MirrorFugue:

Communicating Presence in Musical Collaboration Across Space and Time

Xiao Xiao

B.S. Computer Science and Engineering
Massachusetts Institute of Technology, 2009

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology
September 2011

.....
Author: Xiao Xiao
Program in Media Arts and Science
August 2011

.....
Certified by: Hiroshi Ishii
Jerome B. Weisner Professor of Media Technology
Program in Media Arts and Sciences
Thesis Advisor

.....
Accepted by: Mitchel Resnick
LEGO Papert Professor of Learning Research
Associate Academic Head
Program in Media Arts and Sciences

MirrorFugue:

Communicating Gesture in Musical Collaboration Across Space and Time

Xiao Xiao

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on September 2011,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology

Abstract

This thesis examines the problem of conveying presence across space and time. My work focuses on collaborative music, but findings may be generalized to other fields of collaboration. I present MirrorFugue, a set of interfaces for a piano keyboard designed to visualize the body of a collaborator.

I begin by describing a philosophy of remote communication where the sense of presence of a person is just as essential as the bits of raw information transmitted. I then describe work in remote collaborative workspaces motivated by this view. I apply this philosophy to musical performances, giving a historical perspective and presenting projects in musical collaboration and pedagogy.

Next, I describe two iterations of MirrorFugue interfaces. The first introduce three spatial metaphors inspired by remote collaborative workspaces to display the hands of a virtual pianist at the interaction locus of a physical piano. The second iteration introduces a pianist's face and upper body in the display. I outline usage scenarios for remote collaboration between two users and for a single user interacting with recorded material.

I then present user studies of a MirrorFugue prototype in the context of remote piano lessons. I outline future work directions for increasing the portability of MirrorFugue, enhancing the sense of presence beyond the visual, and expanding MirrorFugue as an augmented piano platform.

Thesis Supervisor:
Hiroshi Ishii
Jerome B. Weisner Professor of Media Technology
Program in Media Arts and Sciences

MirrorFugue:

Communicating Presence in Musical Collaboration Across Space and Time

Xiao Xiao

The following people served as readers for this thesis:

.....

Thesis Reader
Tod Machover
Muriel R. Cooper Professor of Music and Media
Program in Media Arts and Sciences

.....

Ken Perlin
Professor of Computer Science
NYU Media Research Lab

Acknowledgments

I extend my deepest gratitude to the following people; I could not have completed my thesis without your contributions.

To my advisor Hiroshi Ishii for always pushing me to think more deeply and for helping me appreciate the poetry in Human Computer Interaction.

To my readers: Tod Machover for his insightful comments on music and constructive feedback on my work and Ken Perlin for sparkling conversations over coffee and at the piano.

To the members of the Tangible Media Group for allowing me to bounce ideas off you. Thanks especially to Leo for mentoring me since I started as a TMG UROP in 2008 and for suggesting that I apply to be a student in the group. Thanks to Natalia (and Lisa and Sarah before that) for making sure everything ran smoothly.

To Lauren Gust, for being the best UROP ever by helping me design and build two shiny pianos.

To Paula Aguilera and Jonathan Williams for shooting the beautiful video that made my final demo come alive.

To Katie Cornag and Avid for the generous donations of the M-Audio keyboards.

To Noah Jessop for helping me build prototypes and for kind words of encouragement when I really needed them.

To Margaret and Marvin Minsky for inviting me to the musical salons, from which I always returned with stimulating ideas.

To Jason Moran and Bruce Brubaker for sharing their insights on the music learning process during interviews.

To my parents for all the love and support a daughter could ever ask for and for keeping me healthy and sane from afar.

To my piano teacher, Donal Fox, for guiding me along this journey in the world of musical improvisation and self-discovery. This thesis is dedicated to you.

Contents

Prelude

Mirror: a reflective surface
Fugue: a contrapuntal composition in two or more voices
Mirror Fugue: a pair of fugues where each is the mirror image of the other

In baroque music, fugues are often preceded by a prelude to prime the audience for the musical journey to come. So too I open my thesis with this prelude.

I have played the piano since I was four years old. Music has been a part of my life for as long as I can remember, and I often find myself using musical metaphors to understand ideas from other realms of life. My favorite type of musical composition is the fugue. I have always been fascinated by how the complex layers of the form fit together.

The idea for this thesis came to me during one of my daily sessions in the practice room in October of 2008. I noticed the reflection of my hands and the keyboard on the piano surface and thought that it would be rather lovely if it were someone else's hands on the other side. That someone could be from a distant land or perhaps even myself from the past.

In the past two years, I have worked to realize this little whim. Throughout the process, I have tried to weave together ideas of interpersonal communication, learning, and music into a complex, contrapuntal piece, which I present in this thesis.

1: Introduction

We often regard Music as an Art elevated from the prosaic patterns of life, but Music can be seen as a particular form of human expression and communication. Specifically, playing a musical instrument is a way of externalizing ideas using a specialized sonic vocabulary with rhythmic constraints. The expression of musical ideas echoes fundamental human interactions. Ideas may be externalized for the self alone or to an audience, they may be scripted or improvised, and they may be expressed as a monologue or a conversation.

Like other traditional forms of human communication, a musical performance is deeply tied to the physical world on multiple sensory dimensions. While the sound may captivate the primary attention, we cannot forget that music is intricately tied to the physical movements of a performer's body on an instrument situated in a space. These movements include the technique to play notes, expressive motions of the performer, and deliberate gestures at fellow musicians [33]. During ensemble performances, musicians often use each other's gestures to anticipate and coordinate playing. Learning by watching and imitating is crucial for students to acquire new techniques and methods of expression [15].

The connection to the physical world limits music performance and collaboration in space and time. Though recent advancements in musical telepresence networks enable remote performances, rehearsals, and lessons, music network systems have focused on achieving the highest fidelity reproduction of sound. Audio recording enables musical performances to be replayed anywhere at any time, but recorded music is divorced from the physical instrument and performer. Though video can capture the movements of a performer on an instrument, recorded performances have generally existed on screens disconnected from the physical world because of the lack of appropriate spatial reconstruction using the video.

This thesis introduces MirrorFugue, a set of interfaces to convey the presence of a remote or recorded pianist at the interaction locus of the piano keyboard using spatial metaphors. MirrorFugue aims to preserve the physical relationship of a performer's body on an instrument in sharing of musical ideas across space and time in the context of learning, performing, and reflection.

1.1 Thesis Structure

This thesis is divided into five remaining chapters:

Chapter 2: Background and Related Work begins with a philosophy of remote communication where the sense of presence of a person is just as essential as the bits of raw information transmitted. In the context of this point of view, I describe related work in remote collaborative workspaces, networked music systems, and music learning.

Chapter 3: Design and Prototypes first introduces three MirrorFugue interface configurations to display the hands of a pianist and then presents a second design iteration of MirrorFugue that includes a display of the face and upper body. I describe implementation details as well as initial audience reactions to the demonstrations.

Chapter 4: Scenarios outlines two sets of scenarios for MirrorFugue in the context of learning and self reflection. One focuses on two-user synchronous interactions while the other focuses on single-user interactions with recorded material.

Chapter 5: Evaluations presents two informal studies of hands-only modes of MirrorFugue in the context of a remote lesson and concludes that seeing the to-scale keyboard and hands at the interaction locus helps in music learning for novices.

Chapter 5: Future Work discusses future explorations for increasing the portability of MirrorFugue, enhancing the sense of presence beyond the visual, and expanding MirrorFugue as an augmented piano platform.

Chapter 6: Conclusion summarizes the thesis and suggests a new vision for interaction research.

2: Background and Related Work

We inhabit seamless spaces constructed from multi-sensory perceptions of the physical world. To communicate with those around us, we use the physical world as a medium. We may translate our thoughts into speech, to drawings, and to actions immediately perceived by collocated peers. Our peers also perceive additional layers of information derived from our physical presence, such as hand gestures and facial expressions, that add context and meaning to the discourse [11].

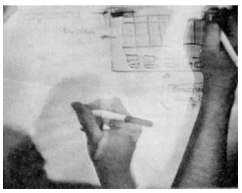
When communicating with others in spaces remote in distance or time, we must choose the channels of information to transmit, which are then reconstructed on the other side like the disjointed perspectives of a cubist painting. For the efficient delivery of raw information across distance and time, we purposely strip away rich layers of context and meaning present in a collocated exchange [12]. An idea may survive the arduous journey across space and time, but traces of the human being who conceived the idea are left behind.

