IOWA STATE UNIVERSITY
Digital Repository

2006

Synesthetic music experience communicator

Lewis Charles Hill II
Iowa State University

Recommended Citation

Hill, Lewis Charles II, "Synesthetic music experience communicator " (2006). *Retrospective Theses and Dissertations*. 3077.
https://lib.dr.iastate.edu/rtd/3077

www.manaraa.com

Synesthetic music experience communicator

by

Lewis Charles Hill II

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Human Computer Interaction

Program of Study Committee:
James Oliver, Major Professor
Kris Bryden
Barry Larkin
Chiu-Shui Chan
Constance Hargrave
Gurpur Prabhu

Iowa State University

Ames, Iowa

2006

UMI Number: 3243559

Copyright 2006 by
Hill, Lewis Charles, II

All rights reserved.

TABLE OF CONTENTS

ABSTRACT

The Synesthetic Music Experience Communicator combines research in the areas of human computer interaction, music technology, and human perception to illustrate the experience of synesthesian mental imagery in response to musical sounds. Synesthesian musicians have reported positive benefits from their augmented awareness of sound in the areas of pitch identification, memorization, composition, and improvisation.

This dissertation attempts to communicate both the experience and performance benefits associated with this cognitive phenomenon. Several virtual worlds have been developed to explore group education, rehearsal, and the artistic transformation of live performances into informative and entertaining visual presentations. Initial inspirations, background research, development stages, iterations, user evaluation, and future directions are discussed.

Two virtual reality prototype systems are overviewed. The first demonstrates chromesthetic translations of real-time Musical Instrument Digital Interface (MIDI) events. The second demonstrates real-time transformations of multi-band Fast Fourier Transform processed audio into visual scenery. Observations and feedback about the initial prototype systems are summarized.

Three systems are proposed which expand the initial prototypes and demonstrate concepts for computer augmented ensemble rehearsals. The Synesthetic Visualizer modules combine real-time audio and MIDI data to demonstrate synesthesia and photism-like imagery. The Computer Augmented Percussion Trainer and Small Ensemble Trainer are proposed to augment musical practices and rehearsals by providing real-time displays of ensemble members' performances and displaying transformations between instrument fingerings.

An introductory user study is conducted to determine which aspects of music are best communicated by the visual displays and to evaluate the potential benefits of this synesthetic approach. The user study asks participants to consider this research in relationship to existing music visualization and education methods. Exhibitions and publication efforts are reviewed. The user study, observations, and exhibitions serve to

validate the core hypothesis of this research. The dissertation concludes by proposing an intelligent interactive synesthetic software agent to facilitate profile driven multimedia content creation.

CHAPTER 1. INTRODUCTION

This dissertation examines an intersection of synesthesia, music, and human computer interaction (HCI). Synesthesia, a cross modal sensory phenomenon, is often experienced as an internal translation of sensory stimuli into mental imagery. Musical synesthesia, some times referred to as colored hearing, involves the real-time transformation of sound into a combination of colors and texture information. This research attempts to demonstrate how synesthesia can provide musicians with an internal visual connection to a performance. This connection can augment the musicians' theoretical knowledge, training, and intuition to give them increased awareness of their performance and of other musicians in the ensemble.

This research presents several methods for algorithmically converting musical sound into meaningful and informative computer graphics imagery, thereby mimicking the experience of synesthesia. This dissertation also proposes systems that translate the mental and perceptive processes of synesthesia into visual displays in order to facilitate group education, musical performance, and ensemble rehearsal. This work will be accomplished through the integration of music principles, technology systems, synesthesia, and human perception.

This work is motivated by personal experience with synesthesia and musical performance. These experiences and observations have revealed potential applications. Technology approaches and specific hypotheses have been formulated to explore these concepts. The research process will be documented through several stages including:

- Discussion of initial concepts for synesthesian-like applications of music technology
- Presentation of general hypotheses for the study
- Survey of relevant principles and research avenues in the areas of human computer interaction, music technology, synesthesia, and perception
- Development of prototype systems and discussion of preliminary observations

- Presentation of three proposed systems including testable hypotheses and evaluation methods
- Discussion of experimental data gathered from a study of the implemented systems
- Discussion of future directions for music technology integration in live performance settings and multimedia software.

This chapter discusses the author's personal observations as a synesthete in music performance settings. The research areas of HCI, music, synesthesia, human perception, and computer visualization are introduced. A general approach to a developing a Synesthetic Musical Experience Communicator is presented. This approach advocates the use of music technology to capture the nuances of a music performance and to convey a multi-sensory approach to musical expression. This chapter concludes by identifying research goals and providing an overview for the remainder of the dissertation.

1.1 Observations in Music Performance

The following are descriptions of three scenarios where music technology application ideas have arisen. These descriptions are given in a narrative form to creatively illustrate the thought process that lead to each discovery. Each scenario describes the role of the musician, gives some sample thoughts about the performances, and presents ideas for technology applications.

Scenario One: *Thursday evening at a gospel choir rehearsal. The director plays piano and conducts the musicians and singers. The drummer, bass player, guitar, steel drum, and choir prepare to start the next song.*

~~ "One… Two… Three… Four…" The musicians begin to play. The choir begins to sing. The drums are hitting nice and the bass sets a solid foundation. The piano fills in and around the choir harmonies. Everything is moving smoothly, though it feels a bit too busy. There seem to be a few stray notes. Where is that dissonance coming from?

The electric bass and the bass drum are slightly out of time on the chorus. How can they meld together a bit more smoothly? Perhaps the guitar could emphasize the downbeat a little bit more. The musicians should emphasize the chorus and hold back on the verse. This is an improvisational setting, but we sure don't want chaos. The vocal soloist could adlib a bit more at the end of the verse to help the transition along.

The choir has a good tape recording that we can review. Perhaps we should step through the chords to find that dissonance. In a great rehearsal this ensemble is like a finely tuned machine. Can I explain these musical relationships with pictures? How can I freeze time and display each individual part as it was played? How can I show the bassist more options for his bass lines? And where did all my stray piano notes end up? ~~

Scenario Two: *Laptop Techno event at a San Diego coffee house featuring the Unidentified Wailing Mammals. Two laptop computers are connected to an audio mixer and a public address system. Two biotech digital music daydreamers take to the 8-inch tall corner stage.*

Figure 1.1 Unidentified Wailing Mammals recording setup including PC Laptop, Mixer, Apple iBook, MIDI Keyboard, Roland Sound module, and speaker setup.

~~ Tick, tick, tick… A series of bleeps and blurps emit from our speakers. The audience watches intently as our two computers generate eclectic sounds. Koan generated musical sequences [SSEYO] merge with human sequenced Reason percussion loops [Propellerhead]. Multi-layered musical sounds weave in and out of space: effected and panned by some quasi-constrained random number engine. Initially the audience is

intrigued by the music. Thirty seconds later… they're off in space staring blankly at the back of two laptop computers.

In a synesthesian mind, this generative music is an unpredictable visual movement. It is like an animated mural or a painting on a tunnel moving through time. Each new sound the computer selects is like a character in a play. Each sonic character wears a unique costume and has a distinct motion. New sounds enter this audio/visual soundscape. The mental images elevate and descend as pitches rise and fall. The notes move about the stage as the sounds pan through the three dimensional space.

Then the thought occurs, could a computer program be written to show the audience what the music looks like as a synesthete experiences it? A graphical display might help engage the audience despite the oddness of these compositions. These generative audio software algorithms operate on music principles to compute 'paths' through user-configured music spaces. The computers generate sounds that are based on the user's preferences. However, the computer cannot fully appreciate the calculations that it has performed or the resulting music. The audience never sees the mathematics. Perhaps a visual display could show the audience the musical and mathematical relationships.

How can I represent the sounds, notes, and progressions of chords? How can I animate the sounds visually? Can I illustrate this music as pictures in motion? Can I represent the individual textures of the sounds so the audience can better hear the contrasting musical parts? ~~

Scenario Three: *Earth Day at the Balboa Park World Beat Center in San Diego. A live hip-hop spoken word band takes to the stage. The drummer and his kit are set between the bass player and the fender Rhodes electric piano. Three emcees stand in the front row armed with microphones and metaphors. A graphics screen sits behind them upon which live video footage of the bands will be projected in the nighttime.*

Figure 1.2 Urban Root at Tecate, CA Spring Festival '02
Figure 1.3 Able Minded Poets at the Cross Border Festival for Peace '02

~~ Deep dropping bass and a heavy funk beat are sprinkled with some ocean-like ambient Rhodes piano. The poet speaks forcefully to convey his inspirations. "Downtown, metal and brick buildings reach endlessly towards the sky". "Nature becomes strategically placed like propaganda." Twelve measures pass and the verse is done. The drummer twists the beat up for a half-measure drum fill. Six syncopated hits for the full ensemble and a voice cries out 'To the change…'

The rhythm section establishes a new groove. This transformation is a twist on the first verse with a lighter backbeat and different chords. There is less bass, more Rhodes piano, and the drums are lightly balanced in-between. The band wrote this section to be vibrant; like a walk along Ocean Beach pier in the daytime. The poet speaks, "We ride these melodies like seagulls coasting on the Sunset Cliff's breeze."

Each groove is like a new neighborhood or a new state of mind. Each song is another transformation of sound, grown from a seed of inspiration into a tasty new fruit. Every sound is another movement of musically expressive spirits seeking to inspire the audience towards a more united world. The musicians transform feeling, through thought, to action, onward into movement, to sound through the air. The audience receives sensation, recognized as sound, transformed through thought into emotional movement, and mental imagery.

Looking out to the audience I see children dancing circles around their parents while artists draw their feelings on community canvases. I feel their energy and

creativity. They clap their hands to show their approval and shout along in a chorus of "freedom". The band plays on. The audience grooves along. The musicians wonder if the audience can feel the emotions that drive the performance? Can the musicians somehow show the audience the moments and observations that inspired these songs? Can the musicians control the visuals while playing their musical instruments at the same time?

As a synesthete sitting inside the middle of the music, experiencing, coordinating and creating, I wonder. What would that fire-charged guitar-solo look like if it were painted into the sky? Or the moving sea blue and midnight purple waves of the Rhodes piano? What of the thunderous drums and warm embracing bass guitar? I ponder all of this while striking the shimmering sizzling cymbals. ~~

Discussion of Music Performance Scenarios

The preceding scenarios differ in setting, instrumentation, and style. The interaction between the musicians, music technology, and with the audience also varies. Several opportunities were identified for visual communication of sounds and music information. This section reviews each scenario and each opportunity is extracted.

As director and a performer in a church-based mixed vocal/musical ensemble, there is potential to use the technology to illustrate style and accuracy issues in rehearsal. Improvisation and spiritually inspired creativity are a part of the gospel tradition. However, the performers must work together stylistically. There is a desire to present the interrelations between the melodic and rhythmic parts being played and sung by the ensemble. It is hoped that the display would facilitate discussion of ideas between ensemble members. Such a system could also help in teaching traditional and unwritten songs to the whole group.

In the laptop techno and generative music scenario, the audience is challenged to comprehend what seems to resemble an alien musical expression. The computer generated sonic landscapes are very different from classical or folk compositions. Generative music presents a unique combination of sounds that follow some common music practices, but break many conventions. While the performers intend the music to

be unique, they also realize that the audience needs a bit of help in connecting with the melodies, rhythms, and synthesized sounds. Also, the performers' constrained laptop interactions do not make a very engaging visual performance. A visual display may resolve both the attention and entertainment issues and give the audience greater insight into the musical mathematics being calculated by the machine.

In the live hip-hop band scenario, the show is already super charged with energy and funky syncopated rhythms. The audience is moving and grooving, engulfed from head to toe in waves of sound. However, the band wants to show the audience some things that they've never seen before. There is a desire to illustrate the emcee's verbal metaphors. The musicians want to show the audience imagery of their music. Further, the charged musical solos inject an emotional and improvisational feel into the music. Here the desire is for a dynamic and artistic presentation that arises directly from the music, while revealing the vocalists' inspirations and reinforcing their metaphors.

In reviewing these three scenarios, a number of potential uses for a synesthetic illustration of music are given. To summarize, a visual representation of sounds and musical information is predicted to be useful in:

- **Presenting** sounds, pitches, and timings in visual displays to facilitate awareness within the group of performers
- **Demonstrating** musical concepts to the audience
- **Entertaining** the audience through visual depictions of verbal metaphors and musical sounds
- **Enabling** musician controlled artistic displays of performance
- **Inspiring** new thinking about musically driven multimedia presentations.

Given these observations and the state of modern computing resources, it seems that a suitable technology system can be assembled to accommodate some of these needs. Developing such a system requires exploration of the related fields of knowledge and relevant resources.

1.2 Introduction of Research Areas

This project is founded upon three major research areas: Human Computer Interaction (HCI), Music, and Synesthesia. Each area is introduced here to establish a basic context and to facilitate discussion of the concept and validation plan for a Synesthetic Musical Experience Communicator. Computer visualization, a field of study related to HCI is also briefly introduced here as the communications medium.

Human Computer Interaction

Human Computer Interaction is an interdisciplinary field of study that has arisen in part to study the use of technology systems. HCI research considers humans operating computer systems in diverse settings. HCI practitioners work to improve or augment current interaction methods. Iowa State University's HCI initiative is focused on five major research areas: information sensorization, mobile/ubiquitous interfaces, intelligent agents, accessibility for non-technical collaborators, and enabling infrastructure [HCI].

HCI is the focus of the second chapter. This document overviews interaction paradigms and design principles for interactive computer systems. HCI design methods emphasize the importance of considering operators' knowledge, cognitive processes, and goals when developing computer system interaction. This dissertation gives examples of how the frameworks can be used to consider the various roles of the musical performers, composers, and the audience.

Interactive computing systems are introduced in a variety of sizes ranging from head-mounted portable computers to large-scale immersive virtual reality auditoriums. Each system offers unique options for user interaction and experiences. The systems' graphic and sonic capabilities enable visualization and musical presentation for single member through large group sized audiences.

HCI research is presented in the key area of data representation. Sonification, the representation of graphical and spatial data as audio information, is described. Multimodal interfaces, those in which interaction is accomplished through multiple

sensory channels, are also presented. Research in human perception is identified to illustrate the unique considerations associated with these types of interactions. Also, forward-looking perspectives on virtual experience sharing are described.

Music

'Music is the art of producing pleasing, expressive, or intelligible combinations of tones [Olson].' It involves the study of sound and its various aspects including time, pitch, and energy. Music study further organizes sound into interrelated structures of melody, harmony, and rhythm [Olson][Helmholtz].

Music can be performed vocally or instrumentally, generated with analog equipment, calculated digitally, and can emanate from natural events. Musical instruments are constructed from diverse materials including wood, metals, strings, membranes, and so on. Each instrument generates unique frequencies and overtones when played. The blending of these overtones and various pitches contribute to the experience of musical sounds. The environment surrounding the musical performance impacts the perception of those tones.

There are many established methods of musical performance, composition, and education. A philosophy and several goals of music education, several key musical concepts, and teaching methods are discussed in the third chapter. Some of the teaching methods leverage human motion and sensory awareness to reinforce learning. Other methods encourage students to develop a deep appreciation for the qualities of the sounds and to explore their creative options while performing.

Technology plays an important role in musical training, interdisciplinary research, performance, and recording. Chapter Three considers:

- Established and emerging applications of music technology
- Electronic music systems for digitally encoding and transmitting music performances
- Technology research and software algorithms for analyzing compositions and extracting musical events from real-time audio

- Algorithmically generated sound
- Technology and software tools that enable groups to experience and contribute to musical performances
- Multimedia technology systems which feature combinations of audio, visual imagery, and musical concepts.

Criteria are presented for classifying and evaluating multi-modal interactive music experiences. Potential for expansions of current methods towards more synesthetic representations are identified. Guidelines for developing effective multi-art applications are also identified. The diverse teaching methods and diverse applications of technology provide useful insights in developing this system.

Synesthesia

Synesthesia is a cognitive phenomenon in which stimulation along one sensory modality causes a perceptual experience to occur in a different sensory modality. Common examples of Synesthesia are color association with numbers, letters, smells, tastes, or flavors. These color associations appear as 'photisms': disembodied, colored, and textured mental images. Some photisms appear to be 'projected' into the real world. Others are mapped to the location of the physical stimuli in the body. True synesthesia is not a learned condition. It is not random, not forced, nor suppressible. It is an involuntary and repeatable sensation that is triggered by specific stimuli [Cytowic][Dann].

Synesthetic musicians and artists report strong connections between physical stimuli and color impressions. Some musicians have reported mental imagery or gustatory response to audio stimulus [Beeli]. These perceptual experiences have been used to aid in music memorization, to identify musical intervals, to serve as creative inspiration, and to aid in performance. Synesthetic impressions from audio can include visual patterns: grids, geometric lines, movement, rotation, and spiraling. Musical impressions can include the perception of dynamic color and texture matched to pitches, intervals, and sound types. Sound characteristics including, panning, reverberation, tone, and envelopes can also impact the synesthetic perceptions [Cytowic].

Studies have been conducted to study the origin and neural basis for Synesthesia. Functional Magnetic Resonance Imaging (f-MRI) studies of Synesthetes have revealed unique neural functioning. There have also been studies of learning affects, memorization techniques, and other cognitive strategies employed in relationship to synesthesia.

As indicated previously, synesthesian musicians and artists have indicated that this phenomenon plays an important role in their artistic practice. The cross-sensory perceptions provide composers with an augmented awareness of the balance and contrasts of the elements of music. Synesthetes report greater awareness of their musical performances. Further, these individuals fine-tune their sound through their augmented perception. There have been a number of explorations designed to communicate the experience of synesthesia to audiences in the past. Synesthetic pairings have also been used to derive compositions from nature and to create group experiences through combinations of musical instruments and visual displays [Harrison][Duffy].

Synesthesia is given deeper review in Chapter Four. Included are examples of photisms, of color-number synesthesia, and multiple theories about the causes and purpose of synesthesia. Some of the technology systems used in past music and art visualization are described. Also, correlated research into sound symbolism and visual meaning is presented to illustrate analogous experiences in non-synesthetes.

Computer Visualization

Computer Graphics (CG) systems facilitate the presentation of visual information to users. CG systems feature interactive devices that transform user's movements into actions inside of the computer system. CG systems provide developers with software interfaces that enable development of dynamic interactive graphical content. This content can include scientific simulations, engineering models, artistic renderings, design related information, text, graphs, sound, tactile information, and so forth. Computer graphics systems can present realistic and dynamic two- and three- dimensional imagery. These graphics can be colored, lit, shaded, textured, and animated [OpenGL]. CG imagery can

be presented on a variety of displays including palmtop, head mounted display, augmented reality, holographic, immersive, and non-immersive virtual reality systems.

Virtual Reality (VR) systems are designed to envelop users into a compelling experience. To present an alternate space that envelops the user. VR systems provide participants with the ability to step into visual simulations. Users are presented with stereo imagery that appears to 'jump out' of the screen. Many VR systems feature user-tracking systems that provide the computer with information about the user's location and orientation. The tracking systems also enable users to interact through physical movement and manipulation of a variety of custom interfaces.

Virtual Reality systems vary in size and the amount of interactivity. Iowa State University's Virtual Reality Applications Center has a variety of immersive and non-immersive VR systems [VRAC]. These systems and a brief description of their suitability for music technology settings are discussed in Chapter Two.

This dissertation works to augment VR systems through the integration of electronic musical instruments and audio signal processing. These systems have been demonstrated to be compelling and engaging for participants. As such, they are ideal formats for music visualization presentations. Chapter Five discusses the use of Virtual Reality systems in the development of two musical technology prototypes and some preliminary observations from those presentations. Initial observations of those prototypes are forwarded into the proposed Synesthetic Musical Experience Communicator.

Review of Research Areas

This section has briefly introduced the foundational areas of this dissertation: Human Computer Interaction, Music, Synesthesia, and Computer Visualization. Human Computer Interaction research facilitates the study of technology in operation and considers future applications of those systems. Music creation involves generating sound through vocal, instrumental, and electronic devices. Music is a very mature field of study with several established teaching methods. Synesthesia is a phenomenon of human

perception in which some individuals experience music in a visual manner. Previous research has been conducted to convey the synesthetic experience, however, this work seeks to augment those methods. Computer visualization provides a means to simulate and interact with a variety of multi-sensory simulations.

Each foundational area is explored more extensively in the subsequent chapters. HCI concepts and technologies are presented in Chapter Two. Principles of Music Education and associated technologies are discussed in Chapter Three. Chapter Four focuses on Synesthesia and its relationship to music and the arts. Chapter Five discusses the use of Virtual Reality systems in the development of prototype music interactions.

1.3 An Expanded Approach

This dissertation presents methods for further integration of music technology into entertainment, education, rehearsal, and performance settings. This dissertation combines principles in the four foundational areas, HCI, music, synesthesia, and computer visualization, to develop informative and creative transformations of music performances and to reinforce the composers vision of a piece of music.

Music technology research provides a means of monitoring and analyzing the musical performance of an ensemble. HCI enables the systematic study of operators working on performance tasks and the effective representation of real-time information. HCI and music technology research do intersect in modern software design. Both areas seek to create intelligent help systems to facilitate learning as well as monitoring users/performers at tasks to improve their skills. Synesthesia provides a glimpse into the cognitive processes of certain musical and artistic individuals. Computer visualization enables the transformation of musical concepts and simulation of synesthetic mental processes as visual imagery.

The core hypothesis of this approach is that combining real time musical performance data, audio signal processing, and musical concepts with representations of the synesthetic cognitive experience in a computer generated visual display can increase musical awareness for composers, performers, and the audiences. This research proposes

that an interactive system built upon these principles can provide individuals with a more intuitive view of the musical concepts applied in composition, performance, and improvisation. Also, that these interactive systems can illustrate nuances of the performance in informative and artistic ways to stimulate creativity.

The intent of the Synesthetic Musical Experience Communicator is to:

- **Inform and educate** the audience about the characteristics of sounds and the complex interaction of musical performances.
- **Improve individual and group performance** of musical tasks by increasing their awareness of their performance in relationship to the ensemble
- **Enable** performers to **directly control** artistic visuals from their musical instrument to augment a live musical performance.
- **Demonstrate** the possibilities of **simulating musical synesthesia** to illustrate cognitive processes and the subtleties associated with such awareness.

1.4 Research Goals and Overview

This dissertation seeks to:

- Document application ideas which intersect the domains of music, synesthesia, human computer interaction, and computer visualization
- Identify previous research efforts in the areas listed above
- Investigate methods for mapping music performances into visual spaces that demonstrate musical concepts and simulate synesthetic experiences
- Develop and study music technology applications to facilitate group education of basic musical concepts, to enhance artistic performances, and to facilitate ensemble rehearsal.
- Increase awareness of Synesthesia and encourage future research that combines art, engineering, mathematics, and human perception.
- Provide a foundation for future research in music technology.

This dissertation discusses the research conducted in the study and attempts to validate the core hypothesis of this work. **Chapter Two** presents background information on human computer interaction and computer visualization systems. **Chapter Three** reviews music education, concepts, technology, and research. **Chapter Four** reviews Synesthesia and its applications in audio-visual arts. **Chapter Five** discusses two prototype systems developed as part of this dissertation. The motivations, observations, and potential future directions for each prototype are described. Initial observations and future work plans for the prototypes are described.

Chapter Six presents a research plan for the completion of this dissertation. Included is a description of three proposed portions of a Synesthetic Musical Experience Communicator: a Computer Augmented Percussion Trainer, a Small Ensemble Trainer, and a Synesthetic Visualizer

The **Computer Augmented Percussion Trainer** is presented as a tool for improving performance skill in a percussion ensemble by translating each member's notes into a scrolling visual display. The **Small Ensemble Trainer** is a tool for visualizing the notes and rhythms played by a musical group. The Small Ensemble Trainer is intended to facilitate communication between musicians and to display the interrelations between individuals' performances. The **Synesthetic Visualizer** will represent musical performances and real-time audio information into an informative multi-sensory experience. The Synesthetic Visualizer graphics will modulate colors, textures, and placement of simulated musical photisms. **Chapter Six** also proposes experimentation methods for evaluating the system.

Chapter Seven documents the implementation of the Synesthetic Visualizer and Small Ensemble Tools. **Chapter Eight** presents a brief user study and evaluation of the entire system. Finally, **Chapter Nine** presents a conclusion to this research, reviewing the literature study, prototypes, implementation, and experimentation in relationship to the core hypothesis, intent, and goals of this project.

CHAPTER 2. HUMAN COMPUTER INTERACTION

Human Computer Interaction (HCI) is a diverse field of study, carefully crafted in recent years to meet the needs of a growing group of interdisciplinary software systems researchers. There are several research areas in which HCI knowledge can be applied. HCI researchers participate in the conceptual development, prototyping, and user testing of computer systems. The intent of such studies is to measure and improve the means by which users complete their software tasks. One method of improvement is to match users expectations and experiences to the interface task, thereby creating a more intuitive means for user interaction.

An HCI practitioner monitoring an operator considers the sequence of operations that are performed, observes the user's attention patterns in performing the tasks, and studies the outcome of those steps. HCI practitioners consider methods for improving user interaction at those tasks. They work to augment current approaches and/or paradigms to increase the use of technology in an area. They consider the impact of technology upon human life and society.

HCI academic programs provide students with design principles and guidelines for effective design. Students are given information about user perception and cognition. Students are also given useful information about the types and variety of computer systems and interfaces. HCI also provides experimental methods for conducting user studies.

Students are encouraged to consider the application of technology systems to improve human performance on a variety of tasks. This dissertation considers applications in the area of music, specifically, communicating musical concepts and performances. Musical data and audio will be processed and converted into animated graphical imagery to simulate Synesthesia, a subject of study in cognitive psychology.

This chapter presents key information relevant to the development of computer interfaces and multi-sensory interactions. Two sets of design principles for developing software interfaces are presented. A basic system/user interaction framework is discussed

in the context of music technology. Interactive computing systems with special potential for large group presentations are described. This chapter concludes with a brief discussion of sonification and multi-modal software systems, followed by an expert's future vision for computer generated art and virtual experience sharing.

2.1 Design Principles for Interactive Systems

Computer operators have expectations and biases when working with software systems. Some expectations stem from users' experiences with computing systems. Other expectations are based on the discipline and research domain of the software application. Designers of interactive systems can leverage these expectations to design meaningful interactions for their participants. Many design principles have been determined through experimental observations of human behavior. These principles are combined into frameworks and multi-stage design processes to assist an HCI practitioner in meeting the needs of their users.

Music technology applications are designed to support study, performance, and composition. The field of music includes many well-defined concepts, methods, and techniques. Computer systems for music education leverage knowledge in the musical domain. Musical notation, instrument layout, and music fundamentals are examples. Similarly, a simulated piano keyboard layout may be used to input musical notes. Simulated analog keyboard sliders are used to adjust parameters. Simulated music notation tools are used to write compositions onto a simulated musical staff. These bits of domain knowledge provide users with software interfaces that are representative of their normal musical practice.

HCI design principles provide the means for developing good mappings between musical practice and software interfaces. Two sets of common design principles for interface and software design are presented in this section. The first set, developed by Donald Norman, considers the thought processes involved in interaction [Dix][Norman].

Norman's Execution-Evaluation model of interaction acknowledges two entities: the user and the system. This model considers the steps involved in interaction between

the user and the system. Initially, the user formulates a task that he/she wishes to execute and performs the action using the system. The user then evaluates the outcome of the action as presented by the system. Norman further divides the interaction model into seven stages:

- Establishing the goal
- Forming the intention
- Specifying the action sequence
- Executing the action
- Perceiving the system state
- Interpreting the system state
- Evaluating the system state with respect to overall goals and intentions.

Norman's design stages can be augmented by modeling the behavior of the system alongside of the user [Norman][Dix]. An expanded model considers the behavior of the system as well as the interface. Abowd and Beale [Abowd] extend Norman's model to incorporate any communications that the system sends to the user through the interface. Here the interactive system is divided into four major components: User, System, Input, and Output.
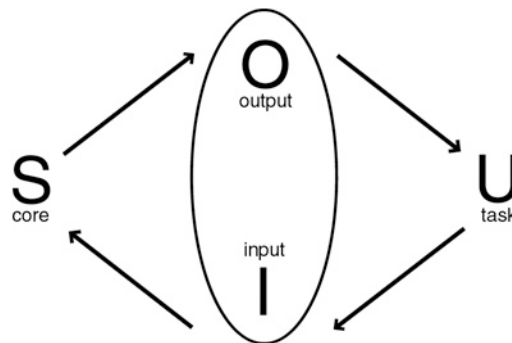
Figure 2.1 Interaction Framework depicting System, User, Input, and Output

Figure 2.1, adapted from [Abowd] depicts the four major conceptual components of an interactive system. The user (U) is the computer operator seeking to complete an interaction *task*. The system (S) is the computing device and contains a representation of the *core* status of the digital simulation. The middle two portions *output* (O) and *input* (I) represent the system peripherals that enable communication between the user (U) and the

computer system (S). Input peripherals are used to send data into the system. Output devices provide the user with knowledge of the system state. Example input peripherals are the mouse and keyboard. Example output peripherals include computer monitors, projectors, and speakers.

Interaction operations proceed in a cyclical fashion. The Input and Output peripherals and devices serve as conduits between the System and the User. Abowd and Beale's model features four main operations: presentation, observation, articulation, and performance. Figure 2.2, adapted from [Abowd], indicates these relationships. *Presentation* is performed by the system, through the output channel, to indicate system status. The user *observes* the system output and formulates tasks. The user performs actions by *articulating* them to the system through the input channel. The system then *performs* the task that has been articulated through the input channel [Dix98]. The input/output channel plays an important role in the interaction model discussed earlier.
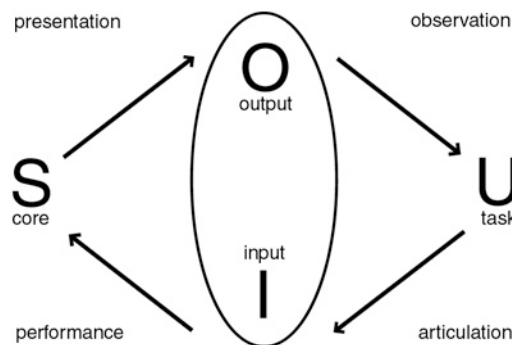
Figure 2.2 Interaction Framework and Translations between Components

Abowd and Beale's framework is useful in analyzing user's patterns when interacting with systems. Fig. 2.2 above considers the case of a single user and a single computer system. It does not give specifics about the system input or output, however evaluating the user's operations in each step can reveal limitations in the interface designs.

When considering musical interaction, the input and output options are expanded to include other devices. Input systems can include mouse, alphanumeric keyboards, piano keyboards, digital musical instruments, music performance encoding sensors,

pedals, knobs, sliders, and microphones. Output systems include computer monitors, projectors, speakers, lighting systems, displays integrated into digital controllers and so on. Musical technology system data includes imagery, audio data, digital midi events, and instructions for digital controls.

In a group music technology setting, users differ in the amount and type of interaction. This discussion considers four types of users: operators, performers, audiences, and composers. Operators interact with the software applications and computer peripherals to control settings. Performers provide their inputs through musical input devices and or sensors. The audience observes the performance through the system output displays. They hear it through the speakers and observe the musicians performances. The audience can also interact through their own vocalizations and movements. A composer can consider the full potential for communicating both an audio and visual artistic experience to the audience.

The four tasks of presentation, observation, performance and articulation can happen through multiple sensory modes. Presentation and observation happen though the audio system and visual displays. Indicators built into electronic musical devices also serve to inform the performers. Audio output from the computer and musical instruments can provide the operators, performers, and observers with information about the musical performance. This information can include multiple channel audio information. Articulation can be performed through standard devices, including mouse or virtual wand. Performers may also use their musical instrument to articulate commands to the system by playing specific combinations of notes or changing control surfaces. Finally, the system can perform a variety of state changes based upon the interface controls, the musical performance, and any additional sensors placed on the performers.

Design Guidelines

General guidelines have been assembled for developing intuitive interactions. Two sets, one developed by Norman and another by Shneiderman, identify some common principles for designers to consider. These guidelines are useful in evaluating

the presentation, observation, articulation, and performance operations in music technology applications. Norman's guidelines are paraphrased below [Norman], followed by examples in the area of music technology. The guidelines are:

- Use existing knowledge
- Simplify tasks
- Increase visibility
- Present correct mappings
- Exploit constraints
- Designing for error
- Standardize controls and operations

These guidelines can be considered in relationship to musical systems. For example, musicians have knowledge about their instruments, musical concepts, and about musical hardware. Software simulations should leverage that information to create displays that follow similar principles, or present the options in a conceptually analogous manner. *Simplifying tasks* means providing access to functions in as few steps as is necessary. *Increasing visibility* means that operations should be readily accessible to operators, or organized such that they can be easily located. Configuring any group oriented visual music applications should be as simple as possible to allow the musician to focus on their performance and normal operation of their instrument.

Presenting correct mappings means that modifying the user interface elements causes intuitive and meaningful changes in the system. For example, twisting a knob clockwise generally signifies an increase in a value. A reverse mapping would be counterintuitive and could lead to incorrect usage. *Exploiting constraints* means that interface elements should have clearly labeled maximum and minimal values. The interface should prevent a user from selecting or modifying a parameter into an invalid value state.

Designing for error will enable users to recognize and fix mistakes. As an example, changing a note or parameter on a display should be easily 'undo-able.' The system could also ask users to confirm any operation that could result in a severe error or result in data loss.

Finally, *standardizing controls and operations* across music technology applications enables users to migrate to new versions of software and to other musical tools with minimal re-training. Standardized operations ensure that operations are performed in a similar manner to other software programs and likewise reduce users' learning curves on new tools. A non-standardized interaction requires the user to learn a new mapping between their intent and the system changes. Users have to refer to help manuals for instruction and can make mistakes while learning the new interface. The extra cognitive effort or learning the new interaction method can distract the users from accomplishing their task. Conversely, a standardized interaction means that users can operate controls successfully by applying their knowledge of similar interactions in other familiar software tools.

Ben Shneiderman [Shneiderman] provides a different perspective from Norman. Shneiderman's principles are summarized below and discussed in relation to music technology applications.

- Strive for consistency
- Enable frequent users to use shortcuts
- Offer informative feedback
- Design dialogs to yield closure
- Offer error prevention and simple error handling
- Permit easy reversal of actions
- Support internal locus of control
- Reduce short-term memory load

Striving for consistency in a music technology application means that similar parameters should have similar dialogs or methods for modification. Also, those interfaces should bear a common resemblance across modules. Audio or note data should have a consistent appearance throughout the application. Further, displays should be consistent with music technology hardware system presentations and interfaces.

*Enabling frequent users to use shortcut*s can reduce the amount of time knowledgeable users require to complete their tasks. An example might include the integration of a user-defined template for music studio software projects or collections of

commonly used items. Another option could relate to automatic routing of input/output signals from computer hardware devices to software controls.

Offering informative feedback means that the system's presentation should clearly communicate necessary information to the user. An example is a volume meter in a recording application. Audio signals that peak too loudly should be clearly indicated such that operators can immediately identify the problem and know exactly how to correct the action. *Designing dialogs to yield closure* means that the user should be able to respond to system alerts efficiently. System dialogs should provide users with the information necessary to complete the operation and perhaps offer the option to fix the condition directly.

Offering error prevention and simple error handling relates to the use of constraints in controls. System dialogs could be presented to alert users to incorrect uses or settings. Error handling can entail implementing an undo/redo capability; it can also involve preventing users from making incorrect routing or data assignments.

Reducing short-term memory load is especially relevant for musical applications. The musician is focusing attention on the audio, instruments, musical concepts, and working to integrate new ideas into compositions. Their attention resources are dedicated to the music. The computer should not force the user to focus on operating complicated controls.

There is much more research into effective interface design for multimedia applications and some specific studies are presented in Chapter Three. HCI provides relevant information in the areas of task-analysis, user testing, audio-visual perception, and cognitive psychology. While, in-depth discussion of those topics is beyond the scope of this dissertation each topic is briefly summarized below and references are provided for detailed study.

Task-analysis research is a detailed process for mapping user intentions into system operations [Redish]. User testing methods are essential for constructing experiments to evaluate the effectiveness of various technology interfaces and to better observe individual variations in task performance [Dumas][Cockton]. Audio perception information characterizes the human auditory system. A sample characteristic is the

number and type of sounds that a human can perceive [Brewster]. Visual perception information gives a deeper description of human visual acuity and visual system response [Sperling]. Cognitive Psychology presents a number of theories about human thought processes and neural mechanisms for sensory acquisition and integration. Engineering Psychology and Human Performance integrate cognitive psychology research with applied methods for studying users attention on complex interaction tasks [Wickens]. Also, the Human Computer Interaction Handbook provides a wide array of useful information for HCI practitioners on input technologies, haptic interfaces, multi-modal interfaces, adaptive interfaces, accommodating a diverse base with a wide range of physical abilities, user centered design and more [Jacko].

This section has presented a summary of design principles and guidelines for application development. A framework was given relating the roles of the user and the system. The primary tasks handled by system input and output devices were described. Music technology application interfaces include a wider variety of devices than desktop systems, many of which reflect traditional musical instruments and can be custom-configured to control parameters in software tools. User roles in group-based multimedia systems were presented as operators, performers, observers, and composers. Two sets of interface design guidelines were given and discussed in relation to situations found in music technology applications. Finally, the areas of task-analysis user testing, audio-visual perception, and cognitive psychology were summarized and references were given for additional study.

2.2 Overview of Interactive Computer Systems

Interactive systems range in size from single-user head mounted displays to multi-person large auditoriums. These systems vary in the levels of interactivity and immersion. Some systems are presented below, including: head mounted display, tablet, desktop monitor, PowerWall, IPT, and auditoriums. Each system is described, pictured, and considered for its applicability in developing a Synesthetic Music Experience Communicator.
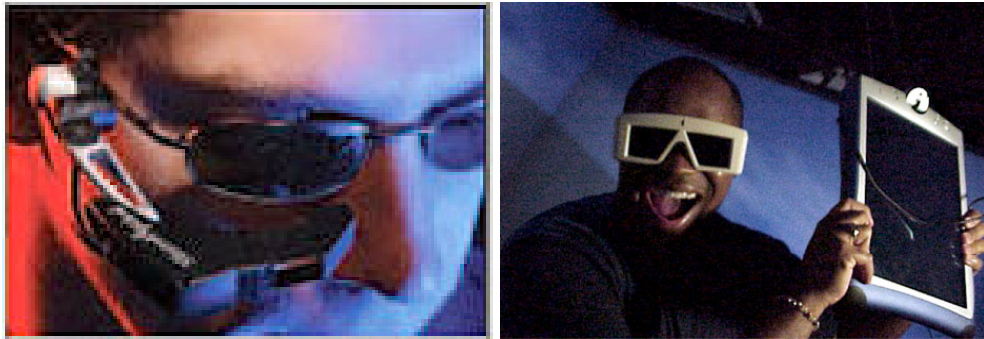
Figure 2.3 Head Mounted Display And Tablet [VRAC]

Figure 2.3 depicts a single-eye Head Mounted Display (HMD) and a tablet computer. A HMD device projects an image in front of one or two of the users eyes. It may block the external world (virtual reality), or project imagery on top of it (augmented reality). Users typically interact with HMD systems through miniature keyboards and mice. HMD systems can be connected to portable wearable computers or full sized systems. Tablet and hand held computers have the same features as desktop systems and provide users with touch-screen or stylus interactions. Both wearable computers and tablet devices include several features that can support multimedia applications. Connectivity options include USB, FireWire, audio input/output, and wireless network interfaces.

