Perception in Graphics and Visualization*, 2006, pp. 119–126.
- [41] S.K. LANGER. *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, 3rd ed. Cambridge, Massachusetts: Harvard University Press, 1957.
- [42] J. LIEBERMAN. "Slink." *Proceedings of ACM SIGGRAPH International Conference on Computer Graphics and Interactive Techniques*, Boston, 2006, p. 144.
- [43] A. LILLIE. flyingpudding.com: Soundsieve. [Online]. <http://soundsieve.flyingpudding.com/>, 2008.
- [44] J. FINEBERG. *Lolita: An Imagined Opera*, perf. F. Beukelaers, dir. J. Clayburgh. GMEM de Marseille et IRCAM. April 2008.
- [45] O. NEUWIRTH and E. JELINEK. *Lost Highway*, dir. D. Paulus. English National Opera. The Young Vic 2008.
- [46] T. MACHOVER. "Hyperinstruments: A Composer's Approach to the Evolution of Intelligent Musical Instruments" in *Cyberarts: Exploring the Arts and Technology*, ed. L. Jacobson. San Francisco: Miller Freeman, 1992, pp. 67–76.
- [47] T. MACHOVER. "Hyperinstruments: A Progress Report." MIT Media Laboratory, Cambridge, 1992.
- [48] J. MAEDA. *Design by Numbers*. Cambridge, Massachusetts: MIT Press, 1999.
- [49] T.A. MARRIN NAKRA. "Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture." Massachusetts Institute of Technology, Ph.D. Thesis, 2000.

- [50] J.H. MAZO. *Prime Movers: The Makers of Modern Dance in America*, 2nd ed. Hightstown: Princeton Book Company, 1977.
- [51] P. MCALLEER et al. “Perceiving Animacy and Arousal in Transformed Displays of Human Interaction.” *2nd International Symposium on Measurement, Analysis and Modeling Human Functions*, Genova, 2004.
- [52] MIT Media Laboratory. Camera Culture: Second Skin. [Online]. <http://cameraculture.media.mit.edu/secondskin>
- [53] MIT Media Laboratory, Opera of the Future. Death and the Powers. [Online]. <http://powers.media.mit.edu/>, 2008.
- [54] MIT Media Laboratory, Opera of the Future. Death and the Powers: Promotional Brochure. 2008.
- [55] MIT Media Laboratory, Opera of the Future. Death and the Powers: Technical Document. March 5, 2007.
- [56] M. MORI. “The Uncanny Valley.” *Energy*, trans. K.F. MacDorman and T. Minato, vol. 7, no. 4, pp. 33–35, 1970.
- [57] W. MORITZ. “The Dream of Color Music, And Machines That Made it Possible.” *Animation World Magazine*, vol. 2, no. 1, April 1997.
- [58] D.A. NORMAN. *Things that Make Us Smart: Defending Human Attributes in the Age of the Machine*. New York: Basic Books, 1994.
- [59] K. OATLEY, D. KELTNER, and J.M. JENKINS. *Understanding Emotions*, 2nd ed. Hoboken, New Jersey: Wiley-Blackwell, 2006.
- [60] Open ACN Project. [Online]. <http://www.engarts.eclipse.co.uk/acn/openacn.html>
- [61] G. PAPAIN. [Online]. <http://www.gillespapain.com/>
- [62] K. PEACOCK. “Instruments to Perform Color-Music: Two Centuries of Technological Experimentation.” *Leonardo*, vol. 21, no. 4, pp. 397–406, 1988.
- [63] Personal Robots Group. [Online]. <http://robotic.media.mit.edu/projects/robots/operabots/overview/overview.html>
- [64] R.W. PICARD. *Affective Computing*. Cambridge, Massachusetts: MIT Press, 1997.

