

EmbodiNet: Enriching Distributed Musical Collaboration Through Embodied Interactions

Dalia El-Shimy^(✉) and Jeremy R. Cooperstock

Centre for Interdisciplinary Research in Music, Media and Technology,
McGill University, Montreal, QC, Canada
{dalia, jer}@cim.mcgill.ca

Abstract. This paper presents EmbodiNet, a novel system that augments distributed performance with dynamic, real-time, hands-free control over several aspects of the musicians’ sound, enabling them to seamlessly change volume, affect reverb and adjust their mix. Musical performance is a demanding activity necessitating multiple levels of communication among its participants, as well as a certain degree of creativity, playfulness and spontaneity. As a result, distributed musical performance presents a challenging application area for the “same time/different place” category of Computer-Supported Cooperative Work (CSCW). In fact, musicians wishing to play together over a network are typically limited by tools that differ little from standard videoconferencing. Instead, we propose leveraging the technology inherent to the distributed context towards meaningfully augmenting collaborative performance. In order to do so without introducing new paradigms that may require learning or that may distract musicians from their primary task, we have designed and evaluated embodied controls that capitalize on existing interpersonal interactions. Further designed to restore the spatial properties of sound that are typically absent in the distributed context, and apply the notion of “shared space” found in CSCW research, EmbodiNet also helps confer a greater level of co-presence than standard distributed performance systems. This paper describes the implementation of EmbodiNet, along with the results of a long-term collaboration and experiment with a three-piece band. The long-term collaboration helped illustrate the benefits of augmenting an artistic form of distributed collaboration, and resulted in a system that not only helped enhance our users’ sense of enjoyment and self-expression, but one that they would also likely use in the future.

1 Introduction

According to Ackerman, one of the challenges central to the field of Computer-Supported Cooperative Work can be described as the “social-technical gap”, a mismatch resulting from the flexible and nuanced nature of human activity when contrasted with the rigid and brittle nature of technical systems [1]. Thus, the author continues, bridging this gap through computational entities (e.g., information transfer, roles, and policies) that are also flexible and nuanced in nature,

is essential to the successful design of CSCW applications. This is particularly crucial for distributed collaborative environments, where participants often suffer from a lowered sense of shared awareness, and a decrease in mutual perception of non-verbal cues (e.g., gaze direction, gestures, posture) [38]. Such a problem has been tackled extensively by conventional investigations of videoconferencing technologies: telepresence systems, shared virtual table environments (SVTEs) and mobile remote presence (MRP) systems have all emerged in a bid to enrich social engagement within the distributed context. However, such systems strive to improve collaborations of a functional nature, or cooperation on specific, work-related tasks among remote participants [21]. In an effort to explore the breadth of human activity that computer-mediated communication can enrich, we became particularly interested in examining the creative, ludic and spontaneous aspects of social interaction within a distributed context. An additional motivation was exploring whether distributed collaboration could improve on its co-present counterpart by leveraging its underlying technology towards further assisting target users in effectively accomplishing the activity at hand. One area particularly suited for such investigations, given its socially and temporally exacting nature, is that of distributed musical performance.

We decided to examine such challenges by taking a user-driven approach to the design of an augmented distributed performance environment. By choosing an application area where communication is strongly driven by creativity, self-expression and spontaneity, we wanted to explore the ways we could better support the “highly flexible, nuanced, and contextualized” aspects of human activity [1]. Furthermore, as Corness and Schiphorst explain, “[p]erformers tacitly know how to pay close attention to bodily cues that accompany movement, as they have consciously developed their awareness of these cues to enable skilled interaction with other performers” [15]. Thus, we hoped that capitalizing on embodied performer-performer interactions would offer the added advantage of enabling musicians to use our system’s functionality without detaching themselves from the higher level task of performance. Finally, by creating a system that allows musicians to experiment with paradigms that traditional performance does not offer, we sought to examine whether the distributed version of a collaborative activity could offer unique benefits of its own.

Our efforts resulted in EmbodiNet, an augmented distributed performance environment that allows musicians to utilize common gestures and behaviours, such as head tilting, body turning and simple motion, as a means of affecting each other’s volume and reverb levels, adjusting audio mixes and experiencing spatialized sound. EmbodiNet was designed for relaxed performance settings that include room for improvisation or experimentation (e.g., loose rehearsals or jams). An example use case scenario for our system would involve geographically displaced friends who wish to play music together over a network, but seek alternatives to traditional videoconferencing that can further enrich their interpersonal interactions. EmbodiNet can currently only support electric or electronic (rather than acoustic) instruments, in order to ensure that the modified audio mix played back through the musicians’ headphones is not overshadowed by the actual sound of their instruments.

To the best of our knowledge, EmbodiNet is the only distributed performance system of its kind that simultaneously: (1) exports the notion of “shared space” from the CSCW domain to the distributed performance context, allowing musicians to perceive local and remote environments as simple extensions of one another, (2) uses shared space as a means to restore the spatialization of musical instruments that is inherent to the co-present context, yet lost in the distributed one, (3) capitalizes on embodied interactions as a means of control, and (4) offers performers the ability to affect one another’s sound parameters through their interpersonal interactions. Together, such properties allow EmbodiNet to confer a greater level of co-presence than traditional solutions for online performance. The results of our long-term study with a three-piece band confirm that musicians found EmbodiNet to be enjoyable and useful, and that they would likely use it again in the future.