Portable systems and HMDs are small enough to be worn by a performer without interfering with their ability to play the instrument. These systems have wireless network capability that can be combined with music and video capture devices to become un-tethered capture systems for their performance data. Also, tablet computers provide intuitive touch screen access to parameters in music applications and are small enough to fit very conveniently into a music studio.

Fig. 2.4 Apple iBook Laptop Computer and Desktop PC in a Home Studio.

Laptop and desktop computers include network connectivity, USB, FireWire, audio input/output, and additional display output connectors. Personal computers and laptops are found in many recording and performance settings. Laptops are portable enough to serve as personal recording studio devices and are great hosts for portable sound/synthesizer modules. The full sized keys and integrated mouse provide familiar interactivity in a convenient package. Laptops are powerful enough to replace MIDI sound modules as seemingly limitless tone generators.

Desktop computers are great for stationary studio setups. Desktop systems have an advantage over laptops in that can support multiple channel audio capture cards. Desktops also typically have larger hard drive capacity and faster processors. Additionally, some audio recording systems are built around PC operating systems and function like an integrated digital recording station with keyboard, mouse, and display.

Fig 2.5 PowerWall and ISU C4 IPT System [VRAC].

Both laptop and desktop systems can generate high quality computer graphics imagery. Some of these systems are powerful enough to simultaneously run software synthesis, graphical simulations, audio processing, and midi monitoring software.

PowerWall and Immersive Projection Technology (IPT) systems provide users with large displays of computer-generated imagery and approach a fully immersive experience. These systems often feature PC interaction tools and interfaces. They also add six degree-of-freedom user-tracking and stereo graphics systems. User tracking enables the system to determine the location and orientation of a user and any tracked system input devices. Stereo graphics and glasses enable users to experience three-dimensional imagery with independent images for their left and right eyes. Users are free to use their bodies to interact with the virtual simulations. In VRAC's C6, a six-sided IPT system, users are completely immersed in the simulations. They can look, rotate, and move freely in any direction and experience graphical imagery. PowerWall and IPT systems vary in size and immersion, but can typically accommodate one or two tracked users and up to ten passive participants.

These systems are coupled with powerful image generating computers, capable of supporting the 20+ frames per second display rates required for a smooth experience. The image generators can be based on a variety of computing platforms including UNIX, IRIX, LINUX, WINDOWS, and OS X. The image generators can be large multi-pipe mainframes such as a Silicon Graphics Onyx 2, or more recently systems have migrated towards clusters of PC's or Macintosh systems. Software tools like VRJuggler [VRJuggler] and CAVELib [VRCO] provide cross platform software development capability and handle all computation aspects of the virtual simulation. These tools connect to tracking systems, generate the graphics displays, distribute data across compute clusters, and communicate with system peripherals.

IPT and PowerWall systems are well suited for group musical education. The size and number of participants allow simulations to be projected into large viewable volumes. As a result, users are placed directly into the middle of simulations and their attention is focused directly upon the graphical imagery.

Fig. 2.6 Alliant Energy Lee Liu Auditorium [VRAC].

Immersive auditorium displays such as VRAC's Lee Liu Auditorium are at the high-end of the scale in terms of the number of participants and size of the displays. This auditorium features a 29-foot wide screen, has seating for 244 participants, and includes a powerful audio system. The auditorium also includes a user tracking system for the operator who stands in front of the display. Participants wear thin polarized glasses that enable them to see the stereo scene. The auditorium images are generated on similar computing hardware and operating systems platforms as immersive IPT and PowerWall systems.

This type auditorium is well suited for presenting a synesthetic musical experience to large groups. Most seats offer a good view of the simulation. There is plenty of room at the front of the stage for an ensemble. The auditorium features networking, audio routing, and the ability to plug additional graphical displays in at the front of the room. Participants are usually passive participants in the simulations however, large group interactions have been implemented through the use of laser pointer devices and colored paddles. In these systems, users can influence the virtual simulation by aiming a pointer at the screen, or by turning the paddle to reveal specific colors.

This concludes discussion of computer systems for music technology applications. All of the systems have potential benefits for application in music technology. HMD and Tablet computer systems are portable and could be used by a

performer while on stage. Laptop systems are convenient and portable, making them suitable for portable studio and sound generation tasks. Desktop systems offer large amounts of data storage, generous computing power, and can be expanded with custom interface hardware. PowerWall, IPT, and auditorium systems are great for immersing medium and large size groups into musical simulations though their options for group interaction are somewhat limited.

2.3 Multimedia/Audio Related Research

Music technology applications share common concepts with other HCI research areas including sonification, multimodal interaction, and artificial intelligence. Sonification applications convert simulation data into audio and graphical imagery. Multi-modal applications accommodate multiple styles of interaction in a single experience. Sonification and multi-modal applications are discussed below. The following discussion includes some principles of sonification and strategies for mapping individual interactions. The section closes with a future perspective on virtual experience sharing, as predicted by Ray Kurzweil, and his vision of autonomous computer generated virtual art experiences.

Sonification, also called non-speech auditory output, offers an advantage in graphical systems for several reasons. First, vision and hearing are interdependent. Humans have limited visual range, but our hearing extends in all directions. "Our eyes tell out ears where to look" and we can direct our visual attention in the direction of interesting sounds. Sounds can also reduce the amount of information that has to be presented on screen, transferring the load from our eyes to our ears. Also, our auditory system is underutilized in compute systems. Some parameters may be represented more intuitively with audio objects [Brewster].

A common example of sonification includes sounds that are modulated by pitch, timbre, and placement in the three dimensional space. In a virtual environment, an application might monitor the intersection between a tracked interaction device and a three-dimensional data set, causing a sound to be played in response to the value at a

given location. Some perception issues can limit the effective use of sonification. Users can effectively determine relative information, changes in pitches, or differences in pitches. However, most individuals do not have a means of directly converting a pitch to a mathematical value. Also, modulating a sound can change its volume level causing some listeners to misinterpret the presented information [Brewster].

Multimodal interfaces are an extension and combination of multiple input devices and multiple output devices. Seemingly human communication occurs with verbalizations, gestures, gaze, movement, and more. Multimodal interfaces seek to expand our options for interaction with computer systems to reflect our natural behavior in interpersonal interactions. Some useful terminology from multimodal interfaces includes feature-level fusion, frame-based integration, and unification-based integration [Oviatt].

Feature-level fusion involves fusing data from parallel input signals. In a synesthetic music application, this data would include all musical instrumentation and real-time audio data. *Frame-based integration* involves matching pattern data to combine semantic information into a common meaning. In this case, all of the instrument data streams can be compared to extract the common aspects of each player's musical contribution. *Unification-based integration* is a means for integrating the meaning of each input mode into a common representation. In music performance, the digital encodings of musical notes, intensity, and timing are the most easily correlated chunks of data. Other issues like tone color and sound envelopes have to be extracted for each individual instrument.

Multimodal signals can be integrated and represented in different ways, according to the type of information being received [Oviatt]. Each musical instrument can and probably should have a different type of representation, based on sound type. Percussive notes have different considerations from sustained notes. Also multiple signals can be used to disambiguate the input information. In an ensemble, multiple musical parts can be fused together to accurately extract important factors like key signature. Another task could be to differentiate individual solo accents from full ensemble hits. Also, musical data streams and real time audio streams can be compared to discover which digital

instrument was responsible for the sounding of individual tones in a combined audio stream.

Looking to the future, a vision for Virtual Experiences was presented by Ray Kurzweil In his book, "The Age of Spiritual Machines." This text proposes a future interactive system where humans have nanoprobes along our sensory pathways. These nanoprobes intercept all of our sensory inputs and motor outputs. In one mode the probes can allow normal transmission of impulses, connecting the human to their normal sensory world. In another mode the human's normal nerve impulses could be switched off, and relayed into and out from a computer simulation. In this way, a human could share the entire sensory experience of another person's life as if they were in that person's body. Further, a person could also provide access to their emotions and secondary responses. This system of 'experience beaming' could eventually be as straightforward as accessing a website and assuming someone's virtual life experience [Kurzweil03][Kurzweil99].

Kurzweil also presents a vision for artistic and musical collaborations. He envisions a new form of visual artistry where virtual-experiences become a dominant software market in the next two decades. This visual art is predicted to range from simulations of real experiences to abstract environments with no correlation to the physical world. Further, Kurzweil proposes dominant digital artists will actually be associations of software-based agents. His text concludes by providing pseudo code over three foundational cornerstones of Artificial Intelligence: recursive algorithms, math-less neural networks, and evolutionary (adaptive) algorithms [Kurzweil99].

This dissertation's intent of communicating the synesthetic experience of music is akin to Kurzweil's concept of an experience beamer. Synesthesia is an internal emotional and cognitive experience in response to various world stimuli. Though the actual experience of the phenomena is somewhat describable, it is difficult to communicate exactly what it is, how it feels, and how it might be used without a dynamic simulation. Also, developing a technology around algorithmic translations of sensory inputs opens the door towards experiential virtual artistry. An adaptive synesthetic visualization system coupled with an adaptive generative music system could provide a limitless source of dynamic computer generated digital art.

To review, sonification, the conversion of data into audio, and multimodal interfaces, applications that feature multiple sources of input information, provide important considerations in tasks that are analogous to synesthetic music communication. This research seeks to leverage the connection between the senses to in a way reverse the process of sonification. Here the goal is to help the audience's eyes inform their hearing about the source and quality of the musical events, much like the synesthetes 'inner mind's eye' converts the music into photisms. Also, the multimodal applications research provides paradigms by which diverse musical signals and data can be integrated to provide a combined view. Mutual disambiguation is a useful principle for use in extracting more precise musical knowledge from multiple audio signals. Ray Kurzweil's vision for experience beamers and virtual computer generated artistry also provides an interesting view of the future of human computer interactivity. As adaptive synesthetic musical communicator program appears to be a solid exploration of experience beaming and is an attempt at a computer-mediated transformation of art.

Chapter Summary

Human Computer Interaction provides frameworks and guidelines for developing effective interactive systems. The frameworks consider the roles of the system and the users, and four tasks: presentation, observation, articulation, and performance. Each task was considered in relationship to this research project. Four user types, Operators, Performers, Audience, and Composer were defined to clarify the roles of active and passive participants in the target applications. Two sets of system design guidelines were described and examples were given for their relevance to musical applications. Also, references were given to additional research in task-analysis, user testing, audio-visual perception, and cognitive psychology.

Several types of interactive computer systems were reviewed for their suitability in visual music communication. Head Mounted Display, Tablet, Handheld, Laptops, Desktop, PowerWall, IPT, and Auditorium sized systems are depicted and their interaction options were described.

The last section of this chapter explored application concepts that are related to group music technology applications. Sonification focuses on the conversion of data into audio and graphical information. Multimodal applications integrate user data from multiple sensory modalities to support more natural interactions, to enable the fusion of multiple streams of data, and to differentiate events though mutual comparison of data streams. Ray Kurzweil's predictions for human experience beaming and computer generated artistic spaces also serve as motivation for this research effort.

This chapter has provided a general foundation for the principles of Human Computer Interaction. These guidelines establish a set of criteria by which to consider and evaluate music technology applications. A description of the interactive systems is useful in considering the best hardware configuration for music technology systems. The sonification and multimodal applications area provide us with additional techniques and motivation for creating synesthetic musical experiences.

CHAPTER 3. MUSIC EDUCATION AND TECHNOLOGY

Developing a software system for musical communication requires an in-depth study of musical concepts. This chapter documents that study, beginning with a general description of music, followed by views on music education. A philosophical view, the fundamental concepts, and a survey of teaching methods are given. Music technologies are discussed, including the role of technology in education. A survey of music technology software is presented. Interactive media tools, electronic instruments, algorithms for music analysis and creation, music applications, educational software tools, and group oriented interactive music systems are discussed. The information and techniques gleaned in this chapter will be incorporated into the design of the proposed systems for synesthetic musical experience.

Music is the creative manipulation of expressive sounds [Reimer]. Performing music is a complex task that engages the audience and the performer in a multi-sensory dialog. Performing music requires cognition, audition, sensory motor control, emotional sensitivity, visual processing and so forth.

Music is an integral part of our human existence. It is studied in many disciplines to gain insight into aspects of our world. Music is a basic form of cognition. Psychologists study human's perceptual response to music. Sociologists and anthropologists examine music to garner insights on our attitudes, class and culture. Historians find that musical compositions help to documents out past and the migration of creative influences. Aestheticians consider the essence of art and how it relates to our intrinsic nature [Reimer].

3.1 Music Education

Music educates our subjective nature as only the arts can. It is thought that all people have some level of musical intelligence and students must be developed in that way. Studying music cultivates the individual's natural responsiveness to sound [Reimer]. This perspective has lead to the creation of music arts programs in our

education system. There are national education standards for dance, music, theatre, and visual arts. The goals of which are to develop students who can appreciate and analyze the arts, and to enable students to define and solve problems in these areas with insight, reason, and technical proficiency [Chosky].

This section of this chapter summarizes the philosophical justifications for music education. Musical concepts are presented and some contemporary teaching methods are described. The discussion of teaching methods will indicate methods to design technology tools for music education.

Philosophy of Music Education

One of the first steps in developing a philosophy for music education is to understand the role that this philosophy will play. One author writes that a philosophy of music education will serve as 'a set of beliefs to guide the group.' Also it is believed that a philosophy will help clear doubts of self-justification, that is the urge to demonstrate the worth of the musical discipline. Reimer further states that a philosophy must offer a convincing basis for valuing music education [Reimer].

Music as Communication

One main principle of Reimer's philosophy of music is that musical communication is simple. This is in contrast to the complicated looking equipment that is used to make music. This appearance of complexity can prevent some from exploring the possibilities of music performance.

Music is viewed as a means of self-expression. In some ways, it is a 'virtual' language of emotions. Music is not fully a language in the communications sense of a language for two reasons. First, the *meaning* of the individual tones is *vague*. Each listener has their own interpretation or emotional experience of the tones. Second, the *arrangement* of the sounds is *not structured* like a language. The emotions are not communicated exactly or explicitly. But, to use an analogy, just as writing and reading

educate reason, creating and experiencing art educate feeling [Reimer] Musical study can educate students in the areas of reason and feeling.

The artist and the audience share the emotional context of music. Thus a philosophy of music also has an aim to involve people in the creation of music. Involvement enables individual experiences, explorations and discoveries of feeling through the act of musical creation [Reimer].

Instilling Deep Value

A philosophy of music must teach students to value music on a personal level. Some methods have been identified for teaching students deep value. It is important to teach good music, specifically music that enhances the student's appreciation for musical concepts. Also, the impact of the music should be emphasized without overemphasizing the details. Students must be exposed to and sensitized to the fundamental elements of music: *melody, harmony, rhythm, tone color, texture and form*. The instruction should **draw attention to the emotional content** of the music through the use of appropriate language. The emotional content should not be labeled for the students; rather they should be encouraged to understand that emotional content in a personal way [Reimer].

Another means for instilling deep value in students is to **involve them in creating art**. There are some useful guidelines for meaningful musical and artistic creation. First, the compositions should have expressive potential. Students should be given the opportunity to be creative in three areas: *actively listening to music, composing through the aid of technology, and creative performance*. Technology can assist the student by letting them explore a variety of sounds. Also, technology can enable the student to explore beyond the limits of their understanding of musical notation. Another guideline is that students should move towards a mindful, not merely feeling-full, **interaction with the qualities of sound**. This also requires that students understand those aspects of sound in which they can be most creatively involved.

A final guideline for teaching value is that the techniques and language of the educator should not distract from the artistic qualities of the sounds of a piece. An

example of a distracting exercise would be drawing pictures of melodies that do not highlight the quality of the sound itself. The point here is that the experience of sound is primary [Reimer].

The preceding guideline seems to be in contrast to the experience reported by synesthetes. Chapter Four discusses synesthesian musical imagery, some of which does not seem to be directly representative of the qualities of the sound. Still the synesthesian musicians report that their imagery adds to the experience of the sound and can serve to improve their performance.

Behaviorist Vs. Cognitive Teaching Methods

There has been a transition from behaviorist based training towards a more cognitive approach. This transition has opened up the doors for the integration of technology learning tools. Behaviorist teaching methods tend to isolate sub-tasks of a skill. Those sub-tasks are then instructed and repeated until well learned. Students are rewarded for their progress at the sub-tasks. A major critique of such methods is that the students only learn the required skills. The behaviorist teaching process does not necessitate deeper learning of the concepts or exploration beyond their skill set [Reimer].

A switch to a cognitive teaching approach has introduced a new goal of *meaningful learning*. This approach calls for the use of higher order behaviors. Cognitive-based educational activities include learning episodes, rehearsals, individual practice, and coaching. These activities are employed to engage thinking, reasoning, and problem solving. Students are encouraged to explore the perceptual, expressive, and synthesizing characteristics of the subject matter in a larger context.

In the context of music education, the cognitive learning approach aims to provide the students with a deepened awareness. Some guidelines for this teaching methodology are [Reimer]:

- Students must know they're studying music by producing it creatively. This is accomplished through a systematic developmental study. They should understand that performance is a creative act.

- Students should understand that technique is a means to an end. They should understand the impact of proper playing style on their performance.
- Musical learning has an impact on artistic execution. Students should know from early on that aspects of their performance like sound dynamics and tone color are interrelated with posture and breathing.
- The student should learn a balance of skills, understandings, and creative decision-making.
- Students should have exposure to a diverse array of musical pieces and a sufficiently diverse and challenging core of practice materials.
- Students should have general music learning beyond their instrument and specific performance repertoire.

Cognitive Curriculum

To accomplish these objectives, Reimer outlines a seven stage curriculum including focusing on five major areas: why (development of values), what (development of concepts), when (development of systematic understanding), how (development of interpersonal, operational, and experienced knowledge), and finally the exceptional phase which overlaps all the other functional areas [Reimer].

The conceptual phase of this curriculum is an area where technology can play the greatest role. An examination of the guidelines for conceptualized music education highlights that role. First, conceptual music education is compatible with present understandings of cognition, musical intelligence, and musical thinking. Also the activities of problem solving, group learning, and drill learning by computers provide deep engagement for students. Second, conceptualized music education is compatible with post-Piagetian child development theories. Third, a conceptual format supports the totality and history of music experience by acknowledging that:

- There are many levels of performance and talent
- Music reflects the diversity of our literature, culture and society
- There are several traditional and non-traditional musical instruments.

Fourth, a conceptual music education focuses on the properties of all modes of music interaction. Finally a conceptual approach considers the roles that music plays in society and the artistic decisions the student can make [Reimer].

Non-Instrumental Music Education

A conceptual approach to music education reveals a need for an expanded music curriculum. A relatively un-serviced group is that of the non-instrumental performing student. This author participated in a general music course in middle school and was able to experiment with active listening, some instrumental performance, and programming/experimenting with electronic instruments.

Some principles for developing such courses are given here. The course should extend beyond the general K-8 music program. The courses should strive to improve the students' musical and aesthetic abilities in several key areas. Students' listening skills should be developed. Students should have experiences with composition, performance, singing, and improvisation. Students should have opportunities to conceptualize the qualities of music including the nature of sound, qualitative characteristics, varieties of music, and music's relationship to art. Another guideline is that students should explore various methods for musical analysis by viewing and creating graphical representations, charts, maps, and diagrams of music. Students should also develop the skills necessary to evaluate performances [Reimer]

Non-instrumental students should learn about the cultural and historical diversity of musical forms. This should include the relationship of today's music to previous works. Students should have exposure to varied styles of contemporary music: jazz, rock, new wave, funk, reggae, reggaeton, afro-beat, salsa, samba, electronica, and so forth.

Finally, these students should have the experience of performing in small group exercises. Those skills may include inventing instruments, learning about various musical notations, group composition exercises, and music theory concepts. Students should have a working knowledge of basic guitar and keyboard techniques [Reimer].

Summary of Philosophy of Music Education

Examining a philosophy of music education reveals the key objectives and potential uses for technology in developing musical knowledge for students. This section has presented music as a means of communication and as a context for emotional exchange. The guidelines for instilling deep value identify the fundamental elements of music that students should have exposure to. Instilling deep value seeks to encourage mindful-interaction with musical concepts and an enhanced awareness of the qualities of sound.

This section has discussed the differences between behavioral and cognitive-based teaching methods. Technology can be applied in the behavioral teaching settings to aid students in mastering the technical skills of performance. Technology can also be applied in teaching cognitive and conceptual skills. Computer systems provide students with the means to create compositions and experiment with a variety sounds. Also, students can use computers as a resource to discover additional information about music theory and their instruments. Computers can also provide access to a diverse selection of compositions and artistic samples.

Finally, technology can play a role in illustrating musical concepts to non-instrumental performing students. While performers have an instrument through which to discover and interact with the qualities of music, the non-instrumental performers do not have that access. Further non-instrumental performers do not have as many opportunities to learn about group musical performances. A computer and composition tool set will enable non-performers to experiment with the characteristics of sound and visualize the musical concepts.

The next sections will define some of the core musical concepts in musical practice and will study some teaching methods in greater detail. Electronic music instruments and software tools are surveyed. Innovative examples are given of applications that enable experts and novices to interact in an audio-visual musical context.

Musical Concepts

The propagation of sound through a gaseous, liquid, or solid medium is essentially a wave-like displacement of matter. Music compositions are comprised of multiple sounds arranged in a creative manner. This section considers the characteristics of sounds and the arrangement of tones in compositions. Some characteristics are inherent measurable aspects of all sounds. Other aspects are idiomatic, a reflection of region, culture, and/or history of the music. The following discussion covers the basic characteristics of sound as well as musical scales, intervals, chords, and rhythmic concepts. Note that these musical characteristics and relationships can potentially be translated into a visual communication of music.

Basic Characteristics

Sounds can be categorized along several dimensions and types of measurements of the waveform. The basic characteristics of a sound include frequency, timbre, intensity, and duration. These characteristics are considered inherent because they are measurable and arise as a consequence of the sound source.

The *frequency* of a sound is also referred to as pitch. Frequency is the rate of vibration of the resonant material or the rate of the cycle of peaks in a propagating wave. There is a relationship between the period (duration) of an oscillation and the wave's frequency. Frequency = 1/Period, measured in cycles per second or hertz (Hz) [Olson] [Helmholtz].

Another quality of a sound is the *timbre*. Timbre is the combination of the multiple frequencies produced by the instrument. These multiple tones, also called harmonics, combine to give the instrument a unique sound and a distinct textural quality. Each instrument has a unique timbre that can be affected by playing technique, environmental conditions, tuning, or even the ambient characteristics of the performance area [Olson] [Helmholtz].

The *intensity* of a tone relates to the volume or loudness of a note. The intensity of a tone varies over time. *Duration* is a measure of the length of time that a note sounds [Chosky]. A more detailed method of characterizing intensity and duration is to consider the overall audio envelope. The stages in an audio envelope are the attack, sustain, decay, and release. Attack is the rate of increase in the intensity from onset to peak. Sustain is the level of peak intensity before the note is released. Decay is the rate of decrease from the release of the note until the cessation of sound.

Scales

Musical systems are based on defined musical tones. The middle A on a piano is defined as 440 Hz. Doubling a note's frequency corresponds to the an increase of an octave. Halving the frequency results in a decrease of an octave. There are 12 half steps in an octave. The ratios between the half steps are defined by mathematical relationships. Musical scales are constructed from a pattern of intervals in an octave. The C-major scale is a common example. Those tones are C, D, E, F, G, A, B and C (the octave).

The interval distance between the tones is what defines the scale. A scale can be found in any key by applying the same interval calculations beginning from the first note (root) of the scale. The notes in each scale have a unique emotional feel and can trigger an emotional response in a listener. A list of common scales and their intervallic spellings is given in Table 3.1.

Table 3.1 Common Scale Names, Degrees, and Notes [Sessions]

Scale Name	Chromatic Scale Degrees (#= sharp, *b* = flat)	Letter Names (Base key of C)
Ionian (Major)	1, 2, 3, 4, 5, 6, #7	C, D, E, F, G, A, B, C
Lydian	1, 2, 3, #4, 5, 6, #7	F, G, A, B, C, D, E, F
Mixolydian	1, 2, 3, 4, 5, 6, *b*7	G, A, B, C, D, E, F, G
Dorian	1, 2, *b*3, 4, 5, 6, *b*7	D, E, F, G, A, B, C, D
Aeolian	1, 2, *b*3, 4, 5, *b*6, *b*7	A, B, C, D, E, F, G, A

Scale Name	Chromatic Scale Degrees (#= sharp, *b* = flat)	Letter Names (Base key of C)
Phrygian	1, *b*2, *b*3, 4, 5, *b*6, *b*7	E, F, G, A, B, C, D, E
Locrian	1, *b*2, *b*3, 4, *b*5, *b*6, *b*7	B, C, D, E, F, G, A, B

Intervals [Sessions]

Here we consider the perceptual effect of two simultaneous or sequential pitches. The interval distance between two simultaneous tones impacts the feeling and quality of the resulting sound. The intervals also have emotional, impact on the listener. There are a number of ways to describe the intervallic distance between tones. Table 3.2 gives a set of common interval names, the number of half steps betweens the notes, an indication as to the quality of the chords, and the mathematical ration between the fundamental frequencies of the notes.

Table 3.2 Interval Relationships [Sessions]

Interval Name	# of Halfsteps	Example	Interval Quality	Ratio
Unison	0	C-C		1/1
Minor second	1	C-D*b*	Dissonant	16/15
Major second	2	C-D	Dissonant	9/8
Minor third	3	C-E*b*	Consonant	6/5
Major third	4	C-E	Consonant	5/4
Perfect fourth	5	C-F	Consonant	4/3
Augmented Fourth	6	C-F#	Dissonant	7/5
(Diminished fifth)	6	*(C-Gb)*	Dissonant	*(10/7)*
Perfect fifth	7	C-G	Consonant	3/2
Minor sixth	8	C-A*b*	Consonant	8/5
Major sixth	9	C-A	Consonant	5/3

Interval Name	# of Halfsteps	Example	Interval Quality	Ratio
Minor seventh	10	C-B*b*	Dissonant	16/9
Major seventh	11	C-B	Dissonant	15/8
Octave	12	C-C		2/1

Consonance and dissonance are two terms relating to the perception of simultaneous tones. The qualities of consonant and dissonant have been described by their perceived characteristics. Dissonant intervals have been described as 'harsh, disturbing, and unresolved'. Dissonant intervals lack a strong indication of the natural fundamental (or root) of the chord. Consonant intervals have been labeled 'sweet, restful, and harmonious.' Consonant intervals do indicate the root of return [Sessions] [Helmholtz]. Moving beyond pairs of notes into more complex chords leads to more detailed analysis of the intervals between individual tones.

Chords

A chord is the simultaneous sounding of two or more tones. The construction and sequencing of chords can be analyzed as a stream of information communicated by the musical composition. As noted above, this information has both emotional and mathematical meaning [Reimer]. Musical chords create expectations in the listeners mind with regards to movement and resolution. Chords also have related properties called harmony and voice leading.

A brief review of chord types reveals the diverse array of tone arrangements available to the composer. Chords are named on the basis of the relationships between the notes. Table 3.3, adapted from Sessions, shows a set of base chord types, symbolic abbreviations, and chord degrees.

Table 3.3 Names, symbolic spellings, and scale tones in common chords.

Chord Name	Symbol	Chord Degrees
Major Triad	Ma	R-3-5
Minor Triad	Mi	R-*b*3-5
Diminished Triad	Dim	R-*b*3-*b*5
Augmented Triad	Aug	R-3-#5
Suspended Fourth Triad	Sus4	R-s4-5
Suspended Second Triad	Sus2	R-s2-5
Dominant Seventh Triad	7	R-3-5-7
Dominant Ninth Triad	9	R-3-5-7-9
Dominant Eleventh Hexad	11	R-3-5-7-9-11
Dominant Thirteenth Heptad	13	R-3-5-7-9-11-13

The function of chords in compositions can be considered in three ways: individually, locally, and globally. *Individual* consideration looks at only the effect of the simultaneously sounded tones. *Local* consideration looks at the previously sounded chords. *Global* consideration examines the chords in relationship to the scale and mood of the piece. Each method of chord consideration is expanded below [Sessions].

Individual chords have characteristics of return, expectation, mood, stability, activity, and conflict.

- The **return** of a chord is an indication of the root the chord. The mathematical relationships between the individual tones indicate the natural fundamentals of the chords and imply a direction of resolution for the chords.

- **Expectation** concerns the intervals between the individual tones in a chord. The set of tones elicit mental predictions for the next likely sequence of chords. Composers can choose from several options to completely or partially fulfill the listener's expectations.

- **Mood** is established through the melodic movement of tones in a composition. This movement can be increasing or decreasing in pitch. Further, melodies

can employ consonant of dissonant intervals in scales to influence the mood of a piece.

- **Stability** refers to the arrangement of the tones in a chord. This term considers the melodic tendencies of the upper tone in a chord in relationship to the lower tone and root of the chord.

- **Activity** is a measure related to stability. As an example, a chord containing the unstable tritone (augmented fourth or diminished fifth) interval has two possible resolutions in differing musical chords.

- **Conflict** arises in chords with three or more notes. The overlap of the tones can cause discomfort or anxiety for the listener. Conflicts also impact the users sense of expectation in the music.

Local chord considerations can be analyzed in terms of local expectation and local fulfillment. Local expectation relates to the comparison of individual tones to the expected root. Listeners compare the tones in two adjacent chords to determine the movement. Fulfillment is a measure of how well the new chord matches the users expectation.

Global chord considerations include movement, ambiguity, and finality. The sense of *Movement* is influenced by the relationship between consonant and dissonant intervals, the arrangements of tones in a composition, and the listener's expectations. *Ambiguity* occurs when a listener is unable to determine the progression of a composition. *Finality* is the sense that a composition is nearing completion. The sequence of chords and melody influence the mood of a piece. The final resolution of a composition typically provides a sense of completion or fulfillment for the listener.

Rhythmic Concepts

The listeners' experience of a composition is also influenced by the dynamics, rhythm, timbre, and texture of the sounds. Above, dynamics were considered in the context of individual sounds. Dynamics can also be measured by considering the volume levels of a composition over time. The timbre of a sound is influenced by the frequencies

and overtones it contains. The texture is the density and interaction between parts. Rhythm is a measure of the timing between sounding of the notes.

Jaquez-Dalcroze identified thirty-four elements of rhythm. These elements have been integrated into a teaching methodology. With the intent of sensitizing the performer to the rhythmic choices they make in interpreting a piece of music. These concepts also influence the performer's body motions, conceptual organization, and execution of a piece of music. Some of the elements and their meanings are summarized below.

- **Time-space-energy-weight-balance** describes the overall placement of emphasis and the rhythmic flow it implies.
- **Tempo and changes in tempo** relate to the overall speed or movement of the musical passages.
- **Dynamics** describe the instantaneous level, increase, and/or decrease of volume, energy, and intensity of music.
- **Articulation** relates to the style with which the notes are played. This can include staccato, legato, portamento and accent. Also relevant are the attack, release, sustain and vibrato of the generated sounds.
- **Measure and rests** govern the organization and sub-division of a piece of music. Rests are the silences between segments of music. Subdivided rhythms impact the feeling of a composition by the changing speed or quality of movement in a piece.
- **Duration** is the length of the tones.
- **Patterns** are groupings of smaller rhythmic elements. Single, duple, triple patterns, as well as rudimentary combinations.
- **Intrinsic beat** is an extraction of tempo from a complex pattern of musical layers.
- **Phrasing** is the process of identifying the breaks between patterns and segments in music
- **Diminution** refers to performing a musical passage at double, triple, quadruple, etc the original tempo.
- **Augmentation** is performing the passage at half the tempo.

- **Rhythmic counterpoint** involves playing contrasting rhythms simultaneously,

- **Syncopation** is the use of patterns that anticipate or retard an existing rhythmic pattern.

Other identified elements of rhythm include complimentary rhythms, unequal measures, unequal beats, polymeters (combinations of two different musical meters), polyrhythms, and rhythmic transformations. An example of a rhythmic transformation is the sub-divisions of a twelve-count phrase into two six-count phrases, four three-count phrases, and so forth. The last listing is rubato: the expressive lengthening or shortening a pattern while maintaining a constant tempo [Chosky].

Additional percussion techniques include sets of standard practice materials, called rudiments. These rudiments are typically short exercises designed to develop mastery over common elements of rhythms. Rudiments are employed to improve general playing technique issues including hand placement and stick/mallet control. Percussion ensembles use rudiments to develop group tempo and to teach marching fundamentals [ISU-Drumline].

Notation Systems

Music is commonly notated on a musical staff (figure 3.1). Tones are assigned note values and are placed onto the scroll. The staff's key signature is denoted by the placement of a clef as well as sharp and flat symbols. Note durations are designated by the appearance of the notes. Some notes use hollow circles, filled circles, stems, flags, and ties to indicated duration. Musical notation also contains markings for dynamics, tempo, and symbols that represent sequencing and facilitate rehearsal.

Figure 3.1 Sample Music Notation [Sharer].

There are several nuances in musical notation that are beyond the scope of this document. However, considering music notation in the context of music education does reveal a potential use for musical technology. The process of converting musical notation into artistic or abstract representations can be beneficial to students. Conceptualizing in this manner is good because staff notation can be difficult to learn [Reimer]. Technology could also be used to help students explore various notation systems.

The Music Notation Modernizing Association (MNMA) is an organization promoting the development of alternate musical notations. The reasons they cite include the bias of current notation towards common key signatures. Currently, notating alternative key signatures requires the use of several accidental signs: sharps, flats, and naturals. Also current notation uses four different clefs, treble, bass, alto, and tenor. Conventional notation makes it difficult to calculate interval relationships among the notes. The MNMA website offers a tutorial on chromatic notation systems [MNMA].

Alternate methods of notation include enumeration as is done in Chinese numbering system. A longer list includes; Fasola (four shape notation system), Andrew

Law's four shape notation, Dearborn letter notation, Harrison letter notation, Tonic sol-fa notation, Shaker letteral notation [Steel].

Piano roll (figure 3.2) is a notation form that is commonly found in software applications. This mapping includes a representation of a piano keyboard, a horizontal scroll of musical measures, and colored blocks indicating the length and volume of a note.
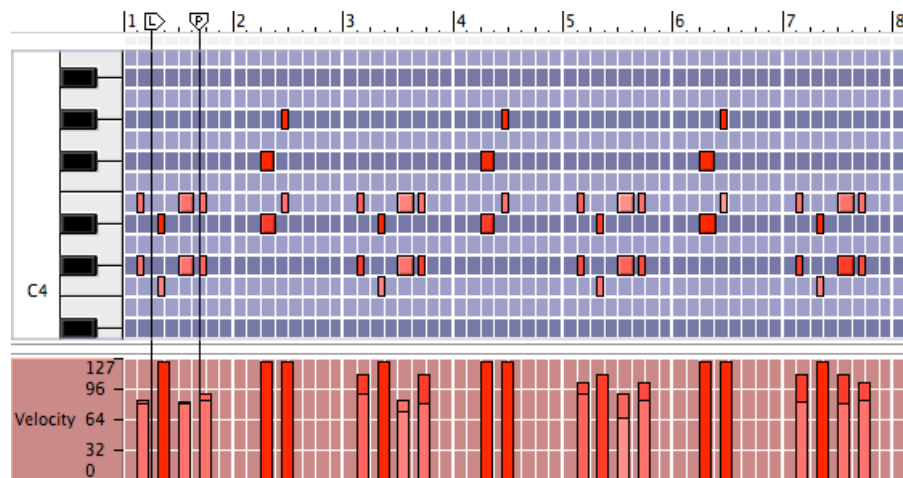
Figure 3.2 Piano notation view in Propellerhead Reason sequencing software.

Chord blocks and tab notation are often used in popular music charts to communicate guitar chord fingerings. Tablature indicates fret numbers for guitar parts on a multiple line diagram. Each line represents a string on the bass guitar, and the number is the fret that should be depressed to generate the correct sound. Chord blocks and tablature are very intuitive because they reflect the exact layout and finger positioning of the instrument.

Summary of Musical Concepts

This section has presented much knowledge about musical structure, representation, organization, and notation. Each concept is an additional means by which a computer algorithm might visualize a piece of music.

In summary, the musical tones in a composition can be organized in three primary dimensions: **horizontally** with respect to time, **vertically** with respect to tones, and **expressively** with respect to dynamics. Horizontal organization of music includes the representation of melody lines. Rhythm in simple terms is the division of sound and silence. Melody includes the pitches and rhythms of a piece.

Vertical organization of sounds in a unit is the comparison of harmony and texture. Harmony arises from the set of sounds that are played or sung simultaneously. The texture is a representation of the characteristics of the sounds. Texture arises from the layering of multiple sounds in a composition. An example of a variance of texture might be guitar and a piano sounds in one part of a composition, and electric bass and violin sounds in another.

The expressive qualities of a composition include intensity and timbre. The factors describing the intensity of a sound or composition includes the volume, dynamic variance, and amplitude of the total sound. The timbre is the tone quality or tone colors (overtones) of a sound [Chosky].

Several other aspects of music have been presented. The **basic characteristics** of sounds were given, including wave propagation, frequency, tone quality, and intensity. **Musical scales**, the selected tones used in a composition, were presented. **Intervals**, the measure of the number of half steps between tones, and an introductory measure of their impact on a listener were given. **Chords** were defined and examples of chords were given. Methods for considering the functions of chords were described. The cognitive effect of chords on the listeners was introduced. A list of **rhythmic concepts** was given, several of which involve perception and cognitive processing by the performer. **Notation systems** were described, including samples of standard notation, piano roll notation, and a list of historical notation methods. Also, some of the perceived limitations and future growth needs of notation systems were described. Finally, musical concepts were summarized in a short synopsis of organization along temporal, pitch, and expressive dimensions.

Several of the concepts presented here are numerical and mathematically related. They are prime candidates for representation through visualization systems. However, to

effectively design a system to support education and rehearsal, the current methods of teaching these concepts must be reviewed. The next section introduces teaching methods that are used to introduce these concepts to students.

Teaching Methods

In this section, a number of teaching methods are reviewed to better understand the role that music technology can play in education. These teaching methods are an extension of the various discussions about the philosophy of music education presented above. Further, methods include a body of 'well defined practice', worthwhile 'goals and objectives', and an integrity of purpose [Chosky].

A set of seven musical behaviors have been identified to bridge between the larger goals of music education and the specific skills that students are to acquire. These seven musical behaviors, also referred to as modes, are organized in to three categories: ends behaviors, means behaviors, outcome behaviors. Ends behaviors are designed to enable students to interact with the music. The ends behaviors include *perceiving* and *reacting* to the music. Means behaviors are designed to develop a 'heightened aesthetic experience.' Means behaviors include *creating*, *conceptualizing*, *analyzing*, and *evaluating* musical compositions. Outcome behaviors are meant to enable students to *value* music [Reimer].

Pestalozzi

Another set of musical teachings can be found in an adaptation of Swiss education reformer Johann Heinrich Pestalozzi's principles. Lowell's Mason, a founder of Boston's Academy of Music, presented an adaptation of Pestalozzian Principles. Those goals included the development of musically intelligent adults, a focus on the quality of music, and improvement of the process of teaching. Mason also sought to begin music education with children at a young age, to focus on multi-sensory engagement, and to stress that learning music required effort. This approach stated that experience precedes theoretical

learning and theoretical understanding arises from that practical experience. Mason stated that most individuals can and should become musically literate. His teaching pedagogy was "an inductive method, exercising actively the reasoning principles of the mind." This is summarized in three key phrases, "sounds before symbols", "principles before rules", and "practice before theory" [Chosky].