- [65] S. PLIAM. "The Chandelier: Toward a Digitally Conceived Physical Performance Object." Massachusetts Institute of Technology, S.M. Thesis, 2007.
- [66] V.S. RAMACHANDRAN and E.M. HUBBARD. "Synesthesia—A Window Into Perception, Thought and Language." *Journal of Consciousness Studies*, vol. 8, no. 12, pp. 3–34, 2001.
- [67] R. RASKAR et al. "Prakash: Lighting-Aware Motion Capture Using Photosensing Markers and Multiplexed Illumination." *Proceedings of ACM SIGGRAPH 2007*, vol. 26, New York, 2007.
- [68] Research and Development at the Institute of Animation. [Online]. <http://research.animationsinstitut.de/45.0.html>
- [69] M. REYNOLDS, B. SCHONER, J. RICHARDS, K. DOBSON, and N. GERSHENFELD. "An Immersive, Multi-User, Musical Stage Environment." *Proceedings of ACM SIGGRAPH*, Los Angeles, 2001, pp. 553–560.
- [70] H. RHEINGOLD. *Virtual Reality*.: Summit Books, 1991.
- [71] W. RICHARDS. *Anigrams: Experiments in Collective Consciousness*. Cambridge, Massachusetts, 2007.
- [72] W. RICHARDS. "Modal Inference." Massachusetts Institute of Technology, Cambridge, 2008.
- [73] W. RICHARDS, J. FELDMAN, and A. JEPSON. "From Features to Perceptual Categories." *British Machine Vision Conference*, 1992, pp. 99–108.
- [74] J.A. RUSSELL. "A Circumplex Model of Affect." *Journal of Personality and Social Psychology*, vol. 50, no. 6, pp. 1161–1178, 1980.
- [75] O. SACKS. "A Neurologist's Notebook: The Mind's Eye." *The New Yorker*, pp. 48–59, July 2003.
- [76] A. SCHNEIDER. Experimental Devices for Performance: TwitchSet. [Online]. <http://andrewjs.com/EDP/twitchset.html>
- [77] S. SEYMOUR. *Fashionable Technology: The Intersection of Design, Fashion, Science, and Technology*. New York: Springer, 2008.
- [78] D. SHIFFMAN. mostpixelsever. [Online]. <http://www.mostpixelsever.com>

- [79] T.S. SHIKLER, R. EL KALIOUBY, and P. ROBINSON. "Design Challenges in Multi Modal Inference Systems for Human Computer Interaction." *Proceedings of International Workshop on Universal Access and Assistive Technology*, 2004.
- [80] E. SJUVE. "Gestures, Interfaces and other Secrets of the Stage" in *Transdisciplinary Digital Art: Sound, Vision and the New Screen*, eds. R. Adams, S. Gibson, and S.M. Arisona. Berlin, Germany: Springer-Verlag, 2008, vol. 7, pp. 301–312.
- [81] E.N. SOKOLOV and W. BOUCSEIN. "A Psychophysiological Model of Emotion Space." *Integrative Physiological and Behavioral Science*, vol. 35, no. 2, pp. 81–120, April 2000.
- [82] F. SPARACINO, C. WREN, G. DAVENPORT, and A. PENTLAND. "Augmented Performance in Dance and Theater." *International Dance and Technology*. Arizona State University, Phoenix, February 1999.
- [83] L. SPIEGEL. "Graphical Groove: Memorium for a Visual Music System." *Organised Sound*, vol. 3, no. 3, pp. 187–191, 1998.
- [84] A. STEGGELL. "Mind, the Gap: Synesthesia and Contemporary Live Art Practice." Østfold University College, Norwegian Theatre Academy, Fredrikstad, 2006.
- [85] J.A. STRICKON. "Smoke and Mirrors to Modern Computers: Rethinking the Design and Implementation of Interactive, Location-Based Entertainment Experiences." Massachusetts Institute of Technology, Cambridge, Ph.D. Thesis, 2003.
- [86] J. SVOBODA. *The Secret of Theatrical Space: The Memoirs of Josef Svoboda*, ed. J.M. Burian, trans. J.M. Burian. New York, New York, United States of America: Applause Theatre Books, 1993.
- [87] F. THOMAS and O. JOHNSON. *The Illusion of Life: Disney Animation*. New York: Hyperion, 1995 (1981).
- [88] E.R. TUFTE. *Envisioning Information*. Cheshire, Connecticut: Graphics Press, 1990.
- [89] E.R. TUFTE. *The Visual Display of Quantitative Information*, 2nd ed. Cheshire, Connecticut: Graphics Press, 2001.
- [90] J.-M. RAVEL. *Visualizing Music: Visuals for Listening*, perf. A.L.P.T. Ying Quartet, dir. T. Machover. University of Iowa. 29 March 2008.

- [91] J.W. VON GOETHE. *Theory of Colours*, trans. C.L. Eastlake. Mineola, New York: Dover, 1810.
- [92] D. WAKIN. “Techno-Alchemy at the Opera: Robert Lepage Brings his *Faust* to the Met.” *New York Times*, p. C1, November 2008.
- [93] S. WILKINSON. “Phantom of the Brain Opera.” *Electronic Musician*, vol. 13, no. 1, January 1997.
- [94] A.P. WITKIN and J.M. TENENBAUM. “What is Perceptual Organization For?”, 1985.
- [95] J. ZILCZER. “Color Music: Synesthesia and Nineteenth-Century Sources for Abstract Art.” *Artibus et Historiae*, vol. 8, no. 16, pp. 101–126, 1987.