2.1 Remote Collaboration

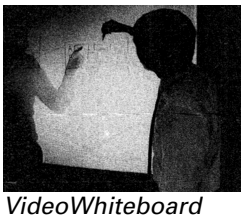
Silhouettes in Myron Krueger's Videoplace

In the 1980's, interaction researchers began experimenting with telecommunications environments that conveyed the presence of collaborators. Myron Krueger's Videoplace displayed users in disparate locations as life-sized, colorful silhouettes in the shared virtual space of a large video wall [17].

The 1990's saw a cascade of remote collaborative workspaces to support physical tasks such as shared drawing. These projects demonstrated the revealing role of a one's face and body movements in remote communication. VideoDraw, the Double Digital Desk, and TeamWorkStation showed the significance of hand gesture visibility in referencing shared



VideoDraw interface showing physical hand (right) and video of collaborator's hand (left)



VideoWhiteboard interface with silhouette of collaborator's upper body (left)

work, coordinating attention, and teaching a physical task [31, 34, 11]. In addition to displaying gesture of the hands, TeamWorkStation and ClearBoard presented video of the collaborator's face, which provided emotional cues such as whether the collaborator is confused or bored [13]. ClearBoard allowed for gaze awareness so that a user can see where the collaborator is looking.

Remote collaborative workspace research also offer insights in how to structure disjoint streams of information using real-world spatial metaphors and conventions of interpersonal space. To make sense of a remote collaborator's presence, Videoplace, Double Digital Desk, and VideoWhiteboard used shadows or silhouettes to represent the collaborator's body. Interfaces like the HyperMirror employed the metaphor of a mirror to display distant collaborators in a common virtual space[21].

To simulate interpersonal spatial relationships between remote collaborators, TeamWorkStation, VideoDraw, and DoubleDigital Desk presented the shared virtual space as if the two collaborators were working side by side, looking over the same area. ClearBoard and VideoWhiteboard presented the shared space as if collaborators were working across from each other.

TeamWorkStation interface with half-transparent overlays of workspaces on the left and video of collaborators on the right

ClearBoard interface with shared drawing space and implied face to face interpersonal space between two collaborators

2.2 Capturing a Musical Performance

Until the invention of recording technology, music could only be heard in live performances. Concerts, rehearsals, and lessons occurred at specific places at agreed upon times, where musicians and spectators, students and masters could meet in person. The communication of musical ideas was intricately tied to the body of the performer. Audiences attended concerts to marvel at the synchronized movements of a symphony orchestra painting lush landscapes of sound, they held their breath at the dazzling displays of pyrotechnics of particularly demanding passages, and breathed longing sighs at lilting chords amplified by the lingering expression on the face of the soloist. Music lived as rich, multi-faceted experiences but was fundamentally constrained by space and time.

Although written scores as instructions to repeatable repertoire have existed in various cultures for centuries, the technology to capture specific musical performances did not begin to flourish until the late 19th century. In 1877 Thomas Edison invented the phonograph cylinder, the first practical sound recording and reproduction device. For the first time in history, the sound of a musical performance could be preserved and played at a later time without the physical presence of the performer. Audio recording continued to develop and proliferate to the extent that in 1966, the celebrated pianist Glenn Gould declared the destined demise of the public concert at the hands of electronic media within the next century [28].

Canadian pianist Glenn Gould [8]

Also popularized in the late 1800's were mechanical player machines, such as the player piano. First commercialized in the 1890's, the player piano captured and played back the notes of a piano performance using punched holes in a paper roll. The recorded roll caused the keys of the player piano to physically move, as if under the invisible hands of a ghost pianist from the past [29].

Player piano from 1985

While early player pianos could only record the notes of a passage without the dynamics and phrasing, the underlying idea of the technology eventually evolved into sophisticated, modern manifestations such as

Yamaha's Disklavier, which uses MIDI to record detailed data about notes played [36].

To this day, audio recording remains the primary technology to capture and reproduce a musical performance. However, neither the highest fidelity sound recording nor the most advanced player piano of today can capture all the facets of a performance, for both types of technologies fundamentally neglect the performer's physical presence. Recently, video recording of music playing has gained popularity, as evident in the thousands of performances shared and watched by users of the ubiquitous video-sharing site Youtube [39]. Though video may capture some views of a musical performance, captured moments are imprisoned behind screens separated from physical reality.

Thus, nearly half a century after Glenn Gould's infamous prophesy, the live concert is still alive and well as the public continues to flock to concert venues big and small to witness the physical act of musical creation [15]. While the clarity and precision of a well-crafted recording may satisfy the probing ear of an astute listener, music as a whole can never be completely divorced from its physical roots. In lessons, the teacher still corrects the student's technique by demonstration. In rehearsals, musicians still cue with glances and nods. In performances, musicians still project emotion through their bodies.

2.3 Remote Musical Collaboration

Many projects have aimed to bridge the gap of distance in musical collaboration by connecting remote players in some sort of shared virtual space. Collaborations include lessons, rehearsals, auditions, and performances. They may be synchronous, where remote parties are participating at the same time, or asynchronous, where collaborators contribute in their own time. Synchronous collaborations may be in real-time, where players are synchronized based on shared cues, or they may tolerate a certain amount of latency due to transferring data over the Internet. In this section, I discuss early work in remote musical

collaboration, more recent high-bandwidth systems, and new models for asynchronous collaboration over the Internet.

2.3.1 Bypassing Latency

The latency of transferring high volumes of data over the Internet has influenced the types of synchronous, remote musical collaborations [1]. Several research groups focused on creating sound-only virtual spaces. TransMIDI communicated music as MIDI data for remote ensemble playing [9]. Similarly, Young and Fujinaga used MIDI to support remote piano master classes [38]. Transjam enabled musical applications that did not require playing "in-time", such as Webdrum, in which remote users collaboratively sequenced drum beats with a grid [3]. Sarkar's TablaNet was designed for real-time collaborative Indian percussion in bandwidth-limited networks by predictively generating drumming patterns through the analysis of previous beats [30]. Due to network limitations, the musical collaborations enabled by these systems necessarily occurred within narrow communication channels.

View of ensemble in Distributed Music Rehearsals project

2.3.2 Synchronous, High-Bandwidth Remote Collaboration

Several groups have engineered systems that support real-time, high-bandwidth remote musical collaborations. Interfaces built upon these have been used for remote lessons, rehearsals, and performances and typically include video in addition to sound. As early as 1997, the Distributed Music Rehearsals project installed teleconferencing environments with high-definition video and 3D surround sound to connect an orchestra in Geneva and its conductor in Bonn [16]. Since 2000, the New World Symphony in Miami has used Internet2, which transmits information across a continent in hundredths of a second, to enable remote rehearsals, master classes, and auditions [26]. During remote rehearsals, the remote conductor is shown projected on a large screen above the orchestra and can be seen and heard in real-time.

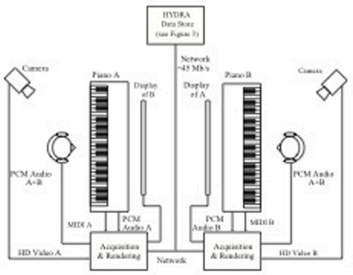


Diagram of the Distributed Immersive Performances project. The cameras are located behind the pianists, but the video is shown on screens in front of the keyboards.

In the realm of remote collaboration on the piano, Zimmermann investigated the playing of classical piano duos over distance in the Distributed Immersive Performances (2008) project, which transmitted audio, MIDI, and video with an over the shoulder view of each player [41]. The MusicPath system for remote piano lessons also displays an over the shoulder view of the student on a screen in front of the teacher to supplement the high quality reproduction of the student's playing on the actuated keys of a Yamaha Disklavier [25]. David Rosenboom's Jazz Crossings concert in April 2011 connected pianos, performers, and audiences in California and New York [4]. Based on Yamaha's RemoteLive technology, the data from the Disklavier pianos, audio, and video streams were combined and presented in real-time in both locations.

While all of these projects have included video streams that communicated useful information about collaborator's physical movements, they focused more on the engineering challenges of transmitting real-time, synchronization-sensitive video over distance rather than seamlessness in interface design and representation of collaborators as emphasized by research in remote collaborative workspaces.