Dalcroze

Èmile Jaquez-Dalcroze's approach included a focus on three musical aspects: Eurhythmics, Solfège, and Improvisation. Eurhythmics begins with the assumption that all musical rhythms are reflective of the "natural rhythms of the human body". Dalcroze studied many human kinesthetic processes, noting that our movements generate physical sensations. These sensations are encoded according to direction, weight, force, speed, changes of mass, and other factors. Our awareness of our movements is a form of kinetic sense. This combination of moving-sensing and feeling is titled kinesthesia [Chosky].

Through kinesthesia, Dalcroze's method views music as an intersection between physiology, psychology, and physics. The automatic happenings, the subconscious aspects of our existence are referred to as automatisms. Dalcroze sought to make students aware of their own kinesthetic sense, in order to improve their awareness of their musical performance. This lead to the development of the 34 elements of rhythm summarized above as well as a set of movements and exercises. Examples of stationary movements include clapping, swinging, turning, conducting, bending, swaying, speaking and singing. Non-stationary movements included walking, running, crawling, sliding, and more [Chosky].

Stages of movement associated with those exercises include preparation, attack, prolongation, and return to preparation. These movements parallel the duration of sounds: attack, sustain, decay and release. Dalcroze's approach sought to develop inner hearing in students. Students learn aspects of music by acting them out through body movements. Students learn to constrain their movements and mentally perform the aspects of music.

In this way, students can internalize beat, meter, rhythm, melody, form, and other aspects of music [Chosky].

Another aspect of Dalcroze's method is Solfège. Solfège is a study in musical notation and rhythm through the use of voice. Notes on a staff are assigned syllables, "*do, re, mi, fa, so, la, ti, do.*" Students perform sight singing of musical phrases. Solfège also has exercises for studying scales, modes, intervals, melody, harmony, dynamics, phrasing, vocal improvisation, posture, breathing control, and other skills [Chosky].

Dalcroze's improvisation aims to teach the use of rhythm and sound to produce music. This method uses a variety of instruments, speech, movement, and visual imagery. The roots of improvisation are found in the transformation 'of a spoken story into movement and sound… or of movement and sound into poetry and story.' Improvisation instructors teach students how to perform and accompany the stories in innovative manners. Students are encouraged to improvise on the basis of their musical knowledge and artistic compulsions. Improvisation is seen as a unifying activity, allowing students to begin to integrate their feelings, thoughts, perceptions, and movements into their performance [Chosky].

There are several other teaching methods that were briefly researched but are not described here. Those methods include the Kodàly Method, the Orff Approach, and Schulwerk. These methods all seek to accomplish musical awareness, to inspire musical thinking, to develop solid technique for players in ensembles, and to encourage the students to appreciate the value of musical practice.

Summary of Music Education

This summary ends the section on music education. The current state of music education reflects an evolved philosophy of music education. These philosophies endorse the perspective that music has value as an artistic and emotive means of communication. As such, all individuals can and should develop musical appreciation in their formalized education. Teaching methods work to instill a deep value in students and a personal connection to music through performance and rehearsal. Educators have been encouraged

to push beyond behavioral-based teaching methods towards a more cognitive approach. Cognitive approaches seek to increase students' mental engagement in the artistic process. Several stages of a cognitive curriculum were identified including the development of values, concepts, systematic understanding, and applicable knowledge. A need was identified for developing musical understanding in students who do not pursue formal training with a musical instrument.

The first part of this chapter has presented several musical concepts. Sound was introduced and some basic quantifiable properties were introduced. Musical scales, intervals, and chords were presented and their impact on the listener was described. Methods of analyzing the chords in a composition were overviewed. A large number of rhythmic concepts were given and methods of percussion education described. Musical notation systems and some potential improvements were identified. A summary of three ways to organize the parts of a composition was described.

The last portion of this section was devoted towards an overview of teaching methods. A set of desired musical behaviors was identified. Pestalozzian principles were summarized including an emphasis on developing musical reasoning. Dalcroze's three-sided approach has been summarized. The first part of that approach, eurhythmics, is based on kinesthetic body movement. The second, solfège, is a method for teaching students about, singing, rhythms, and posture. Thirdly, improvisation works to teach students about combining musical knowledge and expression.

Music education, philosophy, and methods endorse the use of visual and mental imagery to develop musical understanding. The musical concepts presented here have a mathematical basis and are classifiable. The melodic, harmonic, rhythmic, and textural components of music are measurable and can be used in a conceptual representation of a musical piece. Again, all of these musical concepts and teaching methods provide a context into which a Synesthetic Musical Experience Communicator can be designed. Now that the goals, concepts, and techniques of music education have been established, the next task is to examine the technologies available to musicians.

3.2 Music Technology

Digital systems play several important roles in the music arena. These systems facilitate performance, recording, and communication of music. A major aim for music technology research is to encourage expanded adoption in music education and performance. Another aim is for technology to serve as a creative partner or assistant in the process of music creation. The role of music technology is described here including its impact on individual access to music, its use in composing and distributing music, and its integration into the stages of learning.

Technology has had an important historical impact on music access. The introduction of the phonograph gave everyone access to music. Prior to the phonograph and various other music distribution technologies, the only access individuals had to music was live performances in the community or in families [Reimer]. With today's digital devices, individuals have almost limitless portable access to audio materials. For example the author's 20-gigabyte iPod holds 4500 songs. Services like iTunes provide access to preview, purchase, and download songs over the Internet. In some cases, studios have pre-sold and/or released albums over the Internet before they arrived in traditional brick and mortar retail outlets. Technology also provides access to a wide variety of songs, sounds, and experiences [Reimer]. Music DVDs provide better than concert views of live performances and high quality multi-channel audio.

Another perspective on music technology is found in the philosophical root of music education. One perspective states that listening and composing music are activities of the imagination. Music performers make expressive sounds. The notation is not the music, rather the notation is a set of instructions for the performer. The notation is a representation of the sounds. Notation is a great aid in the development and transcription of music, but a technology system allows the student to interact directly with the musical sounds [Reimer]. Technology also enables the publication and sharing of musical ideas and content. An amateur composer can compose, record, and publish to the Internet, instantly reaching a worldwide audience.

Technology can play the role of a learning partner, integrating successfully into a four-stage model of learning. Initially students require assistance with the fundamental concepts of music. Once a student understands a concept, they transition towards self-assistance. In this stage, students are knowledgeable enough to improve their own performance. Eventually the student is able to perform without any assistance [Chosky]. The fourth stage is when students return to self-assistance to develop a new internalization of a concept or task. This fourth stage is where technology can play a significant role in providing students with access to information about their instrument or music theory. This information can help them develop a greater understanding.

The next sections present some of the interactive media applications that are available. The applications are introduced in high level categories and some brief examples are given. Next, digital music instruments are described, including the protocols they use to communicate and some of the benefits they provide for performers. Active areas in music technology research are discussed, including some algorithms used to analyzing musical performance. Finally, some research applications for desktop computers and group settings are presented.

Interactive Media and Software Applications

There are web pages, books, CDs, and videos with examples of playing techniques, fingerings for instruments, practice rudiments, music theory tutorials, techniques for ensemble performance, tips for gig playing musicians, and more. There are software tools that provide assistance with notation, music sequencing, computer assisted instruction, accompaniment, and even web-based teleconferencing for conducting remote lessons.

Computer Aided Instruction

The Society for Music Theory publishes a thorough bibliography of computer-aided instruction tools. These programs are available for platforms for various computer

platforms. The list is too extensive to detail here. Instead, the categories of applications are given. Categories include analysis, composition, conducting, dictation, ethnomusicology, history, jazz, keyboard, MIDI, music appreciation, orchestration, part writing, performance, set theory, sight-reading, singing, terminology, testing of aptitude and knowledge, theory, and tuning [SMT].

Some of the categories listed above are further subdivided. Dictation applications are subcategorized by application-focus into chords, errors, harmonics, intervals, jazz, melodies, modes, and rhythms. Performance applications are available in the areas of jazz, percussion, and rhythm instruction. Theory application focus areas include chords, intervals, jazz, non-harmonic tones, notation, scales, and solfège [SMT].

Certain computer aided instruction applications can be used in combination with digital musical instruments and microphones to monitor student learning and performance. Students can be quizzed through intelligent systems and the software can adapt lessons to the user's performance. Students are also able to query specific information and playing techniques. These software tools are commercially successful and have been proven effective in music instruction. The concepts and goals demonstrated by these desktop tools are applicable and should prove useful to group settings.

CALMA, Computer Assisted Learning for Musical Awareness, is a software tool for lesson authoring. It is tool for creating courseware for music learning for independent learning, group work, and seminar preparation. This system enables users to link CD quality sound, text, graphics, pictures, and documents to build music lessons. This application is well suited for developing instructive tutorials that will draw the student's attention to selected sections of audio [Huddersfield]. Proposals for other learning tools call for the ability to change the speed of playback of a complex piece of music, to isolate and extract individual instrument parts, or to assist in the transcription of specific music passages [Coyle].

HUMDRUM is a toolkit designed to benefit music teachers by providing a set of query tools for text formatted songs. Song formats can include many symbolic formats: "medieval square notation, MIDI data, acoustic spectra, piano fingerings, changes in

emotional states…" and more. Humdrum tools can perform several operations on sets of songs including: visual display, aural display, searching, counting, editing, editorializing, transforming between representations, arithmetic transformations, extraction of information, linking information, generating inventories of songs, classifying, labeling, comparing, capturing live data, even trouble shooting notational errors [Humdrum].

Synthesis and Production

There are several programs available that enable music creation, recording, and live performance. Two examples of very popular music tools today include Propellerhead Reason and Ableton Live. Used together, these enable a personal computer user to execute many of the music production tasks once reserved for professional music studios.

Reason is a virtual music rack similar to today's rack mounted studio hardware. Users are able to create instances of sound modules that simulate subtractive synthesis, graintable synthesis, multi-sampling and percussion. Reason sound modules are configured by presets that are available from the Internet. Users can also create audio effects modules for digital reverb, digital delay, analog delay simulation, distortion, phase/flange, chorus, equalizer, and vocoder effects. Reason includes the ability to route signals through mixers, to split audio and control voltage signals, also, to animate control voltages and system control knobs. Users can configure Reason controls to respond to external control devices including musical keyboards, alphanumeric keyboards, sliders, drum pads or any other MIDI capable device. Reason users can record, sequence, and edit from any MIDI data source. Reason's generated audio can be exported via the Re-Wire protocol directly into other digital audio processing applications [Propellerhead].

Ableton Live enables audio recording, sequencing, effects, and live triggering of audio samples. This application is capable of recording audio from external sources, from Re-Wire enabled applications, from MIDI devices, and from other virtual sound modules. Users can apply digital effects to audio samples, select portions of samples for playback, and trigger their samples from external devices. Users can synchronize audio sessions between multiple computers. Also, their audio sessions can be exported to better than CD

quality audio. Ableton Live was developed by a group of digital music performers to enable real-time remixing of audio materials from an intuitive and efficient user interface. Ableton Live has found its way into the studio and into large concert productions [Ableton].

Other similar audio synthesis and production tools include Cakewalk SONAR [Cakewalk], Digidesign Pro Tools [Digidesign], Logic [Logic], Sonic Foundry Acid, and Sound Forge [Sony]. Also of interest are Virtual Studio Technology (VST) instruments. VSTs are modular instruments, effects, and processing tools designed to work interchangeably in software applications. These tools simulate the operation of actual music hardware and vintage analog equipment. Any person can develop a VST by following the interface rules and data formats.

These audio creation tools represent a significant milestone in music technology. Home users have the ability to experiment with a seemingly endless number of sounds. Creating and recording a composition is possible with commodity pc hardware. These software tools also enable a personal computer to become a live improvisational digital music instrument. These software systems deal with MIDI device protocol and common audio formats. These powerful software simulation and recording tools are suitable for integration into musical visualizations systems.

Summary of Interactive Media and Software Applications

Computer Aided Instruction, Audio Synthesis, and Production Software tools provide musicians with the means to research musical information, to improve playing technique, to experiment with sounds in compositions, and produce finalized music compositions. Today's systems offer endless opportunity for learning and creating. Computer Aided Instruction tools teach several musical concepts. Audio synthesis and production software tools use common protocols with current music audio and performance systems. Both categories of software systems demonstrate the potential uses and benefits of technology in individual performance and education settings. An informative synesthetic musical system would share a common intent with existing

interactive music software systems. The next step is to connect the software systems to the musical performer.

Electronic Music Instruments

This section presents some of the digital instruments available for use in music education. Electronic music instruments simulate real instruments and share design features with their acoustic counterparts. These instruments play an important role in musical learning and exploration by providing students with access to an array of sounds, and by integrating with software tools. A discussion of the communication protocols, musical instruments, and related issues is given below.

MIDI

Artist's performances on digital instruments are commonly encoded and transmitted using a byte-oriented protocol called MIDI. Musical Instrument Digital Interface (MIDI) was first introduced as Universal Musical Interface (UMI) at the National Association of Music Merchandisers conference in 1982. The protocol was presented as a standard to enable communication of musical performance data between digital music devices. The protocol includes messages for status changes and data messages. Status messages relay setting changes for knobs, dials, sliders or similar control surfaces on digital instruments. Data messages communicate the new status of the changed controller. Data messages include fields of device id (ranging from 0-16), channel number (0-16), note number (0-127), velocity (0-127), and others. MIDI allows a controller or hardware device to communicate that the performer has played a note or made a change. MIDI also enables a receiving device to accept commands from another device or sequencer (midi event recorder) that would indicate that a note should be played or a controller setting modified [MIDI].

Computer systems support MIDI through a combination of hardware interfaces and software. MIDI events are used in software audio generation and audio recording

packages described in the preceding section. Most major computer operating systems have a facility for handling MIDI events and tools for creating software applications. AS an example, MidiShare is a software library for developing MIDI-based applications on Macintosh, Windows, and Linux systems. This library can intercept and/or create MIDI events. MidiShare also provides hooks to C++, Common Lisp, Pascal, and Java [MidiShare].

Instruments

Two common electronic music instruments are keyboards and percussion. Figure 3.3 depicts an M-AUDIO MIDI keyboard. These devices are built to simulate a piano keyboard. Individual notes are identified and the intensity and duration of played note are transmitted by MIDI. This keyboard also includes nine parameter sliders, eight parameter knobs. Pitch bend and modulation wheels are also included . Each control is digitally encoded and changes are transmitted over MIDI [Avid].

Figure 3.3 Basic Midi Keyboard Controller with 61 Keys, 9 Knobs, 9 Sliders, Pitch bend, and Modulation control.

Electronic keyboards are available in a variety of sizes and scales. Some models include as few as two octaves of keys. Others are full-sized with eighty-eight weighted pressure sensitive keys. Some consumer keyboards feature keys with integrated lights or displays of keyboard layouts. These displays indicate keyboard status and can be used to teach individuals how to play songs. Keyboard features can include tone generators, rhythms and auto-accompaniment, integrated surround sound, and karaoke video support.

Additional control surfaces include infrared motion sensors, two-axis joysticks, analog touch strips, and analog touch panels [Roland][Casio][Korg][Alesis].

Figure 3.4 shows an electronic drum set. Drum sensors come in many types including gum-rubber coated pads, mesh-head covered simulated drums, mounted piezoelectric, and multi-pad percussion systems. Typically these systems involve the use of an analog triggering device and a digital encoder. Each sensor transmits a unique identifier and intensity level over MIDI. In some systems, the percussion triggers are connected to a 'brain' that encodes the measured signals and generates audio corresponding to the inputs. Some percussion devices include integrated assists to monitor their performance. Some features include a metronome, groove measurements, and/or light up drum pads [Roland][Yamaha][Alesis][Ddrum]. Other percussion devices include xylophones and marimba simulator systems [AltMode].

Figure 3.4 MIDI Electronic Drum Set, Laptop Computer, and Midi Keyboard.

Wind controllers, another class of MIDI instrument, simulate woodwind instruments. Wind controllers include simulated keys that accommodate hand placements for physical woodwind instruments. These controllers also include mouthpieces that perform breath measurement. Wind controllers can work with customized tone generators to create realistic sounding wind instruments. Other MIDI tone generators can also be used to gain access to non-wind instrument sounds [Yamaha][Akai].

MIDI enabled guitars use special sensors to measure and encode individual string vibrations. Some MIDI enabled guitars include fret board sensors that encode exactly the frets that the performer has depressed [Roland]. There are also a variety of foot playable keyboards, digital stomp pedals, and analog pedals that can be easily incorporated into MIDI equipment.

Some MIDI systems provide a collection of sensors and wireless transmission to enable development of custom musically responsive hardware. I-CubeX system sensor types include bend, touch, pressure, light, linear sliders, rotary dials, temperature, and others. These sensors work in a modular way with each one connecting to a sensor box that encodes the signals and transmits them into MIDI software applications. An I-CubeX system can be used to create musical instruments or bodysuits for full-body interaction with music [Infusion]. Another example of a custom MIDI interface device is the Musicpole. This device "wraps a standard keyboard around a visually striking acrylic pole". Touch sensitive strips are arranged in a spiral pattern according to the musical principle of the circle of fifths. Designed for both education and performance, this device enables intuitive soloing and easy transposition of key signatures [Tuttle].

Video Hardware

Another type of MIDI enabled-device is video hardware and software. These systems connect to MIDI devices and render visual images instead of audio sounds. Users are able to select and assign video clips and visual effects transformations to MIDI events. These systems can modify real time video inputs in response to user actions or audio input [Edirol].

Pros and Cons of Electronic Instruments

While electronic instruments are designed to simulate real ones they can have a very non-real feel. Electronic instruments are advantageous for certain settings and limiting in others. Those benefits and limitations are discussed here.

Electronic music instruments enable musicians to work with diverse sound sets. Musicians can express ideas using a variety of MIDI input controllers with any sound available in their sound generator. Electronic instruments integrate well with acoustic instruments. As an example, an acoustic drum set can be augmented with electronic sensors and digital percussion.

MIDI instruments facilitate transcription of performances. A digital representation of the MIDI events can be modified in a sequencer to fine-tune the performer's technique. Electronic instruments can simplify the recording process because they do not require microphones or special sound recording booths. Digital instruments can also bypass analog conversion of audio and they can provide recording engineers with refined control over audio parameters.

The primary complaints against early electronic instruments were authenticity of play and the sound quality. Modern systems are much advanced in realism, often providing CD quality sampled audio. Keyboards have evolved from monophonic to polyphonic with pressure sensitive and weighted keys. Workstation keyboards and sequencers have progressed from single track recording to full auto-accompaniment including rhythm, adaptive chord following, onboard effects, hard drive recording, sampling, and even built in karaoke TV monitor support.

Availability in Educational Settings

Cost is a considerable factor in dealing with electronic music systems. For example, electronic keyboard costs range from ten dollars to thousands of dollars. It is an important goal of this research to determine what types of musical devices will be available in educational settings. MENC, the National Association for Music Education, publishes opportunity to learn standards that focus on equipment availability in classrooms. The year-1999 standards call for the presence of MIDI keyboards and video projector technology in Pre-Kindergarten classrooms on through secondary grade levels. Elementary and higher levels are encouraged to provide multiple MIDI keyboards for students. High schools are encouraged to provide alternative MIDI controllers including

wind, string, guitar, or drum devices. The MENC standards also call for an integration of technology into course materials and rehearsal settings [MENC].

Summary of Electronic Music Instruments

There are several music technology instruments available on the market including piano, guitar, percussion, wind instruments, and a variety of smaller custom controllers. The MIDI protocol enables the integration of these instruments with various audio generators. Music systems have evolved greatly in capability and realism since their inception. MIDI responsive video systems are becoming increasingly common in electronic music settings. Electronic music and video systems play a role in music production and are strongly encouraged for use in the music education process. These finding indicates that a synesthetic music communications system could be used by a variety of musical performers, it would be supported in most well equipped school musical classrooms, and the algorithms it uses could eventually be incorporated into existing MIDI-responsive video hardware.

Current Music Technology Research

Academic and corporate institutions develop algorithms, software, hardware, and interfaces for use in education and entertainment. Music technology researchers apply concepts of music theory, acoustics, human computer interaction and human perception. Many efforts seek to better understand how humans interact with music systems, to investigate new methods of training students, and to develop intelligent systems to improvise to enhance human performers abilities. A graduate-level course in music perception and cognition might address the topics of software tools for music research, grammatical and digital representations of music, beat detection, rhythm, meter, pitch recognition, melody, tonality, context, and expression [Chew05]. Here we consider some algorithms, visualizations, and impact studies associated with music technology. Many of

these algorithms can by leveraged to develop synesthetic visual transformations of real-time performances.

Algorithms

There are several algorithms in use in musical systems. Some algorithms involve the generation of sound. Sound synthesis is often accomplished by adding sound samples and waveform data together to create unique timbers. The addition of fundamental frequencies and overtones resembles the natural creation of harmonics in acoustic instruments. This research project does not focus on sound synthesis, rather, on the real-time analysis of sounds and determination of musical structure. Of specific interest to analysis of real-time music performances are Fourier analysis, beat detection, music transcription, improvisation, and visualization.

Fourier analysis is essentially the inverse of sound synthesis; the goal is to determine the frequencies that comprise a complex waveform. The output of this type of analysis is a set of instantaneous measurements of power spectrums, or intensities, over several frequency ranges.

Beat detection algorithms analyze the Fourier power spectra for peaks and compare adjacent band intensities. Beat detection has been used to extract tempo information from recordings, to synchronize analog sound sources with MIDI sequencers, and to animate graphical imagery [Goto].

Beat detection algorithms are very effective at identifying percussion sounds. For example, a sudden, strong, and staccato spike in the lowest frequency bands typically indicates a bass drum sound. A sudden mid-range tone spanning several frequency bands may indicate a snare drum. A short rhythmic pulsing of high frequencies may indicate a ride cymbal or hi-hat sound.

Advanced beat detection systems employ multiple sophisticated algorithms to extract musical information. Some tools compare power spectra values against libraries of common rhythmic patterns and sounds. Other systems use multiple customized

intelligent software agents that collaborate to perform complex analyses [Scheirer]. Wavelet analysis has also been employed in beat detection [Tzanetakis].

Automated transcription is an advancement of beat detection algorithms. These algorithms extract musical notation from an analog signal recording. The output musical notation should indicate the pitch, intensity, and duration of the notes in the recording. Transcription requires many more bands of frequency information, and it is difficult to extract several simultaneous tones. Recordings of instruments mixed with percussion are difficult to transcribe because the overtones created by percussion instruments cross multiple power spectra. Also the windowing technique employed in transcription can cause long duration sounds to be detected several times [Klapuri]. Another aspect of transcription is determining the home key, or root, of a composition. A key finding algorithm is described in [Krumhansl].

Modeling of musical style, the modeling the sequences of chords or notes in a composition, is another active area of research. One example would be modeling the chords in a jazz progression. There are some chords and progressions that are common to many jazz compositions. There are also several chord substitutions that can be applied to embellish compositions. Rhythm section players employ these substitutions when accompanying instrumentalists. Composers use the substitutions to add color or tension to a composition. Extracting and labeling the chords is meaningful information for the soloist and the ensemble members. Research has been conducted into developing a dynamic generative grammar for jazz chord sequences, basically enabling a computer to evolve a complex jazz progression from a simple one [Steedman84][Steedman96].

Representations of musical style are useful for storing and retrieving songs from a database. One method converts song sections into a musical texture. Two examples of measurable musical textures are the chords and instrumentation of the piece plotted with respect to time [Aucouturier]. Another principle for storing musical textures is Zipf's law. Zipf's law "concerns the frequency of use of different words" in written language. This law is applicable to the musical tones and chords in compositions. Encoding the occurrence and repetition of music tones in a composition creates a unique graph that reveals characteristics of the song [Zanette].

Incremental Parsing is an alternative method of stylistic mapping. This method constructs data tree graphs from a sequence of musical notes. The trees contain the shortest paths necessary to represent a musical piece. The nodes in the tree contain probabilities indicating the likelihood of each subsequent node being traversed. Prediction Suffix Trees are another method of probabilistic modeling. These techniques have been applied in developing a real-time improvisation system titled 'The Continuator' [Lartillot].

The Continuator is an interesting example of an intelligent musical system. The Continuator learns a person's musical style by 'listening' to them play on a MIDI device. Continuator monitors MIDI events and constructs a model of the note sequences and probabilities. Continuator can then improvise around the learned musical style or transpose that musical style based on a second player's input. The Continuator's improvisations inspire new ideas because they are permutations on a learned style. The author also reports that children engage well with the continuator system because its continuations are similar to the their random explorations [Pachet].

Koan, produced by SSEYO, is an example of a *generative music system*. Koan produces unique compositions based on musical rule sets. Users create the rule sets through an intuitive interface. Several parameters can be defined including instrumentation, musical scales, note duration, intervals between notes, panning, audio effects, etc. The resulting compositions are randomly generated explorations of the rule space and identical rule sets result in different compositions each time they are played. Koan's output can range from sparse single tones up to a dense cacophony of sound. Koan can be a source for many improvisational ideas and has been used in professional musical releases. A generative music performance using Koan in the Unidentified Wailing Mammals project is described in the first chapter as part of this author's motivation towards music visualization [SSEYO].

Iannis Xenakis pioneered generative music systems with his research from 1950-1990. His research explored the construction of algorithmic sound by mathematically mutating short sound samples into large-scale musical compositions. His diverse background in civil engineering, architecture, and music provided him with knowledge of

several mathematical operations and musical principles. Hs works combined these areas to create instrumental, electro-acoustic, and computerized musical compositions. Xenakis applied a mixture of deterministic and non-deterministic operators to audio data sources including concepts found in set theory, probability and calculus [Mode]. His acoustic audio sources were as diverse as burning charcoal, airplanes, automobiles, boxes, and nature sounds [EMFmedia]. Xenakis's book, Formalized Music, presents stochastic methods for converting visual sketches into compositions, plotting audio frequencies into vector and volume spaces, and synthesizing compositions from probability and state transition diagrams. Formalized Music describes the use of symbolic logic, algebra, and matrix operations to create and represent compositions. Composition using micro-sounds is also explored in the text [Xenakis71].

Xenakis's work spurred philosophical discussion about the nature of creativity, the history of music, and the interrelationship between arts and the sciences [Bois]. His varied composition approaches inspired the development of a set of fundamental phases for a musical work. Xenakis' stages are paraphrased below [Xenakis71][Harley]:

- **Initial Conceptions**: collecting initial ideas, intuitions, data, and possibly sketches for a composition.
- **Definition of Sonic Entities**: selection of acoustic instruments, electronic sounds, methods of ordering, and assembly.
- **Definition of the Transformations**: determination of how the sonic entities will change over time.
- **Micro Composition**: identifying the relationship between the sonic entities.
- **Sequential Programming**: choosing the schema and patterns of the musical work.
- **Implementation of Calculations**: these calculations cause the animation of the schema.
- **Final Symbolic Result**: conversion of the calculated output into notation, graphs, or other musical representation.
- **Sonic Realization**: performance or synthesis of the musical work resulting in an observable sonic experience.

Xenakis designed the computer program GENDYN as one tool for developing unique algorithmic-based music. With this program, users were able to experience the autonomous evolution of music and explore combinations that reached beyond the limitations of the composer's mind. Philosophically, Xenakis saw creativity as a moment of transcendence where deeply rooted artistic inspiration overcame the limitations and insecurities of the author's ego. A computer program, algorithmically programmed and configured by a human, would generate possibilities that the author's ego or expectations would not normally explore. Xenakis drew inspiration from these computer-generated works, which were used as base materials for other compositions. His editing techniques involved sampling and re-arranging snippets of generative composition [Pape].

Xenakis also explored the notion that true musical creativity of free and abstract forms with computers required a cross-disciplinary skill set he titled the artist-conceiver. Artist-conceivers would require a diverse skill set with training in aesthetic appreciation and training in theoretical ideas. He envisioned a fundamental overlap between the domains of music, mathematics, science, engineering, architecture, biology, and genetics. Xenakis explored this overlap by developing computer applications to convert visual imagery into audio compositions [Xenakis]. Xenakis' music/visual systems are explored later in this chapter. His works also included multimedia presentations called Polytopes that combined sound, light, movement, and architecture [Harley][Matossian]. The Polytopes works are detailed in the chapter on Synesthesia.

 Author David Cope's work on computer synthesis of musical style is a more contemporary reference for these synthesis techniques. His book *Virtual Music: Computer Synthesis of Musical Style* discusses such topics as the composition of style specific music, the importance of patterns, and the impact of structure. The four sections of the book cover his work in musical intelligence and the prior related efforts, the operation of his sophisticated multi-component composition software, expert critique of the generated works, and Cope's consideration of those critiques [Cope].

Grammar Checking

Palestrina Pal is an example of a grammar checker for musical compositions. This application demonstrates the potential of technology to teach a student a specific style of composition. Palestrina Pal compares users' compositions to counterpoint composition rules, enabling students to develop compositions while automatically highlighting rule violations. Users can interactively correct rule violations over time to arrive at a conformant composition. The intent of this system is to free the student to consider the 'higher level issues of musical aesthetics' while sketching ideas [Huang].

Visualization

MuSA.RT (Music on the Spiral Array – Real Time) is a real-time visual transformation of musical information. While music sequencing and composing applications utilize notation and piano roll displays, MuSA.RT employs a higher-level music theory analysis to develop a custom visualization space. MuSA.RT uses a Spiral Array Model to combine real-time detection of chord roots with real-time determination of chord structures. Major and minor triads are presented visually as triangles connecting pitches that are arranged in a spiraling display. Chord types are distinguished by colors and users have the ability to rotate the arrays to change their camera view. MuSA.RT is summarized as an "interactive environment for navigating through metaphorical tonal space" [Chew03].

Another visualization system is The 21st Century Virtual Reality Color Organ. This project, translates 'musical compositions into visual performance.' It is designed for large-scale immersive environments. The visual system consists of a system for converting aural information into a 3D visualization component and a programming environment to support system wide interaction. This system imposes a 12-step color wheel over the circle of fifths. This results in a color mapping where harmonically related keys share a similar coloring. This system also incorporates a timbre-based coloring, by mapping families of instruments to related colors. The system generates 2D colored and

textured planes corresponding to MIDI data files of compositions. The resulting output is a mythical landscape view. The Virtual Reality Color Organ research includes a goal of projecting visualization of multiple distributed musicians [Ox].

A peripheral example of music research is found in visualization for psychological research. Music has been integrated with virtual reality to facilitate rehabilitation of Alzheimer's disease. In one case, music was used to stimulate curiosity and capture the subject's attention. Anti-stress music is employed to relax the listener and reduce their defensive reactions to VR technology [Optale].

Interface/Impact Studies

Two areas of HCI research in music technology are studies of interface effectiveness and the impact of software tools on learning. This section presets research examples in each of those two areas. Several more examples of music interface research can be found in the proceedings of the International Computer Music Conference [ICMC], IEEE Multimedia [IEEE-MM], Computer Music Journal [CMJ], and related publications.

An example of an interface effectiveness study is analyzing custom musical instruments. One example, the SensOrg is a "modular assembly of input/output devices and musical software." The system presents several interface modules mounted on positionable gimbals. SensOrg devices include musical keyboards, sliders, flexi-pads, and buttons. The study measures the usability of the system along four dimensions:

- Learnability
- Ease of use
- Flexibility
- Users' attitudes

The various control surfaces in the SensOrg are categorized according to function and type of feedback. User movements are measured for input task. An analysis of the users' efforts at various tasks and their reported experience of the musical creation process are given [Ungvary].

The impact of technology on music learning is the subject of a study of 59 Malaysian music teachers in training [Abdullah]. The teachers were pre-tested for their attitudes and perceptions of the potential of music technology in education. Subjects were also tested on their ability to write musical notation. The teachers were given a nine-hour course on the Finale music notation program. Subject's abilities were re-tested after the course. Experimenters reported a significant improvement in the teachers' music writing and computer literacy. Subjects also showed an improvement in their attitudes towards music technology [Abdullah].

Another interface study concerns a custom application for drawing music into a colorful user interface. Hyperscore, a novel music composition program from MIT Media Labs, lets users to paint melodies onto a graphical canvas. Each melody is treated as a musical motive that can be copied and arranged onto a larger score. Motives can be scaled and transposed in this system just like imagery in a photo application. Hyperscore ensures that the resulting melodies will be pleasing. The system automatically checks the motives for continuity and the resulting compositions maintain a sense of musical cohesiveness [Farbood].

Studies of children interacting with Hyperscore have found that they engage in "musical processes commonly associated with adult or professional composers." One 10-year-old subject reportedly "engaged with musical rudiments such as note value, pulse, rhythm, and rhythmic hierarchies, pitch, melody, and texture." Also, the composition methods accomplished by using Hyperscore's painting interfaces were far beyond the subject's notation skills and understanding of music theory [Jennings].

A similar study reports the activities of a pair of "musically untrained college students" working with music software tools. The students demonstrated the ability to develop "archetypal tonal melodies" using software tools. Participants were given a set of "ambiguous melodic materials" and were able to modify them into "coherently structured tonal melodies." An analysis of their works revealed an ability to develop balanced phrases, clearly articulated phrase boundaries, harmonic structure, and the ability to resolve cadences (musical phrases).

This study also pointed out some educational implications of student access to such tools. Musically untrained students already have some amount of intuitive musical knowledge. A curriculum (or software tool) for expanding this knowledge should present musical information in a manner consistent with student's understandings. This means the system should point out motives, figures, and phrases. Students should be able to manipulate music at each level of understanding. There should be a variety of representations including "multiple sensory modalities, multiple graphics, and multiple levels of musical structure". Also students should be able to employ methods that help convert their intuitive knowledge into explicit knowledge [Bamberger].

Group-based Musical Applications

These applications are designed to enable musical interaction among multiple individuals. Computer graphics, software algorithms, and custom hardware interfaces are combined to create unique forms of expression. Some of these systems enable novice performers to interact with musical concepts and to 'jam along with' professional musicians. Other systems are designed to facilitate musical rehearsals and lessons. A few representative systems are presented in this section. Additionally, guidelines for multi-art integration projects and a list of measurable characteristics for hybrid systems are given. Also, summaries of some representative applications are given.

A few key principles relevant to multi-art applications can be found in philosophy of music education texts. Multi-art applications have the potential to make each art clearer in that the contrasts between arts reveal the uniqueness of each art. The combination can clarify the underlying principles of each art. Transforming or interrelating art forms can amplify the concepts in each form. Multi-art systems can reveal the interconnectedness of all art forms as a family of expressive methods [Reimer].

Some cautions are also given for these hybrid systems. A multi-art project should not submerge each art by focusing on weaknesses of individual arts. Developers should not over emphasize surface similarities between the arts. The system should not neglect the perceptive process of each art form by reaching for generalized correlations. Finally,

individuals should not employ non-artistic principles to combine the arts [Reimer]. These principles and cautions are designed to preserve the uniqueness of each art form and ensure the meaningfulness of the individual pieces.

Blaine presents a very detailed comparison of multi-art applications in "Contexts for Musical Expression" [Blaine03]. Eighteen systems are characterized according to their interactive aspects. These aspects include "focus, location, media, scalability, player interaction, musical range, physical interface, directed interaction, pathway to expert performance, and level of physicality." These terms are described individually below.

- **Focus** is a measure of the communications in the system. Communications can be between the players or between the players and the audience.

- **Location** is whether performers and the audience are co-located or distributed. In co-location systems, the performers and audience can observe each other's interaction.

- **Media** is the amount and type of audio and visual imagery in the system.

- **Scalability** is the number of participants and observers the system can support.

- **Player Interaction** is the type and number of interactive inputs the performers provide.

- **Musical Range/Notes** measures the number and type of musical tones or sounds that players can generate. Some systems allow users to trigger pre-canned phrases, loops, or melodies. Others give a full range of expressive control.

- **Physical Interfaces/Sensors** are the control mechanisms users are able to use in the system. These can include MIDI systems, computer peripherals, and a variety of custom devices.

- **Directed Interaction** is a measure of type of "group dynamics and social interplay" the system affords. Some systems guide or constrain the performers actions to specific tasks. Other systems are un-directed and support exploratory free play.

- **Pathway to Expert Performance** is a description of the methods, potential, and effort required for a person to master the full expressive potential of the system.
- **Level of Physicality** is a measure of the effort required by the performers. One system can provide several different roles and individual levels of effort.

Each interactive system reviewed by Blaine has a varying level of detail of interaction, yielding unique potential for entertainment and education. The criteria given above establish a framework for considering the experience of novice and expert musicians [Blaine03]. The following applications, Jam-O-World, Augmented Groove, and Beatbugs, provide experiences for a wide range of persons.

Jam-O-World, developed at Carnegie Mellon University's Entertainment Technology Center, is a multi-player world for physical and social musically oriented interaction. The system's predecessor, Jam-O-Drum, embeds multiple percussion triggers into a custom table. A computer graphics simulation is projected on top of the table and each playing station has a localized audio signal. Each performer stands at a location around the table and plays music by striking the tabletop with their hands. The computer simulation animates graphical objects that engage the players in call-and-response and improvisation methods. Users are able to see representations of their notes and also observe system instructions. Jam-O-World is a site-specific installation of advanced Jam-O-Drum tables with turn-able drum pads and increasingly complex visual simulations. The walls of the performance area become part of the experience [Blaine02].

The Augmented Groove system incorporates Augmented Reality, Head Mounted Displays (HMD), and glyph-labeled cards. Users can position the cards to play music and cause changes in musical elements including timbre, pitch, rhythm, and reverb. 3D virtual images are projected onto the surface of the cards in the augmented world. The images change to reflect the state of the music being generated [Poupyrev].

Beatbugs are hand-held custom music interfaces that "record live musical input from MIDI and acoustic instruments." Beatbug players can "transform the recorded materials (samples) in real-time". These palm-sized devices feature a top-mounted

piezoelectric sensor, two antennae shaped flex-sensors, and LED lights. Players can slap the beat bugs to initiate recording, to transfer sounds to other players, and to trigger playback of audio samples. Performers can manipulate the samples by bending the antennae. LED lights indicate the Beatbug's status during the performance. This system works with a MIDI interface, music generation software, recording hardware, and multi-channel audio to provide players with a form of musical interaction. They have been used in live performance settings with a traditional jazz ensemble and compositions have been written specifically for the systems [Weinberg].

Rehearsal Assistance

An example of research to support ensemble rehearsal is the Automated Choir Rehearsal project at Georgia Tech University. This work in progress introduces voice recognition and a projection of a musical score into a choral rehearsal. This system will follow the conductor's vocal commands and the group's performance through the musical score to project the appropriate section of the musical notation in front of the ensemble [GATech].

Another project, Hyperbow, monitors the characteristics of a violin bow during a performance. The system measures bow position and acceleration similarly to the Hypercello [Paradiso] project, but it also incorporates measurements of the downward and lateral strains on the bow stick. The strain forces affect the experience of playing and the sound of the string. The system "enables players (and instructors) to view, inspect, and exploit the subtlety of their bowing" [Young]. Hyperbow and Hypercello are examples of systems to increase the learner's awareness of their performance and to improve their rehearsal process.

Summary of Current Music Technology Research

Music technology research is an active area and a source of information for designing group-oriented musical applications like Synesthetic Musical Experience Communicator. Several algorithms and research areas have been explored in this section.

To review, beat detection and music transcription provide good methods for analyzing the real-time characteristics of audio. Grammatical representations and improvisational applications provide knowledge of how computers represent music digitally. Improvisation and generative music applications enable musical experimentation and give examples of how to configure complex musical parameter options. The interface impact studies are useful in understanding how music-learning systems are evaluated. The impact studies also reveal the teaching potential of learning systems. Visualization applications demonstrate methods of converting musical information into immersive graphical spaces. The Group-based musical applications present systems for large group interaction and entertainment. Several criteria for comparing interactive music applications were given. These criteria are useful in considering how the system's designs impact the audience and a wide range of performers. Finally, the rehearsal assistance tools attempt to improve or augment the rehearsal settings by automating score following and providing additional information to the performers.

