AUDIENCE PARTICIPATION USING MOBILE PHONES AS MUSICAL INSTRUMENTS

A Thesis
Presented to
The Academic Faculty

by

Sang Won Lee

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Music, College of Architecture

Georgia Institute of Technology
August 2012

AUDIENCE PARTICIPATION USING MOBILE PHONES AS MUSICAL INSTRUMENTS

Approved by:

Dr. Jason Freeman, Advisor
School of Music
Georgia Institute of Technology

Dr. Gil Weinberg
School of Music
Georgia Institute of Technology

Dr. Ellen Yi-Luen Do
School of Industrial Design & School of Interactive Computing
Georgia Institute of Technology

Date Approved: May 16th 2012

To my fiancé, Sol Ie

ACKNOWLEDGEMENTS

First, I like to express my sincere gratitude to Dr. Jason Freeman, my advisor, who not only supported me throughout my study at Georgia Tech, but also shared his expertise in music and research unconditionally for this work. I would not have started this research without inspirations he gave me in and out of the classroom and I would not have finished the thesis without his patient guidance and encouragements.

I would like to thank Dr. Gil Weinberg. His insightful questions, which I cannot usually answer promptly, take me one step back and make me think bigger picture. I also deeply appreciate Dr. Ellen Yi-Luen Do's kindness of giving me a series of opportunities to present this raw idea in her class, which completed this research.

I have been blessed with a number of able colleagues who helped me throughout this study. It was very pleasant to collaborate with Yonatan Sasson, who helped me turn a mobile app into a music performance with his talent in clarinet. Ajay Srinivasamurthy read this lengthy thesis draft and fixed loads of grammatical errors for me. Nathan Weitzner has been a valuable lab-mate, giving me ideas on a ride home everyday and helping me develop the simulator with his super power in Max/MSP. Aaron Albin always shared his knowledge in research with encouragements. I also thank all members of MSMT students for their inspiring works and friendly advices.

Finally, I truly thank my family for their unconditional love and endless support, who always believe me regardless of what I am.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

SUMMARY

This research aims at a music piece for audience participation using mobile phones as musical instruments in a music concert setting. Inspired by the ubiquity of smart phones, I attempted to accomplish audience engagement in a music performance by crafting an accessible musical instrument with which audience can be a part of the performance. The research begins by reviewing the related works in two areas, mobile music and audience participation at music performances, builds a charted map of the areas and its intersection to seek an innovation, and defines requisites for a successful audience participation where audience can participate in music making as musicians with their mobile phones. To make accessible audience participation, the concept of a networked multi-user instrument is applied for the system. With the lessons learnt, I developed *echobo*, a mobile musical instrument application for iOS devices (iPhone, iPad and iPod Touch). With this system, audience can download the app at the concert, play the instrument instantly, interact with other audience members, and contribute to the music by sound generated from their mobile phones. A music piece for *echobo* and a clarinet was presented in a series of performances and the application was found to work reliably and to accomplish audience engagement. The post-survey results indicate that the system was accessible, and helped the audience to connect to the music and other musicians.

CHAPTER 1

INTRODUCTION

In a music concert, often the audience has a chance to participate in the performance rather than just to listen to the music; for example, you can find people clapping to the beats, singing along with the music, and waving their arms (or sometimes lighters), probably in a rock concert. These participatory gestures are effective for musicians to engage audience with the music. Hence the musicians sometimes plan those gestures in advance and initiate their participation during the performance, for example, initiating clapping up in the air or holding a microphone out to audience for the chorus. While enabling mass audience participation will limit the musical variability that musicians can have, they sometimes sacrifice their artistic expressions as an alternative of audience engagement by offering a unique experience of music making rather than passive music listening. Eventually, when the audience participation takes place in large scale and becomes a part of the performance, audience feel connected to the music. My research goal starts with this idea to help the audience and the musicians be engaged in the performance through participation.

In previous work, several music pieces have been written to orchestrate audience participation in a music performance, with or without technologies, to achieve similar goals[1-5]. Two common themes found in previous works are accessibility in participation and controlled (or guided) sonic output in music. Easy ways of participation is necessary since each individual audience member has a different musical background and they have no or little chance to practice or rehearse before their participation. Also, it is essential for a musician to turn those large-scale diverse gestures into a music piece. To accomplish these goals, audience participation is often limited to simple musical gestures, such as clapping, or singing the chorus that people already know, as compared to the performance that a professional musician would do. Furthermore, it seems very difficult to make large-scale audience participation both easy and interesting to keep the audience motivated throughout the whole piece.

With these challenges, my research aims at a music piece for audience participation using mobile phones as musical instruments in a music performance. More specifically, using the system, the audience collaborates with the instrumental players and improvises with the mobile music instrument in a traditional audience-stage venue. The mobile application is specially designed to control their improvisation under a networked constraint so that sound from audience will contribute to a piece of music. The research includes developing an accessible mobile music application that allows audience to actually generate sound instead of influencing music indirectly. As most people today carries a mobile phone, I chose to build a musical instrument for audience participation on mobile platform, utilizing the ubiquity and the technology of today's mobile phones.

My hypotheses for this research are as follows; i) The use of a multi-user interactive mobile music application at a live musical performance will offer an accessible audience participation environment and help audiences feel more connected to the music performed. ii) Incorporating elements of social interaction among audience members in the mobile instrument will enable an individual member to have a greater influence on the music and give a sense of collaboration among the audience.

The research begins with the prior knowledge from three different areas - mobile music, audience participation and network music. I wish to follow the spirit of mobile music, where researchers craft a full musical instrument on the mobile platform, utilizing mobility and ubiquity of the devices. I also try to classify successful audience participation works in a musical context according to their characteristics and extract key success factors from those works. Lastly, I apply the concept of collaborative multi-user instrument to fulfill the qualities of successful audience participation so that a wide range of audience members can participate in the performance and enjoy the music they produce.

