Effects of Cross-Instrumental Control on Musical Engagement in a Network Music Environment

Daniel Ethridge
Center for Music Technology
Georgia Institute of Technology
Atlanta, Ga, USA
dethridge7@gatech.edu

Jason Freeman
Center for Music Technology
Georgia Institute of Technology
Atlanta, Ga, USA
jason.freeman@gatech.edu

ABSTRACT

Shape Song is an interactive network music application developed to allow three geographically separated musicians to create music in real time. We wanted to investigate how cross-instrumental control affected musical engagement in a network music setting. A small study was conducted with five participants. Preliminary findings indicate a balance between individual control of an instrument and total shared control of instrument parameters may be an important asset in increasing overall enjoyment, collaboration, and connection to the music and to the other musicians.

1 INTRODUCTION

With the onset of the global pandemic in 2020, the music world was hit hard.[1], [2] The practicality of creating music with others was severed as everyone was isolated. This inspired the development of a new platform, called *Shape Song,* for creating socially distant music in real time with a focus on how to increase musician engagement in the system. It should be noted that this application is not meant as a performance piece. It exists to help connect people who want to make music for themselves.

We used a musical jam session (ie an informal gathering of … musicians playing for their own pleasure[3]) as a model for the core of this application. Three people can participate at a time. A user can create and change a backing track by creating and altering shapes. Each user is represented by a stick figure which in turn represents a synthesizer. The stick figure colors are mapped to synthesizer parameters. (See section 3 for more detail.)

As stated, a main focus of this application was to investigate a potential avenue for increasing musical engagement in this network music environment: cross-instrumental control. Inspired by **[5]**, cross instrumental control is defined by the researchers as the ability to control timbral, temporal, melodic, etc. properties of another participant's instrument or device. This leads to the integral research question of this work: how does cross-instrumental control affect engagement in a collaborative network music environment?

2 RELATED WORK

We adapt the definition of network music from **[8]** and define it as "a setting where a group of musicians, located at different places, interact over a network to play as they would if located in the same room." A field commonly plagued by latency, there have been numerous ideas for how to deal with the issue as described in. The different methodologies focus on various aspects of the experience including perceived synchronization between participants and general feel which is can be defined as "how realistic the session feels in comparison to a conventional music session in the same room" **[3]**.

One such system that utilizes the latency-accepting approach detailed in **[3]** is Hajdu's *Quintet.net*. The system is designed with a server-client architecture and allows multiple users to remotely play specially composed pieces to a remote audience. Rather than minimizing latency, Hajdu's team circumnavigates the problem by giving each musical event a timestamp before sending the event to other clients via the server. Upon arriving at the clients, an event is played after a predetermined delay with respect to the timestamp on the event **[6]**.

In *Monad,* "Users interact with virtual objects in a 3D graphical environment to control sound synthesis." A key aspect of their design is allowing each player to control sound characteristics including but not limited to filtering and temporal properties of every sound. Thus, users can change or delete what others have included. In addition, each user has a self-assigned nick name which "promotes agency and a feeling of presence in the system". These design ideas offer heavy inspiration for making *Shape Song* collaborative**[1]**.

Both of these systems, offer creative variations on the traditional structure of each performer having total control over their own instrument or voice. *Quintet.net* offers audience members the opportunity to submit a survey indicating what changes they would like to hear in the