Summary of Music Technology

This music technology section has considered the historical impact of music technology in providing individuals with access to performances and musical knowledge. Interactive media and software applications provide a vast array of knowledge about several musical techniques and concepts. These tools suit a variety of experience levels. Music software production tools allow individuals to experiment with sounds, build compositions, create notation, and perform high quality digital music. Technology has been used to make music, and concepts more accessible.

Electronic Music Instruments are available to simulate percussion, string, wind, and keyboard instruments. There are also a variety of small sensors that can be used to create custom instruments. MIDI capability provides convenient interconnections between digital music systems and the software tools enable creation of custom musical applications. Recent developments in video hardware and multi-media software enable musical instruments to directly impact visual presentations. Also, education standards committees call for the presence of MIDI and video projection devices in all levels of formalized education. The quality and variety of MIDI capable devices is steadily increasing. Advances in digital music hardware reflect a fusion of musical and visual experiences.

A survey of music technology research indicates several active research endeavors into analysis algorithms, representation, visualization, group applications, and rehearsal assistance tools. Studies of music technology's impact on individual music learning have validated the call for increased usage in education. The group applications and visualizations demonstrate current uses and future directions for music technology research. Group education and entertainment applications are clearly a part of that movement. These applications also seek to provide users with intuitive access to musical concepts.

Chapter Summary

Music is a lifelong artistic study, a form of communication, and a deeply appreciable means of self-expression. Music education has evolved from a behavioral based approach, focusing primarily on proper technique, into a larger cognitive-based paradigm. With cognition, comes the challenge of developing musical awareness in the minds of students, especially those who do not actively perform on a musical instrument.

There are numerous musical concepts, several ways of understanding music. Some concepts like scales and rhythms have mathematical foundations. Sounds and their combinations cause emotional responses in individuals. Compositions are organized as temporal experiences and are analyzable through melodic, harmonic, rhythmic, and

textural means. There are extensive nuances in rhythmic, vocal, and instrumental performance. The methods of teaching music range from full body movement to silent mental exercise.

Music technology provides a means to experiment with sound. The software tools provide access to musical knowledge. Electronic interfaces are available for many musical instruments. Multi-art systems enable humans to interact with music in auditory, visual, and kinesthetic means. Further, music technology can enable the development of improvisational performance software, visual presentations, and rehearsal aids. There are several examples in the literature that demonstrate the benefits of such systems in teaching musical concepts, enabling musical interaction between experts and novices, and augmenting musical performances.

CHAPTER 4. SYNESTHESIA

"Did you know that every number has a color?" I asked a friend during a dinner conversation of several odds and ends in April of 2001. That friend was pursuing a Doctorate degree in Religious studies. She asked me if I had heard about Synesthesia. She performed a Google search and the subsequent conversations opened a new door for self-understanding. That conversation about sound and synesthesia prompted the initial research of this dissertation.

This chapter discusses a most peculiar perceptual phenomenon. *Synesthesia* is defined as an involuntary joined sensory experience. Individuals with synesthesia experience a variety of cross-sensory stimulations, often reported as internally projected mental visual imagery. In the author's case, digits and letters have both a true (perceived) color and an internally experienced coloring. Touch, temperature, scents, and sounds bring color perceptions to mind. Musical compositions are automatically and consistently transformed into internally projected visual experiences. As we will see below, these projections serve as meaningful performance and creativity aids for synesthesian musicians and artists.

Synesthesia is a surprisingly old research topic with publications in the 1800s. There are personal reports of synesthesia and documented artistic contributions dating back to the early Romantic period. This chapter reviews some very informative literature on this curious cognitive experience. Several aspects are presented beginning with the various forms of synesthesia and some experimental observations. Some very divergent historical perspectives are given, followed by contemporary theories on its basis and relevance. Applications of synesthesia in music and visual arts are described.

This chapter ends by presenting a set of sensory correspondences. Characteristics of synesthete audio-visual experiences are reported and compared to non-synesthete perceptions of audio and visual arts. Sound symbolism, the communication of meaning through sounds, is presented as an analogous experience to synesthesia. Finally, the basic perceptual elements of human perception in audio and visual art forms are enumerated.

These correspondences will form the core of the musical-artistic transformations in the Synesthetic Music Experience Communicator.

4.1 General Overview

There are several forms of synesthesia and no two individuals share exactly the same experience. Some of the pairings of senses and the associated phenomena have been identified and defined. The first key definition is *photism* [Harrison]. Photism is the term for the internal visual imagery that Synesthetes experience. Photisms are also referred to as concurrents [Grossenbacher] or chromatisms [Dann]. Chromatisms are colored photisms experienced in response to stimulus. Photisms can be triggered from stimuli such as numbers, letters [Dixon], pitch, temperature, scents, pressure, and more. I.H. Coriat reported cases of colored hearing and colored pain in 1913 and H.S. Langfield reported "a chromaesthetic woman musician" [Dann].

Some synesthetes report experiences of movement in parallel with temporal and spatial cognitions [Duffy]. For example, thinking of the time of day or day of the year becomes a mental journey through an internal calendar space [Cytowic]. These conceptual spatial structures, also called shape- or number-forms, occur in response to enumerated sets [Grossenbacher]. Some number forms can include ladder-like spatial arrangements of year/decade/century concepts [Cytowic]. Synesthetes have report color-coding for sports teams or social-cultural conventions.

Synesthetes report constancy in their photisms. They are neither learned nor repressible. They occur in response to external stimuli. Figure 4.1 depicts Külver's shape visual forms which have been reported in response to synesthesia, hallucinations, and other cross-modal associations [Cytowic]. To a synesthete, photisms are an enjoyable part of life and a means of emotional connection to their outside world. They are quite literally the essence of thoughts [Duffy]. To borrow a metaphor from HCI, Synesthesia is a sort of inner virtual reality, the exception is that there is no software programming involved.
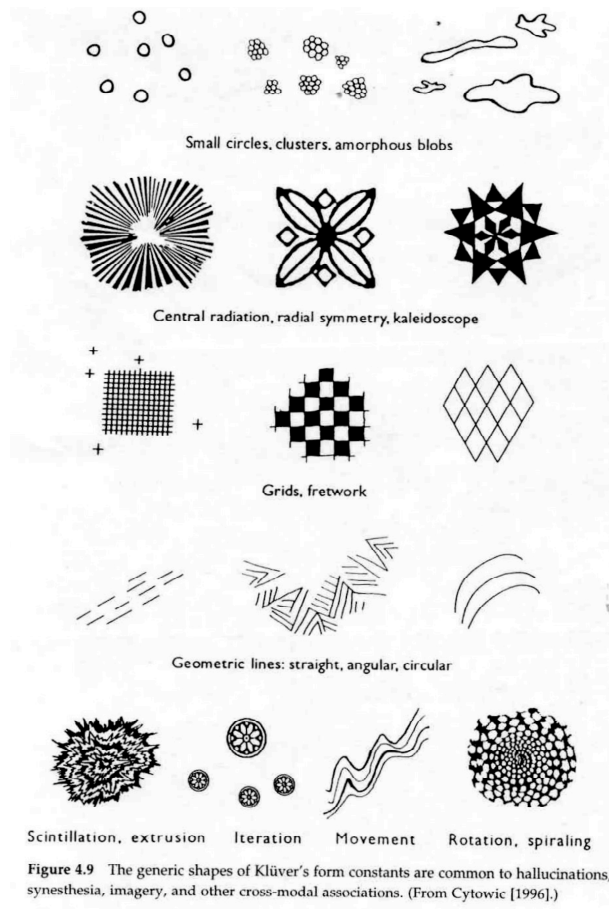
Figure 4.1 Synesthetic Photisms [in Cytowic].

This author's synesthetic color-number pairings are shown in Figure 4.2. The color number pairings are also given in table 4.1 as text.
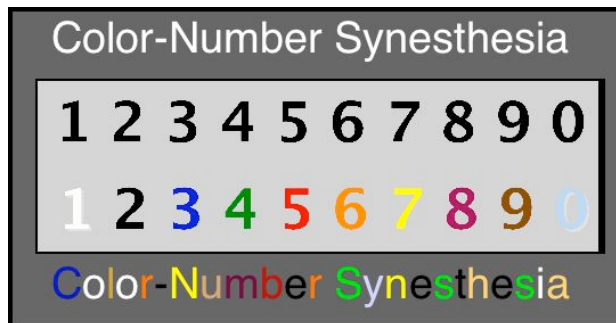

Figure 4.1 The Author's Color-Grapheme Synesthesia

The diagram above presents a normally colored section of text and a number line on the top half of the image. The bottom half indicates the internal cognitive experience of those number forms. These color impressions are perceived as overlapped on top of written text, regardless to the background color of the text. The impressions also occur with internal recollection of text and in response to verbal stimuli.

Table 4.1 Author's Color-Grapheme Synesthesia Pairings as Text

Number	Color	Number	Color
1	White	6	Orange
2	Black	7	Yellow
3	Blue	8	Maroon
4	Green	9	Purple or Brown
5	Red	0	Clear

These synesthetic color pairings have existed for most of this author's life. Just as many authors' of texts on Synesthesia have reported, this author recalls awareness of the phenomena from early childhood. Like other synesthetes, this author did not realize other individuals did not experience this phenomenon until later during adolescent discussions.

4.2 Experimental Research

The American Synesthesia Association holds yearly conferences in the area of Synesthesia, including descriptions, testing methods, artistic implications, and studies [ASA]. There is a growing community of synesthetes with a goal of understanding and quantifying this experience. Some of the experimental methods and observations are given here.

The synesthesia research center at the University of Waterloo, Canada conducts an online survey of color-grapheme synesthesia. Participants select color-grapheme pairings through a web interface [SRG]. Individuals are instructed to repeat their

selections twice for the entire alphabet and number line. Users are also asked to indicate their confidence in the consistency of their selections between rounds. This author's survey results, analyzed by Dr. Phil Merikle [SRG][Dixon], indicate definite presence of color grapheme synesthesia.

Consistency and automaticity are two important measures of Synesthesia. Consistency among individual subject's experiences is necessary because the experimental process relies upon synesthetes' first person accounts combined with measured responses. Consistency supports a non-associative basis for the experience. In other words, individuals are not reporting memorized parings, but rather an involuntary experience.

Automaticity tests measure the time required for synesthetes to identify alphanumerically labeled and colored cards. Synesthetes identify graphemes that are colored congruently with their synesthesia faster than incongruently colored graphemes. The recognition delay caused by incongruently colored cards is considered a form of cognitive dissonance that is not found in non-synesthetes. Additional tests indicate that color-grapheme coloring can be of benefit in identifying [Smilek02a][Palmeri] and memorizing digits [Smilek02b].

An interesting twist on color-number determination is the construction of a large digit from smaller digits. In one study, a large number 5 was constructed from several small 2's. If the subject attended to the local information, the small 2's took on the digit 2's synesthetic color. If the global information, the shape of the number 5 was considered then 5's synesthetic color was perceived [Palmeri].

Measuring the effect of sequential synesthetic stimuli of similar and varying sensory modalities also indicates the presence of synesthetic priming effects. The interpretation of the analysis is complicated by several factors. There appears to be positive association between the experience of a synesthetic photism from an alphanumeric color and the naming of subsequent color patches. Also, negative priming occurs when synesthetically incongruent stimuli are given [Smilek02a][Smilek01].

Studies of the perceptual aspects of synesthesia involve the use of PET scans, functional-MRI, localization tests, and consideration of neural pathways [Smilek01]

[Grossenbacher][Ramachandran]. These studies reveal that synesthesia is not merely a learned association. The research appears split as to whether synesthetic binding occurs before or after awareness of the alphanumeric character. Research also indicates the presence of conceptually driven synesthesia. In situations where an audio (non-visual) stimulus is given, individuals report synesthetic experiences in response to hearing or thinking about words or certain stimuli [Smilek02a].

Research indicates that there may be a genetic component to synesthesia. A study of inheritance and discordant monozygotic twins lead to an interesting insight. One twin experienced photisms for digits while the other did not. Their identical genotype, but incongruent synesthesian expression lead researchers to suspect the cause was an epigenetic event, X chromosome inactivation, or mutation of some key synesthetic gene.

Studies of late onset synesthesia suggest other physiological causes as in the case of acquired synesthesia associated with retinitis pigmentosa. In this study of RP, the patient progressively lost his vision from childhood to age 40. He developed synesthesia as his vision degenerated. In this case the authors proposed three causes. The first theory was the person's brain had begun remapping or areas began cross talking after the loss of vision. Second, the sensation of touch triggered associative visual/tactile memories that were normally not perceived by activity of the visual pathways. Third, the loss of visual stimuli lead to strengthened 'back projections' from somatosensory areas of the brain [Armel].

The causes and reports on synesthesia are as diverse as the experiences itself. Subsequent sections of this chapter give a brief synopsis of the suspected causes of synesthesia. The next section gives an interesting historical look at the wide-ranging interpretations of synesthetic experiences.

4.3 Diverse Perspectives

There are numerous diverse perspectives on the causes and meaning of synesthesia. This section presents a number of these views to provide the reader with a sense of the breadth of research that has been conducted in this area. No single view has

been adopted as perfect or correct. Rather, as science has progressed and greater knowledge has been gained about humans, theories about synesthesia have been refined and re-examined in light of each insight.

In early publications (pre 1900s), synesthesia has been examined from a mystical view. It has been considered a transcendental form of sensory representation. A Theosophical view, founded on the basis of a united view of science, philosophy, and religion, held synesthesia to be an escape from humanity's imprisonment in the five senses. Thus, it was viewed as a libratory form of perception. Occult views held synesthesia to be an intersection of vibrations relating the physical and psychic worlds. Freud's view seemed to imply that synesthesia was an early stage of schizophrenia. It has also been hypothesized as a sympathetic resonance between the senses [Dann],

Coriat and Langfield (1913) offered one of the first physiological explanations, referred to as an "irradiation, neural, sensory reflex theory." It proposes neural crossover and interconnects between adjacent sensory areas. This theory also implies a fundamental interrelation of sensory modes. Lawrence Marks' presents a review of thirty-four studies into correlations between audio-pitch and visual-brightness conducted between 1920 and 1975. One-third of those studies indicated a positive correlation between pitch and brightness, though many subjects were indicating the presence of their mental associations betweens sounds and colors [Dann].

Dr. Gail Martino and Dr. Lawrence Marks (1999) conducted further research on non-synesthete brightness-sound mappings. They found systematic matches between loud sounds and bright colors, low sounds and soft lights. Also they mapped correlations between bright colors and loud sounds, soft colors and low sounds. Martino and Marks concluded that these cross-sensory correspondences were meaningful to all people [Duffy].

Another test of these correspondences was to measure the amount of time people required to play tones on a piano keyboard in response to an audio stimulus. In this experiment, participants were given an audio stimulus (high or low pitched tone) and a visual stimulus (white or black color flashes). Four stimulus conditions were identified: two matched pairings and two unmatched pairings. In the matching stimulus condition, a

white color flash was accompanied with a high tone or a black flash with a low tone. In the unmatched condition, high tones were paired with black flashes or low tones paired with a white flash. Participants were instructed to ignore the visual flash, and play a note on the keyboard solely based on the tone. The experiment found that participants required more time to respond to mismatched stimuli than to matched stimuli, whether the focus stimulus was audio or visual. The results were considered strong evidence for a weak form of synesthesia in non-synesthetes [Duffy].

German psychologist Erich von Hornbostel's work sought to corroborate a unity of senses through studies of scent and color correlations, as well as sound and tactile sensation. Hornbostel offered evidence of sensorial unity in the language of an African tribe. This language had only one word to describe the sensory experiences of "hearing, taste, smell, and touch." Also, the French and Spanish verb 'sentir' (to feel), is noted because it likewise relates to perception in multiple senses. Hornbostel's exploration of sensory unity also considered primitive (prewritten or pre-spoken language) thought. Primitive thoughts were viewed as naturally synesthetic because there were no conceptual terms to distinguish experiences [Dann]. Werner (1943) maintained a theory of a fundamental union of the senses. The hypothesis is that sensory stimulation is first synesthetically experienced and then differentiated. Examples of this type of differentiation can include cognitive processes, correlations with other experiences, and language-communications [Duffy].

Gestalt psychologist Kohler (1947) investigated links between language and images. In one study, participants were instructed to pair the nonsense words 'maluma' and 'takete' with either a rounded figure or a sharply angular figure. The results indicated a strong preference for 'maluma' as the rounded figure and 'takete' as the sharp edged figure [Dann]. This study is similar to the 'kiki' and 'bouba' shape figure research cited by contemporary researchers Ramachandran and Hubbard. In their study participants reported 98% correlation between the names 'kiki' and a sharp edged figure and 'bouba' with a round edged figure [Ramachandran]. This result is also reinforced in literature on sound symbolism, the communication of meaning through sound, a topic that is discussed later in this chapter. True synesthetic language mappings are a bit more abstract.

Wheeler and Cutsforth, a blind synesthete, proposed synesthesia to be a form of thinking or inner speech. This work considered photisms to be unattended mental processes. If a sensory experience was indeed a gestalt, a unified whole of all sensations, then photisms were an interpretation of that experience. Wheeler and Cutsforth maintained that thought was a kinesthetic activity, a function of 'cognitive movement'. In this view, thinking was a form of mental translation through sensory and conceptual spaces. This view espoused that photisms were necessary for thought. More so, photisms were the essence of thought as well as a mechanism for meaning and reasoning. Edward Bradford Titchener, German psychologist and pioneer of introspectionist methods, maintained that all mental processes were based on imagery [Dann].

The concept of synesthetic mental imagery is given support by research into eidetic images. Eidetic individuals are able to conjure 'normal, subjective visual images' inside the mind [Duffy]. Eidetic images are projections of sensory experiences, including hearing, taste, smell, and touch. Eidetics consider the projections as real feeling, beneficial, and comparatively vivid as real experiences. Erich Jaensch's publications of the subject (1911-1930) propose it as a primitive cognitive function. This phenomenon is present from childhood in 2-15% of individuals and persists until adulthood in varying degrees of ability [Dann].

Eidetic abilities fall into three main classifications: *integrate*, *disintegrate*, and *synesthete*. Integrate (B-type) eidetics can visualize almost any concept they think of and can mentally transform properties such as location, color, form, and so on. Disintegrate (T-type) eidetics experiences are more like images. These images are less distinct, and disintegrate eidetics lack the ability to transform the visualizations [Dann].

Jaensch's categorization also included a synesthetic (S-Type). S-types experience a high amount of empathy and perceptual union with the objects. Synesthete eidetics 'feel' an emotional response to their mental visualizations. Jaensch gives a caution that vividness of eidetic and synesthete experiences may lead individuals to prefer the imagined world as an escape from reality. This is due to a variation in the emotional experience of the mental imagery. The integrative eidetics' emotional context for

situations is mostly driven by external circumstances. In contrast, the synesthetic eidetic impresses their own emotional tone on their mental images [Dann].

Jaensch also credits eidetic and synesthetic thought for abstract geometric thinking applied in the development of modern science. This 'photisms as thought' view was the subject of a recent study of synesthesian autistic savant's mathematical skills. Daniel Tammet, age 26, reports mathematical operations as transformations of colorful landscapes and moving shapes [Abc]. Given the task of multiplying two three-digit numbers, Tammet reports that he does not 'calculate' as much as he instantly arrives at an answer through a geometric transformation. Each of the two original numbers becomes a shape in his mind. The two shapes modulate and a third shape emerges. This shape is the synesthetically encoded answer [Guardian]. Other research into musical autism, lightning calculators, calendar calculation, weather forecast recall, and phone numbers can be found in [Dann].

A recent publication by Jose Arguelles, titled 'Earth Ascending: An Illustrated Treatise on the Law Governing Whole Systems (1988)', focuses on a resonant field explanation. In this book, Arguelles proposes that our reality is primarily rooted in resonance and vibration. Arguelles wrote that materialistically dominated views have helped to contribute to a civilization based on materialism and war. In contrast, a resonant field approach to society will lead to a state of harmony and towards 'a multidimensional metaphysical potential.' He holds that synesthesia is key to this paradigm shift. In the synesthetic experience, 'the sense fields interpenetrate and act in harmony with each other.' Arguelles also cites an aboriginal period of early human history where sensory experiences were 'fused in blissful unity.' The introduction of language, focus on abstract concepts and thoughts, and the divorce of art and science have had detrimental effects. Among those effects is "a splintering of knowledge", "the loss of sacred view," and "the splintering of knowledge" [Dann].

Arguelles predicted a holonomic age for society including synesthetic landscapes of computer-assisted light and sound art. This view is similar to romantic poet Charles Baudelaire's, French symbolist Arthur Rimbaud's, artist Kandinsky's, and composer Scriabin's visions that sought to link the senses in a transcendental experience. Some

historical efforts in that direction have included A.B. Klein, who in 1921 designed a color projector to support color music. Klein valued color music as equally important as the hope for religious and philosophical utopia [Dann]. Numerous applications of light and sound art have been published. Leonardo, a journal on tech and visual art is a good source for publications about visual music, kinesthetic art, and synesthetic art [Leonardo].

The last view discussed here relates to synesthesia in infants. One perspective begins by considering the synesthetic unity of the womb. Pre-natal awareness develops in the uterus, an aqueous environment with a typically stable temperature, damped lighting, and muted sounds. Upon birth, human development proceeds through expansion, growth, and experimentation with our environment. Eventually humans experience fulfillment and greater realization of self [Dann]. David Lewkowciz and Gerry Turkewitz's research indicates that babies are born with a common synesthetic sense. This is evidenced by the fact that one-month old infants experienced the same heart rate changes for flashes of light and bursts of sound. Daphne Maurer indicates that before the age of four months, babies' senses have not yet fully differentiated and that their world is an intermingled state of sensory confusion. All of the senses are present and experienced, but not distinguished. Dr Lawrence Marks, Yale University, indicates that one in two thousand individuals will retain partial cross-modal sensory activity as synesthetes [Duffy].

4.4 Overview of Contemporary Theories

The preceding section presented a montage of historical components and perspectives of synesthesia. This section presents all current perspectives on synesthesia in a condensed overview. Six viewpoints about the neural and physiological basis of synesthesia are given. Four classifications of synesthesia are given and two developmental theories are given.

Physiological Basis

Several differing theories about the basis and cause of synesthesia are given below. Each theory is supported by a diverse collection of research. No single theory is held as definitively true. Rather, each theory gives some insight into the operation and function of the synesthetic experiences.

Learned association theories speculate that synesthetes are reporting memorized pairings. The theory proposes that synesthetes' associations were learned at an early age level and that synesthetes they simply do not recall when the learning took place. Recollection of those pairings is theorized to occur on a subconscious level. This is not entirely supported by recent research that indicates physiological differences in synesthete and non-synesthete brains [Harrison].

Sensory leakage theory proposes that synesthesia is a crossover, or muddling, of sensory information that is processed in different regions of the brain. Much of the support for sensory leakage theories is from individuals who have experienced brain injury. Further support comes from the observation that certain ventrally located neurons in the pro-motor cortex seemed to be responsive to visual and/or somasthetic stimulation [Harrison]

Dr. Richard Cytowic's [Cytowic] theory of Synesthesia it that certain brain areas are not properly connected, causing the cross-sensory cognitions to employ the limbic system. His observations are strongly based on observations of blood flow and similarities to migraine headaches. He suggests a re-envisioning of brain function theories to accommodate new findings in the area of synesthesia. Also, his explanations do not seem to accommodate other aspects of synesthesia including an increased disposition towards dejà-vu, clairvoyance, precognitive dreams, color dreams, movement dreams, heightened dream ability, or increased temporal lobe activity [Harrison].

Peter Grossenbacher espouses feed-backwards connections from cortical sensory processing modules to multi-sensory areas. He notes that those areas are bi-directionally connected and that reverse-transmission is normally inhibited. Feed-backwards theory

proposes that the reverse-transmission inhibition can be diminished under certain conditions [Harrison].

Environmentally shaped brain maturation proposes that sensory pathways are selected through the process of exposure. Sensory stimulus causes the strengthening of neural pathways, the absence of stimulation, the apoptosis of inactive pathways. The theory implies that exposure to colored block letters is sufficient to create the neural facility to support color-grapheme synesthesia at a later age [Harrison].

Quantitative vs. qualitative theory asks whether synesthesia is an entirely different phenomenon from 'normal' consciousness? It is possible that the real question is rather "how much synesthetic awareness does a person have?" This theory cites examples of perfect pitch, spectral mapping of sensory experiences, and other seemingly extreme forms of sensory awareness [Harrison].

Posterior Inferior Temporal activation theory is based on differences in blood flow levels in the synesthete brains. Positron Emission Tomography (PET) scans of color and language-processing areas indicate elevated blood flow in the case of auditory processing. Also occipital-parietal junctions indicate elevated reading [Duffy].

Classifications of Synesthesia

There are several divergent references to the concept of synesthesia in literature, arts, sciences, and medical research. Cambridge University Researchers Simon Baron-Cohen and John Harrison provide four primary classifications of synesthesia to clarify the concepts presented in [Duffy]. These classifications are: developmental, metaphorical, acquired, or drug-induced. Each theory is presented and defined.

- Developmental: This form of synesthesia begins in early childhood (4 years). It is not a hallucination or other psychotic phenomenon. It is different from imaginary imagery. It is not induced by drug use. Rather, it is characterized as vivid, automatic, involuntary, and unlearned. This form is also referred to as 'constitutional' or less preferably as 'idiopathic'. The term 'idiopathic' is

often associated with a medical disease and adds a negative connotation of disease to the synesthesian experience.

- Metaphorical: This is the artistic or linguistic expression of one sensory experience in terms of a differing experience. Examples of this would be the 'sticky purple-grey smell' of dry-erase markers, or the 'warm glow of the midnight moonlight'. Other examples are found in paintings and architecture. The music visualization applications presented in the last chapter are also examples of metaphorical synesthesia. Use of synesthetic metaphor in art is not an indication of true synesthesia on the part of the creator.

- Acquired: This is the result of 'neurological dysfunction or other dramatic physical change' resulting from injury, tumor, or other physiological medical condition. Subjects do not report any synesthesia prior to the injury. These forms and the stimulating conditions are less structured than color-grapheme or color-music synesthesia. More commonly, the subjects experience unorganized interruptions of their mental processes.

- Drug-Induced: These are the result of psychoactive drugs. These experiences typically lack organized consistency. While an individual might experience an audio-visual correspondence, the mappings may vary from moment to moment. Also this form of temporary synesthesia diminishes as the effects of the drugs wear off.

Developmental Theories

These theories seek to account for the initial onset of synesthesia in individuals. Of related interest are the theories of learning styles and multiple intelligences. Analysis of learning styles can reveal individual's preferences for information encoding, storage, and recall. Three main styles are identified: visual, tactile, and audio. Analyzing a person's speech patterns can reveal a bias for each style. Verbal conversation responses per type might include common uses of following phrases: "I see" for visual learners, "I feel" for tactile/emotive learners, and "I hear you" for auditory learners [Duffy].

Multiple intelligences relate to the many forms of knowledge a person can acquire. Howard Gardener identifies six forms of knowledge, kinesthetic (body), logical-mathematical, linguistic, spatial, musical, and interpersonal.

The education system emphasizes the learning of factual knowledge. Taken together with the various learning styles, researchers hypothesize that the methods we use to communicate have an impact on the content of our thoughts. The focus on terminology and characterizations of sensations is suspected as a contributing factor in children losing some of their ability to visualize. As the students focus on individual knowledge forms and abstract terminology, the integration of concepts diminishes. Researchers indicate that synesthesia is a holdover from that integrated mode of thinking. Synesthetic thought is seen as an all encompassing or abstracted form of knowledge integration.

4.5 Synesthesia in Music and Visual Arts

Several texts hint at the presence of synesthesia in the arts. The 1889 International conference on physiological psychology featured sixteen papers on synesthesia including the intersection of art and science. Arts with a natural crossover between sensory experiences include poetry, music, painting, and theatre [Dann]. This section gives examples on synesthetic integration in expressive settings.

Music

Erasmus Darwin, grandfather of Charles Darwin, synchronized colored lights with notes on a harpsichord to present his concept of color music. Bainbridge Bishop designed a color organ in 1877 in which playing notes caused colors to be displayed. Arthur Wallace Rimington 1895 developed a color organ in London to project colors with music. In 1889 Louis Favre demonstrated chromesthesia and color music by creating illuminated water fountains. Father Louis Castel's color organ was built with the intent of allowing individuals to hear, see, and 'taste' music. Mary Hallock Greenwalt held a color music concert in 1911. She felt that the color music would remind us of the

'Holy Ghost' in utter sheerness of beauty. The Illuminating Engineers Society, in 1918, referred to color music as the spinal marrow of mankind [Duffy].

Alexander Scriabin's synesthetic musical symphonies were partially inspired by theosophical literature of Helena Blavatsky, Anne Besant, and C.W. Leadbeater. Leadbeater wrote of an auric sensory egg surrounding humans. One document, the Secret Doctrine, contained a mapping between chromatic sequences and powers of the natural world. Louis Wilson is also credited with an occult scale correlating colors and musical tones. Alexander Scriabin's Prometheus: 'Poem Of Fire,' an orchestral poem, included a light line specifically prescribing projected colored light. These color assignments were made using the circle of fifths. Though it was innovative and multi-sensory, these early works were considered by some critics to be distracting [Duffy].

Olivier Messiaen explored his synesthesia in musical compositions translated as The Colors of the Heavenly City, and The Colors of Time. He reported his synesthetic experience of music, but also expressed that the idea of colored music had no importance to the listener. He did add that synesthesia facilitated musical interaction and transformation, though this seemed to be more of a commentary on his internal thought processes than an explicit intent [Duffy].

Iannis Xenakis, a student of Olivier Messiaen, took the multimedia creations to a new level with his Polytopes works: Persephassa (1969), Persepolis (1971), Cluny (1972), and others. These productions combined sound, light, movement, and architecture. The visual effects were tied to the music to establish a contrast in the musical compositions [Harley] and to immerse the participants into moments of artistic truth. These systems provided animated renderings of audio compositions, reflecting his goal of free and abstract intersection between art and science. In Xenakis' DIATOPE system, four 4-watt lasers were shone off of 400 mirrors to simulate shooting stars and bursts of light. 1600 electronic flashes generated rotating and fading imagery. This system involved seven audio tracks and a nine-track digital tape drive providing twenty-five image updates per second [Xenakis85].

The Polytopes were a further realization of Xenakis's varied musical composition techniques and architectural constructions. His works translated swooping curves and

arcs into multi-layered glissandos and structures. Other compositions used probabilistic and stochastic methods to generate algorithmic-based compositions. Xenakis' book, Formalized Music, discusses his varied mathematical approaches to musical composition and representation [Xenakis85].

Xenakis's UPIC (Unité Polyagogique Informatique du CEMAMu) system is another attempt to merge sensory modalities. His goal was to enable anyone to make music by drawing images onto a page by using a special light pen. Those images were then converted into sound and interpreted in a variety of means. Images were treated as 'pressure curves, dynamic envelopes, scores in the time domain, and so on.' Xenakis wrote that this system enabled students to think about musical compositions without the pre-requisite of mastering the nuances of performance [Xenakis85].

Two modern adaptations of UPIC system are the Sonic Scanner and the Graphonic Interface. The Sonic Scanner is a handheld image scanner that has been modified to convert pixel information into sound. This device offers multiple modes of transformation: waveform, spectrum, rhythm, and sampler. Also, three finger sensors and one palm pressure sensor provide expressive control over the audio renderings. In the Graphonic Interface, individuals draw images onto a clear or frosted glass display equipped with a whiteboard tracking system. Pen coordinates are converted into audio and played through a glass-mounted or external speaker [Overholt].

The sensual links between the artist and the medium create multiple means of exploring creativity. Composer Michael Torke also engaged in color music composition. He noted specifically that an E-major chord gravitates towards the fundamental root E. The note E and the chord were synesthetically experienced as the color green. He developed certain works with the titles "Green Symphony" and "Ecstatic Orange". In his mind the compositions were musical rooms, each holding some element of emotional truth. Torke also felt textural experiences from instrumental sounds. French horns are round, trumpets are pointed, flutes sound cottony, and clarinets were sleek as panther's fur. His compositions were not designed to mimic those individual textures. Rather, his perception of the tactile differences between the sounds influenced his artistic decisions.

He consciously used his perceptions of the sound textures to create contrasting patterns by alternating the instruments used over the duration of the composition [Duffy].

Artist Paul Klee developed a visual music system to visualize note lengths and relationships between tones. This system was designed to reveal hidden relationships between the parts of the music. Greta Berman, an art historian at the Julliard School, talks about a common vision among synesthetic composers. Many of them were longing for the ultimate truth: a unity of arts, religious nature and humanity. Synesthesia, an inner-space, is a dimension where everything has found the right place. Senses exist in harmony and become each other.

Not all color music works were well received. Some critics of combined music and art experiences report indifference or incredulity. Some debated the fundamental premise of the analogy, others felt the works were lacking in each individual areas. Some of the music was too alien to the audience or the audio mappings were not clear. The Grove Dictionary of Music is uncertain towards color music, citing too much reliance on special effects. The guidelines given in the last chapter regarding multi media compositions seem especially relevant for future integrative multi-art applications [Duffy].

Visual Arts

Vasily Kandinsky's works are heralded as unique synesthetic transformations. It is written that he was not just painting his inner thoughts or inner psychic states. From his perspective, he was reinforcing some inner state of consciousness. Kandinsky felt he was making the inner element sound more strongly by limiting the external. He also referred to his works of art as *Inner Notwendigkeit*; a sort of inalienable spiritual truth [Dann].

Sir David Brewster, an optician is credited with the development of a kaleidoscope as a synesthetic device. Marcia Smilack, a synesthetic photographer, reported an ability to hear music from images. She took pictures using non-conventional methods by framing her photos on the basis of sounds she heard in the framed image.

Ruth Armer and Georgia O'Keefe made paintings of musical compositions, based on synesthetic impressions [Duffy].

Carol Steen, a contemporary synesthetic artist, tells of an experience of colored sensation with electro-acupuncturist David Myer. In these therapy sessions Myer connected oscillating electrical signals to acupuncture needles. Maximum pain relief was found by tuning the signal's frequency to match a Steen's colored perceptions of their pains [Duffy].

Miscellaneous

Rimbaud's poem *Voyelles* assigns colors to vowels. Marcel Proust references colors in association with a character's name in his writings. French Symbolists discussed synesthesia as a transcendental method of perception. In the 1960s synesthesia was held as an impending advance in human evolution. It was hoped to be an antidote to rationalism and materialism. It was considered as a pathway to positivism and a central theme of a counterculture seeking a path to unity and holistic living [Duffy].

Nobel Prize winning physicist Richard Feynman commented about math equations and arithmetic progressions being mapped to musical scale, geometric progressions mapping to accelerating whoops, and other synesthetic experiences. A scientific application, DNA Music, exploits the observation that the human auditory system can identify two sound patterns at a time, but only one visual pattern at a time [Duffy].

This section will closes by mentioning two well-known synesthete musicians who have had great influence in our time. The first is Billy Joel, and the second is Dizzy Gillespie. Dizzy, a color-hearing synesthete, talked about his colored hearing in meetings at his church.

4.6 Synesthetic Multisensory Correspondences

This section ends the chapter on synesthesia by collecting some synesthete descriptions of their color music perceptions. Also, some non-synesthete information is given about the areas of sound symbolism and perception across the expressive arts. Sound symbolism concerns the linkage between sound and meaning in communication. Perception across the expressive arts is presented as an exploration or the intersection of visual imagery, sounds, and written metaphors for non-synesthetes.

Musical photisms are a coupling of color with shape and/or texture. These are slightly different from color grapheme experiences, where the color occupies the space inside of or adjacent to the character or digits. Some pairings involve the emotional content of musical compositions. Arnold Schoenberg's color-emotion schema, similar to one used by Kandinsky, tied emotions to projected colored stage lights. Passivity was indicated through darker hues such as brown, green, and violet. Increasing activity was represented by red, orange, and yellow colors. Vaporous blue represented celestial apotheosis. It is noted that these color pairings reflected the 'traditional color symbolism' of the time [Dann].

Another approach is per-pitch coloring. Blavatsky's Secret Doctrine assigns colors chromatically to pitches and emotional terms. Table 4.2 gives those parings.

Table 4.2. Blavatsky's Secret Doctrine Pairings

Note	Color	Emotion
C	Red	Power
D	Orange	Energy
E	Yellow	Intellect
F	Green	Sympathy
G	Blue	Devotion
A	Indigo	Indigo
B	Violet	Psychism

Though the coloring terms given above are useful for creating pitch-color assignments in compositions, synesthete experiences tend to be more algorithmic and focused on the textures of the sounds. Michael J. Zeigler (1930) reported synesthete musical experiences to include a mixture of specific hues and geometric terms. Similar to Figure 4.1 above, abstract shape terms included sparks, spots, lines, streaks and zigzags. Geometric shapes included thimbles, tubes, morning glory, spheres, quadrangular blocks, spheres, and bursting masses with jagged splinters. Külver's list adds spirals, lattices, arcs, and honeycombs. Other descriptions include lumpy dough, ribbons, streams, and daggers [Dann].

Synesthetes also note that the appearances of their photisms follow a dynamic rule set. High tones produced small and faint photisms while larger and dark photisms appeared for low tones [Dann]. This author adds that the placement of an instrument, or panning of a sound in an audio mix impact photisms locations. Volume and effects have an impact on the photisms locations as well. These photisms are present for both real (stimulated) and conceptual (internally conceived) audio.

Musical performers have leveraged photisms as an enhancement to audition. Photisms have been reported as internal representations of pitch information for singers and intonation for musicians. Performers experience automatic photism-based comparison between the perceived tones and the intended tones and are able to correct their performance accordingly. In another case, the tones are converted to color information. The performer remembers compositions on the basis of their chromatisms.

A rare case of an experience of flavor in response to hearing musical intervals has been reported [Beeli]. Each tone interval, as described in Chapter Three, relates the distance between musical notes. The experimenters reported that the subject had perfect interval accuracy. A Stroop-test was performed to measure influence of priming on the subject's synesthetic perceptions. The subject was primed with flavors and words describing flavors that were congruent or incongruent with her perceived tastes. The experimenter's theory was that the presence of a priming flavor would impact her synesthetic response time. Priming with actual taste stimulus had an impact on performance, while conceptual (word) priming of a flavor did not. Congruent taste

stimulus priming lead to shorter reaction times, incongruent taste stimulus priming to longer times, and no flavor taste priming lead to average response time. The authors conclude that synesthesias may be used to solve cognitive problems [Beeli].

Sound symbolism, is the study of the direct linkage between sound and meaning. Sound symbolism occurs in four main forms: corporeal, synesthetic, imitative, and conventional.

- **Corporeal sound symbolism** refers to sounds that express the speaker's internal emotional and physical state. Examples include: grunts, coughs, cries, and other unspecified interjections.

- **Synesthetic sound symbolism** refers to the symbolism of non-sounds or the modulation of speech to communicate concepts. Examples include: using vocal pitch to demonstrate size, or intonation to specify the connotation of a word.

- **Imitative sound symbolism** is the representation of a non-vocal sound with language components. Onomatopoeia is a common example of imitative symbolism. Examples include: "whish, tap, bang, and cuckoo".

- **Conventional sound symbolism** refers to the commonly occurring symbolic elements and shared cultural responses to those elements. An example is words with common beginning sounds or letters. The cluster '/gl/' appears in several words: glass, glisten, glow, glaze, and glimpse, and commonly communicates light or shining [Hinton].