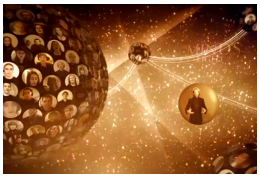
2.3.3 Asynchronous Collaboration

Audio recording technology has long enabled asynchronous remixing of musical performances from the past. While extolling the virtues of recorded music, Glenn Gould described how he could splice together multiple performances of a piece into a single constructed rendition [28]. Bill Evans's 1963 album Conversations With Myself employed overdubbing, where Evans layered up to three individual tracks for each song [7]. These early examples feature techniques ubiquitous in today's music production.

The Internet has spawned new, innovative models of asynchronous musical collaboration. Launched in 2008, the Youtube Symphony Orchestra was assembled by open auditions based on posted videos on Youtube [40]. Winners were invited to perform in a live concert, and audition submissions

were compiled into a mashup video hosted on the orchestra's Youtube channel.

Composer Eric Whitacre used Youtube to assemble virtual choirs comprising hundreds of voices from dozens of countries based on user-contributed performances of his choral pieces [35]. Whitacre compiled the contributed parts into the complete pieces, *Lux Aurumque* (2010) and *Sleep* (2011), released as Youtube videos featuring videos of each individual singer and Whitacre conducting presented in a virtual space.



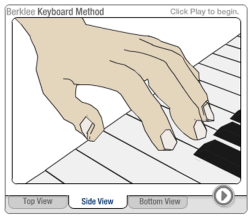
Screenshot of the Virtual Choir "performing" Sleep.

Even asynchronous collaboration over the Internet cannot escape the influence of the physical world. The ultimate goal of the Youtube Symphony Orchestra is to gather talent from across the world in a physical location for a live concert. Though the Virtual Choir's concert is purely virtual, both videos employ spatial metaphors to visualize the singers. For example, the video for *Sleep* uses the metaphor of planets to organize the virtual choristers by country of origin.

2.4 Learning an Instrument

Performing music on an instrument requires knowledge of the music itself, the ability to channel personal expression, and the physical technique to play the notes. On the piano, physical technique includes hand and body posture, proper fingering to play notes, the ability to feel the correct timing, and specialized ways of moving, commonly called "touch", to achieve different qualities of sound.

During a typical, collocated lesson, the student learns technique by watching demonstrations from the teacher. The teacher also watches the student and corrects unhealthy habits, such as tension in the wrists. In a typical lesson, all attention of both student and teacher is focused on what occurs at the instrument. Synchronous remote lessons such as MusicPath mimic this model by transferring video feeds of the student and teacher across distance, but video feeds tend to be shown separate from the space of the instrument.



Animation of hand from Berklee online piano course.

Asynchronous, remote musical pedagogy across the Internet has gained popularity in recent years. Both advanced amateurs and professionals release homemade music tutorial videos on Youtube available to all. The Berklee School of Music hosts comprehensive online courses that teach instrumental performance, music theory, and music production [2]. Berklee's instrumental courses feature videos of instructors, exercises, and recordings designed for remote students to follow along in their own time.

In asynchronous remote lessons, students may follow the teacher to learn technique but do not receive real-time feedback. In Berklee's online piano courses, hand technique is demonstrated in videos and computer animations, but student only post audio recordings and not video for instructor's feedback.


Lewiston's haptic guidance piano with a user's gloved hands

Another approach in teaching physical movement without the presence of a collocated teacher is using haptic guidance, where the system physically moves a user's body to demonstrate correct technique. Grindlay showed that haptic guidance in learning motions associated with percussion performance significantly benefitted the recall of note timing and velocity [10]. Lewiston outfitted a piano with electromagnets that can physically pull the fingers of a gloved pianist so that the pianist can feel what it is like to play a piece [18]. One advantage of haptic guidance is that it occurs at the interaction locus of the instrument. In contrast, watching a demonstration video on a computer diverts a student's attention from the instrument. On the piano, haptic guidance can help in teaching fingering and timing but does not address hand and body posture and different technique of touch.

Learning interfaces for string instruments with more continuous performance motions focus on fine-tuning these movements. A notable example is i-Maestro, which uses motion capture and playback to help violinists visualize and reflect on their playing [27].

3: Design and Prototypes

Recent work in remote music systems have achieved high bandwidth collaborations that include real-time streams of both audio and video. Musicians have also explored new modes of asynchronous musical collaborations across the internet. However, the body of the collaborators remain behind screens separated from physical space. I wanted to design interfaces for musical collaborations that tightly integrate the collaborator's body at the interaction locus of the instrument. Inspired by work in remote collaborative workspaces, I describe two iterations of interfaces for the piano where I integrate the physical presence of a collaborator using spatial metaphors.

I chose the piano as the basis of MirrorFugue interfaces for two reasons. Of the familiar instruments in the Western musical canon, the piano exhibits the most legible relationship between the performer's physical movements and sounds produced. It is also an architectural artifact that affords displays tightly integrated into the space of the instrument. While the piano is often combined with other instruments in ensembles, I focus on symmetric scenarios where pianists are playing with each other. I present analysis of the three most common spatial configurations for collocated collaborative playing which informed my designs.

3.1 Collocated Collaboration on the Piano

Lessons, duets, and duos comprise common collocated collaboration on the piano. During a lesson, the student and teacher usually sit side by side and take take turns playing. Visual attention primarily focuses on the hands, but movements of the arms, shoulders, and feet are peripherally observed and may attract foreground attention from time to time.

For duo playing on separate keyboards, the two pianos are either placed side by side or across from each other. When side by side, the arms and hands of the the partner are ambiently visible. Pianists can also coordinate by deliberately looking sideways. When across from each other, pianists

Spatial configuration of a typical piano lesson: student and teacher are side by side [22]

In a piano duet, the two partners sit side by side [14]

In a piano duo, the two partners sit at different pianos across from each other [20]

can see each other's face and shoulders but not the hands. As with duets, pianists cue with eye contact, head movements, and breathing.

In both lessons and performances, pianists perceive a collaborator's movements of the hands, arms, shoulders, and face. During a lesson physical movements, especially of the hands, are regarded in the foreground. In performance, collaborators tend mostly towards peripheral awareness of each other's movements except for occasional deliberate cues.

3.2 Hands Only Modes

I began by considering interfaces to display the hand movements of a performance. In this iteration, I focused on designing for remote and asynchronous learning for beginners, who are most concerned with learning how to use the hands. I present three configurations based on spatial metaphors—shadow, reflection, and organ.

3.2.1 Design Space

I considered three factors borrowed from remote collaborative workspaces in my designs: interpersonal space, placement of remote space, and orientation of remote space.

Interpersonal Space: Two remote users can be presented as if working side by side (like TeamWorkStation) or working face to face (like ClearBoard).

Placement of Remote Space: The remote workspace can be overlaid directly on the physical workspace (like Double Digital Desk) or located in a separate space. When located in a separate space, the remote workspace can be scaled and aligned or not spatially related at all to the physical workspace.

Orientation of Remote Space: A scaled and aligned remote space can be oriented vertically or horizontally, placed at 90 degrees or 180 degrees to the physical workspace.

3.2.2 Interface Configurations

Shadow Mode projects video of hands directly onto the keyboard, much like the shadows of users' hands in the DoubleDigitalDesk. The orientation of the projection is such that the keyboard from the video lines up exactly with the physical keyboard so that users can tell which keys are pressed down in the projection. This orientation also gives the impression that the collaborator is sitting next to the user. Because this configuration projects directly onto the keyboard, it has the disadvantage that the shadow is not distinguishable when the pianists' hands are playing in the same octave.

Reflection Mode is inspired by the reflective surface on a lacquered grand piano that mirrors the keyboard and performer's hands, this mode shows the collaborator's keyboard and hands as a simulated "reflection". I prototyped Reflection Mode by projecting the mirrored top-down view of another keyboard on the vertical surface in front of the keys, making sure to align the virtual with the physical. The video is distorted to mimic the perspective as seen from the player's point of view. Reflection Mode situates collaborators as if they were sitting across from each other. The orientation of the remote space is 180 degrees to the physical workspace, but in the prototype, it is simulated on the vertical screen in front of the keys.