2 Related Works

Given its interdisciplinary nature, our work draws inspiration from a variety of research areas. While existing systems for distributed performance have naturally influenced our work [13, 14, 26], we were interested in performance environments developed specifically to explore the implications of the network as “a space for being” [34], rather than simply mimic co-present performance. Examples include Barbosa’s Public Sound Project [3], Tanaka and Bongers’ Global String [39], Braasch et al.’s ViMic system [8], as well as the works of Rebelo, Schroeder and Renaud, which emphasize the network as both an acoustic and social medium [31, 32, 35]. A thorough overview of distributed performance environments in relation to our work is presented elsewhere in references [20].

It should come as no surprise, however, that the act of distributing performance over a network would have a strong impact on the nature and level of communication between remote musicians. Renaud, for instance, explains that “[i]nteraction is a real issue in network performance systems as natural visual or sensory cues, such as breathing and gesture, are completely removed from context” [9]. To this, Kapur adds that “[w]aiting backstage to go on, and important aspects of socialisation after a performance, are not the same over a network”, leading to a “loss of society within the band” [26]. In that sense, distributed performance shares a common challenge with Computer-Supported Cooperative Work, a set of activities that also often exhibit a decreased sense of mutual awareness and spontaneous interaction [16, 38]. In fact, we regard distributed performance as a unique application area of the “same time/different place” category of CSCW [25]. As such, the design of systems for distributed performance can benefit from a number of CSCW research topics aiming to facilitate remote collaboration. One such area relevant to our work is that of awareness, described by Dourish and Bly as the ability to know “who is ‘around’, what activities are occurring, who is talking with whom” [16]. Awareness entails a certain level of transparency among remote participants, allowing them to develop a sense of trust and community that, in turn, encourages the playful and creative

sides of interaction that are crucial to successful musical collaboration. Another concern is providing support for the “rich set of social behaviours and cues that we as humans know and share” [2], such as body postures, subtle movements, gaze direction, room acoustics, joint interactions, eye contact and other forms of non-verbal communication [17]. As Sirkin and Ju explain, “[w]e use embodied non-verbal communications such as gestures, body movements, posture, visual orientation, and spatial behavior in concert with our verbal communication to signal our attention, express emotions, convey attitudes, and encourage turn-taking, and... we (perhaps subconsciously) prefer that our technological counterparts follow suit” [38]. An example of a system designed to support such cues, and which directly inspired our display topology, is Hydra, a set of independent communication units, each with its own video display, microphone and speaker. As such, when distributed on a local participant’s desk, Hydra units allow for the spatial and acoustical separation of remote collaborators [37]. Another example from CSCW research that came to influence our work is Ishii’s notion of shared workspaces, conceived as continuous extensions of individual work areas that afford a seamless, two-way transition between collaborative and individual modes of work. In fact, we argue that shared workspaces, as seen in the TeamWorkStation [23] and ClearBoard projects [24], exemplify the philosophy found in the literature on distributed performance of being “in” the network, and described above. Such ideas led to our design of a system configuration that can support the illusion of “shared space”, as described later in this paper. The relationship between CSCW and distributed performance is further expounded upon elsewhere in reference [19].

Finally, Computer-Supported Cooperative Work research illustrates that successful collaboration over a network is contingent not only on resolving technological challenges, but also on the development of interaction paradigms that can support both the complexities and subtleties of cooperative behaviour [1, 7, 11, 12, 33]. However, exporting this notion to musical context begets some interesting implications: musical performance is a temporally exacting activity, demanding multiple levels of communication between the players [14]. Therefore, any novel interfaces aiming to augment or facilitate such an activity should be designed, whenever possible, with the intent of reducing the cognitive load they may pose on their users, and avoid distracting them from the higher-level task of performance. As such, embodied interaction, a notion built on the premise of capitalizing on a “broader spectrum of human skills and abilities” [28], lends itself quite naturally to the design of musical interfaces. Our work exemplifies the definition of embodied interaction provided by Antle et al., who describe such an approach as “leveraging users’ natural body movement in direct interaction with spaces and everyday objects to control computational systems” [5]. In fact, embodied interaction has proven to be a suitable option for the design of many non-utilitarian applications [4, 29], including musical interfaces [5, 6, 15].

In summary, to situate our work within the research areas described above, EmbodiNet is a distributed performance environment that offers musicians the illusion of “shared space”, and allows them to utilize embodied interactions to manipulate sound parameters, with the aim of augmenting and improving a

unique form of online collaboration. The ideas behind EmbodiNet reside at the intersection of CSCW and distributed performance research, two fields we believe share many similar challenges and yet which, with very few exceptions [21], have yet to benefit from a full bidirectional flow of information.

3 System Description

Our performance environment was deployed across three separate locations. In order to provide musicians with the illusion that they could physically interact in relation to one another, we used a “shared space” metaphor, whereby each of the musicians’ local spaces are mapped onto the Cartesian plane such that they border one another without overlapping, creating, in essence, one large seamless area. Such a configuration for three musicians can be seen in Fig. 1. This solution, in turn, allows the *virtual locations* of remote musicians to appear as though located within an extension of each local musician’s space. When applied to a scenario with three musicians, the virtual locations of remote collaborators places them on either side of the local musician. To support this configuration, every location was equipped with two monitors, each displaying a view of one of the remote spaces. To prevent users from falling out of view as they move about their space, ensure reasonable support for eye contact and, in turn, confer a greater sense of mutual awareness, a camera was mounted behind each monitor, thereby maintaining a line of view between the distributed musicians. Tracking of user position and orientation was carried out with Microsoft Kinect units.

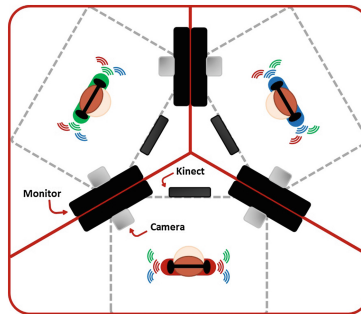


Fig. 1. Mapping of three musician locations to create a sense of shared space.