Further analysis of these common sound clusters reveals similar implications from common assonances. These implications, also called aural images, demonstrate the connection from sounds to meanings. The letter p- in the words pop, ping, pow, patter, and peep often signify an abrupt onset. The letter b- in boom, bang, beep, and boing indicate an abrupt loud onset. The letter r- in the words rip, roar, rattle, and roll, an irregular onset. Each group of these words also have a common amplitude/time curve that is related to the initial syllable. The sounds P, and B have a sharp, immediate onset and imply similarly immediate events. Pl-, Kl-, a more gradual onset, and represent less sudden events. The letter R has a jagged plateau, relating to irregular events [Rhodes].

In the context of a synesthetic music experience, sound symbolism provides a useful means for mapping from the real-time properties of sounds and volume envelopes to conceptual meaning. If sharp sudden sounds have a common meaning in language, it seems reasonable that sharp sudden musical sounds may communicate a similar meaning to the audience. If slow soft words imply moderate experiences in language, then a graphical representation may be developed that successfully communicates the difference between slow and soft accelerations. Sound symbolism research is a potential bridge for non-synesthetes between the characteristics of a sound, emotional content of a sound, and visual correspondences.

A study of cognitive poetics, the expressiveness of sounds, also provides an interesting emotional perspective on listening. Reuven Tsur's proposes that our listening system operates in speech, non-speech, and poetic modes. In speech mode, we extract phonetic content. In non-speech mode we perceive the speech as we would natural noises or musical tones. In poetic mode we extract categorical information and intuitive information about message.

In this poetic mode, we find several conceptual mappings that span physical characteristics, speech tones, and colors. *Front <-> Back* is a mapping of the spectra of Bright Vowels (produced at the front of the mouth) to <-> Dark Vowels (produced deeper in the throat). *Low <-> High,* a measure of pitch and also height, relate to *Thick <-> Thin*, a measure of width. *Heavy <-> Light, Slow <-> Fast, and Wide <-> Narrow* are all correlated conceptual spectra [Tsur].

This concept of spectra is further examined by two hypotheses on how we build our correlations between cognitive spaces. In a Unified Space Hypothesis "a mental representation of space is constructed, so as to integrate out sensor-motor interaction with objects in space." In an Analogous Space Hypothesis, each sense is perceived uniquely and distinctly, though some analogies are common between them [Tsur]. Synesthesia seems to represent an intermediate representation for the unified spaces or a mapping of the correlations between analogous spaces.

Finally, an examination of aesthetics provides a set of criteria to be considered in developing an application to demonstrate audio-visual correspondences. The purpose and

function of art are also presented here. Also, elements in individual art forms reveal potential areas for transforming between modes of perception [Sporre].

The four main types of aesthetic responses to works of arts are:

- **What is it?** This is a formal description of the work.
- **How is it put together?** This is a technical description of the elements of an artwork.
- **How does it appeal to the senses?** This is an analysis of the experience.
- **What does it mean?** This is a consideration of the context and personal relevance of a work of art.

The four main purposes for art are:

- To serve as a creative record.
- To give tangible or experiential form to feelings.
- To reveal metaphysical or spiritual truths.
- To help people see the world in new or innovative ways.

Art has several functions in society. Art is to be enjoyed. It can be a political social weapon. Art can be applied as therapy, and it is an artifact of society. Being able to identify art, learning some relevant terminology, and understanding how individuals' perception affects their response to that art help develop aesthetic response to art [Sporre].

A basic analysis of a work of art in each media includes a consideration of set of particular criteria. This dissertation study has examined criteria in the visual arts of pictures and sculptures. Criteria in the performing arts of music, theatre, cinema, and dance were studied, along with literature language arts. All criteria were reviewed in relevance to developing visual transformations of musical experiences [Sporre].

Some highlights in printed visual arts include the use of line, form, color, repetition, balance, focal area, and deep space. The dimensionality and texture of sculptures hold interest in part because of the three-dimensional nature of graphics libraries. In musical compositions, the form, sound, rhythm, tempo, texture, and ensemble are relevant considerations. In cinema, the shots, cutting, dissolves, focus, and camera

movement are of interest because of the potential of a visual music system to rotate and reveal various elements of the musical space.

Dance also shares the visual elements of line, form and repetition. An analysis of Dance also considers the message or experience that is being communicated and the relationship of the dance to the music. Dance appreciation also entails consideration of the Mise-en-scène, the combined effect of all the elements: settings, lighting, costumes, and movements. Landscape architecture includes a focus on the use of space and the impact of design on attention and movement. All of the afore mentioned arts consider the reaction of the participants. This reaction can be considered in many ways including the impact of the individual elements of the piece, the impact on the viewer's attention, and the effect of the composition as a whole [Sporre].

Chapter Summary

The mind of a synesthete is akin to an augmented inner world. In true constitutional synesthesia, concepts, characters, sounds, and a whole array of sensations are translated into visual photisms. Synesthetes typically report a beneficial aspect to their unique perceptions. Synesthetic artists utilize their perceptions in composing and performing their works. Beyond just colors and movement, their photisms are fundamentally linked to their thoughts, feelings, and actions. Historically, synesthesia has been viewed as a bridge between forms of expression, as a union of sensory experience and internal thought. Multi-art exhibitions of synesthesia have sought to demonstrate underlying connection between the expressive media and the creator's vision.

Synesthetic photisms are very diverse in their appearances, but they do follow an algorithmic consistency. Photisms share common characteristics with eidetic imagery and sound symbolism, two 'non-synesthetic' cognitive phenomena. While the physiological basis for synesthesia is not perfectly understood, researchers are able to classify and study individual experiences. Researchers are also beginning to identify developmental factors in synesthesia and have hypothesized connections to children's thought processes and gestalt sensations.

This chapter has presented several descriptions of synesthesia. Historical perspectives were given. Also some examples were given for synesthesia as a literary or artistic device. It has been noted as a beneficial aid for many. While it is not medically possible to give synesthesia to a person, it is possible to use synesthetic mappings to translate audio experiences and mathematical experiences into informative and entertaining experiences. Also, the aesthetic elements of human art forms provide a set of meaningful criteria to be considered when designing and evaluating a Synesthetic Musical Experience.

CHAPTER 5. PROTOTYPES

5.1 Motivations and Goals

The author's interest in music related software began with early graphics programming and in music ensemble performance. This author has considered potential applications of technology as a performer and/or leader in various music ensembles. Performance experiences include high school and church ensembles, home studio, experimental generative music, laptop techno, and accompaniment for interpretive vocal performances. Some of the potential applications for the technology are to:

- Illustrate aspects of music including melody/harmony, intervals, tempo, timing, grooves and waveforms
- Show the notes and sections of a song on a virtual score
- Indicate how the instruments and parts of music combine and interact
- Communicate chords and music parts while learning songs
- Signal introductions, endings, and tempo changes to the ensemble
- Demonstrate the concept of improvisation
- Demonstrate how music is made in loop editors/sequencers
- Illustrate the process of remixing and re-sequencing motives and rhythms
- Demonstrate applications of synesthesia in relationship to music education

The first music-tech prototypes were developed in the fall of 2003, based in part on application concept sketches in the preceding years. The prototypes are constructed in order to:

- Learn about MIDI, real-time audio, and computer graphics software tools
- Develop and integrate the software and required to capture and display musical events
- Explore methods of analysis and visualization
- Gather feedback from audiences about the various visual modules.

5.2 Implementation

Two prototype applications have been developed and run in Iowa State University's C4, C6, and Alliant Energy Lee Liu Auditorium. Both applications are written in C++ and OpenGL running on top of VRJuggler [VRJuggler]. The first application, MIDIViz, translates real-time MIDI events into chromatically mapped graphical objects. The second application, Illuminator, transforms multiple-band real-time audio power spectra into visual objects.

MIDIViz and Illuminator use custom Mac OS X helper applications for MIDI and audio data acquisition. Those helper applications are Midi-Transmitter and FFT-Transmitter. Also written in C++, these helper applications utilize open-source, cross-platform software libraries. Each helper application connects to the virtual reality application over a simple byte-oriented socket.

Figure 5.1 M-Audio Radium 61 MIDI Keyboard and OS X Host Computer

The first helper application, MIDI Transmitter, is built on top of the MIDI-Share library. MIDI-Transmitter intercepts all messages from MIDI-keyboard devices including note on/off/velocity messages, and knob/slider parameter messages for an M-Audio Radium 61 key-controller. Each note event is formatted into a packet and transmitted to the VR host applications. MIDI Transmitter also has the capability for bi-directional communication, allowing the VR application to trigger MIDI events on the OS X host computer.

The second helper application, FFT-Transmitter, is an adaptation of the audio processing used in the open source program Fluxus [Griffiths]. FFT-Transmitter uses

PortAudio to acquire the host computer's audio stream. That audio stream is collected into a buffer and processed by a FFTW, an open source Fast Fourier Transform library. The resulting FFT is collected and smoothed to produce 16 bands of power spectra. Each band indicates the loudness of a sub-portion of the frequency space in the input audio. The sixteen power spectra bands are scaled into the range 0-127, averaged over several frames, and transmitted via socket into the VR application at approximately 50 times per second.

MIDIViz Application

The MIDI Visualization Application was designed to illustrate basic music concepts in response to real-time MIDI keyboard-data input. Music fundamentals, colored-hearing synesthesia, and some adaptive systems methods are demonstrated. The app was developed with a universal audience in mind and was expected to have value for musical novices and non-instrumental students.

The application features several visual objects that are colored by a chromatic (rainbow-like) scheme. All objects responding to the same pitch share the same color. The objects respond by scaling, spinning, or moving in response to musical events. The visual modules in this system and a brief description are given below.

Table 5.1 Midi VIZ Modules and Description

Module Name	Description
Musical Towers	12 chromatically colored towers that rise and fall in response to musical notes. Each tower is mapped to one of the 12 notes in an octave. The tower height corresponds to the velocity of the played note.
Spinning Blocks	12 chromatically color towers that spin out and scale in response to input notes. The spinning and scaling are velocity sensitive.
Parameter Orb	8 concentric rings that scale and rotate in response to 8 keyboard parameter sliders. The parameters are colored according to the author's color-number synesthetic pairings.

Module Name	Description
Parameter Towers	16 small towers that rise and fall according to 8 parameter knobs and 8 sliders.
Spiral 5-Octave Bell Tree	8-foot wide cylindrical spiral arrangement of chromatically colored trapezoid-shaped blocks. Each full rotation corresponds to an octave. Notes spin in response to played notes.
Colorful Mosaic Wall	Scalable texture mapped wall that speckles with the color of new notes. The speckles disperse and fade to black over time.
Multi-Octave Grid	Grid 12 x 6 grid layout of musical notes. Grid cells feature objects that scale in proportion to note velocities.
Warping Star Field	Linear particle system simulating outer space. Particles accelerate, color, and brighten in response to note inputs.
Musical Planets	6 octaves of planets with drawn orbits. Planets and paths, scale, brighten, and accelerate in response to musical notes.
Music Responsive Worm	20 segment line of blocks. Each block moves and colors in response to a musical event. Playing a sequence of three notes results in a spiraling shaped multi-colored path.
Volume-Based Music Builder	A 30 x 30 x 30 cell volume with a moving cursor. 12 notes are mapped to 6 directions. New objects are created on every incoming note and the cursor moves in the corresponding direction. The resulting shape is a transformation of the composition.
Grid-based Artificial Life Simulation	This Two Dimensional Grid simulates the game of life with musical input. Musical notes cause new cells to appear and spread diamond-like through the population. Fitness is based on the loudest note played. Notes also feature a height-mapped age, contributing to the artistic appearance.
Lindenmayer-Systems Tree	This is a rule-based tree display. Trees are evolved from a character grammar including commands for segments, branches, twists, and flowers. Twist nodes are tied to musical events. They flex and change color in response to musical note input.

Module Name	Description
Sierpinski Fractal Pyramid	Creates a fractal pattern based on note velocity. The bases of the notes are arranged around a clock face, resulting in three-dimensional pyramid.
Geometrizer	Converts incoming notes into graphical objects based on number and range of notes. One note becomes a circle, two a line, three a triangle, etc. Playing notes in a single octave results in an outlined line figure, across two octaves results in a planar figure, and across three octaves results in a volume.
Frequency Display	Displays the fundamental frequency of the played notes. Displays individual notes as colored sine waves on one axis. Presents a summation of all played notes on a parallel axis.

The visual modules from the MIDIViz System are depicted below. Each module is demonstrated in an active state with some notes pressed. Each visual module can be translated, scaled, and rotated through the program source code.
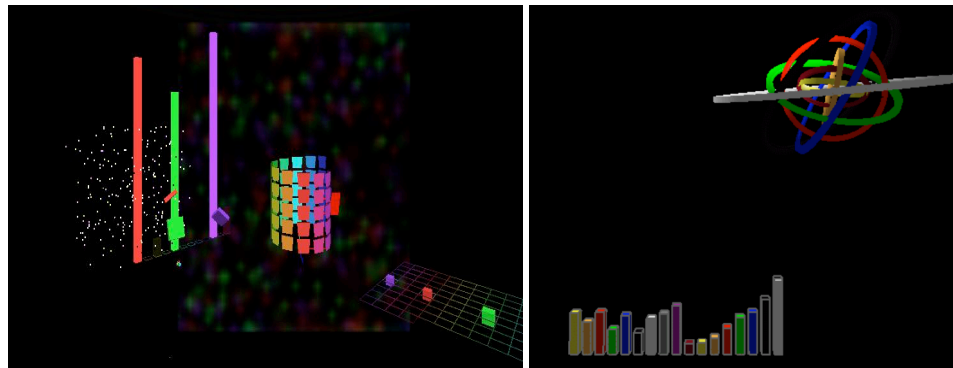

**Figure 5.2 (l.) MIDIViz Towers, Bell Tree, Mosaic Wall, Multi-Octave Grid
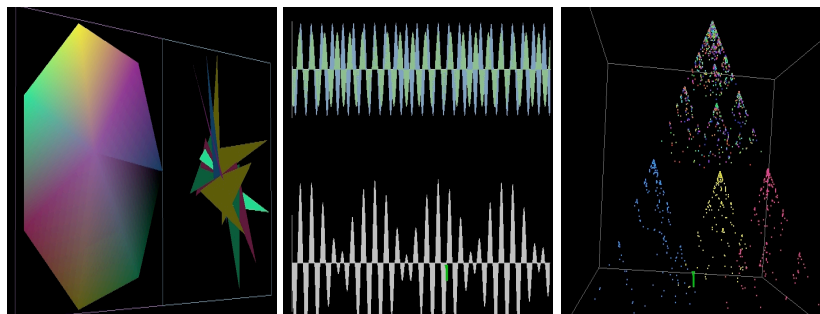(r.) Parameter Towers and Parameter Orb**

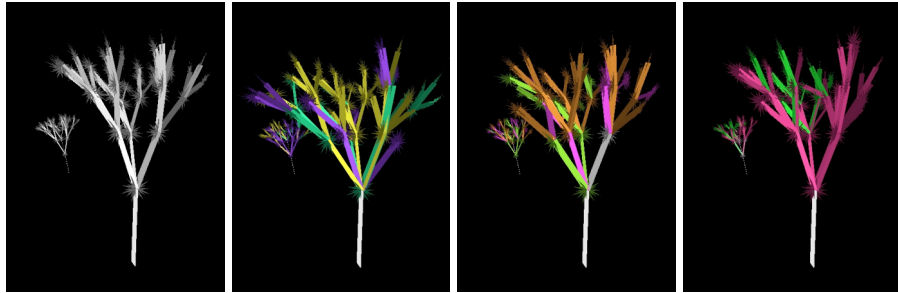**Figure 5.3 (l. to r.) Geometrizer, Frequency Plot, Sierpinski Fractal**

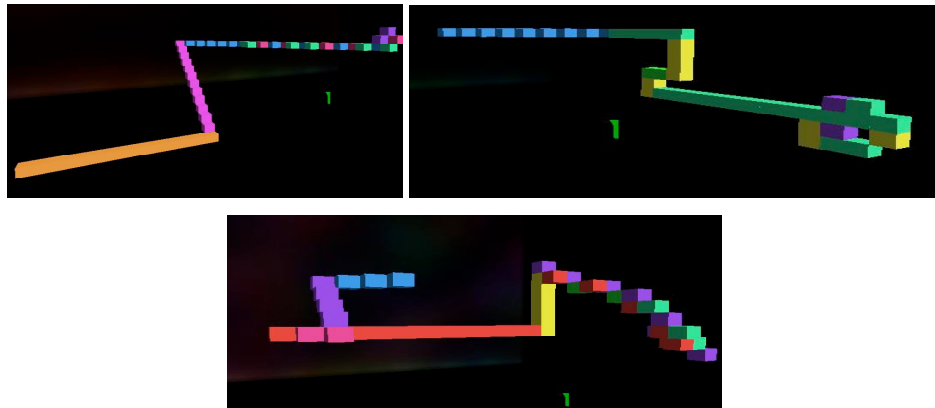**Figure 5.4 Lindenmayer-Systems Trees**
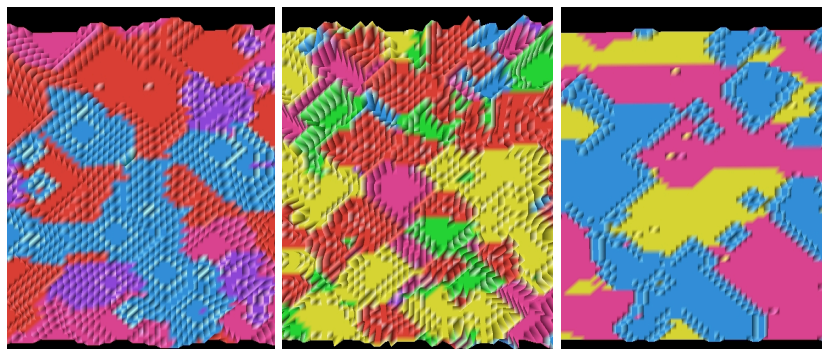

Figure 5.5 MIDIViz Worm

Figure 5.6 Grid Based Artificial Life System

MIDIViz's software design is simple, consisting primarily of frame and draw loops. Music events are received during the frame function, stored in a central table,

distributed to each display node, and routed to the currently displayed visual modules. Each visual module has a method for processing new MIDI events. The modules are arranged into scenes. Scene switching is accomplished by pressing one button on the VR interaction wand or device. Users navigate by pressing a second button on the wand and aiming in direction.

Illuminator

The Illuminator project was designed to explore possibilities of representing real-time audio in an immersive space. It was designed to represent the characteristics of sounds. Illuminator receives 16 measurements of frequency spectrum information and converts them into visual displays. The system also constructs weighted averages of the power values over time to provide smoothed data for some slower modulating visual effects. The modules in this system are described and depicted below.

Table 5.2. Illuminator Audio Modules and Descriptions

Module Name	Description
SpeakerBoxx	Each audio band is converted to a grey-scale spectrum and mapped to a square polygon. Low frequencies are placed on inner bands, high frequencies on outer bands. A color spectrum-coded linear bar display is placed in the middle of the box, indicating the level of each band.
Lollipop	Similar to speaker box. The entire object is color coded according to band intensity. Band corners expand outward with increased power levels. Creates a gentle moving effect.
Frequency History	Converts each time slice of power band measurements into a color-coded linear display. Several measurements are stored providing a long-term historical view.
Musical Water Fountain	This fountain resembles a 16-layer cake. Each layer is colored in a spectrum of blue to white in response to a frequency band. Also, a band issues water particles if its power level crosses a certain threshold. This module was also used to augment a real-time dance performance.

Module Name	Description
Factory Room	Designed to augment a factory-like audio composition with clanging metallic noises and reverberant spaces. 16 discs in the middle of the room move in response to individual power bands and flicker if the power level crosses 90%. The walls and floors are grey-scale colored in according to the slow average value per band. The flickering disks represent the sparks of clanging metal. The walls represent the reverberant echo of the factory space.

The visual modules from the Illuminator system are depicted below. Each module is animated with a sample audio and a couple of examples are given. These representations are scalable and appear person sized in the virtual environment.
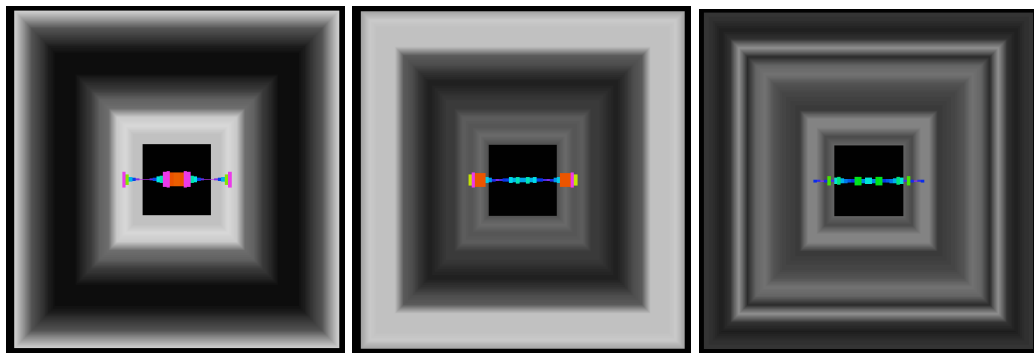

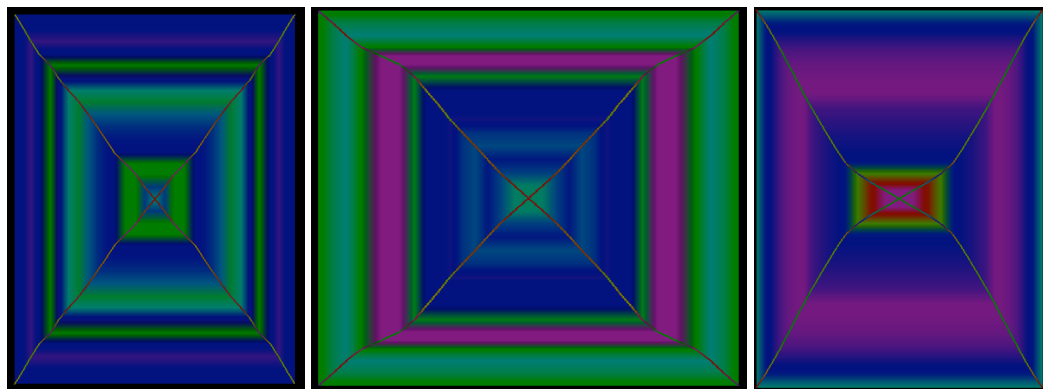**Figure 5.7. SpeakerBoxx with Central Audio Indicator**

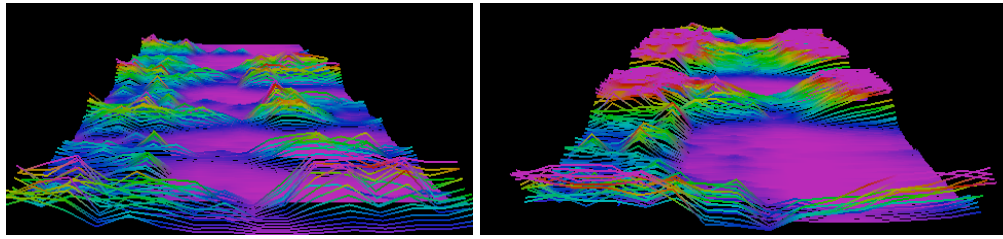

Figure 5.8 Lollipop

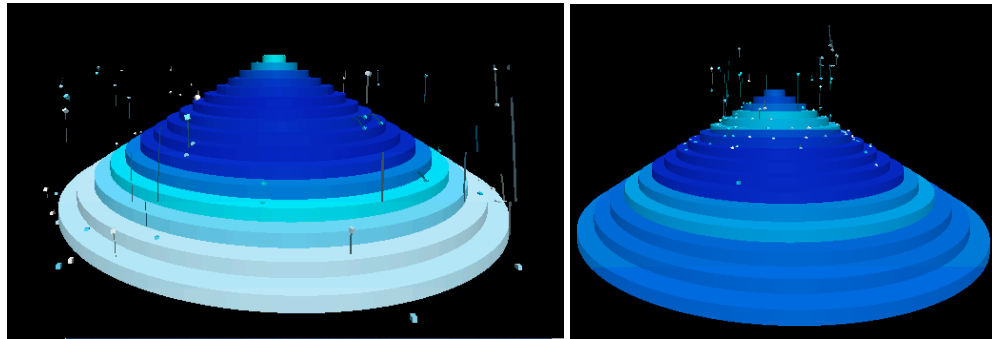**Figure 5.8 Frequency History**
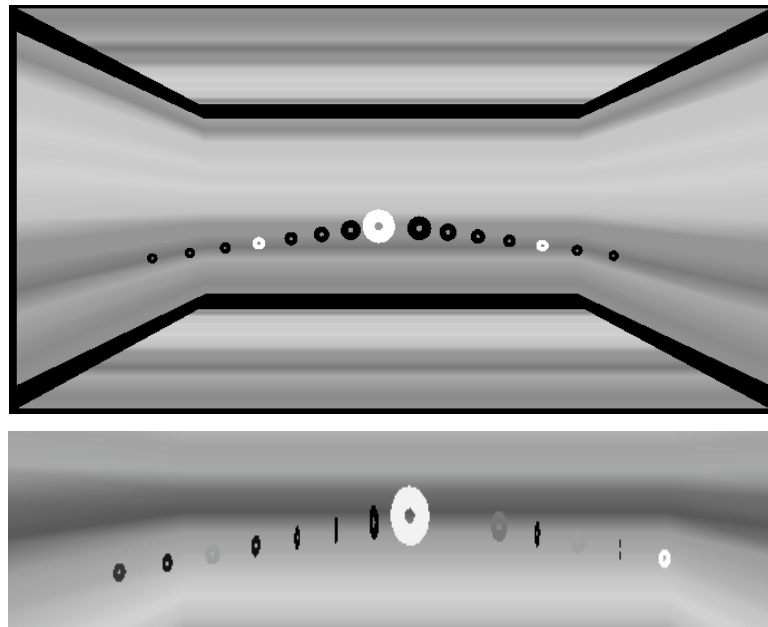

Figure 5.8 Musical Water Fountain

Figure 5.9 Metallic Factory Room

The software design for the illuminator system is also fairly simple. Each module exists separately in a scene. Every draw loop assumes that new audio information is

available. Each module reads the band levels and draws itself dynamically. Some objects do retain a history or a set velocity. Those velocities are attenuated each frame till a zero condition is met. Scene switching and navigation are similar to the MIDIViz Application given above.

Audio Based Computer Game

A third prototype is presently in development. This project is an interactive audio-based computer game. The game explores the use of sound as the primary cue for game interaction. Audio samples are modulated on the basis of pitch and location. Users navigate with arrow keys to locate the randomly placed sources. This game uses OpenAL, an open source cross platform tool for integrating positional audio into applications.

<p style="text-align:center">5.3 Observations</p>

The MIDIViz and Illuminator audio projects have been demonstrated to sample audiences for several Iowa State University Events including First Lego League (2004&2005), VRAC Tour Days, Educational Talent Search summer program (2005), Human Computer Interaction Spring Forum (2004, 2005). Some initial observations of the program are as given below:

- The program appears to be most effective for younger audiences from elementary through high school years.
- Students enjoy the participatory aspects of the event such as playing the keyboard and making the sounds to activate the visuals.
- Older audiences ask more questions about the technology while younger audiences like to explore and play.
- Parents and classmates are excited to see their children and peers interact with the system.
- Students are focused and alert during the demonstrations.

- Students talk to each other about the behavior of the visual modules.
- Students jump and strain to reach the virtual MIDIViz objects.
- Many of the students are active in music lessons and understand many music fundamentals.
- Immersive environments like the C4 and C6 give the students the strongest feeling of being immersed.
- Students were less responsive when the C4 was operated in an open (wide screen) configuration than in a closed (cube) configuration.
- The MIDI Modules do not convey a sense of the long-term musical structure of a piece.
- The MIDI modules do not respond to the type of sound being played by the MIDI instrument.
- The Illuminator Modules do not convey the pitch or tones of the input audio.
- Many students watch their fingers while they play on the keyboard. They are not able to experience the graphics simultaneously.
- Live performance settings require careful navigation and adjustment to ensure proper placement of the graphics.

These are some of the initial observations garnered from working with the system. Note, these were not collected with a survey method; they are simply the recollections of the operator and author.

5.4 Potential Extensions

There are several additional avenues of discovery for the prototypes listed here. These items are listed here in anticipation of the dissertation research study tasks proposed in the next chapter. This listing is not a commitment for future dissertation work. Rather these options will implemented on the basis of desired function. Some of these ideas are noted in response to specific deficiencies observed in the operation and development of the prototypes. Others are in response to information discovered during the literature search.

- **Increased Note Visualization:** This system would benefit from additional representations of real-time music events. Some ideas include converting a sequence of notes to sculptures, textures, and patterns.

- **Combinations of Audio and MIDI:** The audio and MIDI information should be integrated together. Each instrument in an ensemble could be captured with both audio and digital data streams. Those streams should be combined to illustrate both the aural and theoretical aspects of the sounds.

- **Increased Interaction:** An ideal system would enable participants to touch and arrange their melodic notes in the virtual space and create compositions.

- **Aggregate and Long-term Displays** The system should give higher-level long-term representations of the musical performances. This can include items like chords, song segments, dynamics, and texture.

- **Improve Visual Effects** The system includes diffuse colored objects in most locations. These do not exercise some of the advanced rendering, shading, and texturing options in today's graphics tools suites.

- **Dynamic Configuration** of module and placements in the virtual space. This tool is seen as a creative aid for a composer. Composers should have control over modules be placement and routing of MIDI and audio data.

- **Demonstrate Instrument Fingerings.** Instrument fingering positions could be introduced to enable audiences to learn how the performers generate the tones. Also performers in the ensemble could see translations from one instrument to another to accelerate communication.

- **Multi-channel Midi.** Presently the system only accommodates one channel of MIDI data. MIDI capacity per port is 16 channels. Separate Modules should respond to individual MIDI-channels, and possibly multiple MIDI ports.

- **Cultural Components.** The system could incorporate the ability to recognize rhythms, musical scales, or other attributes representative of diverse cultures. The system could incorporate elements of those cultures into the visual displays. Elements might include maps, localized-representations on a globe, and flags.

- **Intelligent Tutoring.** Exploratory adaptive lessons for building compositions in the virtual space or forms of improvisational help could be integrated.
- **Waveform Synthesis.** The application could enable participants to configure waveforms from within the virtual space. Configuration options could include overtones and/or envelope information.
- **Integration with Virtual Design Applications.** This would allow a designer to experiment with interactive synesthetic components in architectural designs. Performers could experiment with stage design for live music acts.
- **Increased Configurability for Modules.** Users should be able to configure modules to respond to MIDI controllers, single notes, chords, or sequences of chords. Interface options should include settings for channel numbers (0-15, omni), and note numbers (0-127, %n, etc). Experiment with coloring modes, static, dynamic, multi color, and ranges.

Chapter Summary

Three prototype applications have been developed to support research in group education and entertainment. MIDIViz incorporates real-time MIDI keyboard notes and controls into an immersive visual space. Illuminator animates real-time audio using sixteen frequency band power levels and custom interactive graphics. The audio based video game attempts to use real-time audio parameters in an interactive gaming challenge.

Initial observations indicate that this system has good potential for engaging younger audiences and getting them interested in the technology. Several areas for future expansion have been identified including increased variety of visual displays, expanded configuration options, long-term displays of the composition, and improved displays of musical concepts. Some of these expansion ideas are forwarded into the dissertation research proposal plan.

CHAPTER 6. RESEARCH PLAN

This chapter presents three potential directions for the final dissertation project. Each proposal performs a transformation of the audio and MIDI data. The first concept is designed to facilitate percussion instruction and rehearsal, the second, to facilitate communication in a small musical ensemble, and the third, to demonstrate Synesthetic musical perception. This chapter begins by reviewing the foundations for these projects and summarizing the research precedent for the proposed systems. Each system is described briefly and testable hypotheses are stated. A list of evaluation methods is given with consideration for the impact of the systems on the composers, performers, and audience. Avenues for publishing these research results are identified, and a schedule for completing the work is given.

6.1 Review of Research

This dissertation began by presenting potential uses for music and visualization in musical settings. These ideas have been developed through literature review and discussions with musicians, educators, artists, and technology professionals. A set of objectives has been defined for this research effort. Those objectives include:

- Demonstrating musical concepts through the use of real-time computer graphics.
- Creating a synesthesia like transformation of a musical performance.
- Increasing performers' and audiences' awareness of musical concepts.
- Stimulating thought into multi-media integration of music and artistic visuals.

Human Computer Interaction principles have been researched to discover guidelines for developing and evaluating interactive systems. Computer graphics and visualization systems have been reviewed with special consideration to the number of participants, level of interactivity, and amount of immersion they afford. Multi-modal applications research has been reviewed to consider proper methods for integrating multiple sets of musical information. Research resources in the area of human perception

have been identified. Forward-looking perspectives on virtual experience sharing and autonomous computer generated art/visualization systems were presented.

Philosophy, principles, methods, and concepts of music education have been researched. The role of technology in music education as well as current research concerning the impact of those systems on participants has been identified. Selected music applications for performance monitoring, rehearsal support, and large group interaction have been identified. Guidelines for developing multi-art systems were presented and analysis criteria for multi-person musical installations have been identified. Research into generative musical systems has been conducted. Applications designed for transforming visual imagery into sound and applications that enable novices and experts to interact musically through the aid of visual imagery were identified.

Synesthesia, a cross-modal sensory experience, has been studied as a means for transforming kinesthetic and sensory information into visual content. Synesthetes experience increased sensory awareness through involuntary cross-modal associations. Synesthetes' experiences are physiologically based, algorithmic in behavior, and analogous to other relevant cognitive phenomena. Several theories on the causes and implications of synesthesia have been identified. Synesthetic artists have reported creative stimulus and performance benefits from their experiences. Several historical experiments into the synesthetic combination of imagery and sound have been reviewed. Also, the related areas of sound symbolism and aesthetic appreciation between the arts have been surveyed to determine the potential use of synesthetic metaphors for non-synesthetes.

Prototype applications have been developed to demonstrate synesthetic and musical concepts. These applications have explored representations of musical information and multi-spectrum real-time audio in immersive graphical environments. The applications have been demonstrated within the university, in teleconferenced conference presentations, and to the local community. Preliminary observations of these systems indicate that they are especially effective in capturing the attention of K-12 students. The prototypes have been successfully used to augment a modern dance performance with visual imagery. Students have been able to interact with the

applications and experiment with the combinations of audio and visual information. The applications do require expansion to effectively illustrate high-level music concepts. A list of future expansions for the prototypes has been constructed.

The combination of initial ideas, in depth research into HCI, music, and synesthesia, and the experiences with the prototype systems have been combined to motivate the next stage of this research. The next section introduces three application concepts that aim to demonstrate the research objectives identified in this dissertation.

6.2 Proposed Systems

Three conceptual systems, demonstrating various aspects of a synesthetic musical experience, are presented here. This research process has suggested potential applications of these systems in education, entertainment, performance, and rehearsal. The systems are all interrelated, sharing common characteristics, technologies, and research areas. Each system combines real-time processed audio and MIDI-data streams into a visual display of performance information. The first system, a Computer Augmented Percussion Trainer is intended to facilitate percussion rehearsal and performance for individuals and groups. The second system, a Musical Ensemble Trainer, seeks to provide visual feedback to a small musical ensemble and support communication of ideas between performers. The third system, a Synesthetic Visualizer, is intended to represent a music performances and concepts as well as entertain and inspire an audience. Each proposed system is described in general terms and a formal experimental hypothesis is stated.

Computer Augmented Percussion Trainer

The Computer Augmented Percussion Trainer would be designed to assist in individual practice and small ensemble rehearsals. The system would convert each percussionist's performance into a scrolling visual display to indicate note velocity (intensity) and timing. In a group setting, multiple players could be presented in parallel.

By illustrating each note, this system has potential to reinforce group tempo, accents, dynamic control, and rudimentary performance. At present, the majority of feedback comes from the conductor or technician. Performers may also watch their movements in a mirror or use audio/video recordings. However, the loudness of an ensemble and the speed of movement can interfere with listening or visual measurements.

Novice percussionists lack adequate dexterity to accurately control their loudness, timing, and strike intensity. Students develop these skills by playing basic patterns, or rudiments, at a slow tempo. Students progress to faster tempos and towards more complex rudiments. This system can potentially help individual students learn rudimentary control by revealing the variations in the individual notes.

A computer augmented percussion trainer can make use of drum triggers or electronic drum pads to measure the striking force of a drum hit. Those measurements can be converted to MIDI using an electronic percussion module and routed to a graphical display. Also, analog audio can be processed to differentiate bass, snare, hi-hat, or tom sounds. Figure 6.1 gives a concept for a visual display, indicating the individual players notes, and volumes relative to the ensemble. A customized glyph could also be used to indicate the individual players average timing and intensity levels.
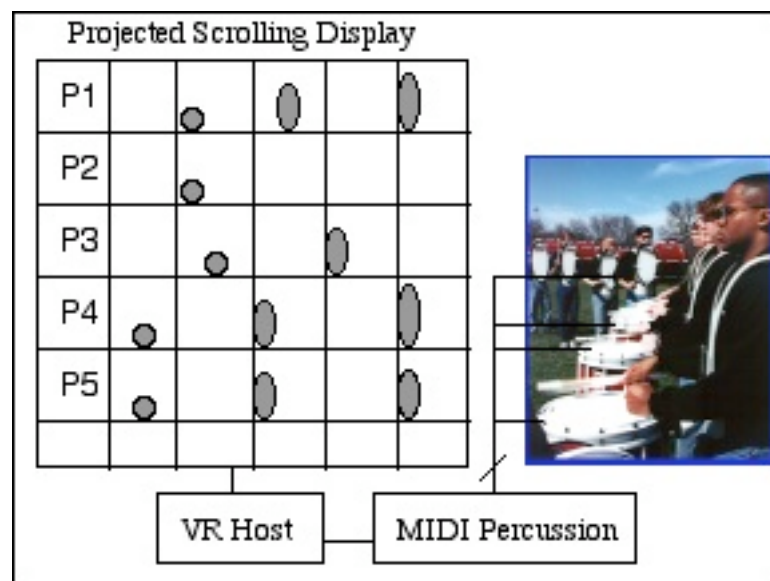

Figure 6.1 Computer Aided Percussion Trainer with Scrolling Note Display

Steps in developing this system would include:

- Interviewing percussion instructors and students about teaching and learning techniques
- Collecting practice materials and rudiments
- Identifying the control issues that are difficult to teach or learn
- Selecting percussion instruments for the study
- Purchasing drum triggers
- Developing the graphical displays
- Studying student's performance abilities with and without visual assist of the system.
- Iterating through the design process to refine the system.

This system can be extended to a full electronic drum set, or to multiple players in an ensemble. There is also potential to augment the performance with an interesting visual presentation for an audience. Looking beyond the scope of this dissertation, additional measurements could be added for stick heights and hand angles through the use of camera monitoring. Foot/heel movements could be tracked to ensure proper mark timing (marching in place).

Hypothesis for The Computer Augmented Percussion Trainer

Music Technology and Computer Visualization can improve percussionists' playing technique by providing an accurate display of individual and group timing and striking force. This tool can also illustrate elements of rhythm including common patterns, syncopation, and polyrhythms.