Initial mockup of Reflection Mode

Early prototype of Reflection Mode

Organ Mode is named because the position of the physical and projected keyboards is reminiscent of the tiered keyboards of an organ. Organ Mode also uses the vertical surface in front of the keys but displays an unaltered top-down video of the other keyboard, aligned with the physical keys. Like Shadow Mode, the orientation of the projected keyboard gives the impression of a side by side partner. This turned out to be the preferred configuration for most users, both anecdotally and in my evaluation (see Section 5:Evaluation).

MirrorFugue in action with Marvin Minsky

Children playing on MirrorFugue. Organ mode is in the front, and Reflection Mode is in the back.

Shadow

Interpersonal Space	Side by Side	Face to Face
Placement of Remote Space	Overlay	Aligned and Offset
Orientation of Remote Space	90 degrees	180 degrees

Reflection

Interpersonal Space	Side by Side	Face to Face
Placement of Remote Space	Overlay	Aligned and Offset
Orientation of Remote Space	90 degrees	180 degrees

Organ

Interpersonal Space	Side by Side	Face to Face
Placement of Remote Space	Overlay	Aligned and Offset
Orientation of Remote Space	90 degrees	180 degrees

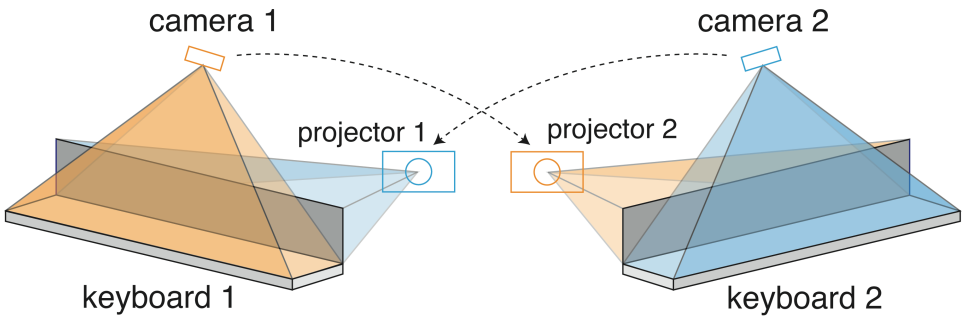
Summary of MirrorFugue design space for Shadow, Reflection, and Organ Modes

3.2.3 Prototypes

I prototyped all three configurations with MIDI keyboards, firewire cameras, and projectors, using the MAX/MSP/Jitter platform to manage video and sound. I mounted the camera directly above the piano keyboard for an overhead view. Before building the screen for the full piano keyboard, I made several smaller scale prototypes to experiment with the spatial configuration of the display. In my first prototype, I used a laptop screen as the display surface for a mockup piano keyboard made from cardboard. In the next prototype, I built a screen for four octaves of the keyboard.

I constructed the full-scaled prototype by making a back projection screen 48" wide and 12" high from a projection fabric called rosco, which I mounted vertically in front of the piano keys. In order to capture the entire piano keyboard, I attached a wide-angle lens to my camera and used a lens-correction patch in MaxMSP to correct for barrel distortion.

Since my project focuses on interface design and not on the technical details of transmitting audio and video over the internet, I set up my



System diagram of hands only prototypes

prototypes in the same room where video of one keyboard was projected on the screen of the other keyboard. Audio from each keyboard was played aloud and not transferred through the network.

I tested the remote communication by transferring 640x480 video at 30 frames per second over gigabit ethernet between two locations in the same building and was able to do so without noticeable latency.

3.3 Full Body Mode

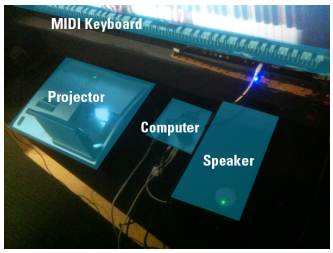
Although we "tickle the ivories" with our fingers, playing the piano involves much more than motions of the hands. Technical facility alone requires control of the wrists, arms, and shoulders to achieve a kaleidoscopic palette of texture and tone. Some passages demand the weight of the entire upper body to summon swelling chords while others call for subtle suppleness of wrist and arm to coax sensual, singing lines. Even movements of the feet, which control the sustain and soft petals, intensify the color of the soundscape. The performer's face and body also channel the emotional stream of music. The ebb and flow of a stately sonata echos in the serene face and soft undulations of a performer's body. The beats of a familiar riff swings in the performer's head bops and foot taps.

Diagram of Full Body Mode with Organ Mode displaying the hands

Diagram of Full Body Mode with Reflection Mode displaying the hands

3.3.1 Prototype

I designed a second iteration of MirrorFugue to display the face and upper body of a pianist in addition to the hands and built a case that borrowed the form of an upright piano. The large surface in front of the keys was made into a back projection display with a thin sheet of tinted acrylic attached in front to mimic the glossy surface of a piano.

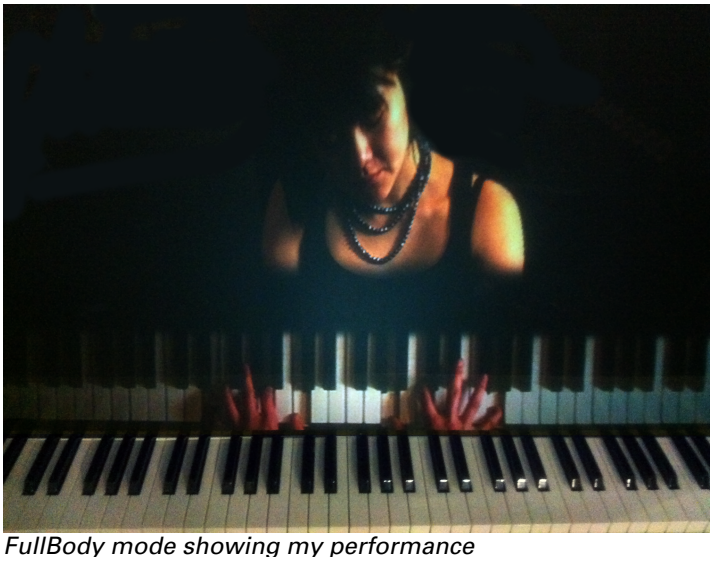


Inside of MirrorFugue case with components highlighted in blue

I used the Organ Mode orientation for the hands because it was overwhelmingly preferred among users who tried out my systems. To display the virtual pianist's face and upper body, I projected a to-scale video where a real pianist would be reflected on the surface of a tall, upright piano. The tinted acrylic on my prototype reflects the body of the user so that the user appears to inhabit the same space as the virtual collaborator.

Although the spatial metaphors used to display hands in Reflection Mode is more consistent with the reflection metaphor to display the upper body, orienting the virtual keyboard in the same direction as the physical makes the hand positions more legible for imitation, which is more suitable for learning applications. Using Reflection Mode to display hands could be a good option for applications where hand position need only be peripherally observed, such as in duet performances.

I was more concerned about user-end interface configuration rather than the seamless integration of content generation. Thus, I built the prototype to simply play back prerecorded video of performances. I used a wireless keyboard and programmed key presses to control playback.


FullBody mode showing my performance

3.3.2 Recorded Performances

I recorded two sets of piano pieces for this prototype: various interpretations of Gershwin's well known jazz tune "It Ain't Necessarily So" by pianist Donal Fox and my playing of a selection of classical passages. Each performance was shot in high definition from two points of view, one of the face and upper body, another of the full keyboard and hands. I discuss each recorded performance and its purpose for the MirrorFugue demonstration.

Classical- Performed by Xiao Xiao

***Toccata* from Bach Partita No. 6**

I performed the opening as an example of an expressive classical piece.

***Sarabande* from Bach Partita No. 2**

I played three versions of this piece: one with both hands, one with left hand alone, and one with right hand alone. This illustrates how MirrorFugue can be used to teach a piece by breaking it down.

Allegretto from Beethoven Symphony No. 7 Arranged for 4 Hands

I recorded an excerpt of the secondo part to illustrate how I can play a duet with myself from the past by playing the primo part with the recording.

Jazz - Performed by Donal Fox

Typical

This was a typical rendition of the tune with the syncopated melody in the right hand and rhythmic chords in the left hand to give the audience an idea of the piece.

Melody Only

The melody is played slowly in the right hand alone to illustrate a teaching scenario.

Left Hand Harmonies Only

Chords in the left hand are played several times on repeat. This demonstrates how a user can practice improvisation over the recorded bass loop.