[1]https://www.bbc.com/news/entertainment-arts-54966060

[2]https://www.ifpi.org/ifpi-priorities/music-industry-response-to-covid-19/

[3]https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000014117

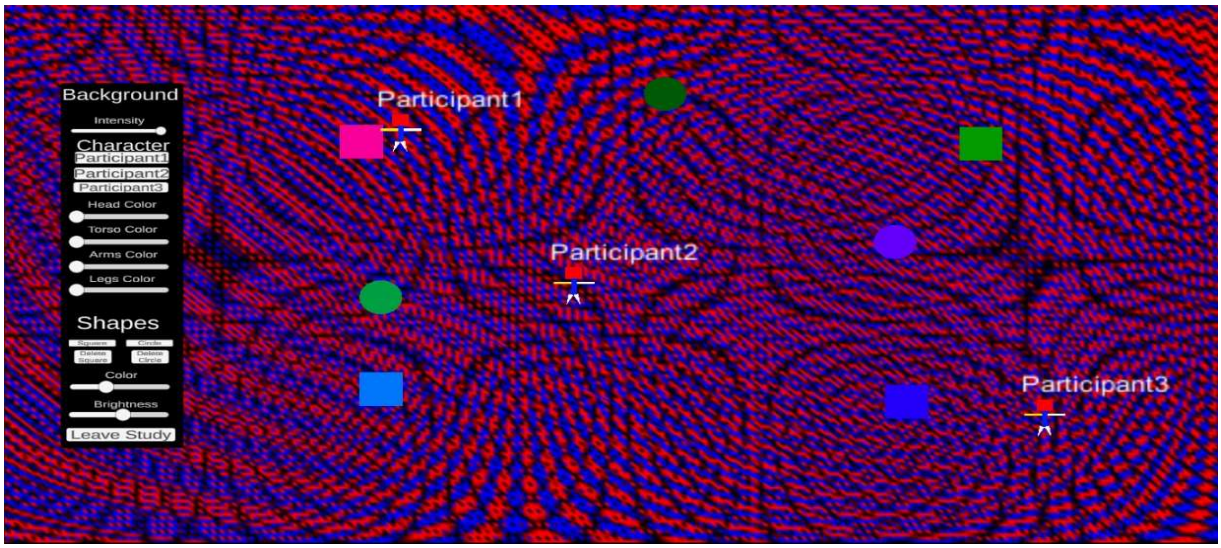

Figure 1: Main Interface View

performance **[6]**. *Monad's* design characteristic of allowing performers to control parameters of every sound present presents an interesting idea**[1]**. This idea is further explored in **[5]** via a Pure Data library entitled bf-pd. Through this novel tool, multiple users on a local network control digital musical synthesizers. Interestingly, the users are able to request parameter changes of others' instruments which introduces a unique extra layer to the musical collaboration.

3 DESIGN AND DEVELOPMENT

The software used to realize *Shape Song* includes Audiokinetic's Wwise[4], the Unity game engine[5], and the Photon Engine[6].

3.1 Interface Design

The user interface was designed to be as simple as possible for a user to understand. Each user has access to the same controls: a background intensity slider (which controls both the brightness of the background seen in Figure 1 and the cutoff frequency of a low pass filter on all audio), a way to select participants, sliders to control the color of the selected stick figure body parts, and ways to instantiate, delete, and modify the color and brightness of shapes.

3.2 Sound Design

The overall compositional process in part takes guidance from **[4]**. In particular, we applied two of Cook's Human/Artistic principles: "programmability is a curse" and "instant music, subtlety later". Despite being discussed outside the context of network music, these principles provided a valuable starting point in the initial sound design experimentation and development. As such, *Shape Song* does not offer endless opportunity but rather a set of tools that can be linked together in several ways to change the overall musical output. The music is easy to initiate; one circle equals a bass line. Add a square or two and a general groove is heard. Through this method of creating the background, anyone, regardless of musical ability, can begin creating music with minimal practice.

The music heard by users can be divided into two categories: pre-composed and improvised.

3.2.1 Pre-Composed Music

The pre-composed music consists of two-measure background melodic lines and one-measure percussion patterns. These patterns are activated when

any user creates a shape with the interface (see Figure 1).

