4QUARTERS: REAL-TIME COLLABORATIVE MUSIC ENVIRONMENT FOR MOBILE PHONES

Nathan Bowen

Moorpark College
Music Department
7075 Campus Road
Moorpark, CA 93021

ABSTRACT

This paper presents '4Quarters,' a new musical multiplayer compositional and mixing tool designed for the iPhone and computer as a controller-plus-software setup. 4Quarters is designed for the general public to serve as an alternative to a consumption-based approach to music (simple playback). It also aims to depart from interactive score-matching games such as Rock Band by offering an open-ended architecture for sound selection. Sounds are based primarily on prerecorded audio files, but playback is non-linear and determined by keypresses in real-time. The ultimate purpose of this software is to provide a structure for making music that is fundamentally decentralized and collaborative in nature, where prepared content, live manipulation, and the final recording of the music is brought to pass by multiple participants. In this paper I describe the setup in terms of network configuration, mapping strategies, visual feedback, and content/retrieval architecture. Observations of the tool in beta performance are also discussed.

1. INTRODUCTION

Within the last decade there has been an explosion of applications and research for mobile phones in musical contexts, frequently with diverse intents, outcomes, and uses of mobile technology. Reports on the community by Levin [10], Behrendt [2], Tanaka [15], Gaye et al [7], Oh, et al [13], and Essl & Rohs [4] underscore the rapid rate of transformation of work, which is in part due to the increased capabilities of mobile devices [9][16]. Here I will focus primarily on mobile music scenarios that depend on multiple mobile devices. Golan Levin's *Dialtones: A Telesymphony* (2001) was one of the first pieces that took advantage of the notion that concert halls are typically filled with people carrying mobile phones, and that these devices can be used for musical purposes [11]. David N. Baker's *Concertino for Cellular Phones and Symphony Orchestra* (2006) likewise acknowledged the ubiquity of mobile phones by encouraging audience members to play their ringtones at specific sections in the piece [18]. Jason Freeman's *Glimmer* (2004) does not make use of mobile phones, but is likewise an ambitious audience participatory piece using glow wands and motion capture [6]. Joshua Knowles and Joo Youn Paek's *Phoneplay* (2007) allowed participants to dial into a central server with keypresses helping to shape a pan-diatonic soundscape [9].

Yet in each of these scenarios the role of the lay participant is rather limited in influence. 4Quarters was likewise conceived to cater to a broad public, but to give each participant a greater influence on the musical outcome, or at least ensure that each participant can receive some feedback that he/she is influencing the music. Successful commercial games Rock Band and Guitar Hero address this issue by providing effective visual and audio feedback, but when all is said and done, the role of the participant is closer to performer rather than composer [12]. 4Quarters likewise provides pre-recorded content for a participant to 'play', but the content is presented as an array of loops to mix and match in a non-linear format.

2. RELATED WORK

Group music making has been a central focus for many mobile music scenarios [8] and network music architectures [1][22], some of which are geared towards ad-hoc jam sessions [5][17][23] while others [3][14][20][21] are situated in formal ensemble settings such as Stanford MoPhO, The Michigan Mobile Phone Ensemble, the Helsinki MoPhO, and the Yamaha Mofiano Mobile Orchestra [13]. Mobile phone instruments such as RjDj and Smule's Ocarina extend the possibilities for social interaction by leveraging locative data and online communities [19]. Within this context, 4Quarters is designed primarily for ad-hoc performance scenarios, though it can be adapted for concert pieces and in the future may include an online repository for generated works.

3. DESIGN CONSIDERATIONS

3.1 Evolution of design

In its original conception 4Quarters was intended to accommodate a wide variety of devices situated in the realm of audience participation. Similar to Joshua Knowles and Joo Youn Paek's *Phoneplay* (2007), a visual projection would show all participants' activities. The initial visual design was intended to correspond to the familiar 12-key interface found on mobile phones. Various types of assignments were conceived to work with the 12-key tactile interface. Preliminary designs anticipated multiple phones communicating keypresses to one central server, with all sound processing occurring on that server. Participants would link in by dialing a phone number, and DTMF keypresses would be captured via Asterisk and route the data to Max/MSP.