Embellished

An elaborate version of the tune to show how MirrorFugue can give users a closer look at expert performances.

3.3.3 User Reactions

Ken Perlin playing on Full Body Mode of MirrorFugue with a virtual Donal Fox

I demonstrated the Full Body Mode of MirrorFugue during two open houses in April 2011 for over 100 guests from sponsor companies of the MIT Media Lab. Reactions were overwhelmingly positive. Many commented that MirrorFugue would be a great pedagogical tool for remote lessons. Sponsors also described the project as “beautiful”, “enchanting”, and “emotionally evocative”. Some sponsors even said that seeing MirrorFugue made them want to learn how to play the piano.

4: Scenarios

In this chapter, I discuss two suites of scenarios for MirrorFugue. Two-Part Invention considers synchronous remote collaborations. Duet for Solo Piano presents scenarios where a single user interacts with recorded playing of the self or others on MirrorFugue. While I classify each section by type of collaboration through the eyes of a human computer interaction researcher, I consider pieces within each suite as a musician, discussing the implications of MirrorFugue for lessons, rehearsals, practice, and performance.

4.1 Two Part Invention: Two Users Across Distance

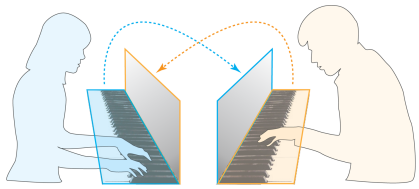


Diagram for synchronous, two-user interactions on MirrorFugue

Synchronous real-time musical interactions between two people can be divided into two categories: lessons and duets. MirrorFugue can be used as a platform for remote piano lessons. To contextualize MirrorFugue's role in a remote lesson, I begin with a brief discussion of the piano lesson and a musician's learning process. My overview is derived from interviews on the music learning process with three expert pianists (see Appendix) and my own 20-year history of piano lessons, including detailed journal entries of my own learning process from the past two years.

4.1.1 Structure of Piano Lessons

During a lesson, the student and the teacher usually take turns playing the piano. After the student plays, the teacher gives feedback on various dimensions of the student's performance, ranging from low level technical details such as hand posture and finger usage to high level expressive considerations. The teacher often demonstrates passages on the instrument for the student and may also give real-time feedback while the student

plays a passage, occasionally singing, clapping, or playing along with the student to illustrate a point.

While the overall structure of a lesson remains fairly consistent independent of the student's skill level and genre of music, the topics of focus shift as the student progresses. Playing interpretations of repertoire and playing original improvisations both require learning a diverse technical, intellectual, and expressive vocabulary involving both the body and the mind. A performer must gain physical command of the instrument by mastering how to produce a diverse array of tones. At the same time, the performer must understand the music intellectually and emotionally to render each expressive phrase.

Absolute beginners who are learning their way around the piano are most concerned with the proper hand position and typical fingerings for simple pieces and basic building blocks like scales, chords, and arpeggios. As the student acquires a larger technical vocabulary, the student begins to focus on employing these skills to express musical ideas.

4.1.2 Remote Lessons with MirrorFugue

As evident from the analysis of a traditional piano lesson, it is crucial for the student and teacher to see how each other plays during a lesson. Existing interfaces for remote piano lessons may include two-way video feeds, but the visual streams are presented in a screen disjointed from the space of the instrument. MirrorFugue extends remote lesson interfaces by presenting the body of the collaborator in full scale at the interaction locus of the keyboard.

4.1.3 Duets with MirrorFugue

Synchronous real-time duo playing includes both the rehearsal and the public performance. Though the two differ in presence of audience and in that breaks for self-evaluation are allowed during a rehearsal, interface requirements for the two classes of scenarios are practically identical. Both rehearsals and performances require players to make eye contact for

specific cues and for players to maintain peripheral visual awareness of each other's movements in addition to concentrated listening of each other's playing.

While MirrorFugue could be used to supplement remote duets, the lack of eye contact in both the hands-only and the full-body prototypes renders it non-ideal. MirrorFugue could better serve as a research tool used in experiments to answer questions about the nature of synchronous collaborations by isolating which parts of a collaborator's body is visible. These experiments may inform the future design of interfaces to supplement remote musical performances. For examples, the hands-only modes of MirrorFugue could be arranged in a collocated setting where the two keyboards are placed back to back so that the two users can make eye contact. This setup could help evaluate the question of whether seeing a collaborator's hands contributes to synchronization and anticipation. The keyboards could also be arranged so that users cannot see each other to investigate how collaborators compensate when eye contact is unavailable.

4.2 Duet For Solo Piano: Single User and Recorded Material

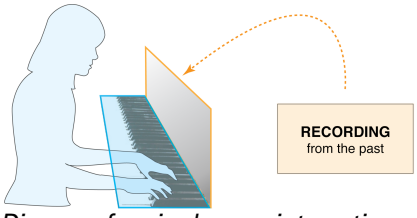


Diagram for single user interactions on MirrorFugue

Another class of applications on MirrorFugue is based on interactions with recorded material, which may be of the self or of another pianist from the past. Audio recording on its own has become an indispensable tool for both musicians and aficionados alike. Both aspiring and professional musicians listen to recordings to learn about and from others with different styles of playing. Musicians, especially those at high calibers, habitually record themselves to evaluate their own playing from a third person point of view.

MirrorFugue adds a visual dimension to the recorded performance and attempts to link a performance from the past with the present moment using spatial metaphors. In the most recent prototype, the face, body, and

hands of a pianist are presented at full scale on the surface of the piano. The technique of the pianist is visible from the hands, arms and shoulders. At the same time, the face and body of the pianist project emotional expression. The past pianist's sense of presence is amplified by the scale of the video and the implied interpersonal space between the recording and the user.

A user may interact with a recording of the self or of another. For both types of interactions, the user may choose to watch or play along. I describe specific usage scenarios and discuss how they can supplement practicing and self-reflection.

4.2.1 Recording of Self

Users can record and examine their own performances using MirrorFugue. The to-scale video of the hands gives students an outside perspective to their own playing. The integration with the physical keyboard facilitates self-evaluation during practice sessions by enabling students to seamlessly transition between playing and reflecting.

Users can also play along with recorded passages of their own playing. Classical piano repertoire beyond the absolute basics involves complex coordination of the hands, and it often takes weeks for a student to be able to play a new piece in its entirety. With MirrorFugue, a student practicing a classical piece can record a portion of the piece and supplement the rest while playing along with the recording. For example, the student can record one hand's part and practice the other hand's part with the recording. This allows the student to experience the big picture of a piece at an earlier stage in the learning process. Playing along with recorded material can also make practice more engaging by supplying musical context for technical drills.

Jazz students can use MirrorFugue to supplement improvisation practice. Jazz piano often requires one hand to hold a steady pattern while the other improvises. Beginners may lack the physical coordination to keep the

Playing a duet with the self on MirrorFugue

pattern steady. MirrorFugue can be used as improvisation training wheels, allowing a student to record a chord progression or bass line and improvise over the looped recording.

4.2.2 Recording of Others

Users can interact with the recording of another pianist, such as a teacher or other expert. Watching an expert performance at the piano can support learning by imitation. When learning a classical piece, a student can review expert performances on MirrorFugue to learn the fingering and technique of difficult passages. The student can slow down the video to better understand the hand motions and mirror the movements of the expert on the physical keyboard.

MirrorFugue can also supplement learning by ear. Jazz students often learn to play improvised solos from recordings of musicians they admire. Because improvisations are rarely written down, students must figure out all the notes by ear, a difficult task especially for a beginner whose listening is not yet so developed. Using MirrorFugue, students can supplement what they hear with what they see and play along with the expert recording to learn the passage.

Users can also play duets with the expert recording. A jazz pianist can practice improvisation over a recorded performance, using the display to better anticipate upcoming passages in the piece. A classical pianist can practice parts of a piece “karaoke style”, with the expert performance playing in the background.

5: Evaluations

In this section, I detail user studies of MirrorFugue in the context of the remote lesson. After prototyping the hands-only modes of MirrorFugue, I conducted two informal user studies on MirrorFugue's effectiveness as a remote lesson interface. I first conducted a pilot study to determine whether seeing the hands in the display helps musicians think about music. With the pilot study, I also looked to identify the pros and cons of each configuration, and to establish which, if any, stands out over the others. I then conducted an informal user study to measure the effectiveness of the winning system from the pilot study against two other systems in the context of remote learning for novices.