The audiovisual design concepts used in this application largely are inspired by Payling et al's *Hue Music* **[9]**, Vickery's discussion of audiovisual cross modal correspondences in **[10]**, and Hope's discussion of graphic scores in **[7]**. Hue music demonstrated a novel process of translating color hues in images to soundscapes. Vickery discussed the possible cross-modal correspondences between the auditory and visual realm. Hope's discussion provided interesting ideas for incorporating movement and animation into *Shape Song.*

3.2.2 Improvised Music

Participants can improvise music their synthesizer which is represented by their stick figure in *Shape Song*. Different

[4] https://www.audiokinetic.com/products/wwise/

[5] https://unity.com/

[6] https://www.photonengine.com/

	Participant 1	Participant 2	Participant 3
Condition 1	Own avatar colors and intensity slider	Own avatar colors and squares	Own avatar colors and circles
Condition 2	Own avatar colors, participant 2's avatar colors, intensity slider, and squares.	Own avatar colors, participant 3's avatar colors, squares, and circles.	Own avatar colors, participant 1's avatar colors, circles, and intensity slider.
Condition 3	All avatar colors, intensity slider, squares, and circles.	All avatar colors, intensity slider, squares, and circles.	All avatar colors, intensity slider, squares, and circles.

Figure 2: Organization of what each participant could control in each of the three conditions.

synthesis parameters (such as LFO rate, noise level, etc) are mapped to and controlled by the color of the various body parts of each stick figure. Any user can select the stick figure of any other user (including themselves) and change the colors which in turn changes the synthesizer parameters. This means that each user has access to the synthesizer parameters of every other user. (It should be noted that the study limits this control to varying degrees in each condition via instructions to the participants.)

Each user may only play their own synthesizer, and this is done by pressing the "J" key on the keyboard. Synthesizer volume is mapped to horizontal position of the stick figure, and pitch is mapped to vertical position. The position of the stick figure can be adjusted using the WASD keys on the keyboard. There are eight possible pitches to play which lie in the range from A2 to D4. All notes are part of the D minor pentatonic scale. a part of the D minor pentatonic scale.

4 STUDY

We conducted a small study to investigate how cross-instrumental control might affect engagement in a network music setting. Five participants were able to be recruited in for the study. The first three participated in the first session, and the second two participated in the second session. (Note that in the second session, the first author took the role of participant one as laid out in figure 2.) Both sessions of the study included five to ten minutes at the beginning for learning the system. The author stepped through each aspect of how the system worked. Participants were able to practice using the system to become comfortable with it.

Both sessions of the study included three conditions as described in figure 2. The first session went through the three conditions in order while the second progressed through them in reverse order. Each condition lasted for three to five minutes. After each of the three conditions in both sessions, the participants participated in a focus group to see how each of the three conditions presented differing levels of engagement as measured through concepts such as enjoyment, frustration, connection, and amount of listening.

5 RESULTS

Of the five participants, three indicated that they enjoyed using the system the most in the second condition, one in the first condition, and one participant remarked that they could not choose as they "enjoyed testing out different aspects." Participants noted that the constraints offered in the second condition helped to focus their attention and as a result the constraints seemed to be something of a guiding force. The participant who preferred the third condition noted a liking for chaotic and unexpected music and later acknowledged that the third condition offered such music. The same participant noted that condition one "sounded flat" and "felt less structured", noting there was "no overall density to control." In the recording of the first session, this participant could be seen doing the same things over and over in the first condition, hinting at the fact that there was not enough to do.

There was also ample talk of connection both among the participants and between each participant and the musical output. One participant noted that they were connected to the sound the most in the first condition due to being able to more readily decipher how their actions were influencing the sound. The opposite was true for the third condition for the opposite reason. Another participant noted that they felt most connected to other participants in the 2nd, and all the participants remarked that they believed the second condition was the best for musical collaboration. In it, "there was a sense of feedback for any action." Perhaps because of this, there was generally a more favorable view of the musical output as a whole in the second condition. One participant remarked, "[in the second condition,] I had more interest in what was going on musically." Other participants noted better synchronization, which also serves as a potential reason for the better musical output.

6 DISCUSSION

It must be noted that *Shape Song* is not without limitations. Bits of formative feedback received from the participants included a desire to have a verbal communication system. The participants mentioned in the first session that speaking to each other may increase enjoyment of the system. There were other complaints that the synthesizer could not be heard when multiple squares and circles were playing their respective patterns, so more work can be done with balancing audio.

