MirrorFugue: Communicating Hand Gesture in Remote Piano Collaboration

Xiao Xiao and Hiroshi Ishii
MIT Media Laboratory
Tangible Media Group
75 Amherst Street, Cambridge, MA 02142
{x_x, ishii}@media.mit.edu

ABSTRACT
Playing a musical instrument involves a complex set of continuous gestures, both to play the notes and to convey expression. To learn an instrument, a student must learn not only the music itself but also how to perform these bodily gestures. We present MirrorFugue, a set of three interfaces on a piano keyboard designed to visualize hand gesture of a remote collaborator. Based their spatial configurations, we call our interfaces *Shadow*, *Reflection*, and *Organ*. We describe the configurations and detail studies of our designs on synchronous, remote collaboration, focusing specifically on remote lessons for beginners. Based on our evaluations, we conclude that displaying the to-scale hand gestures of a teacher at the locus of interaction can improve remote piano learning for novices.

Author Keywords
Gesture Interfaces, Computer Supported Learning, Computer Supported Collaborative Work, Music Learning Interfaces, Tangible User Interfaces

ACM Classification Keywords
H.5.3 Group and Organization Interfaces

General Terms
Design, Experimentation

INTRODUCTION
The creation of music is inextricably tied to the physical gestures of a performer on an instrument. Performance gestures include both the technique to play the notes and the movements imbuing expression to the sound [17]. These gestures play an integral role in music learning [7]. Learning by watching and imitating is crucial for students to acquire new techniques. Additionally, watching music being performed helps students process what they hear.

The physicality of sound creation limits musical collaboration in space. Though advancements in musical tele-presence networks enable remote performances, rehearsals, and lessons, music network systems have historically focused on achieving the highest fidelity reproduction of acoustic sound.

TEI'11, January 22–26, 2011, Funchal, Portugal.

Figure 1: Photographic mockup of Reflection mode

This paper examines ways of communicating hand gesture in remote collaboration interfaces on the piano, focusing specifically on how these interfaces can aid remote lessons. We outline related work in remote music networks, remote collaborative workspaces, and music learning interfaces. We then describe three MirrorFugue interfaces, which take inspiration from research on bodily presence in remote collaborative workspaces. We detail evaluations of our systems and discuss their results.

RELATED WORK

Remote Music Systems
Many projects have aimed to bridge the gap of distance in collaborative music by connecting remote players in a shared virtual space. Until recently, the latency of transferring high volumes of data over the Internet has influenced these systems' designs [1]. Several groups have focused on creating sound-only environments. TransMIDI [4] and the piano master classes of Young and Fujinaga [20] employed MIDI data instead of audio. Transjam [2] enabled musical applications that did not require playing "in-time".

Attempts at remote musical collaboration with traditional instruments have included video as well audio. Zimmermann's Distributed Immersive Performance project [21], designed to investigate distance playing of classical piano duos, showed over the shoulder video of the remote partner on screens in front of the players. The MusicPath system [13] for remote piano learning displayed an over the shoulder view of the student on a projection screen in front the teacher. Both Distributed Immersive Performance and MusicPath displayed the body-level but not hand-level gestures of the performer. For Distributed Immersive Performance, expressive body gestures helped players synchronize musically. Gestures conveyed in MusicPath allowed a teacher to comment on the full-body elements of technique (such how to use shoulders and upper arm) useful for more advanced students.

The Distributed Musical Rehearsals project [8] installed teleconferencing environments in studios incorporating high definition video and 3d surround sound to create the impression that the musicians in Geneva and the conductor in Bonn were rehearsing in the same room. Both the conductor and the musicians' full body-gestures were visible to the other side.

Interfaces for remote music collaboration have largely been disembodied. Those that did communicate gesture have not explored spatial arrangements to tightly integrate the presence of the remote collaborator into physical space. They have focused on body-level gestures and not the detailed gestures to play specific instruments.

Remote Collaborative Workspaces

Interfaces for telepresence and shared workspaces often employed optical metaphors to portray the body of a remote collaborator in a meaningful manner. VideoPlace [9], DoubleDigital Desk [18], VideoWhiteboard [16], and Shadow Communications [10] used shadows or silhouettes to present remote parties. Interfaces like HyperMirror [12] used the metaphor of a mirror to display all remote parties in a common virtual space.

Remote collaborative workspaces have borrowed interpersonal spatial relationships from collocated interactions to make sense of the connection between disparate locations. Some interfaces, such as TeamWorkStation [6], VideoDraw [15], and Double DigitalDesk presented the shared space as if the collaborators were working side by side, looking over the same area. Others, like as ClearBoard [5] and VideoWhiteboard, presented the shared space as if the collaborators were working across from each other. Both simulated common workspace configurations in collocated collaboration.