Small Ensemble Trainer

The Small Ensemble Trainer concept converts real-time music performance into a large-scale visual display. It is designed to facilitate communication between players in

the ensemble and to reflect a synesthesian musician's internal thought process while performing. The display will feature pitch, timing, and intensity information. Additional information might include high-level technical information for the benefit of the ensemble. Pitches, chord types and intervals would be displayed. Also, translations of fingerings cold be given for other instruments, namely between MIDI-keyboard, bass, and guitar. The display could be piano-roll or sectioned into individual instruments. Figure 6.3 gives a sketch of the system, indicating the notes for each performer in a small ensemble.
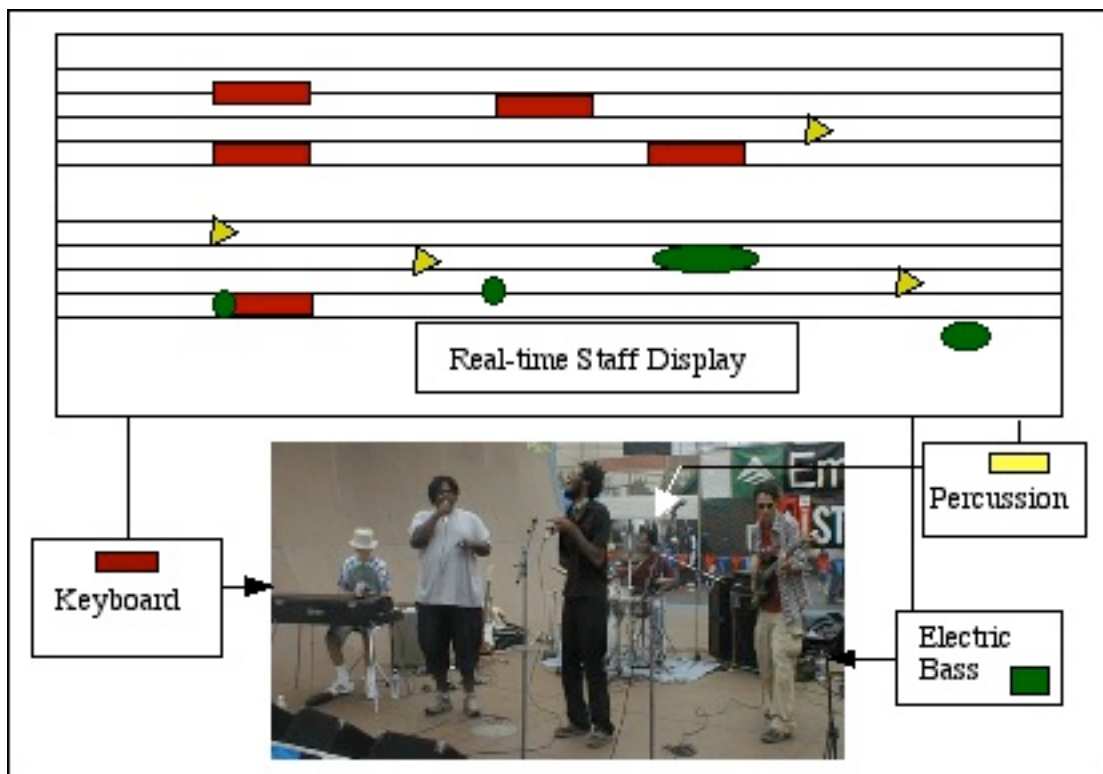
Figure 6.3 Depiction of Proposed Small Ensemble Trainer with Scrolling Real-time Display of Notes and Fingering Translations

Recording and review of musical events would be a useful feature for this trainer. This system would enable the group to discuss the rhythms, voicing, and balance between the various players. Individual improvisations like drum fills, embellishments, and chord inversions could be seen in context with other players. It is predicted the observers would

gain a deeper understanding of the nuances in the performance and improve their ability to work as an ensemble.

Expansions of this system could include customized views for each performer presented in individual HMD or monitor displays. Such a system would enable performers to look at the other players and acquire their musical notes in a manner similar to a group-based digital collaboration.

Hypothesis for a Small Ensemble Trainer

The Small Ensemble Trainer can facilitate communication of musical ideas, awareness of group dynamics, and development of group cohesiveness in improvisational multi-performer settings. This system can help performers by providing a real-time display of individual performances, displaying pitches, durations, and translations between different musical instruments. This system can also help composers in voicing parts, selecting tones in chords, and identifying new options for improvisations.

Synesthetic Visualizer

The Synesthetic Visualizer would create a visual display from MIDI data and real-time audio analysis of keyboard and percussion instruments. This display would focus on blending the multi-spectrum audio together with the pitch and intensity information found in the MIDI events. The system would attempt to recreate a synesthete experience of music by mapping audio parameters including texture, volume, sound envelope (attack, sustain, decay, release), panning, and reverberation effects into a tunnel or cylindrically oriented display. The glyphs would animate over time to simulate the effect of short-term memory recall of recent music events. Figure 6.2 depicts a sample visual display.
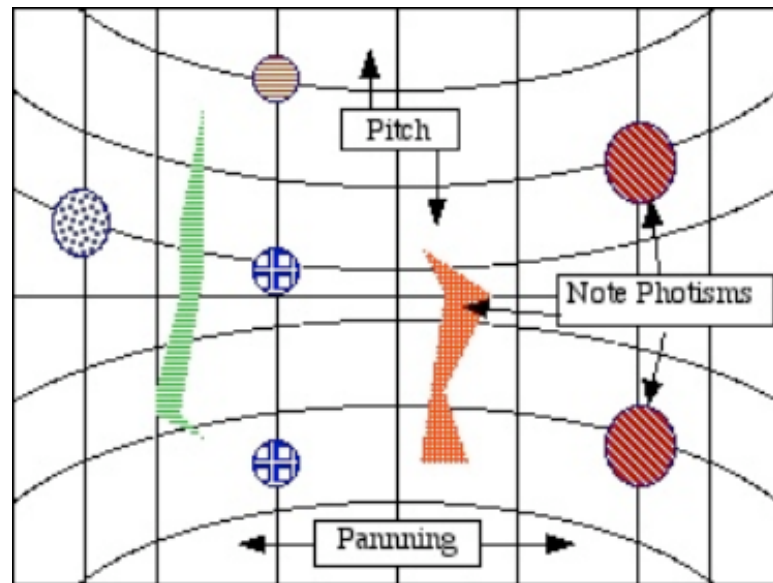
Figure 6.2 Synesthetic Visualizer Depicting Pitch and Note Relationships with Synesthetic Photisms

Performers could have control over the visual glyphs by choosing from an array photisms as described in the literature on synesthesia. MIDI device sliders and knobs could also be mapped to modify the color, texture, shape, and scale of the visual effects, giving the performers simultaneous control over visual translations of the sounds. Also, the system could be designed to automatically develop profiles for MIDI events by sampling and storing power-spectra for individual sounds.

This system can leverage the L-systems work done in the prototype to animate natural looking artifacts. Adaptive algorithms could be included to enable dynamic computer-driven visual interpretation of the musical events. Also, images and text strings could be loaded in as backgrounds and texture information.

Hypothesis for the Synesthetic Visualizer

A Synesthetic Visualization system can transform the real-time performance of multiple musical instruments through a combination of synesthesian-like visualizations and music principles to:

- *Provide an artistic and informative display for the audience.*
- *Serve as an inspiring visual creative partner for the performer.*
- *Provide composers with an expanded means of conveying their creativity through combined audio and visual presentations.*
- *Reveal music principles and nuances of a performance that may not be noticed by developing students and musically untrained individuals.*

Summary of Proposed Systems

The Computer Augmented Percussion Trainer and Small Ensemble Trainer seek to assist rehearsal and practice by giving the performers a visual translation of their performance. The trainers intend to inform the players of the nuances in the music and to assist them in improving their individual rehearsals and group awareness. These two systems are intended to represent how a synesthetic person might perceive and think about music. The trainer can demonstrate the use of synesthetic perception to monitor individual performance. Also, the trainers can help to communicate the nuances of the performance through the ensemble to serve as a source for creative inspiration.

The Synesthetic Visualizer aims to provide an artistic display for communication with the audience. It also seeks to provide the performer with direct control of the visual objects from their musical instrument. This system attempts to demonstrate the algorithmic transformation of sound as a synesthete experiences it. Further, it has the potential to serve as an autonomous creative and dynamic partner for improvisation.

The three proposed systems could be integrated together to simultaneously educate an audience and inform an ensemble. In this scenario, the Synesthetic Visualizer would be projected to the general audience while the Small Ensemble Trainer or Percussion Trainer are presented directly to the musicians on monitors or HMD systems.

6.3 Initial Experimental Plan

There are several options for experimental study of these systems. A complete list of prototype and potential systems includes: MIDIViz, Illuminator, Computer Augmented Percussion Trainer, Synesthetic Visualizer, and the Small Ensemble Trainer. Study tasks could involve measurement of several factors related to the amount of learning, impact, engagement, and rehearsal benefit. Examples of study tasks include:

- Determining which visual modules (transformations) best communicate the basic musical concepts of rhythm, pitch, harmony, texture, and tone.
- Studying the impact of the system with respect to various participant roles (performer, controller, audience, and composer), level of training (novice, expert), and graphics system configurations.
- Determine the effectiveness of the synesthetic mappings between sounds and visuals.
- Measure the emotional impact of the visual transformations on the performer and audience.
- Measure the performer accuracy with and without the visual assists in the percussion trainer.
- Survey the impact of the small ensemble trainer on communication between players.
- Measure the time that performers and audiences focus on various visual modules and/or their instruments.
- Measure them impact of the visual displays on the audience's awareness of musical nuances.
- Gauge the intuitiveness of the interfaces for configuring synesthetic visual elements.

The methods for gathering this information may include:

- Video taped interviews and rehearsals
- Likert scale surveys of subject's attitudes and perceptions of musical content.
- Computer instrumented user attention tracking in the immersive environment

- Computer instrumented recording of performances with and without the percussion or musical ensemble training visual displays
- Observation

The final study and survey methods will depend on the committee's feedback about which of the proposed applications will sufficiently demonstrate the objectives laid out in this dissertation. The final title and scope of the dissertation will also impact the study tasks to be performed. Also the projects must be chose with consideration for the author academic training and research background. It is necessary that he be sufficiently credentialed to propose, conduct the experiments and state results for the systems listed above.

CHAPTER 7. IMPLEMENTATION

This chapter documents the software development tasks performed after to the proposal presentation. These tasks include the development of the synesthetic visualizer modules and prototype ensemble training displays proposed in the Chapter Six. The synesthetic visualizer modules combine real-time audio and MIDI information together. Synesthetic photisms are animated and advanced mappings of color and sound are explored. The Computer Augmented Percussion Trainer and Small Ensemble Trainer displays combine the audio and note information to support and/or enhance a music ensemble's performance.

This chapter begins with a discussion of the data management techniques used to correlate the audio and MIDI data. The next section describes the synesthetic visualizer and ensemble support and their underlying algorithms. This chapter closes with some preliminary observations of audience responses to these modules and their early demonstrations.

7.1 Audio/MIDI Data Management

The first step in creating these modules was to develop a data structure to correlate the frames of real-time audio and the MIDI events. This structure is required because the MIDI and audio data are received separately and asynchronously. Multiple frames (typically 2-4) of audio are received in the span of a single frame. All queued MIDI and audio events are read once per frame update. The data structure stores and synchronizes approximately 300 samples of audio and MIDI events. This storage enables each visual module to access a few seconds of the combined music history.

The data structure consists multiple frame-records, each one representing the audio and MIDI state for each received audio frame. Each frame-record includes an array of 16 floating-point values to store the real-time audio power spectra and an integer array for 127 MIDI-note intensities. The system's read/write cycle proceeds by first reading all MIDI events, second, updating a MIDI status record for that time step, and third

processing the audio queue. The audio event data is inserted into a new audio frame and the current MIDI status is copied forward into the newly created frame.

Iterating over the data is straightforward and employs two integer indices. The first index provides access to the individual frame records, and the second index points to the individual note or audio spectra entries. Each visual module described in this section iterates over the full length of the audio and MIDI history and transforms the measurements into visual glyphs. A more sophisticated data structure would employ scalable data vectors, thread-safe iterations, and long-term data storage. However, this data storage design is adequate for this proof of concept study.

Data Management Limitations

This stage of development has revealed a major shortcoming in the audio and MIDI event transmission system. This shortcoming negatively impacts the performance of the ensemble support applications and is a result of early project decisions. The system was initially conceived to respond only in an immediate mode to incoming MIDI events.

The real-time audio was added on as a secondary development stage and follows its own independent data path on a separate socket connection. So, fusing the data streams together into a combined history revealed some significant design limitations. The first limitation is that neither the MIDI nor the audio events are time stamped. The MIDI-Transmitter and Fast Fourier Transmitter processes run independently and there is no common clock between the events. Secondly, because the events and audio frames are oddly interspersed and separately transmitted, it is difficult to tell precisely when the events occur in relationship to each other. Further, this audio/MIDI history does not support multiple MIDI channels and does not scale well to multiple audio or MIDI devices.

All of these issues stem from limitations in the underlying software architecture. There are some software engineering remedies to this situation. First is to develop or utilize a full MIDI implementation including timestamps and network distribution. Second, the audio and MIDI messages need to be interlaced and ordered in the receiving

data structure. Third, the MIDI and audio histories should be accessible through a data manager that can be iterated efficiently, dynamically scaled, and also archived/retrieved.

The design limitations have been worked around to prototype and study the ensemble support tools. While the timing disparities do not stop the modules from functioning, some of the events are incorrectly displayed. Namely, all of the events received in a frame loop are treated as if they happened simultaneously. This can result in events appearing closer or further apart than they actually occurred.

This performance issue has influenced the experimental plan for this research. The performance of the ensemble modules was not accurate enough to be demonstrated in a real ensemble setting. Rather the concept is presented and evaluated for its future potential if the performance issues were remedied.

7.2 Synesthetic Visualizer Modules

These modules combine MIDI, real-time multi-spectrum audio information, and Synesthesia research. The resulting visual objects respond to the music, modulating and shifting in real-time. Shape, color, and placement change in response to the combined data streams. Also, these modules incorporate expanded coloring schemes. Each module is depicted below followed by a description of the underlying algorithm, data inputs, and coloring methods.

Audio Modules

The first two modules do not utilize MIDI data, but they do implement concepts taken directly from the list of synesthesian accounts of photisms as depicted in Chapter Four.
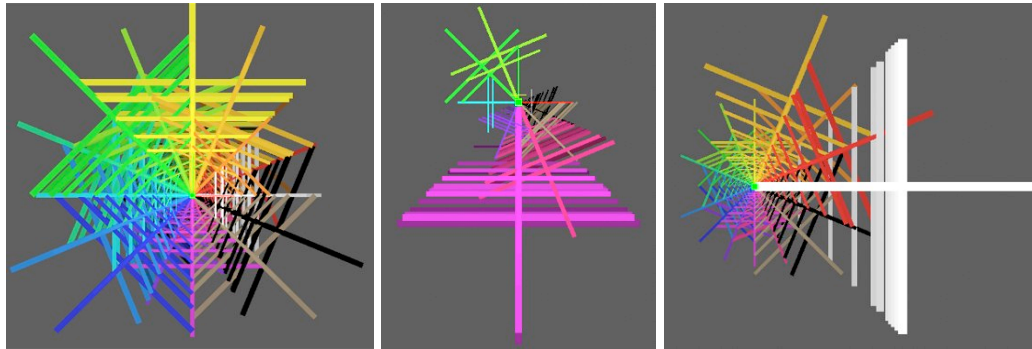
Figure 7.1 Central Radiations

Figure 7.1 above depicts scintillation, extrusion, and central radiation. This display consists of multiple layers, each layer corresponding to a step in time. Each plane is drawn according to samples of the audio in a given frame. The most recent layer is the nearest, and older layers are drawn further and further into the screen. Each layer consists of sixteen colored '+' signs that are distributed about the center point. The scale and translation of each '+' is directly proportional to the intensity of that EQ value. The colors come from a defined array of 16 values ranging from white to black and through an chromatic spectrum of color hues. The three images in figure 7.1 above indicate different responses to sounds. Sounds with high frequency bias range activate the lighter colored elements, and low frequency sounds activate the darker colored elements.

The Multi-Color ribbon, below, also uses the same coloring spectrum as above, but instead of mapping to individual scintillations, this module maps the intensity of each audio band to the thickness of individual bands. The Multi-Color Ribbon produces distinctly different images for different sound sets and sound variations are clearly visible. Audio samples are displayed with respect to time. The most recent sound samples are placed at the center of the glyph and propagate horizontally outward. The top and bottom of the figure are mirrored for visual effect.
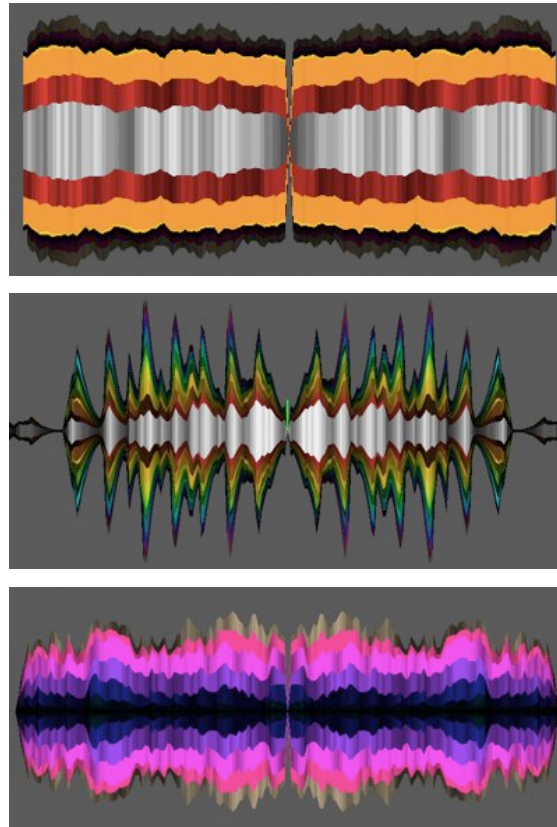
Figure 7.2 Multi-Color Ribbon

The dynamic characteristics of the sound are revealed through modulations of color and band thickness. In general, the bright and high pitch sounds result in brighter colors and lower pitch sounds take on cooler tones. The time propagation effect and result in curiously interspersed shape forms. Layered, contrasting, and rhythmic sounds are clearly visible. As an example, a drumbeat consisting of bass drum, snare drum, and cymbal sounds is translated into an alternation of three different color patterns. The propagation leaves a visual trace, helping the view recall at the sequence of rhythmic events. Also the envelope characteristics of the sounds are clearly indicated as well as a shift in sound tone over time.

Audio/MIDI Hybrid Modules

The next three modules combine both MIDI and audio signal data. The Blue-Matrix, Figure 7.3 was the first attempt to create shape photisms that modulated in response to the sound. This module features multiple rows of colored polygons that are mirrored about the vertically and horizontal axis. This module adapts Xenakis' technique for plotting sound frequencies against intensity on a 2D graph [Xenakis]. Xenakis' approach also proposed the addition of a third axis to display multiple instantaneous 2D graphs over time.
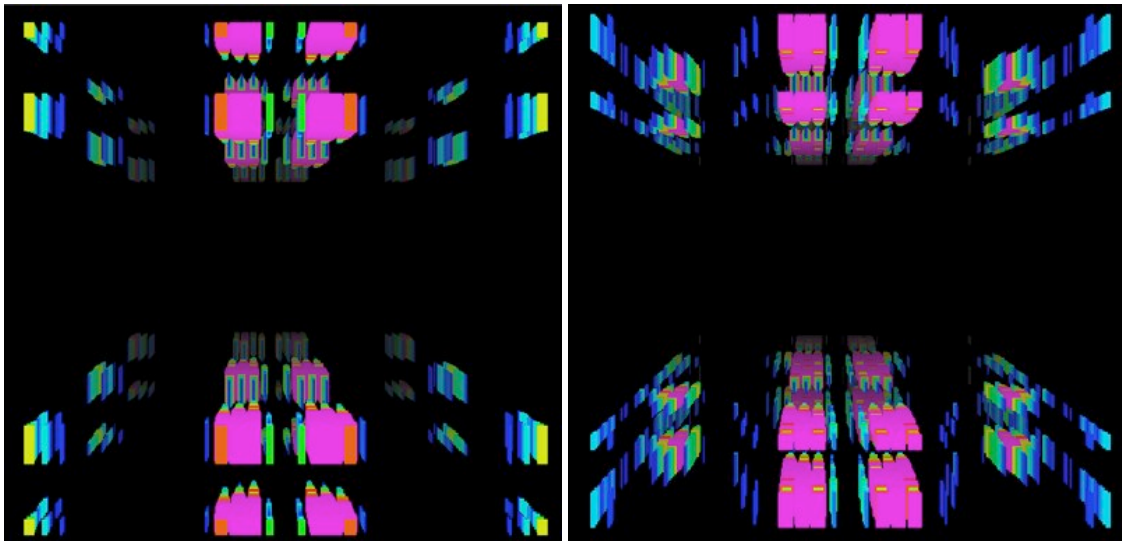
Figure 7.3 Blue Matrix

In the Blue Matrix, the more recent audio samples are located in the foreground and the older samples fade into the background over time. The vertical distance of a row from the center point corresponds to the pitch of the active MIDI events. The horizontal distance from center corresponds to the frequency of the audio spectrum. Low frequency bands are located in the horizontal center of the chart and high frequency bands towards the outside. The polygons' horizontal size corresponds to the intensity of the MIDI notes, and the polygons' width to the intensity of the audio frequency band.

Each polygon is colored using the same color spectrum as the musical river and lollipop from the Illuminator audio modules. Each polygons' color can range from dark

purple at a low intensity value, through violet, indigo, blue, green, yellow, orange, and red and then a lighter violet color at the highest intensity value.

This module's graphics are similar to synesthetic blobs and clusters. The sound of a MIDI note causes clusters of color to appear in the virtual space. Low sounds cause a dense cluster in the center of the screen. High pitch sounds create polygons along the outside edges of the screen. Melodic runs on the keyboard result in a stair case effect and rhythmic patterns cause the rows of data to pulse on and off, resulting in alternating gaps of polygon and not polygon regions.

The next module, titled EQ MIDI shapes, is shown in Figure 7.4. This module is similar in concept to the Blue Matrix depicted above in that the MIDI note pitch determines to the number and height of the rows. Rows are reflected symmetrically with low pitch being placed at the vertical center of the screen, and low audio spectrum at the horizontal middle. Multiple time steps are drawn and the visual history extends and fades forward into the scene.
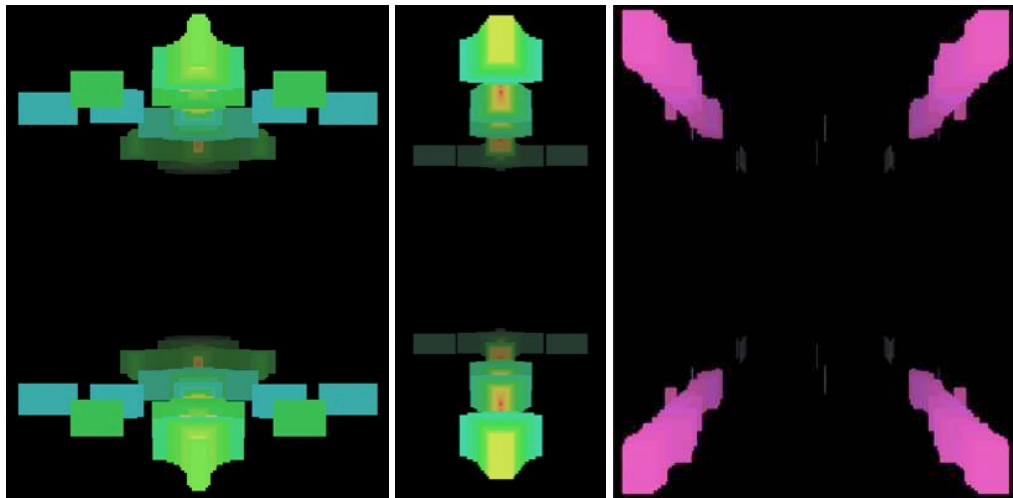

Figure 7.4 EQ MIDI Shapes

The EQ MIDI Shapes are constructed from multiple layers of flat polygons. The horizontal placement of the polygon is determined by a weighted average of the EQ band intensities. Each polygon is drawn near the position of the highest-level EQ peak. The width of the polygon is determined by the total sum of all the EQ-bands. The color of the

peak is driven by a weighted summation of the color spectra used in the multi-color ribbon. One time slice of a middle pitched sound event will appear as a centralized mid-to-dark colored object. Conversely, high-pitched sounds appear as thin brightly colored object on the perimeter of the screen.

The EQ MIDI shapes perform a distinct dynamic translation of sounds. The algorithm for computing the polygon locations results in a dynamic shuffling of the object positions. These synesthesia-like photisms appear to dance in response to the music, and fading time trails give hints about the history of the music. This module also exhibits the staircase-like shifting in response to melodic runs as the Blue Matrix above. Contrasting sounds create alternating photisms with varying colors and widths.

The last Synesthetic Visualizer module is titled Pitch Sound Arcs. These modules generate polygons that travel along hyperbolic trajectories. The polygons step one unit along the parabolic curve per audio frame. This module also uses reflective symmetry for visual effect. The origin of the arcs in the +x, -x, +y, and –y direction is controlled by the MIDI pitch of the notes. Low pitch notes originate form the center of the diagram and higher pitches further from the center.
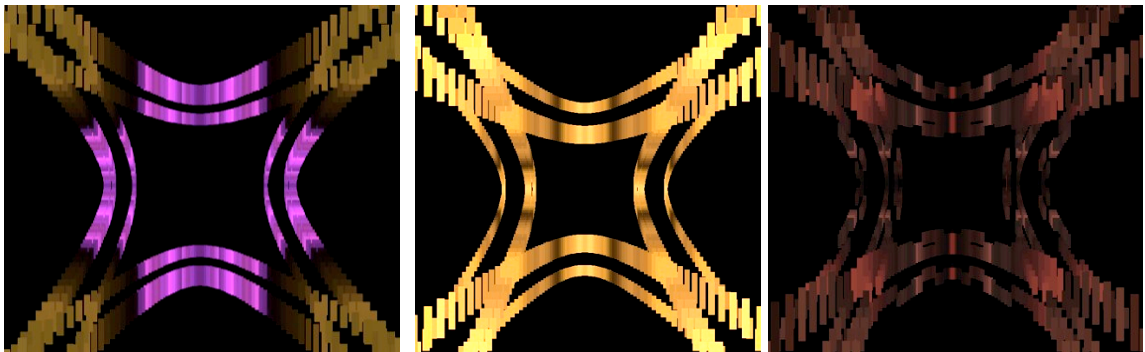

Figure 7.5 Pitch Sound Arcs

The multi-color ribbon colorings from above are applied to the pitch sound arcs as follows. Each real-time audio band is associated with a set color. The colors of the arcs are determined by a weighted calculation of all of the real-time audio values. The colors of the polygon arcs' modulate in response to changes in the audio input. The colored

bands ripple through various colors and appear to propagate in a wavelike fashion from the central source.

Music Mandala

The Music Mandala, Figure 7.6, is a hybrid display that combines both audio and MIDI information. The central portion of the display features concentric colored rings. The rings are colored using the same frequency to color associations as the multi-colored ribbon. The arrangement of the colors is similar to a clock face. Low-frequency bands are mapped to the 3 and 9 o'clock position. High-frequency bands are located at the 12 and 6 o'clock position. The brightness of the bands is driven by the intensity of the audio at that spectrum. The most recent time steps originate from the outside of the disc, and collapse towards the center. Also, a solid line indicates the loudest audio-band in a given time frame.
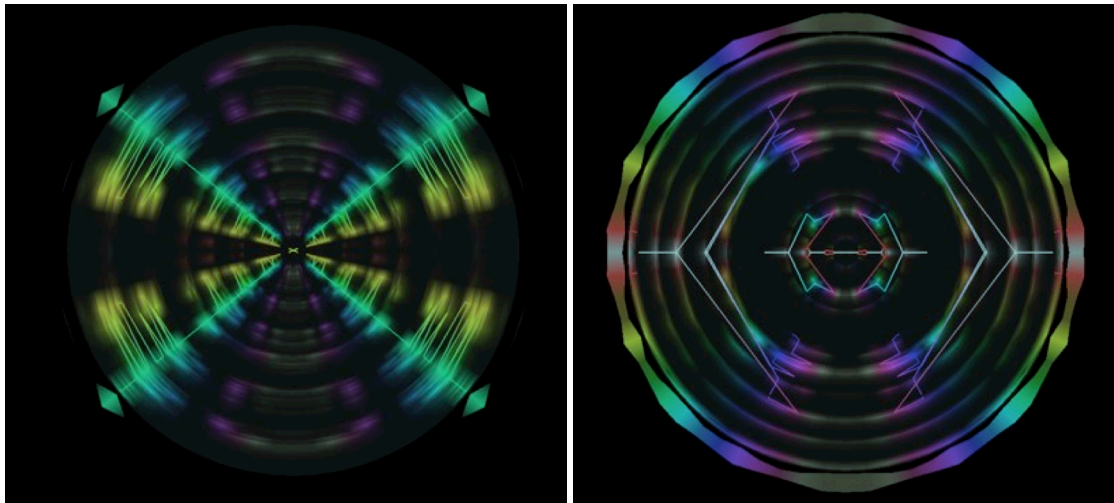
Figure 7.6 Audio Portion of the Music Mandala

Fig 7.7 shows the background of the Music Mandala. This background features horizontal colored lines that are driven by MIDI events. The lines originate from the upper and lower border of the Mandala and collapse towards the vertical center of the screen. The lines are colored according to the chromesthesia mappings used in the MIDI-

Viz Application. The intensity of the color is proportional to the loudness of the MIDI events. The colored bands are interspersed when multiple notes are sounded, resulting in the appearance of a blended color.
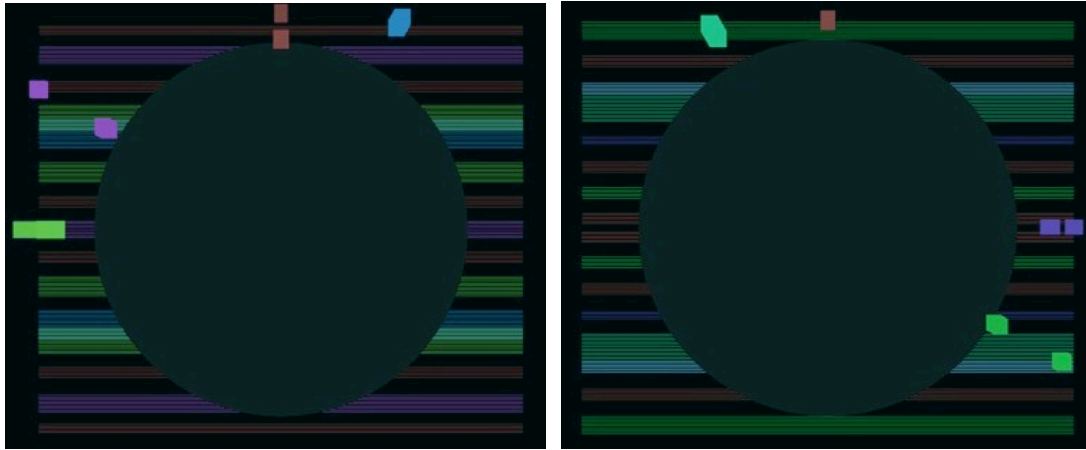
Figure 7.7 Music Mandala with MIDI Data. Cm7 chord, FM7 chord.

Figure 7.8 Music Mandala with Audio and MIDI Data

The border of the Mandala incorporates a ring that grows and pulses in response to the audio bands. The same mapping and coloring are applied to the concentric audio rings. Conceptually, this outer ring functions like an emitter for the audio spectrum rings which in turn collapse towards the center point of the mandala. The border also features twelve musical note emitters that are arranged in accordance to the 'Circle of Fifths.'

Rectangular objects emit from the nodes, scaled and colored in proportion to the intensity of the played notes.

The combined audio and MIDI coloring result in a slowly moving tapestry of musical data. The inner portion displays the audio characteristics of the sound. The MIDI events in the outer segment add a soothing backdrop of color and interpret changes in the chord structure of the song.

Summary of Synesthetic Modules

The modules presented in the first half of this chapter represent the initial steps towards a Synesthetic Music Visualizer. Common themes across all of the modules include colored photisms that are modulated or panned in response to the pitch, frequency spectra, and time of the musical events. The modules are reflective of the cross-modal associations found in the synesthesia literature. The modules are also an adaptation of Xenakis' method of plotting frequency ranges against note intensities. Specifically, radial symmetry and movement are reflected in the Multi-Band Ribbon. Amorphous 'blobs' and clusters are animated in the Blue Matrix and EQ MIDI Shapes. The Pitch Sound Arcs implement circular and geometric line photisms. The Music Mandala exhibits kaleidoscope-like symmetry and demonstrates scintillation and extrusion.

These initial modules could be generalized to implement an expanded Synesthetic experience. A visual module could be designed to enable selection of multiple photism attributes, as well as flexible data and color mappings. Currently no chord or rhythmic analysis is performed. High-level analysis could be integrated to further augment the visual presentation. One possible scenario could change the type of visual photism to reflect the harmonic quality of the chords. Chapter Nine gives a further description of a hybridization of all of the visual elements presented here and identifies future expansions of the Synesthetic Visualizer concept.

7.3 Ensemble Support Modules

These two modules demonstrate potential uses of technology assist in ensemble rehearsal. First proposed in Chapter Six, these modules capture a real-time music performance and translate it into a visual display. Designed to illustrate aspects of a synesthetes' experience of musical events, these tools aim to increase awareness of individual performances and facilitate communication between ensemble members.

The first tool, the **Computer Augmented Percussion Trainer** (CAPTAIN) is proposed as a means for improving the performance of a percussion ensemble. This system concept includes multiple percussion instruments fitted with MIDI triggers. A scrolling visual display is set in front of the ensemble to reveal the timing and intensity of the musical notes.
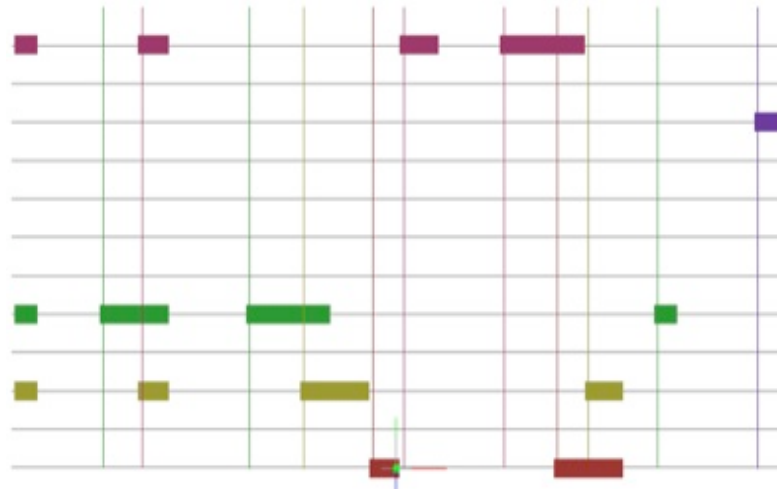
Figure 7.9 Computer Augmented Percussion Trainer

Figure 7.9 includes a screen capture of the prototype CAPTAIN display. Each line represents one percussion instrument and one MIDI channel. Notes are drawn as rectangular bars. New notes appear at the right hand side of the screen and scroll to the left. The horizontal width of the bars represents the duration of the notes, the amount of time between the MIDI NOTE_ON and NOTE_OFF messages. The vertical height of the

bars corresponds to the intensity of the note. Also, vertical lines are drawn indicating the onset of each note, enabling an intuitive visual comparison of event timings.

The **Small Ensemble Trainer**, shown in Figure 7.10, expands on the concept of the percussion trainer. A scrolling display presents all of the recently played notes across a multiple octave piano-roll. As above, the notes appear on the right hand side of the screen and scroll to the left. Bar heights are proportional to the intensity of the MIDI event and bar lengths are proportional to the duration of the note.
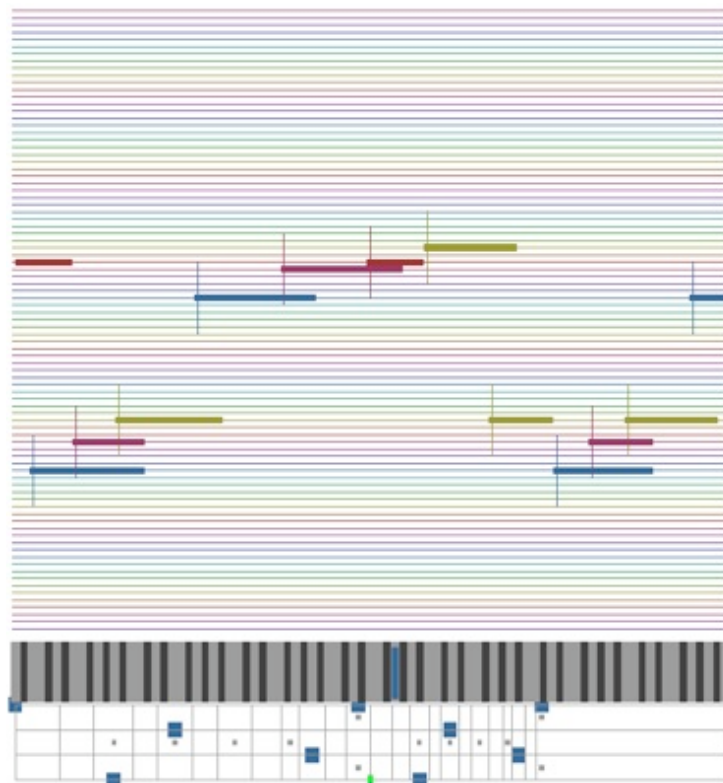
Figure 7.10 Small Ensemble Trainer

A full musical notation staff has not been implemented for this prototype. Instead a simple Chromesthesia colored staff has been created. Each note is mapped to a unique color just as in the MIDI VIZ modules. This coloring enables the audience to readily identify which notes are of the same pitch or that share an octave relationship. However it does not allow actual note identification as a musical staff would.

An additional objective of the Small Ensemble Trainer is to facilitate communication of musical ideas between ensemble members. A piano keyboard simulation and a bass guitar fret board display are incorporated to display all active notes. The scrolling display, keyboard simulation, and bass fret-board use the same Chromesthesia note coloring.

The last feature of this module is a scrolling print of the real-time audio information. The audio information is intended to increase ensemble awareness about group dynamics. This ribbon of data reveals the correlation between the MIDI events and the dynamic volumes.

The ensemble support displays are very preliminary and have some significant limitations. Data recording and review options are not supported. There is no true musical staff. There is also currently no distinction between multiple MIDI channels or instruments. Also, the data management limitations mentioned above cause some event timings to be displayed imprecisely.

Despite the shortcomings, these prototypes do provide a working display for conceptual evaluation of the small ensemble trainers. Early demonstrations of this system have shown that it can effectively communicate information about the number, timing, and duration of musical notes. The next section discusses the initial demonstrations of these displays to audiences.

7.4 Initial Observations

The Synesthetic Visualizer modules respond clearly and distinctly to the notes. The presence of multiple vertical rows of photisms conveys the fact that multiple notes are being sounded. However, the Synesthetic modules communicate very little information about pitches or harmonic structures. The Circle of Fifths in the Music Mandala does reveal some of the harmonic structure of the tones, especially chords, though the audience must know to interpret it as such. Observing the Mandala operating during the development has been interesting in that it presents both informative and

entertaining imagery. The background of the mandala seems well suited for illustrating legato bass lines or chord roots.