5.1 Pilot Study

5.1.1 Method

Five amateur pianists (4 men, 1 woman, aged 20-35) participated in the initial study. The skill levels of these users ranged from beginner to advanced, with a variety of backgrounds from completely self-taught to trained in the classical and jazz styles. For this study, I presented the participants with a keyboard in each of the three configurations in random order. I sat at the corresponding keyboard, which did not display the hands of the study participant. For each configuration, I played some chord progressions for five minutes and asked the user to improvise a melody over the chords. After all three improvisation sessions, I debriefed with each participant in an informal interview.

5.1.2 Results

All but one of the users indicated that seeing the partner's hands helped them with "listening and synthesizing sound" and to "better anticipate what is coming next by seeing where the hands are heading". The one user who disagreed, an advanced pianist trained in classical and jazz, indicated that for him, improvising together involves a highly trained ear and does not require the extra help of seeing hands. These results seem to suggest

that seeing a partner's hands could help beginner and intermediate pianists who are learning to play together with others.

Of the three setups, Organ Mode was most preferred among almost all participants. All of the participants said that Organ allowed them to clearly see what the remote partner is playing and that the location of the image is not distracting. Some participants remarked that the implied spatial arrangement "feels almost like having someone sitting next to you playing with you", which made having the displayed video feel "non- invasive". Because of the implied spatial arrangement, participants expressed that the system does not require the addition of eye contact to make sense because "when someone is sitting next to you playing, you don't often look at them". All the participants agreed that Organ Mode is best for remote learning ("because the student can see and follow exactly what the teacher is doing") and for watching a recording ("because it's in the same space as the keyboard but you can also easily join in").

All except for one user (self-taught classical pianist) liked Shadow Mode the least, pointing out that players must be playing at least an octave apart for one's hands to not obscure the shadow hands. Some called the setup "distracting and chaotic". The one user who preferred Shadow expressed that he liked seeing the remote partner's hands in the same place as his own.

Several users found Reflection Mode confusing, citing the "extra cognitive load of having to flip the image in one's head to make sense of it". One user (intermediate classically trained pianist) mentioned that the Reflection almost begs for more of the partner's body to be shown "to make sense of the spatial configuration of someone sitting across from you".

5.2 Learning Study

I designed an informal study to evaluate Organ Mode in the context of remote learning for novices. I chose this scenario because MirrorFugue can be especially beneficial for potential students of piano who do not have

regular access to a teacher and those who start learning piano through watching recordings. I asked the following research questions:

- Is seeing the hands of an instructor more helpful than seeing abstract indicators of notes for novice students learning a piece?
- Is having visual aid at the locus of interaction of the piano keyboard preferable to visual aid on a separate screen?

5.2.1 Method

To answer these questions, I recruited 10 absolute novices in piano (7 men, 3 women, aged 19-33) and taught them three simple melodic sequences on Organ and two other interfaces, each designed to answer one of the questions. One setup involved projecting a small colored dot in front of the key corresponding to one played at a second keyboard. Similar to the Moog Piano Bar [23], which includes an LED in front of every key to help pianists visualize a piece, this setup indicated what is being played using abstract symbols. The other interface displayed the same image as the one shown in Organ on a 24-inch monitor situated behind the keyboard where it is easily glance-able by the user. This simulated the configuration of when users try to learn a new piano piece by watching a video of a performance—where the visual aid is not spatially related to the piano keyboard.

I asked each user to first try to play something on the keyboard, to verify that they are indeed absolute beginners, and then to learn a randomly selected melody ("Twinkle Twinkle Little Star", "Row, Row, Row Your Boat", "Frère Jacques", or "I'm a Little Teapot") on each of the interfaces selected in random order. For each

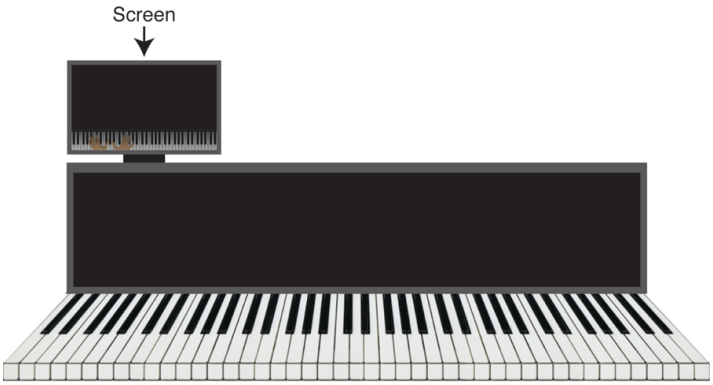


Diagram of Screen Mode


Dots indicating which keys are pressed at the corresponding keyboard

song, I began by playing the melody once through and then taught the piece in 3-5 note segments until the student could play the entire piece once without mistakes. All of the melodies contained between 16-21 notes, and repeats were eliminated from those whose original versions contained them. At the end of the study, I asked participants what they thought of each interface and to rank them by usefulness. I also videotaped each learning session to determine how long it took for users to learn each melody.

5.2.2 Qualitative Results

Most users found Organ the most helpful in learning (7 out of the 10 ranked it first, 1 ranked it a close second). Users said that Organ Mode was "very easy to follow", "very direct", and "easy to get the hand position and finger use correct". One user described how on Organ, he noticed the teacher using different fingering from his for a part he was having trouble with and changed the fingering to the teacher's, which made playing much easier. Other users described how Organ Mode was "good at allowing students to anticipate the next position of the teacher's hand".

	Screen	**Dots**	**Organ**
First	1	2	7
Second	4	5	1
Third	5	3	2

Ranking of MirrorFugue interfaces among users

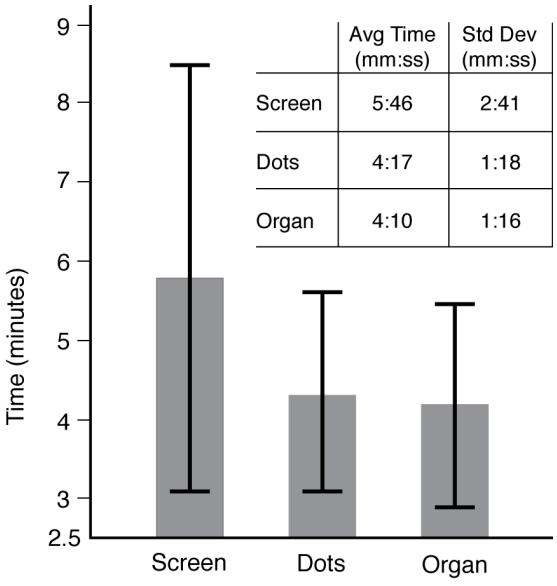
Two users found the abstract Dots the easiest to learn from because of the "very little visual processing involved" and because it "puts music into easy to understand patterns". However, one of these users suggested that the Dots "may be the easiest for following notes but may not be so good for the long term because it does not teach correct hand usage".

Only one user preferred the Screen Mode because she felt that she learned the fastest on it. However, this user actually took three more minutes (7 minutes 9 seconds) to learn the melody using Screen than on the other two interfaces (Dots: 4 minutes 54 seconds, Organ: 4 minutes 41 seconds).

5 out of 10 users ranked Screen as last place because of the "lack of reference frame between the image and physical keyboard" so that "it was difficult to tell where the teacher's hand was". "The difference of scale and the lack of correlation between the image and the physical" also contributed to the "increased visual processing load", making it the most difficult to learn from.

Dots was the worst for 3 users because "it made individual notes more individual" and "made learning into a game of follow the dots", which "made it difficult to remember sequences". Some users also felt that while the dots were the easiest for determining what notes the teacher was playing, "dots detached the lesson from actual musicality".

Two users found it "difficult to determine which keys were being pressed" in Organ Mode and listed it as least favorite. These two users both found it easier to tell which keys were pressed in Screen Mode than Organ Mode even though both displayed the same image, suggesting that their difficulties could be attributed to the fact that the resolution of the video was perhaps too low to be projected on such a large surface. In fact, one of the users who preferred Organ suggested that we should "increase the resolution of the video so that the projection in Organ Mode is more clear".