The arrival of the iPhone introduced new possibilities for data input and interface considerations, such as customizable interfaces with sliders, additional buttons, and colors. Latency with wifi was also much less of an issue, and subsequently plans to use DTMF and Asterisk as the network protocols were abandoned in favor of OpenSoundControl and a wireless LAN. The switch was also made to prepare for the likely transition where most phones worldwide will be smartphones.

3.2 Network configuration

The main software environment is designed in Max/MSP with iPhone/iTouch serving as controllers, sending user input via OpenSoundControl over a wireless network. TouchOSC is the current client app, which sends and receives data to and from Max (see Figure 1). Outgoing messages from the iPhone/iTouch control the following sonic parameters: sound file selection, looping, volume, panning, and EQ. These messages each have a corresponding visual indication on a projected screen that can be seen by all participants. Audio processing is handled within Max, with the stereo image output ideally routed to external speakers.

Data received on the client side provides visual feedback reinforcing the interface correspondence between the main video projection of the Max patch and the screens viewed on the phone.

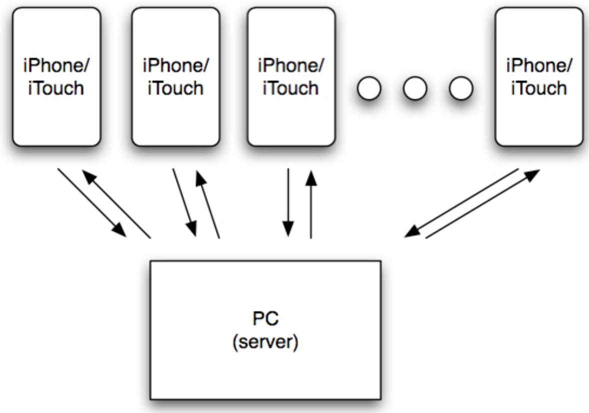

Figure 1. Up to twelve iPhones send and receive OSC messages to/from a server via a wifi local area network.

3.3 Server Interface

The 12-key concept as an interface orientation has remained as a carryover from the initial design. Keypresses from a 12-key mobile phone map onto a 4x3 grid interface (numbered 1-12), with the master patch showing four iterations of this grid. Each of the four grids represent sound layers, with each containing twelve sound files available for playback, making for a presentation of 48 samples at any one time. These are visually delineated into uniquely colored quadrants in the main Max patch (see Figure 2). Those colors become helpful indicators in situations where multiple participants work to control one layer of sound. For instance, 'team red' may have one person in charge of sound file selection, another person controlling volume, and a third person shaping the EQ. Ideally that patch is projected on a large screen so that all participants can see clearly what is going on in the other layers/quadrants.

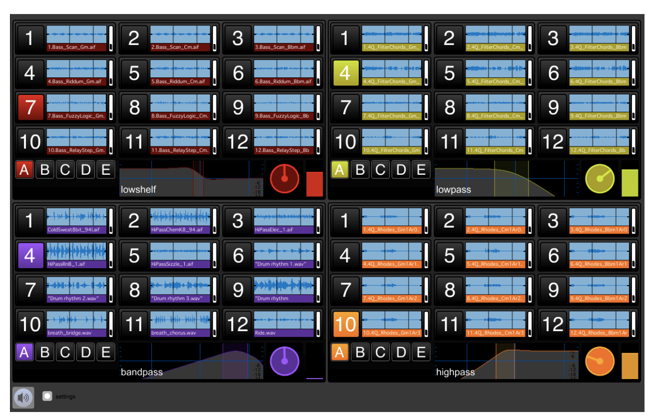

Figure 2. Master interface that all participants see.