Different interfaces showed varying degrees of a remote collaborator's body, depending on the task and space for which the systems were designed. TeamWorkStation, VideoDraw, and Double DigitalDesk focused on actions over a shared workspace and only showed the hands and forearm of a remote collaborator. Interfaces like ClearBoard and HyperMirror, which tried to simulate face to face communication showed the entire upper body of the collaborator. Some interfaces, like VideoPlace and Shadow Communications tried to convey an abstract sense of full body presence by displaying life-sized silhouettes.

Unlike interfaces for remote music collaboration, remote collaborative workspace projects have developed a nuanced vocabulary to communicate the gesture of remote partners using optical metaphors of shadows and mirrors and spatial metaphors of face to face and side by side discourse. Whether the interface supported sharing of hand, arm, or full-body gesture depended on the task.

Music Learning Interfaces

Learning to play a musical instrument involves not only gaining proficiency in playing the correct notes but also developing expressivity on the instrument. Several commercial piano keyboards include built-in interfaces for note learning. Yamaha [19] and Casio [3] both manufacture keyboards with keys that light up from underneath, and the MOOG Piano Bar [11] uses LED's in front of the keys to indicate the notes of a piece. The Yamaha Disklavier actuates the keys themselves. All of these systems cue what keys to play but do not show any relationship between keys and hand gestures. MusicPath, which displays the body-level gestures of a remote student, obscures much hand-level gestures due to the angle of its video.

Learning interfaces for instruments with more continuous performance gestures focus on fine-tuning these bodily movements. A notable example is i-Maestro [14], which uses motion capture and playback to help violinists visualize and reflect on their playing.

DESIGN

In music, a *fugue* is a contrapuntal composition in two or more voices; a *mirror fugue* is a set of two fugues where each is the mirror image of the other. We designed MirrorFugue to enable musicians from separate spaces to play together through a metaphorical mirror space that transmits gestures associated with making the music.

We chose the piano as the basis of our interfaces because of its popularity among musicians and because of the clear relationship between a performer's hand gestures and notes played. Playing the piano well involves not only hand movements but also movements of the wrist, arms, shoulders, and even feet. We focus on remote lesson interfaces for beginners, who are most concerned with learning how to use the hands.

We designed three remote lesson interfaces: Shadow, Reflection, and Organ. We first outline the design space and then describe the three interfaces.

Design Space

We considered three factors borrowed from remote collaborate workspaces when designing our systems: interpersonal space, placement of remote space, and orientation of remote space.

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

Table 1: Comparison of design choices

Figure 2: The Shadow, Reflection, and Organ configurations

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 overlayed directly on the physical workspace (like Double DigitalDesk) 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 (like MusicPath).

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.

Interface Configurations

Shadow

Like Double DigitalDesk, this system uses shadows of hands projected directly onto the keyboard to convey the partner's presence. In our implementation, we project a video of the partner's hands instead of a strict silhouette.

The orientation of the projected video gives the impression that the remote partner is sitting next to the user. The projected keyboard from the remote video lines up exactly with the physical keyboard, and users can tell from the projected video which keys on the partner's side are pressed down. While this interface shows the remote partner's hands directly at the locus of interaction, it has the disadvantage that the shadow is not distinguishable when the two players' hands are at corresponding locations on the keyboard

Reflection

This system is inspired by the reflective surface on a lacquered piano where the keyboard and a player's hands are mirrored. Instead of showing the actual reflection of the present instrument, this system displays the reflection of the remote partner's hands. We prototyped Reflection by building a vertical back-projection surface directly in front of the keys onto which we projected the mirrored top-down video of another keyboard, distorting the image to simulate the perspective as seen from the player's point of view and making sure to line up the reflected keys with the physical keys. The implied interpersonal space is that of two players sitting across from each other. The remote keyboard is aligned and offset and is perceived to be displayed at 180 degrees in relation to the physical keyboard. In actuality, the remote keyboard is displayed on a vertical surface

Organ

Like Reflection, this configuration also uses the vertical display surface in front of the keys. Unlike Reflection, it displays an unaltered top-down image of the remote keyboard, with the bottom of the keys in the projection lined up with the physical keys. Like Shadow, the orientation of the projected video gives the impression that the remote partner is sitting next to the user. We call this configuration Organ because the position of the projected and physical keyboards is reminiscent of the two offset keyboards on an organ. The remote keyboard is shown at 90 degrees to the physical for better visibility to the player.