Obviously, the results of this study are preliminary largely due to the small and statistically insignificant number of study participants. That being said, through the conversations with the study participants, a specific theme came to light that potentially offers insight into the research question at hand.

That theme is the balance between chaos and individual control. In the first condition, participants were limited in what they could do and felt restrained. None of the participants enjoyed the musical output from condition one the most and was at one point described as “definitely not as good [as the other two conditions]”.

The third condition was enjoyed to a degree and was described as offering fluid and enjoyable interactions and as being more exciting. It was equally met with remarks such as “More control in three didn’t correspond to better music” and “Everyone being able to control everything bred absolute chaos.”

The second condition was met with the most favorable reviews. One participant said, “once we caught on to things that we could do together, it made some really cool changes to the sounds.” In the recording of the second session, the third participant is seen moving around the screen and playing while the colors are being changed by the third participant. This points to the collaborative nature of the second session. (Note that a user cannot control the color sliders while moving around the screen.)The constraints mentioned in the second condition helped to “create more of musical piece” and that it bred the strongest sense of synchronization and togetherness.

These findings provide an interesting look into how the design of future network music systems could be designed. Since the engaging face-to-face element of a jam session is lost in network music, designers will need other ways to create an engaging space. These preliminary findings suggest that allowing mixed control over parameters in a system could be a promising way to create an engaging space.

7 CONCLUSIONS AND FUTURE WORK

Too much individual control has the ability to create a stale environment whereas too little can breed unintelligible chaos. There seems to be a point in the middle where there exists an environment capable of fostering engagement in the music making process.

It would be beneficial to do more development on the system such as balancing the audio more and adding in other mediums of communication, namely speech. More participants for a study would be incredibly beneficial to gather more data and potentially solidify the preliminary data here. A final next step would be to add support for more varied music such as allowing more user input and power to actually compose the lines they hear or to change the style altogether.

REFERENCES

[1] Cem Çakmak, Anil Çamcı, and Angus Graeme Forbes. 2016. Networked Virtual Environments as Collaborative Music Spaces. Griffith University, Brisbane, Australia.

[2] Alexander Carot and Christian Werner. 2007. Network Music Performance – Problems, Approaches and Perspectives. Budapest, Hungary.

[3] Perry Cook. 2017. 2001: Principles for Designing Computer Music Controllers. In *A NIME Reader*, Alexander Refsum Jensenius and Michael J. Lyons (eds.). Springer International Publishing, Cham, 1–13. DOI:https://doi.org/10.1007/978-3-319-47214-0_1

[4] Luke Dahl, Florent Berthaut, Antoine Nau, and Patricia Plenacoste. 2017. bf-pd: Enabling Mediated Communication and Cooperation in Improvised Digital Orchestras. In *Music Technology with a Swing*, 454–467. DOI:https://doi.org/10.1007/978-3-030-01692-0_30

[5] Georg Hajdu. 2005. Quintet.net: An Environment for Composing and Performing Music on the Internet. *Leonardo* 38, 1 (2005), 23–30.

[6] Cat Hope. 2017. Electronic Scores for Music: The Possibilities of Animated Notation. *Computer Music Journal* 41, 3 (September 2017), 21–35. DOI:https://doi.org/10.1162/comj_a_00427

[7] John Lazzaro and John Wawrzynek. 2001. A case for network musical performance. In *Proceedings of the 11th international workshop on Network and operating systems support for digital audio and video - NOSSDAV '01*, ACM Press, Port Jefferson, New York, United States, 157–166. DOI:https://doi.org/10.1145/378344.378367

[8] Dave Payling, Stella Mills, and Tim Howle. 2007. HUE MUSIC – CREATING TIMBRAL SOUNDSCAPES FROM COLOURED PICTURES. Montréal, Canada.

[9] Lindsay Vickery. 2018. SOME APPROACHES TO REPRESENTING SOUND WITH COLOUR AND SHAPE. (2018).