3.1 Features

The current implementation of EmbodiNet encompasses five unique features:

- **Dynamic Volume:** As one musician moves *towards* or *away* from another’s virtual location, both can experience each other’s instrument sounds as gradually increasing or decreasing in volume.
- **Dynamic Reverb:** As one musician moves *away* or *toward* another’s virtual location, both can experience each other’s instrument sounds as gradually increasing or decreasing in reverberation, or “reverb”.

- **Mix Control:** A local musician can change the mix of his instrument with those of the remote musicians by tilting his head in the direction where he wants to concentrate the sound of his own instrument. The remote instruments continue to be heard in either left or right headphones as appropriate to their direction.
- **Track Panning:** A local musician can isolate each of the tracks of the remote musicians by changing his body’s orientation.
- **Musician Spatialization:** A local musician can experience the remote musicians’ instruments as spatialized sound sources within his own, local space.

3.2 Graphical User Interface

EmbodiNet supplements shared video with a simple graphical user interface (GUI), seen in Fig. 2, appearing on a computer monitor positioned in front of each musician. Not only does the GUI give the musicians complete control over the system features, it also provides simple yet effective dynamic visual representations of the state of their performance at a glance, in an effort to further increase their level of mutual awareness.

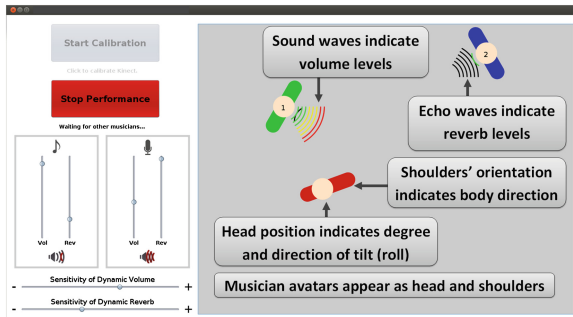


Fig. 2. Graphical user interface, which includes a control panel and animated graphics. The local musician’s avatar is in red (Color figure online).

3.3 Configuration

EmbodiNet’s hardware configuration can be seen in Fig. 3. We opted to create a simple yet stable setup using analog cameras connected directly to Panasonic BT-LH1700W production monitors that were, in turn, located on either side of each computer monitor displaying EmbodiNet’s GUI. As described earlier, each location includes two monitors, with a camera mounted behind each to maintain a reasonable line of sight across the distributed musicians. Each musician’s instrument, along with a microphone for verbal communication and singing, are plugged into a Roland Edirol FA-101, an audio capture interface. The signals are then processed through SuperCollider, an open source environment and programming language for real-time audio synthesis and algorithmic composition,

where they are adjusted in accordance with the system features described above. The audio streams from SuperCollider are subsequently shared among all three locations through Jacktrip, a tool designed for multi-machine network performance over the Internet. To further reduce delay and guarantee sound stability, a real-time kernel is used on all machines executing Jacktrip, and a Local Area Network (LAN) was created to connect them through a Netgear ProSafe 8 Port Gigabit Switch. Finally, each musician is able to hear his own individual mix through a pair of Sennheiser HD Pro 280 closed headphones.

We measured the end-to-end latency between locations resulting from our hardware and software configurations to be approximately 16 ms. Although Carôt et al. have argued that the maximum delay tolerated by musicians can depend on an ensemble’s style, performance speed and rhythm [10], the “Ensemble Performance Threshold” of 25 ms is commonly regarded as a value beyond which distributed musicians begin to experience difficulties remaining in sync [36]. As such, we note that, while much research in network musical performance has focused on decreasing latency to levels that musicians could tolerate, our aim was not to replicate such results, but rather design and evaluate interactions that could further augment distributed performance environments where latency can be considered a non-issue.

The musicians’ position and orientation data was captured using a Microsoft Kinect, and sent to our SuperCollider software via OpenSoundControl (OSC) messages.

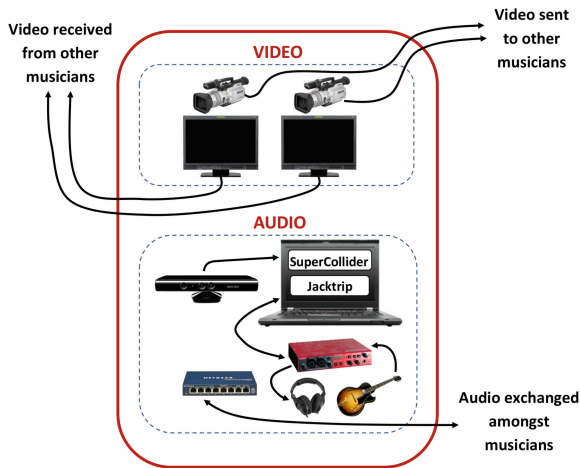


Fig. 3. Hardware configuration for EmbodiNet.

4 Long-Term Collaboration

EmbodiNet evolved through a series of prototypes and formal user tests, a process depicted in Fig. 4. In the interest of space, we will not detail our previous

efforts here, although they are described elsewhere in references [18,20]. In this section, we instead describe the a long-term collaboration and experiment (highlighted in red in Fig. 4) that led to the current implementation of EmbodiNet described in this paper.