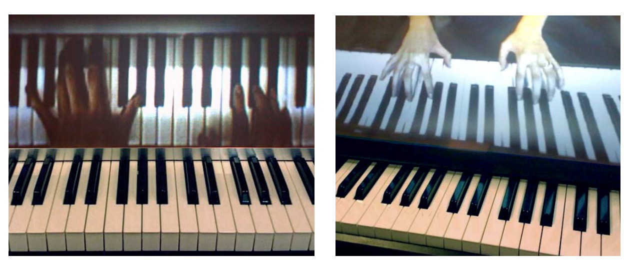

Figure 3: Prototypes of Organ and Reflection modes

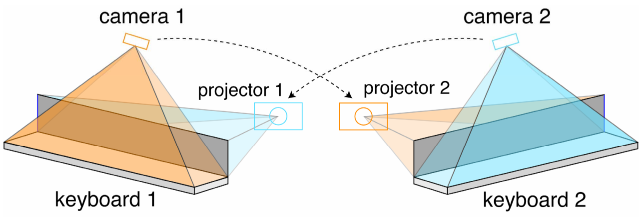

Figure 4: System Diagram

IMPLEMENTATION

We prototyped all three configurations with MIDI keyboards, wide-angle cameras, and projectors, using the MAX/MSP/Jitter platform to manage video and sound. We tested the remote communication by transferring 640x480 video at 30 frame per second over gigabit ethernet between two locations in the same building and were able to do so without noticible latency.

For our user studies, we placed two piano keyboards in the same room arranged to simulate a remote situation where two people cannot see each other but can hear each other and what is being played.

If our interfaces were deployed in actual remote locations, latency of video transfer over the Internet should not be a problem as long as the video is synced with the sound because student and teacher generally take turns playing during a lesson.

USER STUDIES

We conducted a pilot study to determine whether seeing the hands of a remote pianist helps musicians think about music, to identify the pros and cons of each configuration, and to establish which, if any, stands out over the others. We 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.

Pilot Study

Method

Five amateur pianists (4 men, 1 woman, aged 20-35) participated in our 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, we presented the participants with a keyboard in each of our three configurations in random order. The investigator (first author of this paper), who is an advanced level amateur pianist, was situated at the corresponding keyboard, which was not displaying the hands of the study participant. For each configuration, the investigator played some chord progressions for five minutes and asked the user to improvise a melody over the chords. After all three improvisation sessions, the investigator debriefed with each participant in an informal interview.

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 and other advanced musicians, 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, the Organ configuration was most preferred among almost all participants. All of the participants said that the Organ interface 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 the Organ setup 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 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 (self-taught classical pianist) expressed that he liked seeing the remote partner's hands in the same place as his own.

Several users found the Reflection interface 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 interface 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".

LEARNING STUDY

We designed an informal study to evaluate the Organ interface in the context of remote learning for novices. We chose this scenario because we envision that our system 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. We asked the following research quetions:

- 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?

Method

To answer these questions, we 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 our questions. One 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, which includes an LED in front of every key to help pianists visualize a piece, this setup indicates what is being played using abstract symbols. The other interface displays 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 simulates 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.

We 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 song, we began by having the instructor play the melody once through and then taught the piece in 3-5 note segments until the student can 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, we asked participants what they thought of each interface and to rank them by usefulness. We also videotaped each learning session to determine how long it took for users to learn each melody.

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

Table 2: Rankings of learning interfaces among users

Figure 5: Dots indicating what keys are pressed at the corresponding keyboard

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 setup because she felt like she learned the fastest on it. However, this user actualy 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 user who preferred the Organ setup suggested that we should "increase the resolution of the video so that the projection in Organ mode is more clear".

Quantitative

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

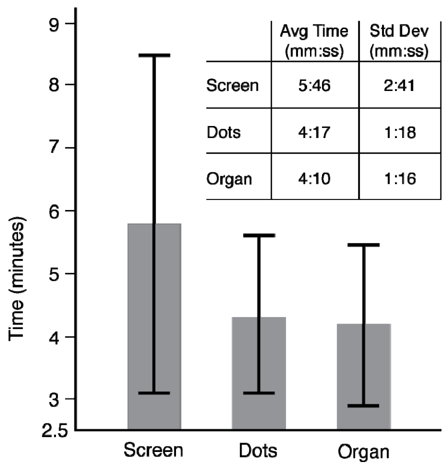

Table 3: Users took about 1 minute 30 seconds longer when learning melodies on Screen mode compared to both Dots and Organ modes.