	Avg Time (mm:ss)	Std Dev (mm:ss)
Screen	5:46	2:41
Dots	4:17	1:18
Organ	4:10	1:16

Users took about 1 minute 30 seconds longer when learning melodies on Screen mode compared to both Dots and Organ modes.

5.2.3 Quantitative Results

Since we conducted an informal study, our quantitative results are not statistically significant but do suggest interesting hypotheses that can be tested in future studies. On average, users took about 1 minute 30 seconds longer to learn a melody using Screen mode than both Dots and Organ mode, which had very close average learning times. Participants learned melodies with better hand and

finger usage for both Screen mode and Organ mode over Dots mode. Using both Screen and Organ mode, all participants played with correct hand position, and all but one (for Organ mode) and two (for Screen mode) employed correct finger usage. Using Dots mode, seven users played with only one or two fingers in Dots mode, two used the whole hand but made up very awkward finger crossings, and one asked the instructor what fingers to use for each key.

5.3 Study Discussion

Through the pilot study, I discovered that users seem to prefer Organ Mode (side-to-side interpersonal space, aligned and offset image). In the learning study, I verified that displaying video of hands playing the piano at the locus of interaction of the keyboard seem to be the most helpful for novices learning from a non-physically present teacher.

Despite displaying hand gestures of the teacher, Screen mode seemed to be the most confusing for students while learning melodies. The difficulty of Screen mode can be attributed to two factors: the lack of relationship and the difference in scale between the physical and remote spaces. While the difference in scale likely contributed to the problem, most of the users commented on the lack of reference frame in Screen mode, which suggests that having no relationship between the two keyboards made it difficult for students to figure out where to play.

While Dot and Organ had almost the same average learning time, users learned to play with better technique on Organ when teacher's hands were clearly visible. Almost all of the students used only one or two fingers to play in Dots mode. In contrast, almost all of the students used all the correct fingering when using Organ. Organ also has the advantage that students can anticipate when the teacher's hand is moving to another position by watching the motions.

While Organ mode was necessary for users to maintain correct hand and finger usage, Dots had the advantage that it involved the least amount of

visual processing for users to determine what note to play. The fact that, on average, users took almost exactly the same amount of time to learn melodies on Dots mode and Organ mode could suggest that it took longer for them to process what notes they are supposed to play in Organ mode. On the other hand, users could figure out instantly what notes to play in Dot mode but forgot them more easily when dots are not present.

For teaching novices on remote learning interfaces, it seems to be important to show both what notes to play and the gestures associated with playing them.

6: Future Work

In the future, I would like to design and evaluate the effectiveness of a tablet-based MirrorFugue, examine even more effective means for enhancing sense of presence on the existing MirrorFugue, and explore a larger space of what can be displayed on the shared virtual space of MirrorFugue.

6.1 Portable Presence

Mockup of MirrorFugue for an iPad

Based on preliminary user feedback, Full Body Mode seemed to be quite effective at evoking the presence of the recorded pianist. However, it requires the construction of a custom display and is limited to use with electronic keyboards. I am interested in designing ways to convey presence in a more portable version of MirrorFugue.

Closeup of MirrorFugue for iPad

In the original MirrorFugue, the virtual keyboard is displayed next to the physical at full scale. The position and scale of the virtual keyboard automatically creates a legible relationship between the keys of the physical and the virtual. One research question to consider in designing a more portable version of MirrorFugue is whether full scale of the virtual keyboard is essential for pedagogy and presence or if another way of establishing the physical-virtual relationship can compensate for the lack of full scale.

I can investigate this question by displaying video of a pianist's hands and upper body on a tablet that can be placed on a piano's music stand, a design that massively increases the interface's portability and practicality. Currently, I have two ideas for establishing relationship between the tablet's keyboard and the physical keyboard that I would like to test in the future.

One idea is to overlay a live video of the physical keyboard over the virtual. This way, users can see their own hands in the space and scale of the virtual keyboard to figure out where the virtual hands are playing. The other idea involves the use of a player piano, such as the Yamaha Disklavier, that can move the keys of the physical piano in synchrony with the virtual piano, giving the virtual pianist a ghostly body. The first idea can be deployed on any acoustic piano while the second requires a player piano with a MIDI interface. I am very interested in whether either or both designs could be an effective supplement for a tablet-based MirrorFugue.

6.2 Enhanced MirrorFugue

Due to constraints in available resources, I was not able to experiment with the effect of sound localization in evoking spatial metaphors. In the future, I would like to study how the placement of sound sources can affect how users perceive the remote or recorded performance.

I would also like to experiment with configurations of the display to enhance the sense of presence over the current prototype. For example, I can move the projection screen further back into the piano case, which may give the displayed pianist a heightened sense of realism from the impression of depth. Moving the vertical display also frees up space for a horizontal display in front of the keys. I could project the keys in Organ Mode at a horizontal orientation, which would better match the perspective of the original video. This layout would also allow installation of an overhead camera for the keyboard to attain a more realistic simulation of reflection mode. I could then better compare Reflection and Organ Modes both with the addition of an upper body.

6.3 MirrorFugue as Augmented Reality Platform

MirrorFugue can be seen as a custom display built into the surface of the piano. Though I have been using it thus far to explore interface configurations to convey a performer's sense of presence, content for the display is not thus limited. A further space for interactions opens if I modify the display to allow for multi-touch input.

Because MirrorFugue can be seen as a sort of shared space—between multiple collocated pianists or between a user and a remote partner—it can be used as a platform to prototype shared augmented reality interfaces. For example, two collocated pianists playing a duet can see each other's reflection on MirrorFugue's surface, where eye contact is supported. When they play music, shapes may bubble out from the keyboard and float around them. They can then manipulate the shapes to remix their past playing. This interaction example extends to the case where the two collaborators are located remotely. Instead of seeing each other in the reflection, the users can see each other in the display, much like with ClearBoard. They can still collaborate with their keyboards and virtual objects in this configuration.

Because it is constrained to the specific application domain of music, MirrorFugue may be a useful tool to study shared augmented reality interfaces on a small scale. The results of these studies may inform the design of more general shared augmented reality interfaces.

7: Conclusion

In this thesis, I introduced a philosophy of communication in which the details of human presence form essential threads in the fabric of an interpersonal discourse, even when stitched across the folds of space and time. I applied this philosophy to the specific domain of musical collaboration to design a set of interfaces for the piano using a vocabulary of spatial metaphors from the physical world to display the hands and body of a remote or recorded performance. I presented scenarios and discussed evaluations of my interfaces in the context of a remote piano lesson.

While my work in the past two years has largely focused on how presence can support specific musical tasks, I had a realization when I completed the final version of MirrorFugue and turned it on for the first time this past April. It dawned on me that the presence of the ghostly performer on the surface of the piano can stir certain powerful emotions in the spectator, much like the way the seamless layers of a fugue come to bloom in the heart and mind of the listener beyond the technical details of the contrapuntal puzzle.

A great musical performance feeds both the mind and spirit, for it distills and delivers essential expressions of the human experience. I wonder if more computer interactions should strive to be similarly evocative, where facets of ourselves are reflected back at us so that we become more mindful at the immensity of our existence.

Coda

In this final cadence at the fatigue of this fugue, I leave you with another personal perspective on this project.

For me, MirrorFugue is at its core a display for displaced moments, where the glimmering facets of a piano performance reflect and reverberate, indifferent to the rigid rules of space and time.

Seen through this lens, MirrorFugue is an artifact for contemplation and conversation with the past. In 10, 20, 50 years, I would like to play a duet with myself at my current age on MirrorFugue. I wonder how I would feel to meet myself from the past. Perhaps my children and grandchildren could also play a tune with me at their age. I wonder what it would be like for them.

Through the slow glass of MirrorFugue, I hear a continuous conversation, a perpetual canon, ad infinitum.

A: Appendix

A.1 Interviews with Expert Musicians

The design of single-user interactions on MirrorFugue was influenced by interviews that I conducted with expert pianists on their learning process and practice methods. I interviewed three musicians: a classical pianist, a jazz pianist, and one known for combining both genres. While the interview results are not statistically significant, they did suggest interesting directions for research. Three themes from the interviews were especially relevant to our designs: physicality, social playing, and staying fully engaged during practice. I discuss each theme in detail and summarize its relevance to MirrorFugue.