Throughout the evolution of EmbodiNet, we had noted that the ‘one-off’ nature of traditional formal user experiments did not provide us with an opportunity to test the effects of small, iterative changes to our system on a regular basis. Furthermore, we questioned whether feedback from first-time users might be biased by the novelty of the system. Thus, after we implemented the beta version of EmbodiNet, which, as seen in Fig. 4 had come to include the Dynamic Volume, Track Panning and Musician Spatialization features, we were motivated to elicit feedback that went beyond simple novelty effects and initial impressions. Inspired by Grudin’s views on the importance of long-term system evaluations within CSCW research [22], and the success of such a methodology within the contexts of both remote collaboration [27] and musical performance [15], we sought to combine the benefits of quantifiable, repeatable user studies and the rich feedback inherent to participatory design by merging elements of both methodologies into a long-term testing and collaboration cycle.

As a result, we invited a band consisting of a 25-year-old guitarist, a 26-year-old keyboardist—both of whom also alternated lead and backup vocals—and a 22-year-old bassist for a series of performance sessions with our beta system. All three were male, and had performed together approximately once per week for almost two years. An introductory brainstorming session was first held, allowing us to showcase our existing system features to the band members, and discuss our vision for the long-term collaboration. Subsequently, we organized weekly meetings that combined formal, quantitative tests with informal, yet in-depth, qualitative discussions.

Since our goal was to “to discover rather than to verify” the effects of each system feature on these various aspects of performance, we knew that the qualitative experiment framework proposed by Ravasio et al. was most suitable to our needs [30]. Therefore, we employed both of their techniques of separation/segmentation and adjection/intensification to design a number of sessions, each focusing on a different feature of the system through an A/B/A-style test, where musicians performed once without the feature, once with the feature, then once again without the feature. At the beginning of each session, musicians were asked to select their base volume levels collaboratively until they reached a satisfactory mix. It is those base levels that our features would subsequently affect during condition B. Each condition lasted approximately 15–20 min, or the time it took the musicians to play through three songs. Musicians were not required to carry out any specific tasks under each condition, only to perform songs of their choice, while voicing to one another or to the test instructor any feelings or concerns they may have throughout the session.

The participants also completed post-condition questionnaires tailored to assess three performance criteria that our previous work, as described elsewhere in Ref. [18], had shown musicians in general tend to deem valuable. Namely, these were enjoyment, creativity and self-expression. Position and orientation data was

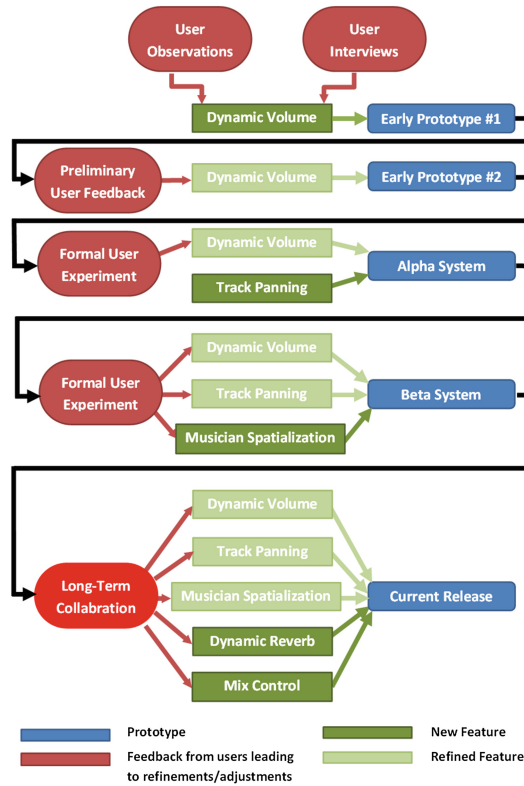


Fig. 4. User-Centered evolution of EmbodiNet (Color figure online).

collected throughout, along with video footage and audio recordings. After the formal test component of each session, an open discussion in the style of a non-leading interview was held. Musicians were loosely probed about their approach towards the performance and their feelings about the system, and encouraged to provide criticisms, along with suggestions for improvement.

4.1 Session 1: Musician Spatialization

Our first session with the band was designed to focus on the Musician Spatialization feature, whereby the sounds of remote instruments are perceived as emanating from the correct spatial position within the musician’s local environment. This feature helps mimic the spatialization effects naturally experienced in a co-present setting, where performers can easily perceive the distance and direction of other instruments surrounding them based on their position and orientation. In that manner, Musician Spatialization was designed to restore some of the natural acoustic dynamics that are lost in typical distributed performance.

Unlike other system features, no explicit gesture is required to activate Musician Spatialization: as long as the feature is enabled, audio from remote musicians will continue to be rendered. However, our post-test discussion with the musicians revealed that the “passive” nature of the feature had somewhat confused them. The guitarist, for instance, explained:

“I could tell there were changes happening when there were changes happening, but I really had difficulty at times making sense of it.”

Although its mapping was discussed with them before the performance, the musicians continued to look for a “triggering” gesture that would allow them to control the effect. After additional explanation regarding the feature’s passive nature was offered, the musicians reflected further on their performance, and subsequently indicated that they would be inclined to try it again in light of their new understanding. We suspect that the explanation of the feature we had originally provided lacked sufficient clarity, seeing as the very same implementation of Musician Spatialization was eventually met with more success when the musicians were given another opportunity to test it in Session 5, described below.

Analysis of position and orientation data did not reveal any significant changes in behaviour when Musician Spatialization was used.