In the master patch, each number represents a sound file, and the file's name and waveform diagram are visually placed adjacent to that number. Additionally, a small fader is placed next to the waveform diagram (see Figure 3).

When in play, each waveform diagram has a scrollbar, which becomes a significant feature in coordinating synchronization between sounds. If sound files are groove-oriented and follow a particular BPM setting, synchronizing the timing of playback must be handled manually. By design there is no tempo map or beat grid that synchronizes playback, with the intent to engender musicianship and social cooperation to get the timing right.

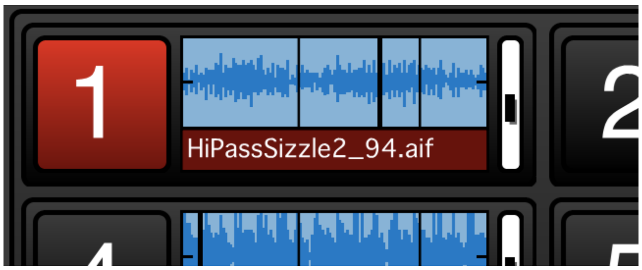

Figure 3. Close up of one sound file 'frame,' with waveform diagram, scrollbar, file name, and fader.

3.3 Audio Content

In terms of physical design, a stereo image comes from the server laptop or PC, ideally connected to speakers. This of course means that the phones do not physically produce sounds, but serve only as controllers. And this of course inherently creates limitations for the physical spaces and places where 4Quarters may be used. So the *mobile* aspect of mobile phones is not leveraged to its greatest potential here. Concert halls, living rooms, or classrooms are the most likely spaces for implementation, as they may have built-in audio systems, decent speakers, and projectors. Though the ability to have a phone act as the server is possible and certainly attractive [9], tethering the phones to a computer is already a necessity for the visual component.

Audio content is comprised of a series of pre-recorded audio files that are then available for real-time playback, looping, and rapid file selection. The framework is open in that a user can drag and drop audio files into sound banks. Twelve sound files may be ready-at-hand for each bank, and up to five banks of twelve files are available to be called up and swapped in and out (labeled A-B-C-D-E). Thus, 240 total sound files can be accommodated.

Because users frequently want to make things happen rather than sit back and listen, the best sound files tend to be short in duration, loopable, and composed to fit 'hand to glove' with other files. One possible format for composition is to have each bank from all four layers correspond in some fashion, with global formal structure characterized by parametric change from one bank to the next. This might be a way to suggest or imply formal guidelines for groups to use in improvising when selecting and synchronizing layers of sound.

3.4 Performance Assignments

There are three distinct roles within one layer/color that players can take: 1) file selection (including swapping sound banks), 2) volume/panning, and 3) EQ. Three corresponding pages of a custom TouchOSC layout form the phone interface (see Figure 4).

The first role, file selection, allows a user to choose from a maximum of twelve files at any one time. File selection and playback is controlled entirely by keypress. As a default pressing buttons triggers only one sound file at a time, though a poly feature allows multiple files to playback at once. A stop button and loop button are additional controls, as well as buttons controlling sound bank selection.

Volume and panning are dedicated to the second page of the TouchOSC layout. Here one plays the role of mixer, adjusting the volume levels of the various sound files, and controlling global volume. One can opt to affect both panning and the master volume on the touch screen via slider, or one can toggle a button to use accelerometer data. Panning (X axis) and global volume (Y axis) data is routed to Max, which in turn send messages back to the phone to activate the sliders visually on the phone.

Basic EQ is controlled on the third page with an XY slider. The X axis affects the peak frequency and the Y axis is mapped onto the Q. There are six filters (with allpass as a default) for modifying timbre. Like volume, EQ can also be controlled via accelerometer data.

Figure 4. Three iPhone/iTouch layouts control sound file selection, volume/panning, and EQ.

If desired, one player can take on all three roles. With four colors, this means that a minimum of four players can operate all available control parameters, and a maximum of twelve can play. In the scenario of one player controlling file selection, volume, and EQ, it can become difficult to swipe from one page to the next on the TouchOSC layout. Hence, on the first page (file selection) there are buttons that can switch on the tilt options for panning, volume, and EQ.