The Computer Augmented Percussion Trainer and Small Ensemble Trainer prototypes effectively demonstrate the intent of the modules. Both are able to receive and display real-time note information. The Small Ensemble Trainer reveals some information about the timing and pitches of notes. It also presents a working translation between keyboard and bass guitar instrument layouts. However, there are some serious technical limitations that prevent them from truly accomplishing the goals set for the systems. Those issues are related to the timestamps, frame update rate, and the distributed nature of the VR software.

The trouble with timestamps is related to the distributed nature of the VR applications and the software architecture used to transmit the MIDI data to the VR host. A decision was made early on to omit timestamp data from the MIDI data stream. This decision was made because the MIDI modules were initially conceived to respond immediately to musical events. Each note activates visual imagery in the subsequent VR draw cycle. The timing of the MIDI events was not particularly important because all MIDI events were transmitted, received, and rendered as quickly as possible.

The ensemble support modules were not an initial focus for this research. Music events, especially percussion, happen more frequently than the graphics update rate of thirty to sixty hertz. Extended timestamps are required to accurately display the information. This limitation was not corrected for two reasons. First, in a finalized implementation, the ensemble trainer displays would run on a single host without the additional overhead of network transmission. Second, the ensemble support modules are fundamentally different from the synesthetic photism–based focus of this project. Re-architecting the MIDI event messages and any core architecture issues was note necessary to evaluate the potential of the ensemble trainer modules.

Chapter Summary

Several visual modules have been implemented to demonstrate the concepts proposed in Chapter Six. The displays fall into two categories, Synesthetic Visualizers and Ensemble Support Tools. The Synesthetic Visualizer modules merge the real-time audio, MIDI events information, music visualization concepts, and reports of the experience of Synesthesia. Simulated photisms include multi-spectrum ribbons, amorphous blobs, modulating color translations and more. Initial presentations of these modules have revealed that they do begin to communicate the experience of Synesthesia to audiences. The Synesthetic Visualizer modules react consistently and distinctly to a variety of musical inputs.

The Ensemble support modules provide a real-time display of musical event timing, pitch, and intensity. Translations between keyboard and bass guitar fret-board are given to facilitate communication within the ensemble. Demonstrations of these modules have provided some insight to audiences and performers. They appear to effectively demonstrate the concept of an ensemble support tool despite inaccuracies in the event timings.

Initial observations and feedback about the modules have indicated preliminary strengths and weaknesses in this Synesthetic approach to music communication. The next stage in this research is to conduct a user study to further analyze audiences' perceptions and experiences. The next chapter presents an introductory user evaluation of these visual modules, their effectiveness in communicating musical concepts, and their potential for use in rehearsal and performance settings.

CHAPTER 8. EXPERIMENTATION

Over its three-year course of development, parts of this system have been demonstrated to several small and large (200+ participant) audiences through campus outreach events, arts performances, and live exhibitions with music ensembles. While no formal studies have been done with the large audiences, much useful information has been gained through audience comments and informal observation. This information has contributed to the project's development. It has also has indicated some key issues to be explored through a user study.

This chapter presents an initial study of the Synesthetic Music Experience Communicator system in operation. This study is intended to collect data about participants' experiences, to determine what aspects of musical sounds are being communicated to them through the visual imagery, and to gather their estimations of the systems potential in entertainment, education, and rehearsal settings.

This testing consists of demonstrations of the visual modules to small focus groups. A survey is administered and user responses are collected. This chapter identifies the objectives and design of the experiment. The presentation method, selection of the participants, and survey questions are described. The participant group is described. Their survey and written responses are statistically analyzed and summarized. Experimenter observations of the user test group are given and a final discussion of the results is presented.

8.1 Objectives

This study seeks to provide partial validation of the objectives for the Synesthetic Music Experience Communicator. This initial study will measure a variety of factors related to the system and seeks to gain preliminary insight about audiences' experiences. This study intends to:

- **Measure** audiences understanding of the representations of the musical sounds including: pitch, harmony, timing, intensity, and texture.

- **Determine** the audience's preferences for color mappings, types of geometric transformations.
- **Measure** the audience's emotional responses to the system modules.
- **Gather** participant's estimations of this system's potential in education, entertainment, and rehearsal settings.

This experiment is not intended to provide conclusive evidence for or against the use of this type of system. Rather, the intent is to gauge the participant's general responses to the various visual transformations and to identify areas for detailed future exploration.

8.2 Experimental Design

Key information about the experiment and associated survey forms is highlighted here. This study is planned for approximately 20 participants and requires one hour-long session to complete. Participants are not compensated monetarily for their participation and are free to exit the study at anytime. Participants are informed that they will be asked to view and evaluate visual modules that respond to musical sounds. Further, they will be given the option to play the musical instruments and/or make vocal sounds to active the visual modules. Participants are informed of the slight risk of personal embarrassment from performing with a musical instrument. However, individual musical ability is not the primary concern for this study.

The background survey asks participants about general information including age, gender, occupation, educational level, and majors. Participants are also asked to estimate their experience and/or training in music, arts, and technology. Specifically, participants are asked about their number of years experience with instrumental and vocal music performance, music ensemble leadership, music composition, and electronic music software. Participants are also asked about their visual and performing arts, computer graphics, and virtual reality.

This experiment divides the visual modules into four groups according to the type of data that drives each display and/or its function. The four module groups are:

Group A. MIDI-Only

Group B. Audio-Only

Group C. Audio-MIDI Hybrid

Group D. Ensemble Support Modules.

Participants are shown visual modules from each category and are asked to complete a short survey about their experiences. Most survey questions are rated on a 5-point scale. Each module group survey includes space for written responses. Subjects are free to ask questions and may request to revisit modules for clarification. Participants are given a short break between each category of visual modules. The questions from each survey section are summarized below and examples of the visual modules are shown for reference.
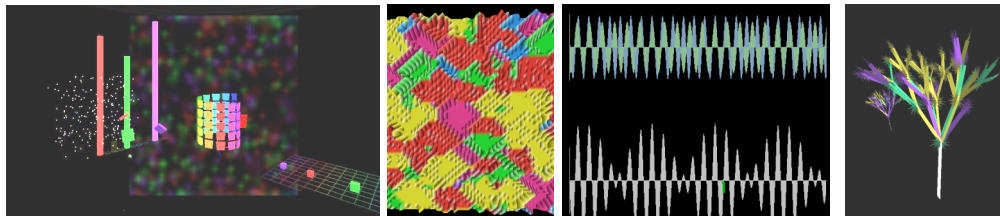
Figure 8.1 Group A MIDI-Only Visual Modules

Group A, the MIDI-Only modules are depicted in Figure 8.1. These modules demonstrate Chromesthesia, the unique pairing between colors and pitch classes. These modules ask participants about their understanding of the pairings between colors and tones. Participants are also asked to estimate the impact of the rotation, translation, and scaling transformations. This section also asks users how clearly the modules communicate pitch, harmony, and texture of the sounds. Subjects were initially asked to rank the modules that most fully captured their attention. This question was subsequently removed because of difficulties correlating the names of the modules with the visuals they had seen.
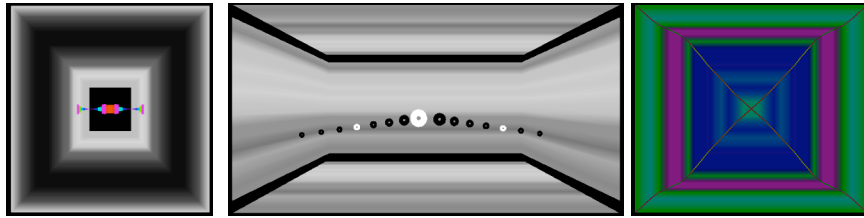
Figure 8.2 Group B Audio-Only Modules - Set 1

Group B, the Audio-Only modules are depicted in Figures 8.2 and 8.3. The Set 1 modules respond only to them most recent audio frame whereas the Set 2 modules incorporate audio history information. This section of the survey asks participants about their ability to discern the visual translations of low, middle, and high frequency sounds. Subjects are asked to rank their preference for the spectral colorings (multi-color, grey scale, and single color).
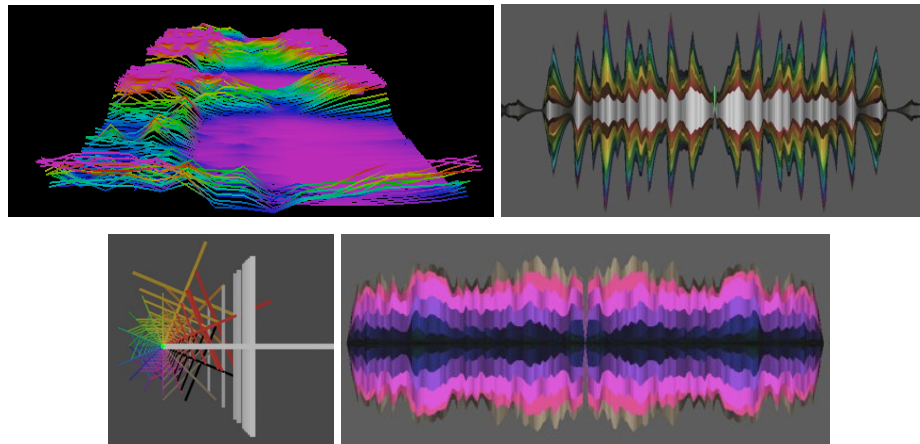
Figure 8.3 Group B Audio-Only Modules - Set 2

For the second set of modules (Figure 8.3), participants are asked how well the visual features reflect the flow of the music over time and whether the modules communicate changes in the input sound. Users are also asked about how clearly the modules communicate pitch, harmony, and texture.

Group C, the Audio_MIDI hybrid modules are shown in Figure 8.4. These modules most closely resemble the proposed Synesthetic Visualizer modules. Graphics placement, appearance, and coloring is driven by a combination of MIDI event pitch and

audio characteristics. This part of the survey asks how clearly the modules communicate pitch, harmony, and texture. Subjects are also asked how effectively the modules communicate changes in color in response to the input sound.
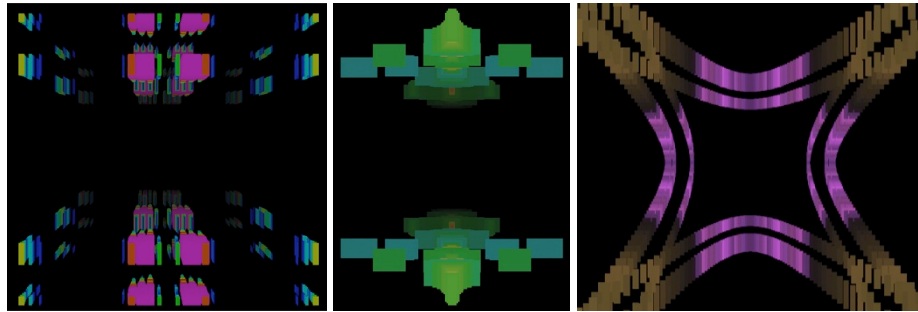
Figure 8.4 Group C Audio-MIDI Hybrid Modules

Group D, the Ensemble Support Modules are shown in Figure 8.5. This part of the survey asks the participants about the transformation of the musical events into a time-based display. Subjects are asked how effectively the presentation reveals information about the relative timing, intensity and number of notes. Subjects then rate the Small Ensemble Trainer and Computer Augmented Percussion Trainer's potential for use in rehearsal and practice settings.
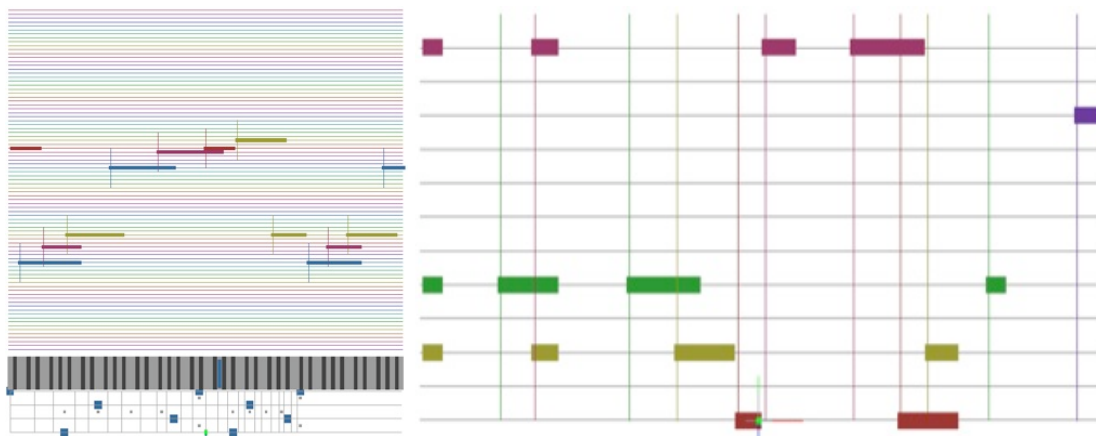
Figure 8.5 Group D Small Ensemble Support Modules

The last part of the survey asks subjects about their overall perception of the modules and the system as a whole. They are asked to estimate and elaborate about whether these modules are fundamentally different from existing music visualizers found in WinAMP or iTunes. Subjects are asked to estimate the benefit of the system in education, entertainment, and rehearsal settings. Subjects are asked if this would be an interesting visual addition to a performance. Also, whether this system has any emotional impact on them and whether the musical experience is made more accessible. Finally participants are asked to estimate the age range that they feel would best appreciate this type of system.

8.3 Description of the Participants

This study targeted a general audience of varied musical, artistic, and technological backgrounds. Survey announcements were forwarded through campus departments via email, online social communities, and flyers posted in campus music, art/design, and virtual reality departments. Twenty-five applicants were chosen to participate out of approximately fifty responders. Three sessions were held in a medium sized conference room with four, thirteen, and eight participants attending the sessions.

The participant group consisted of 16 males, 8 females, and 1 unspecified gender with an average age of 23.5 years old ($\sigma = 4.62$). Occupations were reported as follows: 22 students, 3 academic professionals, and 7 industrially employed. Education-wise, 12 had some college education, 8 held associates or 4-year degrees, and 5 had master's degrees. Majors fields of study included: Music Performance and Education, Engineering, Human Computer Interaction, Computer Science, Art and Design, Digital Media, Business, English, as well as other scientific and liberal arts fields. Participants reported considerable musical experience and less visual and performing arts experience. Table 8.1 gives a detailed view.

Table 8.1 Participant Experiences

Area	Number of Years	Standard Deviation (σ)
Instrumental Music Performance	8.5	7.3
Vocal Music Performance	6.5	8.3
Music Ensemble Leadership	3.0	4.9
Music Composition	2.4	4.1
Music Software Applications	1.4	2.6
Visual Arts	2.9	4.8
Performing Arts (Dance/Theatre)	2.4	4.2
Virtual Reality	0.8	1.4
Computer Graphics Applications	3.0	3.4

Participants indicated instruments with which they had novice or greater ability. Those tallies are as follows (out of 25): 18 (64%) had novice of greater ability with piano, 5 (20%) woodwind, 7 (28%) with brass, 7 (28%) with string, 4 (16%) percussion, and 3 (12%) reported other instruments of guitar, bass, and voice.

Participants were also asked to indicate the types of music and sound software applications they had used. The majority had experience with mp3 players and music visualizers. Those results are given in Table 8.2.

Table 8.2 Music/Sound Software Experiences (n=25)

Type of Software	Number	Percentage
MP3/Music Visualizers	20	80%
MIDI Composition, Sequencing, and Notation	10	40%
Audio Editing	9	36%
VJ/Multimedia	7	28%

Overall, the participant group has a strong amount of musical experience including some instrumental and vocal performance, leadership and composition. Some

participants had experience in visual arts, performing arts, and computer graphics applications. Participants reported very low amount of experience with virtual reality.

In addition to studying the entire population, the data have also been considered by splitting the population group into two sections. The division is made on the basis of the sum total of their number of years of musical experience. That figure was determined by summing the number of years of instrumental and vocal performance, ensemble leadership, music composition, and music software applications. The upper and lower halves of the population are given in Table 8.3, along with the average amount of music experience for the split upper and lower groups.

Table 8.3. Years Music Experience for Upper and Lower Experience Populations

Factor	Lower Half (n= 13)	Upper Half (n = 12)
Age	21.7 ($\sigma = 3.4$)	25.5 ($\sigma = 5.1$)
Instrumental Music Performance	3.9 ($\sigma = 4.3$)	13.5 ($\sigma = 6.8$)
Vocal Music Performance	0.9 ($\sigma = 2.5$)	12.6 ($\sigma = 8.2$)
Music Ensemble Leadership	0.9 ($\sigma = 1.8$)	5.2 ($\sigma = 6.2$)
Music Composition	0.6 ($\sigma = 1.3$)	4.3 ($\sigma = 5.4$)
Music Software Applications	0.5 ($\sigma = 0.9$)	3.0 ($\sigma = 3.7$)
Visual Arts	1.8 ($\sigma = 2.78$)	4.1 ($\sigma = 6.3$)
Performing Arts	0.2 ($\sigma = 0.4$)	4.8 ($\sigma = 5.2$)

The upper and lower half musical experience group data will be used to further investigate certain analysis performed in the survey results section of this chapter. Specifically, any interesting, unclear treads, or significantly different deviations will be compared between the upper and lower groups. It is anticipated that increased musical experiences will impact the individuals understanding of musical concepts and their interpretation of the visual modules.

8.4 Survey Results

Most of survey results are reported according to module group as described in the experimental design. Median and deviation values are reported for most survey questions. The following results pertain to the entire subject group unless otherwise indicated. Also, one set of survey measurements concerning pitch, harmony, and texture are compared between the first three module group conditions A, B, and C.

Group A. MIDI-Only Chromesthesia Modules

The first three questions asked subjects about the clearness of the mapping between colors and tones, the impact of the movements/rotations, and the impact of the scaled objects. Participants were asked to indicate their answers on a 5-point scale. The survey questions, means, and standard deviations are given below in Table 8.4.

Table 8.4 Chromesthesia Questions

#	Question	Mean	Std. Dev.
A.1	How clear was the mapping between the **colors and tones** (1 = not very, 5 = very much)?	4.0	1.0
A.2	How much impact did the **movement and rotations** have on you (1 = not very much, 5 = very much)?	3.8	0.9
A.3	How much impact did the **scaled** objects have on you (1 = not very much, 5 = very much)?	3.8	1.0

The next three questions about the Group A modules asked participants how well the graphics communicated the pitch, harmony, and texture of the music. These answers are compared against the Group B and Group C results. The comparison is given after discussion of the other Group B and Group C questions.

Group B. Audio-Only Modules

These modules transformed only real-time audio input into visual imagery. Participants were asked to estimate their ability to differentiate low, middle, and high frequency tones. They were asked how well the modules conveyed the flow of music over time, and how well the modules reflect the change in input sounds. Table 8.5 shows these three questions and the results.

Table 8.5 Real-time Audio Questions

#	Question	Mean	Std. Dev.
B.1	How well were you able to see the difference in **visual correspondences** between the low, middle, and high frequency tones (1 = not very well, 5 = very well)?	4.1	0.8
B.4	How well do the traveling lines give you a sense of the **flow of the music over time** (1 = not very well, 5 = very well)?	4.4	0.9
B.5	How well do the layers in the Multi-color Ribbon reflect **changes in the input sounds** (1 = not very well, 5 = very well)?	4.4	0.7

Question B.2 asked subjects to rank their preference for color mappings of the audio-spectra (1 = favorite, 3 = least favorite). The three choices were MULTI-COLOR, GREY-SCALE, AND SINGLE-COLOR spectrum. MULTI-COLOR spectrum includes a range of colors for a single variable, similar to what is found on a weather temperature map. GREY-SCALE varies from black to white. An example of a SINGLE_COLOR spectrum would range from dark-blue to white as seen in the water fountain. Subjects demonstrated a clear preference for the MULTI-COLOR spectrum with an average rank of 1.4 (σ = .71). SINGLE-COLOR and GREY-SCALE were closely ranked at 2.2 (σ = .75) and 2.4 (σ = .63) respectively.

Question B.3 asked subjects how well the walls in the Metal Room conveyed the sense of echo in a resonant chamber (1= not very well, to 5 = very well). The average

result was 3.0 (σ =1.14). This question was asked because the metal room was specifically designed to reflect an audio recording used in a virtual reality augmented dance performance. Participants had a pretty neutral response to this question. One explanation for this difference is that the music itself conveys the context through which the audience interprets the visualizations. In the performance, the music included loud metallic factory sounds. In the user test, generic sounds were played.

The last three questions in this section ask about pitch, harmony, and texture. The answers to these questions are given after the Group C modules.

Group C. Hybrid MIDI/Audio Modules

This section asked about the photism-like modules that responded to the input audio. The photisms were placed in the three dimensional space in response to the notes played on a MIDI device. These modules were implemented to demonstrate the proposed concept of the Synesthetic Visualizer. Questions C.1, C.2, and C.3 pertain to pitch, texture and harmony.

Question C.4 asks participants if the visual module with the arcs accurately conveys a color-shift in response to changes in the tone of the sound (1 = not very well, 5 = very well). Participant's results averaged 4.0 (σ = 0.9).

Group A. B. & C. Pitch, Harmony and Texture

The three questions about pitch, harmony, and texture were intended to measure participants' experience of these musical attributes in response to the sounds. Each group of modules handled the music data differently. To review, Group A modules responded to MIDI information only, Group B to real-time audio, and Group C to both real-time audio and MIDI data simultaneously. The results between the three sections are compared to consider differences between the three separate approaches.

Participants were asked in questions A.4, B.6, and C.1 "How well did the colored objects communicate information about the **pitch of individual notes** to you (1 = not

very well, 5 = very well)? Questions A.5, B.7, and C.2 asked "How well did the colored objects communicate information about **harmony** to you (1 = not very well, 5 = very well)?" Questions A.6, B.8, and C.3 asked "How well did the colored objects communicate information about the **texture of the sounds** to you (1 = not very well, 5= very well)?" The results and deviations are listed in Table 8.6.

Table 8.6 Pitch, Harmony, and Texture Answers for Group A, B, and C

Measurement	A – MIDI	B – Audio	C - MIDI and Audio
Pitch	3.8 (σ = 1.0)	3.9 (σ = 1.0)	3.5 (σ = 1.1)
Harmony	3.2 (σ = 1.3)	3.1 (σ = 1.0)	3.2 (σ = 1.0)
Texture	2.8 (σ = 1.4)	4.1 (σ = 1.0)	3.7 (σ = .74)

The data indicate that participants reported an equal response to pitch for both the MIDI-Only (mean=3.8) and Audio-Only (mean=3.9) modules. Their response for the hybrid MIDI-Audio modules was lower. All three questions shared approximately the same standard deviation. With respect to harmony, participants' responses averaged 3.1-3.2 across all three conditions. There was a greater deviation in responses about harmony in the MIDI-Only modules.

In response to the texture question, the participants voted the lowest for the MIDI-Only modules, while scoring almost full point higher for both the Audio_Only and the Audio_MIDI hybrid modules. The texture questions also had a greater variance in comparison to the questions about harmony and pitch. Above average for the Group A and below average for the Group C responses.

The differences in deviations of the results here prompted a further investigation as to the source of the variations. As indicated in the description of the population, the test participant results were split into two groups (n=12 and n=13) according to years of combined musical experience. The upper group had significantly more musical experience than the lower group as shown above in Table 8.3.

Differences were observed in the two groups' responses to the questions. Of specific interest were the cases where the difference in means was close to the deviation.

Table 8.7 below shows the participants' responses to the pitch, harmony, and texture questions. The table also features the scores for the more and less musically experienced groups.

Table 8.7 Split Population Answers to Pitch, Harmony, and Texture Questions

	Lower Group		All		Upper Group	
Measurement	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
A. Pitch	3.46	1.13	3.68	0.95	3.92	0.67
A. Harmony	3.46	1.27	3.20	1.26	2.92	1.24
A. Texture	2.85	1.57	2.84	1.40	2.83	1.27
B. Pitch	4.15	0.90	3.92	1.04	3.67	1.15
B. Harmony	3.08	1.12	3.12	1.01	3.17	0.94
B. Texture	4.08	0.95	4.08	1.00	4.08	1.08
C. Pitch	3.69	0.85	3.48	1.05	3.25	1.22
C. Harmony	3.23	1.24	3.20	1.04	3.17	0.83
C. Texture	3.62	0.65	3.72	0.74	3.83	0.83

From this table, potentially significant differences are found in the subject's responses to the Group A Modules' Pitch and Harmony questions. In the Group A question about pitch, the upper group responses for pitch were 0.5 higher than the lower group with a smaller deviation for the upper group. Also, with respect to harmony, the more experienced group showed a response that was 0.5 lower than the less experienced group, though the deviation on this question was much higher in general.

In the Group B questions, the upper and lower groups agreed on their responses to the harmony and texture questions. Group B answers for pitch were 0.5 lower for the more musically experienced group than the upper group. In the Group C modules, the more experienced participants gave lower rankings for the pitch and fairly even scores for harmony and texture.

Group D. Ensemble Support Tools

The participant group was asked to evaluate the ensemble support tools and their potential to improve music rehearsals. The prototype Computer Augmented Percussion Trainer (CAPTAIN) and Small Ensemble Trainer (SMET) were presented to the groups and some music was played on the MIDI keyboard. Subjects were specifically directed to consider the modules as if they were projected on a screen or in a 2D real-time display in front of an ensemble. The timing inaccuracies were demonstrated and the potential remedies were discussed.

The first question about the Small Ensemble Trainer focused on the piano and guitar displays. This question asked if the simulated keyboard and fret-board layouts present **informative** translations between instruments (1 not very informative, 5= very informative)? The population response to this question was 4.50 ($\sigma = 0.7$).

The next three questions asked participants about whether the displays helped them to determine the relative timing, intensity, and number of notes played. The final questions asked them to estimate the displays' potential to facilitate small ensemble rehearsal. The results of those questions for both CAPTAIN and SMET are presented in Table 8.8.

Table 8.8 Ensemble Support Tool Question Responses

Question	SMET		CAPTAIN	
	Mean	Std. Dev.	Mean	Std. Dev.
Does the display help you determine the **relative timing** of the notes (1 = not really, 5 = yes definitely)?	4.36	0.91	4.48	0.65
Does the display help you determine the **intensity** of the notes (1 = not really, 5 = yes definitely)?	4.08	0.95	3.88	1.05
Does the display help you determine the **number of notes** played (1 = not really, 5 = yes definitely)?	4.32	1.03	4.72	0.61

Question	SMET		CAPTAIN	
	Mean	Std. Dev.	Mean	Std. Dev.
Rate the potential of this display to **facilitate (small ensemble/percussion) rehearsal** (1 =low, 5= high)?	4.40	0.76	4.68	0.56

The participant group gave high marks for the question about the both displays of relative timing. Scores were lower for the displays' ability to represent note intensity measurements. Subjects indicated that the percussion trainer was better at representing the number of notes that had been played. They also indicated a greater potential for the percussion display to facilitate rehearsal. There were not many major differences in responses between more and less musically experienced groups for most of the questions with one exception. The more musically experienced group saw greater potential for facilitating musical rehearsal reporting means of 4.58 and 4.23.

General Perceptions

The last set of survey results are from the subjects' general perceptions of the system in operation. Subjects were asked three sets of questions in this section. The first set asks if the modules displayed in this study were fundamentally different from those found in common music visualizers like WinAMP and iTunes. They were also asked whether this system would make an interesting visual addition to a live music performance, whether they experienced an emotional impact from the visual modules, and whether this system made music more accessible to non–musicians. Table 8.9 summarizes those results.

Table 8.9 General Perceptions Averages for Diverse Factors

Question	Mean	Std. Dev
Do these modules represent a **fundamentally different way of visualizing music** than is currently used in mp3 visualizers like WinAMP or iTunes (1 = not very different, 5 = very different)?	3.90	0.83
Do you feel this would provide an **interesting visual addition** to a live music performance or an arts performance (yes=5 or no=1)?	4.76	0.83
Please estimate the amount of **emotional impact** these modules have on you (1 = low, 5 = high)?	3.32	1.11
Do you feel this system makes the musical experience **more accessible** to non-musicians (1 = not very much, 5 =very much)?	3.88	1.05

These four factors listed above give good insight into subject's overall perceptions of the system. First, the subjects do indicate that this Synesthetic approach is a somewhat different way of visualizing music with an average score of 3.9. There was strong agreement (4.76) that this would provide an interesting visual addition to an artistic performance. They agreed to a lesser extend that this system could help make the music experience more accessible for non-musicians. The more-musically experienced subjects had a slightly higher estimation of accessibility than less-musically experienced subjects with averages of 4.1 ($\sigma = 1.2$) and 3.7 ($\sigma = 0.95$) respectively. Also, the subjects indicated a pretty neutral emotional impact for the modules.

The second set of questions asked participants to estimate the potential benefit of this type of system in education, entertainment, and rehearsal. Table 8.10 indicates those results.

Table 8.10 Benefit Estimates for Education, Entertainment, and Rehearsal

Question	Mean	Std. Dev.
Please estimate the potential benefit of this system for **education** (1 = low, 5 = high).	4.24	0.93

Question	Mean	Std. Dev.
Please estimate the potential benefit of this system for **entertainment** (1 = low, 5 = high).	4.48	0.96
Please estimate the potential benefit of this system for **rehearsal** (1 = low, 5 = high).	4.20	0.71

Here the more musically experienced group saw greater potential for education than the less trained group, averaging 4.4 ($\sigma = 0.8$) and 4.1 ($\sigma = 1.0$) respectively. Also, the less musically experienced half of the population gave higher potential for entertainment than the more experienced group, averaging 4.6 ($\sigma = 0.7$) and 4.3 ($\sigma = 1.2$). These differences are within the standard deviations but may still indicate some relationship.

The last question asked participants to specify an age range they felt would best appreciate this type of application. Participant responses ranged from 0 to 100+ years old with an average low of 6.7 years old and a high age of 66.3 years. The most musically experienced sub-group indicated a higher upper-age range than less trained at 73.2 and 60.0 years respectively. This concludes the survey results of the study. The last section of this chapter integrates these survey results with the other findings in the study.

8.5 Written Responses

Subjects had the opportunity to elaborate on their answers in each section. Key comments from those written comments are given here. Comments are organized by survey section and the respondents' language is preserved as closely as possible.

Group A. MIDI Only
- The bell tree, with its spiraling orientation of musical notes, reflected the aural events.
- The L-systems trees were of greater visual interest than the smaller scaled box objects.
- The frequency plot did a good job of communicating the mathematics of pitches.

- Multiple participants requested the addition of additional labels, octave, and register information.
- Objects that demonstrated persistence were seen as more effective than those that flashed briefly and/or returned to an inactive state.
- More movement would have had a greater impact.
- Playing a more familiar tune would provide a stronger association.

Group B. Audio Only

- Sometimes it was difficult to tell whether the visual modules were reacting to pitch, volume, or texture.
- The echo chamber would not have been identified as such without the explanation. (This perhaps is due to the music that was played along with the module).
- The modules show the flow of music for short time periods only. It would be nice to have a longer history an/or the ability to review.
- Some sounds appeared to trigger both high and low frequency response.
- Traveling lines and multi-color ribbon were effective though some users wanted to see differently oriented motions (horizontal vs. vertical).

Group C. MIDI and Audio

- These modules were more difficult to comprehend. Hard to determine what was being measured and how.
- The arcs did show the texture, harmony, and pitch of the sounds, though the color variance could be a bit greater.
- One subject reported that the hybrid-audio MIDI shapes provided the most information. The other modules had too many parts to easily distinguish the various parts.
- These modules best displayed the changes in pitch (across multiple octaves).

Group D. Ensemble Support Tools

- A single note position on a fret-board was preferred to multiple positions. Too many were difficult to extract.
- The modules need faster processing speed, better resolution, as well as storage and retrieval ability.
- A metronome, common notation staff, and note duration markings would clarify these modules.
- The ensemble trainer doesn't show how multiple instruments interact with each other and it is unclear how well it will scale to larger ensembles. Too many notes at one time could be hard to distinguish.
- This module would not be so suitable for an ensemble with fully scored music and would typically not be available for an actual performance.
- This module could facilitate an ensembles learning. It is a good concept for facilitating rehearsal if the timing can be related between all instruments.
- The Small Ensemble Trainer with a pre-programmed score integrated could have much potential.

General Perceptions

The first question in this section asked participants to evaluate whether these modules were a fundamentally different way of visualizing music than was found in commercial mp3 visualizers. The following are the participants' elaborated comments.

- The color in these modules depends on pitch, as opposed to randomly chosen in certain visualizers. Meter is not used in mp3 visualizers. The visualizers also use the entire amplitude of the sound, or they focus on the beats and intensity, less so than the actual notes and pitches.
- In the case of WinAMP/iTunes, the aural and visual are not connected, and do not interact. It is like two people talking with out listening to each other. With these modules, there's a mutual conversation, a relationship.
- These modules have a more creative outlook on musical depth, pitch, and sound.
- These concepts were a unique departure from current programs.

- The main difference is the indication or mapping of notes to colors.
- It is easier to distinguish pitch, tone, and intensity than in WinAMP/iTunes.
- These modules are more descriptive of the physical sound waves that music produces. This should help people intuit the concept of sound waves more easily.
- The correspondence between the sounds and images is much more compelling. The relationships are clearer, but still very abstract. Better representation of intensity and harmony.
- The modules are creative, but the differences aren't entirely apparent. The person understands that the modules distinguish notes and rhythms, but the intensity seems to be covered.
- Some techniques are similar but this is geared more towards being an instructional tool. These modules appear to represent more information and give more attention to discrete notes/sounds. These modules show/bring out more data about the music.
- The tree and the ensemble display really stand out (as being different).

The last section of the general perceptions survey invited participants to report any additional comments pertinent to the other general survey questions. Participants wrote:

- Look forward to seeing them in action at a show.
- Perhaps in the harmony questions …[modules you]… could include feature some discordant sounds to show the difference.
- This might be a good learning tool for early stages. Might help younger children, to help them speed up the learning process when learning a new instrument or how to read music. Exciting instructional potential.
- This program would be a good bridge for people who currently play by ear.
- Very interesting and informative to watch. This participant would have liked more time to experience the modules and better understand the questions.
- The motion of objects seems to register more than only colors to reflect the input sounds.

8.6 Experimenter Observations

Participants were encouraged to explore the visual modules through the piano keyboard or by making vocal sounds to activate the modules. Some participants experimented just to find out what type of visual responses the system would give. Some participants made extreme and odd sounds to trigger various visual effects from the multi-spectral ribbon. Other participants were interested in trying out the music instruments.

Some portions of the initial experimental design did not proceed very smoothly. The number of modules and display durations had to be adjusted for each group. One survey question was omitted after the first group of subjects was tested. In that question, participants were instructed to rank the MIDI_Only modules according to their preference. Unfortunately, users had seen too many visual objects and they were not able to accurately name them. Reference sheets had been created to help them identity the modules but using them was too time consuming.

The choice of music played during the experiment has an impact on the participant's understanding of the graphics modules. One subject commented that a familiar melody or repeating patterns would help the audience understand the visual display. This comment can be considered from at least two different perspectives.

First, the overall research approach was to make visual modules that responded to a wide variety of input sounds without any prescribed notions about the type of music that would be played through the system. This approach closely resembles the author's experience of synesthesia. Each sound is mapped to a distinct set of colors or motions. With unpredictable sound input, all subjects can do is observe the instantaneous response of the graphics modules.

From a second perspective, the audience predicts the motion of the graphics in response to their expectation of the music's progression. Without a familiar melody, the audience does not know what visual events to anticipate or where to look for them in the visual scenes. The pairings used here may or may not be suitable for or some audience members. Worse yet the modules could be counterintuitive resulting in significant

cognitive dissonance. It is possible that the audience may never fuse together the musical events with the visual translations.

In regards to the overall experiment, it seems that the vocal and/or keyboard exploration gives the audience a better chance to explore the system than passive observation. The smaller group from the first test session did the most exploration with the musical instruments. Those group members were very comfortable around each other and had performed together in a music group. A smaller test group size may encourage individuals to experiment more directly with the modules and hopefully increase their understanding of the transformations.

8.7 Discussion

The results of the study reveal several bits of key information about the audience's experience and their preferences for the audio-visual transformations. This section summarizes the general results of the study. The design of the study, participant group, survey results, and the written comments are all combined to provide an integrated view of the knowledge gained in this process.

The majority of the **participants** were undergraduate and graduate students. They had diverse educational backgrounds and were experienced with instrumental and vocal music performance. The group had a fair amount of music ensemble leadership, music composition, computer graphics, and visual arts experience. Subjects had a diverse set of instrumental skills. The majority indicated novice or greater ability on the piano and other ensemble instruments were well represented. Subjects had experience with software applications relevant to the study including music visualizers and music composition tools. It appears that the group was well qualified to understand the content and concepts of the audio-visual transformations presented in this study. Further, they are close to the target audience age group for an augmented live music performance.

The **survey data** and **user comments** provide useful feedback about this synesthesia-based approach to music visualization. Participants indicated that the chromesthetic pairing of color and tone were clearly understood. The moving, rotating,

and scaled objects had some positive impact on the participants. The visual correspondences between low, middle, and high frequency sounds and the visual graphics were quite clear. The audio modules provided a strong indication of the flow of music over time and reflected the changes in the input sounds. Participants indicated a preference for multiple-color spectrum mappings. Participants also noted that the hybrid MIDI-Audio modules were effective in conveying a color shift in response to input sounds. It was revealed that the metallic room, though designed to reflect the ambient impression of a mechanical factory, did not convey such when the audio source did not provide the appropriate contextual queues.

The **three-way comparison** between the MIDI_Only, Audio_Only, and the Audio_MIDI hybrid modules revealed that participants' observation of pitch, harmony, and texture varied across the three conditions. Participants reported perceiving the most pitch information from the Audio_Only and MIDI_Only modules. The combined data sources in the hybrid modules appear to interfere with participant's detection of pitch information.

Participant responses were neutral with respect to harmony detection with all three-module groups receiving approximately the same average scores. There was a significant difference in the detection of texture information. The MIDI_Only modules revealed the least amount of texture information while the Audio_Only modules revealed far more. The hybrid modules revealed texture information effectively.

The **ensemble support modules** were extremely well received and many comments suggested their applicability to the real world. Participants strongly indicated that the modules provided informative translations between musical instruments. These modules were effective at revealing the relative timing and number of notes being played. The modules were a bit less effective at revealing the intensity of the modules. Overall, the concepts were predicted to have high potential for facilitating rehearsals. User comments indicated the percussion trainer had more potential for realistic use in a rehearsal setting. The concern was that a small instrumental ensemble would create too much information to be presented effectively.

Subjects reported that these modules were a **different way of visualizing music** than was applied in current music visualizers. User comments indicated that this approach was more concerned with revealing the actual notes and information content of a performance. Further, subjects indicated that this approach had strong potential for use in educational and instructional settings. Subjects felt very strongly that these visual modules are suitable for entertainment and would be an interesting visual addition to a live performance. The displays are moderately effective at making music more accessible to non-musicians. The modules are also suitable for a very diverse range of participants.

Subjects did not report a very strong **emotional impact** in response to the visuals. This may be due to the type of music that was played during the tests or it may have to do with the correspondence to the visual graphics. One way to isolate this factor would be to give the participant greater control over the sounds being played. Another way would be to change the appearance of the photisms to match their individual perceptions.

This concludes discussion of the results of this study. These findings have indicated certain strengths and weakness of this synesthetic approach. The study has also indicated potential uses for other evaluation methods in similar projects. This chapter concludes with a discussion of those expanded methods.