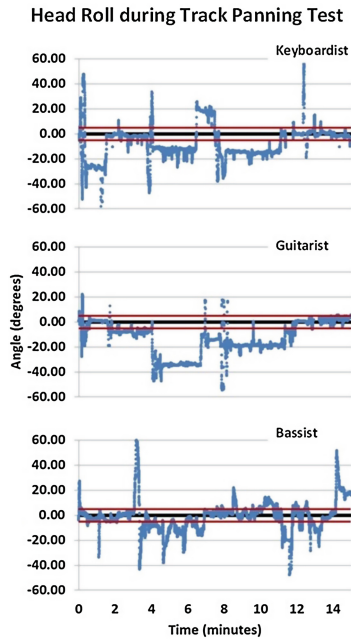


Fig. 5. Head roll, or tilting, data for all three musicians when Track Panning was used. The red line represents the threshold of ± 5 degrees, beyond which the feature was activated (Color figure online).

4.2 Session 2: Track Panning

The second session focused on the Track Panning feature, again through the form of an A/B/A test, followed by a discussion. At the time of testing, Track Panning had been implemented as a function of head roll, and the Mix Control feature had not yet been conceived of. As seen in Fig. 5, orientation data from the formal tests indicates that, while all three musicians experimented with the feature, the keyboardist and guitarist felt more inclined to sustain their interaction for longer periods of time. The guitarist, in particular, regularly isolated the bass track by turning his head to the left, and explained later that it helped him maintain his rhythm.

During the post-test discussion, some of the musicians criticized the head-tilting gesture of the Track Panning feature, noting that it would feel more “natural” to turn one’s *body*, rather than tilt one’s head, towards the virtual location of another musician on whose track they wanted to focus.

Nonetheless, the musicians did appreciate the practical aspect of the function. For instance, when asked to envision use case scenarios for such a feature during performance, the keyboardist explained:

“Well, mid-performance, say there was a part in the song where a few people were harmonizing together, if I could turn to the screen and we could hear each other better that way, like that would be practical for sure.”

The musicians suggested that the head tilting gesture would be better suited to listening closely to a mix, as musicians often do in a studio setting, leaning their heads into one headphone at a time. This gave rise to the idea behind the Mix Control feature, whereby a local musician could listen to his own instrument gradually being mixed with either of the remote musicians’ one at a time, simply by tilting his head in the direction corresponding to the remote musician’s virtual location. Musicians had an opportunity to test this new feature, along with an updated version of Track Panning in Session 5, as described below.

4.3 Session 3: Dynamic Volume

The third session included an A/B/A test of the Dynamic Volume feature. Analysis of position data, shown in Fig. 6, revealed that the use of this feature generally helped encourage all three musicians to increase the range of space they covered, rather than maintaining a fixed location, as they were inclined to do otherwise.

During the post-test discussion, the musicians also expressed their interest in controlling another aspect of their sound beyond volume level, namely reverberation. This was considered as a suitable addition to enhance creativity, allowing the musicians to experiment with different sounds. According to the musicians, an increase in reverb when moving further away from each other’s virtual locations could further enhance their feeling of shared space, giving them a more concrete sense of dimension due to the “echoing” nature of this effect.

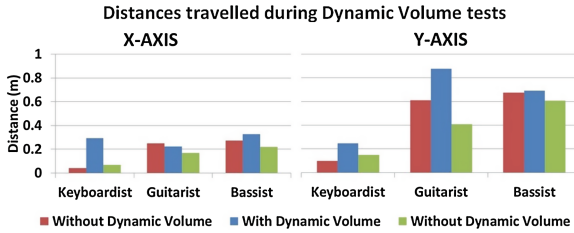


Fig. 6. Distances travelled by all musicians during dynamic volume tests.

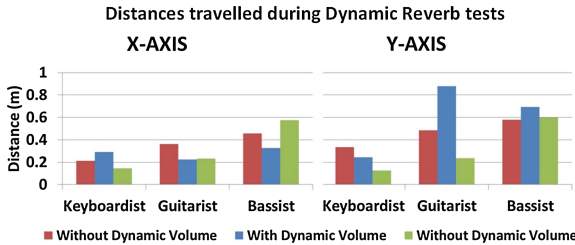


Fig. 7. Distances travelled by all musicians during dynamic reverb tests.

4.4 Session 4: Dynamic Reverb

We held an interim session where the musicians were invited to experiment with reverb used to simulate rooms of different sizes, and help design the overall effect. Subsequently, the fourth session was centered on the A/B/A testing of the newly implemented “Dynamic Reverb” feature.

Similar to the earlier Dynamic Volume feature, Dynamic Reverb helped increase the interpersonal interaction between musicians, and generally encouraged them to take full advantage of the available space (see Fig. 7). Furthermore, in the post-test discussion, the musicians revealed that they were quite pleased with the feature, with the guitarist stating:

“I felt that it kind of reacted how I would have wanted it to. It felt a bit like I was able to use it and predict how it was gonna be. It was cool.”

Furthermore, the guitarist added that while this feature did not necessarily serve a utilitarian purpose, it had an overall positive impact on the performance’s aesthetics:

“I thought it sounded great, like I just liked the sound. A bit of wetness... It doesn’t really have so much utility so much as it is just an aesthetic thing, it feels natural to have it on”

4.5 Session 5: Freestyle

The fifth session was a “freestyle” performance: the musicians were simply asked to jam for an hour, selecting which features to turn on or off throughout according to their needs. This session also provided the musicians with the opportunity

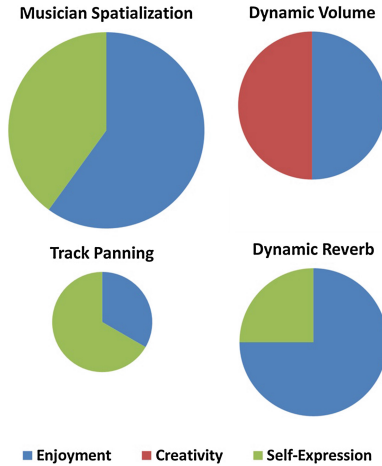


Fig. 8. Effect of each feature on performance criteria. The size of each pie represents the total number of times, across all criteria, that an improvement was marked.

to test the newly-implemented Mix Control feature, as well as the new versions of Track Panning, now a function of body orientation rather than head tilt. The performance was again followed by a discussion, where musicians provided their opinions of the overall state our system had reached as a result of our on-going collaboration.