4. OBSERVATIONS

So far 4Quarters has been presented in classrooms and as an installation at SEAMUS. In observing how participants use 4Quarters, most of the visual focus is spent negotiating the phone interface and seeing how it corresponds to the shared screen projection (see Figure 5). As buttons are the main feature for making things happen (enabling and disabling tilt features are also controlled by using buttons), users tend to zone in on their own device if they are not watching the screen. But as there has yet to be any seasoned performers with 4Quarters, this may simply be the result of people getting oriented with the instrument rather than indicative of some shortcoming in design.

The current controlling parameters for this environment are fairly basic: keypresses and accelerometer-driven gestures constitute the type of physical input used to drive musical effects. In beta testing, most users tend to hold the phone like a remote control and do very little by way of physical gesture. As this is a work in progress, the eventual goal is to incorporate more sophisticated features and algorithms to help the user experience become truly visceral and physical, with rich expressive possibilities.

Additionally, the project must be streamlined to accommodate a broad level of people with varying abilities and exposures to music, computers, and mobile devices. At present the setup is far too thorny for the intended demographic of users. It requires a certain aptitude with networks, and TouchOSC is far from an ideal app because of the need to upload a layout. But once everything is set up, the actual use of the system is fairly intuitive, and participants have been able to navigate the functionality with little explanation. Next steps include the development of a custom iPhone app as well as a one for Android.

In spite of its limitations, perhaps the most promising aspect has come from observing students who supply their own audio files to load into 4Quarters. Rather than writing a piece for an ensemble to perform, I solicited a few students to prepare audio files for 4Quarters with only a few instructions: agreed-upon BPM and tonal center. In bringing these files into one session, a highly engaging improvisatory jam session ensued. Each participant knew how his own files were meant to interact, but the fun came in the pleasant surprises when mixing content.

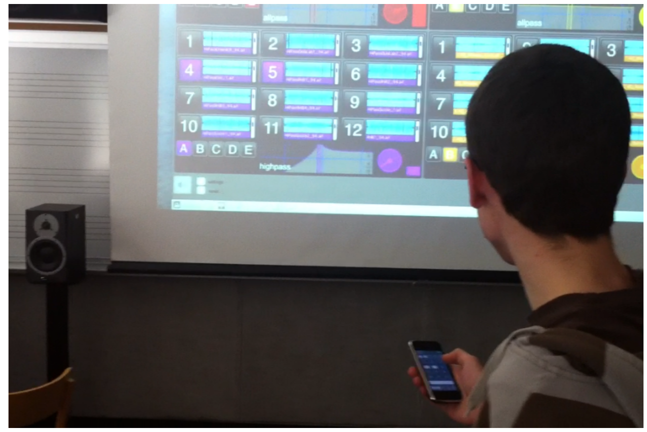

Figure 5. 4Quarters in a classroom setup at SEAMUS.

5. CONCLUSIONS

The use of mobile phones for group music making is fast becoming standard practice. This paper has introduced the concept and design of 4Quarters as a tool for improvisation and group composition. By linking a non-linear and open format for musical content to the mobile platform, and in making the control of that format decentralized, 4Quarters situates itself as a viable option for empowering musicians and non-musicians to make music creatively and collaboratively.

6. REFERENCES

[1] Barbosa, Á. “Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation,” in *Leonardo Music Journal*, *13*, 2003, 53-59.

[2] Behrendt, F. *Handymusik: Klangkunst und ‘Mobile Devices.’* Epos Music Series, Electronic Publishing Osnabrück, 2004.

[3] Dahl, L., & Wang, G. “Sound Bounce : Physical Metaphors in Designing Mobile Music Performance,” in *Proceedings of the International Conference on New Interfaces for Musical Expression,* Sydney, Australia, 2010, pp. 178-181.