8.8 Alternative Study Approaches

Different types of studies could be employed to provide greater insight into users' perceptions of the visual modules. A more psychologically oriented study could perform a tightly controlled investigation of individuals' responses to isolated visual events. A synesthesia research approach could be used to develop full profiles for individual's color and shape preferences. Giving an audience member or musical performer such a test could enable more precise configuration of the color spectra, note-pitch correspondences, and visual effects. The result would be a custom visual presentation that is uniquely tailored to each performer or audience member.

The system could be studied in a true performance environment with multiple musicians connected to multiple display modules. Another option would be a single

virtual environment with multiple visual elements portraying multiple instruments performing simultaneously. In such a complex scene, users' attention may shift from object to object in response to the changes in music. Users focus shifts could be measured through the use of head or eye tracking systems. Conversely, laser pointers or ambient crowd response might be used to measure participant's responses to visual effects.

As user comments have suggested, longer audio histories will enable viewers more to re-scan recent music events. It would also be interesting to study observers' visual scan patterns for those extended histories. Sound features like echoes, repetitions, complex figures, and/or changes in sound level could cause users to rescan portions of the visual image.

User attention tracking could also provide feedback to drive an evolutionary system. Visual module parameters could be adapted in response to both user attention and the incoming music. In that type of two-way adaptive multimodal system, the performer adapts to the graphics and the graphics adapt to the performers.

CHAPTER 9. CONCLUSION

Developing the *Synesthetic Music Experience Communicator* (SMEC) has truly been an enlightening experience for this author. This project fuses together a technology-centric education, a lifelong love of music and computers. This project leverages the author's personal experience of synesthesia. It also incorporates insights from his cultural heritage and spiritual upbringing.

From early childhood experiences of banging on pots and pans in the kitchen onward, sound has always been accompanied by a visual presence. This experience of synesthesia continues to prompt musical exploration. The consistent and dynamic transformations between sound and mental imagery continue to inspire new compositions and sonic landscapes. Musical memories are encoded and recalled with the same visual accompaniment. New genres and sounds bring new visual sensations to mind.

Early experiences with Apple IIe graphics programming and music composition software sparked an interest in developing synchronized experiences of sound and imagery. Observations in music education, rehearsal, and performance settings have prompted inquiry into ways of facilitating and enhancing music rehearsals. Finally, the discovery of synesthesia and its associated literature have suggested expanded means of illustrating this experience.

This dissertation has sought to advance current techniques in the realm of multi-modal research. This work has explored many ways of representing musical concepts in computer graphics systems. It has recombined and incrementally expanded prior concepts. Additionally, music performance experiences, reflections of synesthesia, inspirations from nature, creative collaborations, and numerous discussions have served to refine the approach.

9.1 Chapter Review

This dissertation has documented several key aspects in the development and exploration of a Synesthetic Music Experience Communicator. The process is illustrated

over several stages including initial concept development, background research, prototype development, proposal, implementation, audience evaluation, and conclusion. The early chapters establish the relationship between this work and current research in the related background areas. The middle chapters document the prototypes, proposal, and development plan for the hybrid audio/MIDI modules and ensemble support tools. The last chapters document the hybrid and ensemble support displays, and user evaluation system. Each chapter is summarized below.

Chapter One, **Introduction**, began by presenting initial observations for potential areas of integration between technology and music. Included among these observations are the uses of technology to augment a musical performance, to improve communication within an ensemble, to teach musical concepts, and to demonstrate synesthesia. The key research areas of human computer interaction, music, and synesthesia were introduced. Objectives for this research were listed and general hypotheses were given.

Chapter Two, **Human Computer Interaction**, established fundamental principles relevant to in the development of an interactive music experience. A variety of computer hardware, interaction methods, and display platforms were considered for their potential use in music performance and rehearsal settings. Research into audio perception as well as multimedia and multimodal systems were given. As a whole, this chapter served to establish the means by which this project would be developed and to show its connection with existing HCI research.

Chapter Three documents an intensive review of **Music Education and Technology** literature including music concepts, teaching methods, and a sample philosophy of music education. Music technologies such as interactive learning tools and software systems for composition and live performance were described. A variety of electronic music instruments and selected music technology research projects were presented. The main contribution of this chapter was to illustrate the connection between this dissertation and existing music education methods and technology research.

Chapter Four presented several aspects of **Synesthesia**. Individuals' experiences of synesthesia were presented to provide the reader with a general understanding of this

phenomenon. Other aspects of synesthesia including experimental research results, diverse historical perspectives, and divergent contemporary theories were listed. The chapter then provided a view of synesthesia inspired works in visual arts and music. Some of the more common synesthetic correspondences between sound, emotion, and sensation were given. Taken as a whole, this chapter provides guidelines for implementing synesthesia-like audio-driven imagery. The chapter also hints at the fundamental relationships between human cognition, perception, and expressions.

Chapter Five documents the **Development of Virtual Reality Prototypes** that were built to test initial concepts for this work. The goals and implementation of the chromesthesia-based MIDI-only and audio-only modules are discussed. Both sets of modules were developed incrementally and were guided by feedback from early public demonstrations. Both sets of modules progressed from simple scenery towards increasingly complicated transformations. The MIDI Viz modules demonstrate some musical and mathematical concepts as well as Chromesthesia, the pairing of colored pitches and tones. The Illuminator! modules explore transformation of multiple power-spectra processed audio into unique regions of visual displays. The chapter ends by presenting some initial observations of these modules in operation and by listing several areas for potential future expansion.

Chapter Six presents a **Research Plan** for developing the synesthetic visualizer modules and the two ensemble support displays. The Synesthetic Visualizer concept combines the musical event data and the real-time audio information to emulate synesthetic photisms. The Computer Augmented Percussion Trainer and the Small Ensemble Trainer translate aspects of the musical performance into a scrolling visual display. The Small Ensemble Trainer presents a transformation of a note-fingerings between piano and bass guitar.

The dissertation proposal version of this chapter included a schedule for completion of the research plan. Experimental and evaluation methods were listed and avenues for publication were also listed. Those items have since been expanded into the next two chapters on implementation and experimentation.

Chapter Seven discusses the **Implementation** of the modules presented in the preceding research plan. Several synesthetic visualizer modules were developed to demonstrate the concept of animated and modulated photisms. Each module takes a slightly different approach to combining the MIDI and audio information. Proof of concept versions of the percussion trainer and small ensemble modules were developed. Regrettably, the MIDI communications architecture and performance limitations did not enable a very precise display of the musical events. This chapter discusses the performance issue and other initial observations of the synesthetic and ensemble modules. Inspite of the limitations, these modules were deemed sufficiently functional for conceptual evaluation.

Chapter Eight presents user **Experimentation** that was conducted to gather user impressions of the Synesthetic Music Experience Communicator. The motives, rationale, and experimental design are given. A description of the participants' backgrounds and arts-related experiences is given. User survey responses and written comments are examined, compared, and discussed. Certain factors considered in light of the participants' level of musical experience. This chapter concludes with a description of alternative and expanded approaches to the user study.

9.2 Evaluation of Hypotheses and Goals

This dissertation's core hypothesis, intent, and research goals are reviewed here. The background research, prototypes, implementation, and user study results are incorporated to demonstrate in-depth exploration of the techniques proposed by this research.

Core Hypothesis

The core hypothesis of this dissertation is that real-time musical performance data, audio signal processing, visual transformations of music concepts, and representations of the cross-sensory phenomenon of synesthesia can be combined in a

computer generated visual display to increase musical awareness for composers, performers, and the audiences. This research proposes that an interactive system built upon these principles can provide audience members a more intuitive view of the music concepts that are applied in composition, performance, and improvisation. Also, that these interactive systems can illustrate nuances of the performance in informative and artistic ways in order to stimulate creativity.

Sixteen virtual scenes have been developed to demonstrate key concepts from Synesthesia. The virtual scenes implement Chromesthesia (colored hearing), visual photisms, and linear displays of real-time musical events. Musical concepts including intervals, pitch classes, circle of fifths, texture, harmony, timing, and more have been integrated in the various displays. Artificial life systems including fractal trees and a dynamic life simulation have also been explored.

Each visual module varies in complexity. Some of the modules are in early prototype stages and others have gone through multiple iterations. The system has been presented in educational settings where elementary through early-college students have been able to explore and interact. Concepts have also been demonstrated in educational forums. Portions of the system have been used to augment dance and musical performances.

Feedback from live performances and exhibitions has helped to identify areas for future investigation and recombination. User testing has been conducted to collect audience responses to the system. Those responses have indicated that this effort has been successful in simulating the Synesthetic pairings between color and pitch, between real-time audio sound and dynamically modulated imagery.

The ensemble support concept displays have been evaluated favorably for their potential to translate real-time musical events into a visual display of note intensity and relative timing. User testing has revealed several possible future directions for this type of real-time musical display. User concerns about scalability as well as some performance limitations provide key insight into potential scalability issues. Suggestions have been collected for additional features and applications beyond the realm of sound and music translation.

The user group saw high potential for use of this system in education, entertainment, and rehearsal. Split-population analysis indicated that the audiences' amount of musical experience affected their perception of the musical events. The experimental data and demonstrations of this system have confirmed that a Synesthetic approach can successfully augment a musical performance in an informative and entertaining way. The clear and consistent translation of musical events into visual imagery encourages performers to creatively interact with the system and explore their range of expressive musical input.

Intent of Research

The stated intentions of the Synesthetic Music Experience Communicator have been implemented, demonstrated, and studied. The system has **informed and educated** audiences about certain aspects of sound. The **potential for improving individual and ensemble performance** has been demonstrated, especially for novice or non-formally trained musicians. Performers have been able to **directly control** and influence the visual scenery in live performance settings. Audiences have been able to **experience visual simulations** of the experience of musical **Synesthesia**.

Goals

The specific goals stated at the outset of this research have been accomplished to various degrees. Among those goals was to document new applications ideas that intersect the areas HCI, Synesthesia, and music. Background research, prior research projects, and techniques have been investigated. Synesthesia music perception, musical concepts, music software applications and hybrid music-graphics experiences have been identified.

The SMEC application has been used successfully in education and entertainment. The Small Ensemble Trainer and Computer Augmented Percussion Trainer have been evaluated favorably for their potential benefit in rehearsal settings.

SMEC has been demonstrated to several audiences ranging from elementary school through post-collegiate age. Students have been able interact with the system and to identify the various translations between sound and imagery. Parts of this system have been demonstrated to research colleagues at open forums as well as in virtual teleconference presentations. Discussions and audience comments have reflected both their interest in this area of research and their consideration of the approach for future projects. Finally, this dissertation process has equipped the author with the knowledge and technical base to conduct future research in the areas of computer graphics, Synesthesia, and music.

9.3 Partial Record of Public Exhibitions

First Lego League (2004, 2005, 2006). The MIDI and real-time audio modules were demonstrated to middle school audiences in Iowa State University's Howe Hall Alliant Energy Lee Liu Auditorium. Each year, two to three sessions of over two hundred students viewed the stereo projected graphics on a twenty-nine foot wide screen. Students were allowed to play the keyboard and/or make vocal sounds to trigger the visual modules.

HCI Open Forums (2004 - 2006). The MIDI-VIZ modules were introduced to the ISU community at the Virtual Reality Applications Center's (VRAC) HCI Open Forum in 2004. The real-time audio modules and augmented dance performance modules were introduced in 2005. The Synesthetic Visualizer and Ensemble Support modules were discussed in 2006. Keynote speaker Ray Kurzweil commented on the importance of this work in Synesthesia and the representation of sensory experiences.

Assisted Living 2 (April 2005). Real-time audio modules were incorporated into a virtual reality augmented dance performance. Computer graphics imagery was projected from the front of the stage onto the rear stage wall. Dancers wore a flexible wireless sensors system to measure their knee and arm positions. Those signals were used to

control other aspects of the virtual scenery. This project was a collaborative effort between VRAC and Co'Motion Dance Theater.

Des Moines Television Interview (May 2006). Des Moines IA NBC affiliate WHO TV broadcast an interview about this research project. This interview was filmed in the VRAC conference room and featured the visual modules projected on the conference room wall.

Jumbies CD Release Party (August 2006). A live musical performance by an Ames area Caribbean band, the Jumbies, was augmented with the SMEC visual modules. This six-member band included keyboards, drums, percussion, steel-pan, bass guitar and lead guitar. The visuals were projected from an in-house projector, onto a screen located above the drum set. Several audio and hybrid audio/visual modules were used. The system was controlled from a laptop located 10 feet from the corner of the stage. Single channel audio was used along with operator controlled MIDI inputs.

9.4 Future Directions

This Synesthetic Music Experience Communicator is a collection of proof of concept modules. The concepts of colored-pitches, modulating photisms, and mental representations of music events have been successfully demonstrated. Still, there are more features to be added, more uses to be considered, and other sensory modes to be represented. This section looks at potential future directions for this project, specifically in the areas of: software architecture, performances, sensory transformations, and applications to other domains.

Software Architecture

A key area for improvement is in the handling of MIDI events and real-time audio data. Some of these features are already implemented in commercially available software

packages, but were not included in the VR prototype system's design. Currently the system responds to all incoming MIDI events on a single channel. One MIDI channel can support up to sixteen separate instruments, however these visual modules do not explicitly test the channel number on the MIDI events. Supporting multiple MIDI channels change will require additional data routing and a good interface for configuring the MIDI listen channel for each visual module.

Also, the system requires more than one channel worth of audio data. Multiple audio channels will enable stereo- or multi- channel based photisms. Ideally the system could provide individual audio signals for each instrument in the ensemble. Multiple audio channel support will require data handling, configuration options, multiple channel audio interface and/or multiple computer hosts. Also, the digital audio processing code should be expanded beyond the current sixteen-band limit. Increasing the number of power spectrum bands will give a more accurate picture of the sounds, resulting in more refined visual modules.

This system will also benefit from a nice cross platform client-based user interface. A Java Swing-like interface can be used to configure the data streams and visual modules. Configuration options can be sent to the SMEC application through command architecture on another socket.

The timing issues encountered in the ensemble support tools can be remedied through a more robust MIDI scheme. These changes can include the use of full timestamps, longer memory buffers with storage and retrieval, and incorporation of metronome/clock information. User comments have suggested that these modules need the ability to display a real-time musical score and to track the ensemble's progress through that score.

Visual Features

System modules should be re-designed to support multiple switch able color schemas. Currently each module performs a look-up in a hard-coded table. This could be replaced with a smart color manager and a generalized color lookup procedure. This

procedure would enable quick switching and/configuration through the user interface. Another appealing feature is that of generalized background patterns. Users could be given multiple options for dialing up ambient scene changes.

The inclusion of multiple audio streams can open the way for stereo or 3D placed photisms. Further, each note's duration and envelope characteristics can be more accurately tracked through the use of isolated channels. One visual scene can contain multiple visual modules, each responding to a unique combination of audio and MIDI signals. Another approach is to have a single computer system handling the audio, MIDI, and graphics rendering for each instrument on stage. This solution may require multiple screens and projectors. It would also enable an operator to visually remix the renderings of the individual instruments.

Another potential expansion is to incorporate a learning function to build profiles of the instruments. Envelopes, and frequency spectra could be stored for each MIDI channel and/or note. Then when a note was sounded, the cached profile could be visually rendered.

Visual Performances

The public exhibitions and live demonstrations of this software have indicated some important system considerations. The tempo, complexity, dynamics, and texture of the music have a serious impact on the visual graphics. Too much information or volume can saturate the modules and vice versa. The input audio volume has to be attenuated to maintain a proper level of visual intensity. Some MIDI filtering might also help to reduce the amount of data overload For example, fast drum rolls or individual cymbal hits may be less visually important than the heavy accents or repeated patterns.

In addition to simply doing more performances with these modules, it would be interesting to incorporate 2D imagery and video feeds. These inputs could be color modulated, placed, or shaped by the real-time musical signals. Incorporating the visual graphics with the stage lights could also add another dimension to the integrated system.

Finally, audience interactions by way of laser pointers, paddles, and motion sensing might also be used to influence the graphics. An adaptive algorithm could evolve the visual scene on the basis of the audience's attention or the musical sounds being played.

Expanded Applications

The Small Ensemble Trainer and Percussion Trainer modules could be used to support intelligent musical tutoring. In this type of computer could choose lessons for the student, display them on the screen, and then plays them in conjunction with the student. The software could potentially monitor the student's progress to identify trouble spots. Those trouble areas can be presented to the student interactively. The system can then vary tempos, volumes, and ordering of lessons to progress the student through the materials. This system could also incorporate a camera to measure and encourage proper stick location, grip, and stick trajectory.

Audience comments have suggested potential use for the synesthetic visualizer in helping autistic children or stroke recovery patients. The hope is that the combined sensory experience will capture their attention and help to re-focus their sensory experiences. Also, suggested were applications for the hearing impaired. This type of visual module could be used like a closed captioning for the musical events.

Other aspects of Synesthesia could be demonstrated with this approach. Beyond sound, there are other sensations that could be represented in the visual domain. Color-grapheme perception, sensations, and emotional states could be represented. A full-body model of a virtual person could be layered with localized information from other types of sensors or with simulated data. Example sensory data includes temperature and pressure sensors on the feet and hands. Individual's heart rate, pulse, and blood pressure could be constantly displayed on the virtual model.

A cognitive science approach could even incorporate a mapping of the brain areas and neural pathways involved in certain stimuli. The resulting system begins to approach a Synesthetic Life Emulator and may have specific benefit for medical profession, for

health and status monitoring of remote workers, and to improve understanding of Synesthetic cognitive processes.

9.5 Working Towards A Synesthetic Partner

This dissertation closes with a vision statement for a future application in creative media. This vision reaches beyond augmenting a live musical show and beyond translating sound into visual imagery. Here the computer is envisioned as a creative improvisational partner for multi-modal experiences. This hypothetical computer observes how the user creates music, graphics, documents and photographs. Software algorithms analyze the way the user develops ideas and assembles concepts. The computer then develops correlations and profiles for a user's audio and visual pairings. It models the way a person's visual and musical creativity interact to predict and suggest multi-modal accompaniments for their digital creations.

For example, in a music composition program, the computer could ask the user to draw a representation of the image that a sound brings to mind. Similarly the computer could ask the user to play a melody corresponding to a visual picture or a through a virtual scene. Once the profile is developed, the computer can begin to improvise around the user's ideas. This creative system would operate similarly to an improvisational music program operating off of a probabilistic model of a music performance. Except in this case the system could suggest and create accompanying content across multiple sensory domains.

This digitally based Synesthetic Partner could learn the user's associations for sounds, words, music, and imagery. Once trained, the digital partner would become a reflection of the user's creativity. This partner could inspire the user's creativity and generate new media combinations that bear the essence of the user's creativity. Several of these auto-generated accompaniments could be computed in parallel. User feedback in the form of eye tracking or scored evaluation could be used to help the computer decide which pairings the user has found to be most pleasing. Conversely one individual's digital creations could be sent to another computer and translated according to a different

user's preferences. Similarly, the same computer could load a different individual's translation profile to enable a pseudo tele-collaboration.

This digital Synesthetic Partner becomes an extension of the users creativity. It reflects the cognitive processes behind Synesthesia and the automatic mental associations that Synesthetes leverage in their creative efforts. That process is made explicit and observable for the benefit of performers, content creators, students, and audience members.

Video and image samples of this project can be found through Iowa State University's Human Computer Interaction web site (http://www.hci.iastate.edu) or through the author's website (http://www.elchill.com/smec).

BIBLIOGRAPHY

[Abc]   Abc News Internet Ventures, "Savant Gives Window To World of Autism," http://abcnews.go.com, visited June 2005.

[Abdullah] Abdullah, M., and Ang, M., "Computer-Aided Music Instruction: An Implementation Of The Use Of Notation Software In A Local Teachers Training College," *Proc. of First International Music Technology Conference in Southeast Asia (MusTech Asia'2000).* Serdang, pp. 22-24, 16-18 March 2000.

[Abowd] Abowd, G.D., and Beale, R., "Users, Systems, and Interfaces: A Unifying Framework for Interaction," *Human Computer Interaction HCI'91: People and Computers*, Vol. 6, Cambridge University Press, Cambridge, pp. 73-87, 1991.

[Aucouturier] Aucouturier, J-J. and Sandler, M., "Using Long-Term Structure to Retrieve Music: Representation and Matching," *The Second Annual Symposium on Music Information Retrieval*. Indiana University, pp. 1-2, 2001.

[Akai]   Akai Professional, "Akai Professional," http://www.akaipro.com/, visited October 2005.

[Alesis]   Alesis. "Alesis.com," http://www.alesis.com, visited October 2005.

[AltMode] Alternate Mode, "Alternate Mode and KAT Percussion," http://www.alternatemode.com/, visited October 2005.

[Armel]   Armel, K., and Ramachandran, V., "Acquired Synesthesia in Retinitis Pigmentosa." *Neurocase'99* Vol. 5, Oxford University Press, pp. 293-296, 1999.

[ASA]   American Synesthesia Inc. "ASA Homepage," http://synesthesia.info/, visited November 2005.

[Avid]   Avid Technology, Inc., "M-Audio Radium 61 61-Key USB MIDI Controller," http://www.midiman.com/products/en_us/Radium61-main.html, visited October 2005.

[Bamberger] Bamberger, J., "The Development of Intuitive Musical Understanding: A Natural Experiment." *Psychology of Music* 31(1), pp. 7-36, 2003.

[Beeli]   Beeli, G., Esslen, M., and Jåncke, L., "When Coloured Sounds Taste

Sweet," *Nature*, Vol. 434, March 2005 pp. 38.

[Blaine02]   Blaine, T., and Forlines, C., "JAM-O-WORLD: Evolution of the Jam-O-Drum Multi-player Musical Controller into the Jam-O-Whirl Gaming Interface," *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, Dublin, Ireland, May 24-26, 2002.

[Blaine03]   Blaine, T.. and Fels, S., "Contexts of Collaborative Musical Experiences," *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*, Montreal, Canada. pp. 129-133, 2003.

[Bois]   Bois, M., **Iannis Xenakis: The Man and His Music**, Boosey & Hawkins, London, 1967.

[Brewster]   Brewster, S., "Nonspeech Auditory Output", in **The Human-Computer Interaction Handbook**, eds. Jacko, J. and Sears. A., Erlbaum Associates, New Jersey, 2002.

[Cakewalk]   Cakewalk Software, "Cakewalk SONAR Home Studio," http://www.cakewalk.com/Products/HomeStudio/default.asp, visited October 2005.

[Casio]   Casio Computer Co. LTD, "Welcome to Casio Music Site," http://music.casio.com/, visited October 2005.

[Chew03]   Chew, E. and François, A., "MuSA.RT: Music on the Spiral Array. Real-Time," *ACM MultiMedia'03*, Berkley California, USA. 2003.

[Chew05]   Chew E., "Engineering Approaches to Music Perception and Cognition," http://www-scf.usc.edu/~ise575/, visited November 2005.

[Chosky]   Chosky, L., Abramson, R. M., Gillespie, A. E., Woods, D., and York, F., **Teaching Music in the Twenty-First Century, Second Edition**, Prentice Hall, New Jersey, 2001.

[CMJ]   MIT Press, "Computer Music Journal Website", http://mitpress2.mit.edu/e-journals/Computer-Music-Journal/, visited November 2005.

[Cockton]   Cockton, G., Lavery, D., and Woolrych, A., "Inspection-Based Evaluations," in **The Human-Computer Interaction Handbook**, eds. Jacko, J. and Sears. A., Erlbaum Associates, New Jersey, 2002.

[Cope]   Cope, D., **Virtual Music: Computer Synthesis of Musical Style**, MIT Press, Cambridge, Massachusetts, 2001.

[Coyle]   Coyle, E., "Digital Tools for Music Education, DiTME," Digital Media Centre Homepage. http://www.dmc.dit.ie/ditme/index.html, visited October 2005.

[Cytowic]   Cytowic, R. E., **Synesthesia: A Union of the Senses**, MIT Press, London, England, 2002.

[Dann]   Dann, K. T., **Bright Colors Falsely Seen: Synaesthesia and the Search for Transcendental Knowledge**, Yale University Press, New Haven & London, 1998.

[Ddrum]   Ddrum, "ddrum- Acoustic and Electronic Products since 1983," http://www.ddrum.com/home.htm, visited October 2005.

[Digidesign]   Digidesign, "Digidesign – The Industry Leader in Digital Audio and Sound," http://www.digidesign.com/, visited October 2005.

[Dix]   Dix, A.J., Finlay, J.E., Abowd, G.D., and Beale, R., **Human Computer Interaction, Second Edition**, Prentice Hall, Europe, 1998.

[Dixon]   Dixon, M., Smilek, D, Cudahay, C., and Merikle, P., "Five Plus Two Equals Yellow," *Nature* Vol. 406, 27 July 2000. pp. 365, Magellan Magazines Ltd. 2000.

[Duffy]   Duffy, P., L., **Blue Cats and Chartreuse Kittens: How Synesthetes Color Their Worlds**, Times Books, New York, 2001.

[Dumas]   Dumas, J., "User-Based Evaluations", in **The Human-Computer Interaction Handbook**, eds. Jacko, J. and Sears. A., Erlbaum Associates, New Jersey, 2002.

[Edirol]   Edirol Corporation, "Edirol Video Presenting Products," http://www.edirol.com/products/vpresenting.html, visited October 2005.

[EMFMedia]   EMF Media Website, "EMF Media Past and Future Classics Iannis Xenakis: Electronic Music," http://www.emfmedia.org/index.html, visited December 2005.

[Farbood]   Farbood, M., Pastor, E., Jennings, K., "Hyperscore, A graphical Approach to Composing Music." *IEEE Computer Graphics and Applications, Special Issue on Emerging Technologies*, January 2004,

pp. 50-54. 2004.

[GATech]   Georgia Tech Music Department, "Music Technology Research," http://www.coa.gatech.edu/music/technology/research.htm, visited November 2005.

[Goto]   Goto, M., "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds," *Journal of New Music Research*. Vol. 30, No. 2, pp. 159-171, Swets & Zeitlinger, 2001.

[Griffiths]   Griffiths, D., "Fluxus Website," http://www.pawfal.org/Software/fluxus/, visited 11/15/2005.

[Grossenbacher]   Grossenbacher, P., and Lovelace, C., "Mechanisms of Synesthesia: Cognitive and Physiological Constraints," *Trends in Cognitive Science*, Vol.5, No. 1, Jan 2001 Elsevier Science, Ltd. pp. 36-41. 2001.

[Guardian]   Guardian Newspapers Limited, "A Genius Explains. Feb 12, 2005," http://www.guardian.co.uk/, visited November 2005.

[Harley05]   Harley, J., "Leonardo On-line Iannis Xenakis Bibliography and Discography," http://mitpress2.mit.edu/e-journals/Leonardo/isast /spec.projects/Xenakisbib.html, visited December 2005.

[Harley]   Harley, J., **Xenakis: His Life in Music**, Routledge. Taylor and Frances Group, NY, 2004.

[Harrison]   Harrison, J., **Synaesthesia: The Strangest Thing**, Oxford University Press, New York, 2001.

[HCI]   Human Computer Interaction Website, "HCI Home", Iowa State University, http://hci.iastate.edu, visited October 2005.

[Helmholtz]   Helmholtz, H., **On the Sensations of Tone**, Dover Publications, New York. 1954.

[Hinton]   Hinton, L., Nichols, J., and Ohala, J., "Introduction, Sound-symbolic Processes", in **Sound Symbolism**, eds. Hinton, L., Nichols, J., and Ohala, J., Cambridge University Press, NY, 1994.

[Huang]   Huang, C. Z., and Chew, E., "Palestrina Pal: a grammar checker for music compositions in the style of Palestrina," *Proceedings of the Fifth Conference on Understanding and Creating Music (UCM5)*, Caserta, Italy, November 27-30, 2005.

[Huddersfield]   The University of Huddersfield, "Computer Assisted Learning for Musical Awareness: CALMA Handbook", http://www.hud.ac.uk/mh/music/calma/calma.html, visited October 2005.

[Humdrum]   The Humdrum Toolkit, "The Humdrum Toolkit: Software for Music Research," http://csml.som.ohio-state.edu/Humdrum/, visited October 2005.

[Huron]   Huron, D., "Voice Denumerability in Polyphonic Music of Homogeneous Timbres," *Music Perception*, Summer 1989. Vol. 6, No.4, 361-382.

[ICMC]   International Computer Music Conference, "ICMC 2005," http://www.icmc2005.org, visited November 2005.

[IEEE-MM]   IEEE Computer Society, "IEEE Multimedia", http://www.computer.org/portal/site/multimedia/, visited November 2005.

[Infusion]   Infusion Systems LTD., "I-CubeX Online Store: The Ultimate Controller," http://infusionsystems.com/, visited October 2005.

[ISU-Drumline]   Iowa State University Drumline, http://www.music.iastate.edu/org/drumline/, visited October 2005.

[Jacko]   Jacko, J. and Sears, A., **The Human-Computer Interaction Handbook**, A., Erlbaum Associates, New Jersey, 2002.

[Jennings]   Jennings, K., *"*Computer-Mediated Music Composition - A Case Study." *Special Issue of Education and Information Technologies (EIT) Learning with ICT for the 21st Century*. December 2004.

[Klapuri]   Klapuri, A., Virtanen, T., Eronen, A., and Seppanen, J., "Automatic Transcription of Musical Recordings," *Proceedings of Consistent & Reliable Acoustic Cues Workshop, CRAC-01*, Aalborg, Denmark, September 2001.

[Korg]   Korg USA, "Korg.com", http://www.korg.com/, visited October 2005.

[Krumhansl]   Krumhansl, C. and Schumuckler, M., "A Key-Finding Algorithm Based on Tonal Hierarchies." *19th Annual Mathematical Psychology Meeting,* Cambridge, MA. 1986.

[Kurzweil03]   Kurzweil, R., "Promise and Peril," in **Living with the Genie: Essays**

on Technology and the Quest for Human Mystery, Eds. Lightman, A., Daniel S., and Dresser, C., Island Press, Washington. 2003,

[Kurzweil99] Kurzweil, R., **The Age of Spiritual Machines: When Computers Exceed Human Intelligence**, Penguin Group, New York, 1999.

[Leonardo] MIT Press, "Lenoardo – the MIT Press," http://mitpress.mit.edu/catalog/item/default.asp?ttype=4&tid=11, visited November 2005.

[Lartillot] Lartillot, O., Dubnov, S., Assayag, G., and Bejerano, G., "Automatic Modeling of Musical Style*." International Computer Music Conference*, La Havana, 2001.

[Logic] Apple Computer, Inc. "Apple Logic Pro," http://www.apple.com/logicpro/, visited October 2005.

[Matossian] Matossian, N., **Xenakis**, Kahn & Averill, London. 1986.

[MENC] The National Association for Music Education, "MENC: The National Association for Music Education," http://menc.metapress.com/, visited October 2005.

[Midi] MIDI Manufacturers Association, "Midi Manufacturers Association," http://www.midi.org, visited October 2005.

[MidiShare] MidiShare, "MidiShare," GRAME Computer Music Research Lab, http://midishare.sourceforge.net/, visited October 2005.

[MNMA] Musical Notation Modernization Association. "MNMA Homepage," http://www.mnma.org/, visited on October 2005.

[Mode] Mode Records, "Iannis Xenakis Profile," http://www.mode.com/profiles/Xenakis.html, visited December 2005.

[Norman] Norman, D. A., **The Design of Everyday Things**, Doubleday Books, New York, 1990.

[Olson] Olson, Harry F., **Music, Physics, and Engineering**, Dover Publications, New York, 1952

[OpenGL] OpenGL Organization "OpenGL Overview," http://opengl.org/about/overview.html, visited on February 2006.

[Optale] Optale, G., Capodieci, S., Zara, P., Gamberini, L. and Riva, G.,

"Music-enhanced Immersive Virtual Reality in the Rehabilitation of Memory-Related Cognitive Processes and Functional Abilities: A Case Report," *Presence,* Vol. 10, No. 4, MIT Press, 2001.

[Overholt]   Overholt, D., "The Sonic Scanner and the Graphonic Interface," *Proc of International Computer Music Conference*, Miami, Florida, Nov. 2004.

[Oviatt]   Oviatt, S., "Multimodal Interfaces," in *The Human-Computer Interaction Handbook*, eds. Jacko, J. and Sears. A., Erlbaum Associates, New Jersey, 2002.

[Ox]   Ox, J., and Britton, D., "The 21st Century Virtual Reality Color Organ," *IEEE Multimedia*, vol. 07, no. 3, pp. 6-9, July-September 2000.

[Pachet]   Pachet, F., "Playing With Virtual Musicians: the Continuator in Practice," *IEEE Multimedia* 9:3, pp. 77-82.

[Palmeri]   Palmeri, T., Blake, R., Marois, R., Flanery, M. and Whetsell, W., "The Perceptual Reality of Synesthetic Colors," *Proceedings of the National Academy of Sciences*, Vol.99 No.6. pp. 4127-4131. March 19, 2002.

[Pape]   Pape, G., "Iannis Xenakis and the 'Real' of Musical Composition," *Computer Music Journal,* 26:1, pp 16-21, MIT Press, 2002.

[Paradiso]   Paradiso, J., and Gershenfeld, N., "Musical Applications of Electric Field Sensing," *Computer Music Journal*, 21(2): pp. 69–89. 1997.

[Poupyrev]   Poupyrev, I., "Augmented Groove: Collaborative Jamming in Augmented Reality," http://www.mic.atr.co.jp/~poup/research/ar/ agroove.htm, visited November 2005.

[Propellerhead] Propellerhead Software, "Propellerhead Reason 3.0," http://www.propellerheads.se/index.cfm, visited October 2005.

[Ramachandran] Ramachandran, V., and Hubbard, E., "Hearing Colors, Tasting Shapes," *Scientific American*, pp. 52-59. May 2003.

[Redish]   Redish, J. and Wixon, D., "Task Analysis," in **The Human-Computer Interaction Handbook**, eds. Jacko, J. and Sears. A., Erlbaum Associates, New Jersey, 2002.

[Reimer]   Reimer, B., **A Philosophy of Music Education, Second Edition**, Prentice-Hall, New Jersey, 1989.

[Rhodes]   Rhodes, R., "Aural Images," in **Sound Symbolism**, eds. Hinton, L., Nichols, J., and Ohala, J., Cambridge University Press, NY, 1994

[Roland]   Roland Corporation U.S., "Welcome to Roland U.S.," http://www.rolandus.com/aboutus/index.asp, visited October 2005.

[Scheirer]   Scheirer, Eric D., "Tempo and Beat Analysis of Acoustic Musical Signals," *Journal of Acoustics Society of America* 103:1 (Jan 1998), pp 588-601.

[Sessions]   Sessions, Peter L., **Functions of Chords**, P & R Press, New Hampshire, 2002.

[Sharer]   Sharer, T., "Lift Every Voice and Sing," Wayne State College, Nebraska. http://academic.wsc.edu/education/sharer_t/liftvoic.html, visited October 2005.

[Smilek01]   Smilek, D., Dixon, M. Cudahy, C., and Merikle, P., "Synesthetic Photisms Influence Visual Perception*" Journal of Cognitive Neuroscience*, Vol. 13. No. 7, MIT Press, pp. 930-936.

[Smilek02a]   Smilek, D., and Dixon, J., "Towards a Synergistic Understanding of Synesthesia." *Psyche* 8(01). Association for the Scientific Study of Consciousness. January 2002.

[Smilek02b]   Smilek, D., Dixon, M., Cudahy, C., and Merikle, P., "Synesthetic Color Experiences in Influence Memory" *Psychological Science*, Vol. 13. No. 6. Nov. 2002. American Psychological Society, pp. 548-552.

[SMT]   Society For Music Theory, "Computer-Aided Instruction: Bibliography," http://www.societymusictheory.org/index.php?pid=36, visited October 2005.

[Sony]   Sony Media Software, "Sony Media Software – Home for Vegas, Sound Forge, and ACID," http://www.sonymediasoftware.com/, visited October 2005.

[Sperling]   Sperling, C., "The Information Available in Brief Visual Presentations", **Psychological Monographs: General and Applied**, Vol. 74. No 11, American Psychological Association, Washington, 1960.

[Sporre]   Sporre, D., **Perceiving the Arts: An Introduction to the Humanities**, Prentice Hall, NJ, 2003.

[SRG]   Synesthesia Research Group, "Synesthesia Research Group," University of Waterloo, http://www.synaesthesia.uwaterloo.ca, visited November 2005.

[SSEYO]   SSEYO "SSEYO Koan Pro," http://www.sseyo.com/products/koanpro/, visited November 2005.

[Steedman84]   Steedman, M., "A Generative Grammar for Jazz Chord Sequences." *Perception* 2:1 52-77. 1984.

[Steedman96]   Steedman, M., "The Blues and Abstract Truth: Music and Mental Models," in **Mental Models In Cognitive Science,** eds. Garnham, A. and Oakhill, J., Mahwah, NJ: Erlbaum 1996, 305-318.

[Steel]   Steel, W., "Samples of Musical Notation," University of Mississippi, Music Department. http://www.mcsr.olemiss.edu/~mudws/notes/, visited October 2005.

[Tsur]   Tsur, R., **What Make Sound Patterns Expressive**, Duke University Press, Durham NC, 1992.

[Tuttle]   Tuttle, R., "The MUSICPOLE MIDI Controller," http://www.themusicpole.com/pr.html, visited October 2005.

[Tzanetakis]   Tzanetakis, G., Essl, G., and Cook, P., "Human Perception and Computer Extraction of Musical Beat Strength," *Proc. Of the 5^{Th} International Conference on Digital Audio Effects*, Hamburg Germany. 2002.

[Ungvary]   Ungvary, T., and Vertegaal, R., "Cognition and Physicality in Musical Instruments." in **Trends in Gestural Control of Music**, eds. Wanderly, M.M. and Battier, M., IRCAM, Centre Pompidou, 2000.

[VRAC]   Virtual Reality Applications Center, "VRAC Homepage," Iowa State University, http://www.vrac.iastate.edu, visited February 2006.

[VRCO]   VRCO, "VRCO," http://www.vrco.com, visited February 2006.

[VRJuggler]   VRJuggler Homepage "VRJuggler-Base/Main/About," http://vrjuggler.org, visited February 2006.

[Weinberg]   Weinberg, G., and Driscol, S., "*iltur*- Connecting Novices and Experts Through Collaborative Improvisation," *Proceedings of New Interfaces in Musical Expression. NIME 2005,* pp 17-22. Vancouver, Canada.

2005.

[Wickens]   Wickens, C. and Hollands, J., **Engineering Psychology and Human Performance: Third Edition**, Prentice-Hall, New Jersey, 2000.

[Xenakis71]   Xenakis, I., **Formalized Music: Thought and Mathematics in Composition**, Indiana University Press, Bloomington In. 1971.

[Xenakis85]   Xenakis, I., "Music Composition Treks," in **Composers and the Computer**, edited by C. Roads, Los Altos, Ca. Kauffman, 1985.

[Yamaha]   Yamaha Corporation of America, "Welcome to the World of Yamaha," http://www.yamaha.com/, visited October 2005.

[Young]   Young, D., "The Hyperbow Controller: Real-Time Dynamics Measurement of Violin Performance," *Proceedings of 2002 Conference on New Interfaces for Musical Expression (NIME-02),* Dublin, Ireland, May 24-26, pp 1-6. 2002.

[Zanette]   Zanette, D., "Zipf's Law and the Creation of Musical Context," ArXiv, Cornell University, http://arxiv.org/pdf/cs.CL/0406015, visited October 2005.