Having had the chance to re-visit Musician Spatialization in light of their improved understanding of its functionality, the feature proved to be popular with the guitarist and keyboardist, who were able to finely control it, now that the mapping had been made clearer to them. When asked whether they would use the system in a scenario where they could not be physically co-located, all three agreed that the features would be quite beneficial in facilitating distributed collaboration. The keyboardist, for instance, stated:

“I think it’s like, if we’re doing something like jamming in different cities, any sort of software that has extras like that, would be fun... it could be a means to prolong your jam if it’s getting boring or something. You could try different sounds or just mess around with it. But there’s a practicality to the features too.”

Throughout all sessions, the musicians had also been providing feedback on improving the overall sound of the system, recommending preferred volume and reverb levels, and suggesting means to reduce any distortion. By Session 5, all of them were very pleased with how far the system had evolved, describing the sound as far “smoother” and more pleasant to the ear than it was at the start of our collaboration. For instance, when asked how they would gauge the changes in sound quality based on their previous suggestions, the guitarist explained:

“It’s definitely come a long way in terms of the quality of the sound that’s coming through my ears. So that’s the idea, I guess. It sounds good, so that’s good.”

4.6 Additional Aggregated Results

Analysis of Post-condition Questionnaires. As noted earlier, musicians also completed post-condition questionnaires during each of the A/B/A tests. The questions were designed to assess a number of factors, such as the musicians’ perceived sense of enjoyment, creativity and self-expression. Responses were tabulated and analyzed to determine the number of musicians for whom each of the system’s features helped improve the factors listed above. As seen in Fig. 8, all features helped contribute to increased levels of enjoyment, with Musician Spatialization and Dynamic Reverb performing best in that regard. Furthermore, Track Panning contributed to an improvement in the musicians’ sense of self-expression. Overall, however, creativity appeared to be the factor that least benefited from our system features, increasing only when Dynamic Volume was in use.

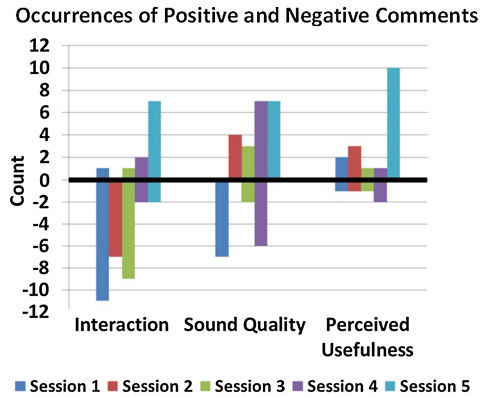


Fig. 9. Occurrences of positive and negative comments made under each major category in post-performance discussions.

Analysis of Post-test Discussions. All of our post-test discussions with the musicians were recorded and transcribed before a Qualitative Data Analysis (QDA) was performed. During a repeated coding process, comments were labelled and grouped, until three major categories emerged: Interaction, Sound Quality and Perceived Usefulness, and comments under each were tagged as being “positive” or “negative”.

As seen in Fig. 9, the number of positive comments for each category slowly improved throughout the sessions, with a particularly sharp increase in the “Interaction” and “Perceived Usefulness” categories seen during Session 5. We believe this to be, in large part, due to the nature of the session itself, as musicians were given the opportunity to try out the system features after all the feedback and suggestions they provided had been incorporated. In contrast, Fig. 9 also shows a steady decrease in the number of negative comments made for all three

categories. Together, these results indicate that we were successful in systematically incorporating the musicians' feedback into our system design. In the end, the band members found the system that evolved from our weekly sessions to be a vast improvement over its predecessor.

4.7 Discussion

Our long-term deployment with the three-piece band was not only beneficial in allowing us to fine-tune EmbodiNet's existing features and introduce new ones, but it also helped us better understand the effects of embodied interaction on distributed performance. As Figs. 6 and 7 illustrate, the musicians made greater use of their local spaces when the Dynamic Volume and Dynamic Reverb features were in use. This is a marked improvement over traditional distributed performance systems, which, much like standard videoconferencing systems, do not encourage, or in some cases even support, a greater level of movement. Since Dynamic Volume and Dynamic Reverb also allow musicians to mutually affect one another's volume and reverb levels, their use also marks an increase in the sense of interplay among distributed musicians. In addition, Track Panning helped facilitate performance by helping the musicians maintain their rhythm with one another, while Dynamic Reverb added an aesthetic dimension to their sound while helping to reinforce the notion of a shared space. We also note that, with the exception of Musician Spatialization requiring a second explanation, the musicians experienced no difficulty in understanding and using our system features, and required little to no training. This helped further illustrate that, by designing them to capitalize on common gestures, embodied interactions can help enhance rather than detract from the higher-level task at hand.

While co-present performance, much like face-to-face communication, will always remain the gold standard, the musicians expressed that in a scenario where they were displaced, our system features would very much entice them to partake in an activity that they, like many other musicians, would otherwise likely not consider. Nonetheless, our long-term collaboration with the musicians also uncovered a number of shortcomings. First, some of EmbodiNet's features can only be experienced passively by seated musicians and, therefore, alternative controls must be designed to better accommodate such participants. In addition, with the exception of Dynamic Volume, the features did not necessarily help the musicians feel more creative. As we consider creative engagement to be integral to the musical experience, additional work is required in order to determine how EmbodiNet can better support such a quality. Finally, having established the practical aspects of seamless volume, reverb and mix adjustments, we would also like to explore the effects of providing musicians with controls of a more abstract or artistic nature.