[4] Essl, G., & Rohs, M. “Interactivity for Mobile Music-Making,” *Organised Sound*, 14(02), 2009, pp. 197-207.

[5] Fabiani, M., Dubus, G., & Bresin, R. “MoodifierLive : Interactive and collaborative expressive music performance on mobile devices.” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, Norway, 2011, pp. 116-119.

[6] Freeman, J. *Glimmer*. http://distributedmusic.gatech.edu/jason/music/music_

technology/glimmer/ Posted 2004. Accessed March 16, 2012.

[7] Gaye, L., Holmquist, L. E., Behrendt, F., & Tanaka, A. "Mobile Music Technology: Report on an Emerging Community," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, France, 2006, pp. 22-25.

[8] Kim-Boyle, D. "Network Musics – Play, Engagement and the Democratization of Performance," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Genova, Italy, pp. 3-8.

[9] Knowles, J., & Paek, J. *3001: Collaborative Musical Gameplay*. http://gophoneplay.com/nime/ Posted 2007. Accessed March 16, 2012.

[10] Levin, G. *An Informal Catalogue of Mobile Phone Performances, Installations and Artworks*. http://www.flong.com/texts/lists/mobile_phone/ Posted September 2001. Accessed March 14, 2012.

[11] Levin, G. *Dialtones (A Telesymphony)*. http://www.flong.com/projects/telesymphony/ Posted 2001. Accessed March 16, 2012.

[12] Miller, K. "Schizophonic Performance: Guitar Hero, Rock Band, and Virtual Virtuosity," in *Journal of the Society for American Music*, 3(04), 2009, pp. 395-429.

[13] Oh, J., Herrera, J., Bryan, N. J., Dahl, L., & Wang, G. "Evolving The Mobile Phone Orchestra," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Sydney, Australia, 2010, pp. 82-87.

[14] Schiemer, G. & Havryliv, M. "Pocket Gamelan: Tuneable trajectories for ying sources in Mandala 3 and Mandala 4," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, France, 2006, pp. 37-42.

[15] Tanaka, A. "Mobile Music Making," in *Proceedings of the International Conference on New Interfaces for Musical Expression,* Hamamatsu, Japan, 2004, pp. 154-156.

[16] Tanaka, A. "Mapping Out Instruments, Affordances, and Mobiles," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Sydney, Australia, 2010, pp. 15-18.

[17] Tanaka, A., & Gemeinboeck, P. "A Framework for Spatial Interaction in Locative Media," in *Proceedings of the International Conference on New Interfaces for Musical Expression,* Paris, France, 2006, pp. 26-30.

[18] Wakin, D. "Horns Up. Bows Ready. Cellphones On." in *New York Times,* http://www.nytimes.com/2006/10/03/arts/music/03cell.htm Accessed March 16, 2012

[19] Wang, G. "Designing Smule's iPhone Ocarina," in *Proceedings of the International Conference on New Interfaces for Musical Expression,* Pittsburgh, USA, 2009, pp. 303-307.

[20] Wang, G., Essl, G., & Penttinen, H. "Do Mobile Phones Dream of Electric Orchestras?" in *Proceedings of the International Computer Music Conference,* Belfast, Ireland, 2008.

[21] Wang, G., Essl, G., & Penttinen, H. "MoPho – A Suite for a Mobile Phone Orchestra," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Genova, Italy, Paper 209. Retrieved from http://nime2008.casapaganini.org/documents/Proceedings/Performances/209.pdf

[22] Weinberg, G. "Interconnected Musical Networks: Toward a Theoretical Framework," *Computer Music Journal*, *29*(2), 2005, pp. 23-39.

[23] Weinberg, G., Beck, A., Godfrey, M., & Walk, T. "ZooZBeat : a Gesture-based Mobile Music Studio," in *Proceedings of the International Conference on New Interfaces for Musical Expression*,Pittsburgh, USA, 2009, pp. 312-315.