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.

DISCUSSION

Study Reflections

Through the pilot study, we discovered that users seem to prefer the Organ configuration (side-to-side interpersonal space, aligned and offset image displayed at 90 degrees). In the learning study, we 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 alignment 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 the performance gestures of the teacher 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 forget 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.

SCENARIOS

Our evaluations focused on synchronous playing by two remote users. In the interviews following the pilot study, we hinted at other scenarios our systems can be used for. We discuss these in more detail.

The Shadow, Reflection, and Organ systems could be beneficial when used in asynchronous playing scenarios. In this case, asynchronous playing involves a user playing with a pre-recorded sound and video of hands. Asynchronous scenarios can be separated into one-user and two-user modes. A single users can record oneself as part of practice. Users can record one hand of a difficult piece and play the other hand along with the first hand to get a sense of what the whole piece sounds like. They can also record a section of a piece, such as a particular chord progression or rhythmic sequence, play it on loop on our device, and play along with the recording to practice improvisation and coordination. When more than one user is involved, one user can record a piece or a segment of a piece for another user, who can then listen and watch the recording, learn to play using it, or play along with it.

Although all of these usage cases are currently possible using a sound-only recorder, displaying hands at the keyboard along with playing sound can be beneficial for two reasons. First, as found in our evaluations, watching hands play as a supplement to hearing music can help users process and learn music better. Second, being able to see and hear a composite piece can be more fun for users and encourage them to play more.

Our systems can also be useful for collocated applications, such as rehearsals, improvisational jam sessions, and group lessons. In group rehearsals and jam sessions, our systems can give novice users a head start in learning to play with others by providing them with an extra means of coordination and anticipation with others. In group lessons, where one teacher teaches a room of several students each on their own keyboard, our systems can serve as a means to help the teacher better monitor the playing of each student.

Our designs, especially Mirror and Organ modes, can also support usage cases with more than two users displayed at a time in the form of semi-transparent overlays, a technique pioneered by interfaces such as TeamWorkStation. This way, users can visualize multiple people playing in the same space and can also play along with the group.

FUTURE WORK

In the future, we plan to investigate how our configurations perform in more scenarios, to incorporate other sensory modes in interfaces for remote piano learning, and to design interfaces that share gestures of more than just the hands.

To continue evaluation of our system, we would like to test the effectiveness of combining Dots mode with Organ mode in a learning scenario. In addition, we plan to design studies to explore how our configurations can help with single user practicing and with lessons for more advanced pianists.

As advanced piano technique involves gestures of the wrists, arms, shoulders, and feet, we plan on designing remote communication interfaces that make visible gestures of more of the body. We are also interested in exploring ways to integrate eye contact for coordination of remote playing.

CONCLUSION

We designed and prototyped three interface configurations to display to-scale hand-level perforance gestures of piano playing at the locus of interaction of the piano keyboard, which we collectively call MirrorFugue. The three spatial metaphors—Shadow, Reflection, and Organ—were inspired by research in collaborative workspaces. We compared our designs in a pilot study and concluded that the Organ configuration was the most helpful for users to better process music when playing. Our results from a second user study seem to verify that showing hands at the keyboard can help novices learn to play simple melodies more quickly and more effectively. Finally, we discussed several different scenarios for which our designs can be used that we plan to evaluate in the future.

ACKNOWLEDGMENTS

We are grateful for the support of the Tangible Media Group and the Things that Think Consortium at the MIT Media Lab. Thanks especially to Leonardo Bonanni and Cati Vaucelle for their invaluable advice. We would like to extend a special thanks to Avid for its general donations that made this project possible.

REFERENCES

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. 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.
3. Casio. http://www.casio.com/
4. 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.
5. 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.
6. 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.
7. 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.
8. Konstantas, D. 1997. Distributed musical rehearsal. In *Proceedings of the 1997 International Computer Music Conference*. San Francisco, California: International Computer Music Association, 279-282.
9. 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.
10. 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.
11. Moog Piano Bar. http://www.moogmusic.com/.
12. 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.
13. MusicPath: Networking People and Music. http://musicpath.acadiau.ca/.
14. 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.

15. 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.
16. 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.
17. Wanderly, M. 2001. Performer-instrument interaction: applications to gestural control of sound synthesis. PhD Thesis, University Paris VI.
18. Wellner, P. and Newman, W. 1992. A desk supporting computer-based interaction with paper documents. In *Proceedings CHI 92 Conference,* 587-592.
19. Yamaha. http://music.yamaha.com/.
20. 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.
21. 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.