Through EmbodiNet, we also hoped to examine how the shortcomings of distributed collaboration could be resolved in a manner that not only bridges the gap between the co-present and distributed contexts, but also serves to further enhance those aspects of the activity at hand that our users had deemed

most valuable. The underlying basis of telepresence research, the most prominent example of the “same time/different place” category of CSCW, has been to engender, as best as possible, a feeling of co-presence through the support of the non-verbal cues and gestures that are typically poorly conveyed between remote participants. The problem with such an approach is that it can only, at best, mimic co-location. Within the context of distributed performance, a parallel methodology is perhaps best illustrated through the breadth of research that aims to decrease latency and increase bandwidth as a means of facilitating musical collaboration. We argue, however, that the goal of distributed systems should not stop at simply mirroring their co-present counterparts. Distributed collaborative environments must, by nature, introduce a certain level of technology to offer their users support over even the most basic aspects of cooperation. As a result, we question whether participants stand to benefit from developers leveraging the technology at their disposal towards *augmenting* the activities these systems afford.

In our case, supporting audio sharing, the most elementary aspect of network performance, meant equipping each location with tools that musicians do not require under normal circumstances. With the addition of a Microsoft Kinect, we were able capitalize on the computing power necessary to make distributed performance a possibility, in a bid to present the network as a unique and appealing medium in its own right to the less technologically inclined musicians. In the end, our goal of augmenting an existing activity in a manner that utilizes existing, well-understood embodied interactions helped the musicians perceive distributed performance as a activity in which they would likely partake again, as expressed by members of the three-piece ensemble at the end of our long-term collaboration.

5 Conclusions

The on-going trend in the “same time/different place” category of Computer-Supported Cooperative Work has been to support, as best as possible, practices that might engender the feeling of co-presence (e.g., telepresence systems). In other words, the goal of such systems is to mimic co-present collaborative environments, in large part through support of the non-verbal cues and gestures that are often poorly conveyed between remote participants. We argue, however, that the goal of distributed systems should not simply stop at mirroring their co-present counterparts. Instead, such systems can leverage their underlying technology towards augmenting, and in turn, perhaps even improving, existing activities.

Through our experience with EmbodiNet, we tested this philosophy within the context of distributed performance, a domain we view as a unique extension of the “same time/different place” category of CSCW systems. EmbodiNet augments distributed performance by capitalizing on embodied interpersonal interactions among remote musicians. Through the use of five unique features, Dynamic Volume, Dynamic Reverb, Track Panning, Mix Control and Musician

Spatialization, musicians are able to seamlessly create and alter individualized mixes mid-performance, simply by moving around their space. Furthermore, by responding to changes in position and orientation, our performance environment allows musicians to utilize its features without having to detach themselves from the primary task of music-making.

The latest implementation of EmbodiNet was the result of a long-term collaboration and experiment with a three-piece band. Questionnaires collected throughout the collaboration proved that our system helped enhance the musicians' sense of enjoyment and self-expression. Position and orientation data indicated that the musicians took advantage of EmbodiNet's features, leading to an increased sense of interplay and spontaneity in spite of their remoteness. Furthermore, qualitative analysis of our discussions with the band members has shown that they found EmbodiNet to be practical, and that they would likely use it again in the future. By augmenting distributed performance, EmbodiNet, helped present such a domain as different from its co-present counterpart, yet appealing in its own right. As a result, we believe that the design of CSCW systems may benefit from a similar examination of how the collaborative activities they support could come to offer advantages that cannot be offered in traditional co-located contexts.

References

1. Ackerman, M.S.: The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. *Hum.-Comput. Interact.* **15**(2), 179–203 (2000)
2. Adalgeirsson, S.O., Breazeal, C.: MeBot: a robotic platform for socially embodied presence. In: *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 15–22 (2010)
3. Àlvaro, B., Kaltenbrunner, M.: Public sound objects: a shared musical space on the web. In: *Proceedings of the International Conference on Web Delivering of Music*, pp. 9–15. IEEE Computer Society Press (2002)
4. Antle, A.N., Corness, G., Bevans, A.: Springboard: designing image schema based embodied interaction for an abstract domain. In: England, D. (ed.) *Whole Body Interaction. Human-Computer Interaction Series*, pp. 7–18. Springer, London (2011)
5. Antle, A.N., Corness, G., Droumeva, M.: Human-Computer-Intuition? Exploring the cognitive basis for intuition in embodied interaction. *Int. J. Arts Technol.* **2**(3), 235–254 (2009)
6. Bakker, S., Antle, A.N., Van Den Hoven, E.: Embodied metaphors in tangible interaction design. *Pers. Ubiquit. Comput.* **16**(4), 433–449 (2012)
7. Bannon, L.J., Schmidt, K.: CSCW: four characters in search of a context. In: Bowers, J.M., Benford, S.D. (eds.) *Studies in Computer Supported Cooperative Work*, pp. 3–16. North-Holland Publishing Co., Amsterdam (1991)
8. Braasch, J., Valente, D.L., Peters, N.: Sharing acoustic spaces over telepresence using virtual microphone control. In: *Audio Engineering Society 123rd Convention* (2007)