A.1.1 Physicality

In both classical and jazz, watching the physical movements of a performer helps with understanding both technique and expression. The cross-genre pianist described how he understood the way a student produced a certain quality of sound by watching the "wave motion" of her hand. The jazz pianist noted how watching Thelonious Monk perform piano with erratic dance movements helped him figure out how music "resonated" in his own body. MirrorFugue makes visible the gesture of a piano performance.

A.1.2 Social Playing

All three pianists emphasized the importance of playing with and for others. The classical pianist suggested that "there's a kind of thinking which only happens with other people present" that is useful in understanding a complex piece of music. The jazz pianist described how many of his most profound learning experiences have occurred on stage.

Playing with others can be a source of new musical ideas. The jazz pianist explained how he developed his skills by playing with better musicians. Playing with others can also make practicing more fun. The cross- genre

pianist cited his most enjoyable early memories of music learning as those from social experiences.

Musicians often use audio recording to simulate social playing. The classical pianist frequently records himself to critique his playing, replicating a third person perspective. The jazz pianist commented that playing with a recording can mimic the feeling of playing with others though a recording can never completely replace the intricacies of interaction with real people. MirrorFugue simulates the presence of a remote or recorded collaborator.

A.1.3 Staying Engaged

All three stressed the necessity of staying fully engaged whether during practice or performance. Csikszentmihály terms the state of full engagement as "flow" [6]. People commonly experience flow in goal-directed systems where one's skills match the challenges at hand.

To maintain awareness of the goal when learning a piece, all three experts prefer to start with the big picture. The classical pianist likes to "make a sketch" with the "right gestalt" and then come back to fill in details. The cross-genre pianist begins a piece by analyzing its harmonic structure and finding where technical difficulties may occur.

Balancing one's skills and challenges during practice involves making sure tasks are neither too difficult to be discouraging nor too easy to be dull. When learning improvisatory passages by ear, the jazz pianist initially used pre-written transcriptions to verify his playing. When practicing difficult passages, the cross-genre pianist varies the material to avoid rote repetition.

MirrorFugue helps amateurs stay aware of the big picture and maintains musical interest while the user practices technical details.

Bibliography

1. Barbosa A. 2003. Displaced soundscapes: a survey of network systems for music and sonic-art creation. Leonardo Music Journal, Volume 13, Issue 1, 53-59.

2. Berklee Online Music School. http://berkleemusic.com/.

3. Burk, P. L. 2000. Jammin' on the web: a new client/ server architecture for multi-user performance. Proceedings of the 2000 International Computer Music Conference. San Francisco, California: International Computer Music Association, 117-120.

4. Calarts News. http://calarts.edu/news/2011-mar-31/jazz-crossings-links-venues-both-coasts-live-transcontinental-teleconcert/.

5. Casio. http://www.casio.com.

6. Csikszentmihalyi, M., Flow: The Psychology of Optimal Experience, Harper Perennial, 1991.

7. Evans, B. (1963). Conversations with Myself [Audio CD]. New York, NY: Verve.

8. Glenn Gould Website. http://www.glenngould.com/.

9. Gang, D., Chockler, G. V., Anker, T., and Kremer, A. 1997. TransMIDI: a system for midi sessions over the network using transis. In Proceedings of the 1999 International Computer Music Conference. San Francisco, California: International Computer Music Association, 283-286.

10. Grindlay, G. The Impact of Haptic Guidance on Musical Motor Learning M.S. Thesis. MIT Media Laboratory, 2007.

11. Ishii, H. 1990. TeamWorkStation: towards a seamless shared workspace. In Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work (Los Angeles, CA, October 1990). CSCW '90. ACM, New York, NY, 13-26.

12. Ishii, H. 1998. The Last Farewell: Traces of Physical Presence. In Interactions v4., ACM, New York, NY, 55-56.

13. Ishii, H. and Kobayashi, M. 1992. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. In Proceedings of

the SIGCHI Conference on Human Factors in Computing Systems (Monterey CA, May 1992). P. Bauersfeld, J. Bennett, and G. Lynch, Eds. CHI '92. ACM, New York, NY, 525-532.

14. Jürg Hanselmann: Pianist and Komponist. http://www.juerghanselmann.li/.

15. Klemmer, S. R., Hartmann, B., and Takayama, L. 2006. How bodies matter: five themes for interaction design. In Proceedings of the 6th Conference on Designing interactive Systems (University Park PA, June 2006). DIS '06. ACM, New York, NY, 140-149.

16. Konstantas, D. 1997. Distributed musical rehearsal. In Proceedings of the 1997 International Computer Music Conference. San Francisco, California: International Computer Music Association, 279-282.

17. Krueger, M. W., Gionfriddo, T., and Hinrichsen, K. 1985. VIDEOPLACE—an artificial reality. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Francisco, CA). CHI '85. ACM, New York, NY, 35-40.

18. Lewiston, C. MaGKeyS: A haptic guidance keyboard system for facilitating sensorimotor training and rehabilitation. PhD Thesis. MIT Media Laboratory, 2008.

19. McPherson, A. and Kim, Y., Multidimensional gesture sensing at the piano keyboard. In Proceedings of the 29th ACM Conference on Human Factors in Computing Systems (Vancouver, BC). CHI '11. ACM, New York, NY.

20. Martha Argerich Recordings. http://www.argerich.jp/.

21. Miwa, Y . and Ishibiki, C. 2004. Shadow communication: system for embodied interaction with remote partners. In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (Chicago IL, November 2004). CSCW '04. ACM, New York, NY, 467-476.

22. Molokans Around the World: http://www.molokane.org/.

23. Moog Piano Bar. http://www.moogmusic.com/.

24. Morikawa, O. and Maesako, T. 1997. HyperMirror: a video-mediated communication system. In CHI '97 Extended Abstracts on Human

Factors in Computing Systems: Looking To the Future (Atlanta GA, March 1997). CHI '97. ACM, New York, NY, 317-318.

25. MusicPath: Networking People and Music. http://musicpath.acadiau.ca/.

26. New World Symphony- Internet2. http://www.americanorchestras.org/technology/internet2.html.

27. Ng, K. 2008. Interactive feedbacks with visualisation and sonification for technology-enhanced learning for music performance. In Proceedings of the 26th Annual ACM international Conference on Design of Communication (Lisbon, Portugal, September 22 - 24, 2008). SIGDOC '08. ACM, New York, NY, 281-282.

28. Page, Tim (ed), The Glenn Gould Reader, New York: Alfred A. Knopf: 1984.

29. The Player Piano Page. http://www.pianola.com/.

30. Sarkar, M. and Vercoe, B. 2007. Recognition and prediction in a network music performance system for Indian percussion. in Proceedings of the International Conference on New Interfaces for Musical Expression . NIME '07. ACM, New York, NY.

31. Tang, J. C. and Minneman, S. L. 1990. VideoDraw: a video interface for collaborative drawing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People (Seattle WA, April 1990). J. C. Chew and J. Whiteside, Eds. CHI '90. ACM, New York, NY, 313-320.

32. Tang, J. C. and Minneman, S. 1991. VideoWhiteboard: video shadows to support remote collaboration. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology (New Orleans LA, April 1991). S. P. Robertson, G. M. Olson, and J. S. Olson, Eds. CHI '91. ACM, New York, NY, 315-322.

33. Wanderly, M. 2001. Performer-instrument interaction: applications to gestural control of sound synthesis. PhD Thesis, University Paris VI.

34. Wellner, P. and Newman, W. 1992. A desk supporting computer-based interaction with paper documents. In Proceedings CHI 92 Conference, 587-592.

35. Whitacre, Eric. Virtual Choir. http://ericwhitacre.com/the-virtual-choir/.

36. Yamaha. http://music.yamaha.com/.

37. Yamaha RemoteLive. http://prweb.com/releases/2011/04/prweb5224404.htm/.

38. Young, J. and Fujinaga, I. 1999. Piano master classes via the Internet. In Proceedings of the 1999 International Computer Music Conference. San Francisco, California: International Computer Music Association, pp. 135-137.

39. Youtube. http://youtube.com/.

40. Youtube Symphony Orchestra. http://youtube.com/user/symphony/.

41. Zimmermann, R., Chew, E., Ay, S. A., and Pawar, M. 2008. Distributed musical performances: Architecture and stream management. ACM Trans. Multimedia Comput. Commun. Appl. 4, 2 (May. 2008), 1-23.