9. Carôt, A., Rebelo, P., Renaud, A.B.: Networked music performance: state of the art. In: Audio Engineering Society Conference 30th International Conference: Intelligent Audio Environments (2007)
10. Carôt, A., Werner Fischinger, C.: Towards a Comprehensive Cognitive Analysis of Delay-Influenced Rhythmical Interaction, pp. 473–476. MPublishing, University of Michigan Library, USA (2009)
11. Carstensen, P.H., Schmidt, K.: Computer supported cooperative work: new challenges to systems design. In: Itoh, K. (ed.) Handbook of Human Factors, pp. 619–636 (1999)
12. Carstensen, P.H., Schmidt, K.: Computer Supported Cooperative Work: New Challenges to Systems Design (1999)
13. Chafe, C., Wilson, S., Leistikow, A., Chisholm, D., Scavone, G.: A simplified approach to high quality music and sound over IP. In: Proceedings of the COST G-6 Conference on Digital Audio Effects, pp. 159–164 (2000)
14. Chew, E., Sawchuk, A., Tanoue, C., Zimmermann, R.: Segmental tempo analysis of performances in performer-centered experiments in the distributed immersive performance project. In: Proceedings of International Conference on Sound and Music Computing (2005)
15. Corness, G., Schiphorst, T.: Performing with a system’s intention: embodied cues in performer-system interaction. In: Proceedings of the 9th ACM Conference on Creativity and Cognition, pp. 156–164 (2013)
16. Dourish, P., Bly, S.: Portholes: supporting awareness in a distributed work group. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 541–547 (1992)
17. Eisert, P.: Immersive 3-D video conferencing: challenges, concepts, and implementations. In: Proceedings of the International Society for Optical Engineering, pp. 69–79 (2003)
18. El-Shimy, D., Hermann, T., Cooperstock, J.R.: A reactive environment for dynamic volume control. In: Proceedings of the International Conference on New Interfaces for Musical Expression (2012)
19. El-Shimy, D.: Exploring user-driven techniques for the design of new musical interfaces through the responsive environment for distributed performance. Ph.D. thesis, McGill University (2014). http://www.cim.mcgill.ca/~dalia/papers/el-shimy_thesis.pdf
20. El-Shimy, D., Cooperstock, J.R.: Reactive environment for network music performance. In: Proceedings of the International Conference on New Interfaces for Musical Expression (2013)
21. Fencott, R., Bryan-Kinns, N.: Computer musicking: HCI, CSCW and collaborative digital musical interaction. In: Holland, S., Wilkie, K., Mulholland, P., Seago, A. (eds.) Music and Human-Computer Interaction. Springer Series on Cultural Computing, pp. 189–205. Springer, London (2013)
22. Grudin, J.: Why CSCW applications fail: problems in the design and evaluation of organizational interfaces. In: Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work, pp. 85–93 (1988)
23. Ishii, H.: TeamWorkStation: towards a seamless shared workspace. In: Proceedings of the 1990 ACM Conference on Computer-supported Cooperative Work, pp. 13–26 (1990)
24. Ishii, H., Kobayashi, M., Grudin, J.: Integration of interpersonal space and shared workspace: clearboard design and experiments. *ACM Trans. Inf. Syst.* **11**(4), 349–375 (1993)

25. Johansen, R.: *GroupWare: Computer Support for Business Teams*. The Free Press, New York (1988)
26. Kapur, A., Wang, G., Davidson, P., Cook, P.R.: Interactive network performance: a dream worth dreaming? *Organised Sound* **10**, 209–219 (2005)
27. Lee, M.K., Takayama, L.: “Now, I Have a Body”: uses and social norms for mobile remote presence in the workplace. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 33–42 (2011)
28. Levisohn, A., Schiphorst, T.: Embodied engagement: supporting movement awareness in ubiquitous computing systems. *Ubiquit. Learn.: Int. J.* **3**(4), 97–112 (2011)
29. Loke, L., Robertson, T.: Design representations of moving bodies for interactive, motion-sensing spaces. *Int. J. Hum.-Comput. Stud.* **67**(4), 394–410 (2009)
30. Ravasio, P., Guttormsen-Schar, S., Tscherte, V.: The qualitative experiment in HCI: definition, occurrences, value and use. *Trans. Comput.-Hum. Interact.* 1–24 (2004)
31. Rebelo, P., Renaud, A.B.: The frequencyliator: distributing structures for networked laptop improvisation. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 53–56 (2006)
32. Renaud, A.: Dynamic cues for network music interaction. In: *7th Sound Music and Computing Conference* (2010)
33. Rodden, T., Blair, G.: CSCW and distributed systems: the problem of control. In: Bannon, L., Robinson, M., Schmidt, K. (eds.) *Proceedings of the Second European Conference on Computer-Supported Cooperative Work*, pp. 49–64. Springer, Netherlands (1991)
34. Schroeder, F., Renaud, A.B., Rebelo, P., Gualda, F.: Addressing the network: performative strategies for playing apart. In: *International Computer Music Conference* (2007)
35. Schroeder, F., Pedro, R.: Sounding the network: the body as disturbant. *Leonardo Electronic Almanac* **16**(4–5), 1–10 (2009)
36. Schuett, N.: *Interconnected musical networks: bringing expression and thoughtfulness to collaborative music making*. Ph.D. thesis, Stanford Center for Computer Research in Music and Acoustics (2002). https://ccrma.stanford.edu/groups/soundwire/publications/papers/schuett_honorThesis2002.pdf
37. Sellen, A., Buxton, B., Arnott, J.: Using spatial cues to improve videoconferencing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 651–652 (1992)
38. Sirkin, D., Ju, W.: Consistency in physical and on-screen action improves perceptions of telepresence robots. In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2012*, pp. 57–64 (2012)
39. Tanaka, A., Bongers, B.: Global string: a musical instrument for hybrid space. In: *Proceedings of Cast 2001: Living in Mixed Realities* (2001)