Inspired by the lessons from previous works, I developed *echobo*, a mobile musical instrument application for iOS devices (iPhone, iPad and iPod Touch). Audience with those devices can download the app and play at the concert to make music in a collaborative manner. Participants use their own mobile phones, which they bring to the performance, as primary musical instruments. I presented a six-minute music piece for *echobo* and a clarinet, twice in a classroom setting for evaluation and once in a music concert setting. *Echobo* is a blend of two words, echo, which represents the state of being empathized with others, and chobo (초보), which means the first step or a novice in Korean.

While there are a number of performances and research that realize audience participation at a music performance in computer music context, there are very few examples where audience were able to participate as musicians and play a musical instrument in a large scale with the aid of technology. To the best of my knowledge, this is the first attempt to develop a large-scale audience participation environment that allows audience to actively make a sound with their mobile phones.

CHAPTER 2

RELATED WORK

As this research is conducted in the intersection of mobile music and audience participation in musical context, I hereby summarize brief history of both areas and explore their intersection to find some room for innovation. To provide an accessible environment for non-musicians to participate, the application follows the paradigm of collaborative network music instruments.

History of Mobile Music

As the research uses mobile devices as the primary musical interface, it is essential to go through the development of computer music with mobile devices to explore the characteristics of mobile music. In the last decade, researchers and artists have explored the potential of mobile devices in computer music. The definite scope of mobile devices is vague due to many kinds of new emerging products but, in general, mobile devices are expected to be portable and wireless. Mobile devices include but are not limited to cellular phones, PDA, smart phone, and other miscellaneous devices such as a tablet (i.e. iPad), a portable media player (i.e. iPod Touch) and e-reader (i.e. Kindle).

Many researchers indicated that *Dialtones (A Telesymphony)* by Golan Levin [6] is the first work that incorporated mobile phones in music performance (Figure 1). It was premiered in the concert at The 2001 Ars Electronica Festival, Austria. The musicians on stage performed a music piece by ringing mobile phones that audience brought to the concert. The devices would play pre-composed ringtones that people downloaded prior to the concert. The piece showed opportunity of mobile phones as an element in computer music, utilized the ubiquity of mobile phones and inspired many researchers in mobile music.

Figure 1 Dialtones – A Telesymphony

While *Dialtones* is the music composed of only ringtones, researchers tried to develop a musical potential of mobile phones by implementing sound synthesis on the platform. The first music making system implemented in mobile phone was *PDa* by Geiger[7]. The work covered issues that mobile devices face to generate synthesized sounds. Geiger tried to port Pure Data (PD), which is a computer music patch language, into PDA. The author noted that the major obstacle for real-time sound synthesis in mobile device was the performance of processors due to its slow emulation of the floating-point. By avoiding floating point calculation, the author could speed up performance of Pure Data in the device. However, the author also suggested that programming PD on the device was cumbersome and another way of controlling PDa is needed due to the limited size of user interface and screen on the device. There was a similar effort to export an existing audio library, STK Toolkit, for the mobile platform Symbian OS by Essl and Rohs[8].

Pocket Gamelan was a noteworthy work of early mobile music, done by Schiemer and Havryliv [9]. In the project, mobile phones, which were implemented to make audible sounds through their speakers, were mounted in a specially devised pouch attached to a cord. Some players swung phones in the pouch while others operated phones to change tones of swinging phones (Figure 2). Bluetooth network enabled interaction between flying and hand-held phones. Even though it was developed at early stage of mobile music, the work utilized diverse aspects of mobile music that current researchers often explore; it included sound synthesis on a mobile platform, accessibility and most importantly, it utilized the mobility (light weight and network capability) of devices by introducing a notion of flying sound source.

Figure 2 Pocket Gamelan

Several researchers focused on utilizing the variety of built-in sensors embedded in mobile phones and developed sensor-based interfaces for music making on mobile platform. Tanaka developed a sensor extension for PDA to capture not only grip pressure but also gestures and motion in 3D space for a participatory music streaming system [10]. Geiger also combined his sound synthesis effort on mobile phone with the development of a touch screen controller on a mobile device[11]. Essl and Rohs utilized new sensors like camera, accelerometer and magnetometer embedded in mobile phones, to use it as a music controller in a music performance [12, 13]. On the other hand, Shepard leverages mobile phones and GPS to allow individuals to record urban sounds with location tag and to facilitate soundscape composition[14]. Recently, a similar soundscape application, *UrbanRemix* by Freeman et al., was developed with a web-based social composition interface [15].

Another big leap in mobile technology was made by the appearance of the iPhone, which facilitated mobile music application development. It brought a number of applications like iBand, PocketGuitar or Scratch and led to the commercial expansion of mobile music[16]. There are two key success factors in iPhone becoming an attractive media in computer music. First, the iPhone offers computational speed, storage and I/O capability comparable to that of a computer. It is equipped with a powerful CPU, graphics, high quality audio pipeline, and several sensing technologies including accelerometer and multi-touch screen. Second, Apple opened App Store and provided software development kit (SDK), which enabled third party programmers to develop their own applications with full support of mobile technology. Now the choice of platform and devices is broadened by other mobile operating systems, such as Android and Windows 8. Yet, mobile music applications are mostly built on iOS with the advantage of its superior audio API.

Ocarina by Wang exemplifies the capability of iPhone in development of mobile music application[17]. It is a virtual realization of a wind instrument of the same name. The design of *Ocarina* integrates technologies capable on iPhone such as multi-touch, microphone, accelerometer, real-time synthesis and interactive graphics. *Ocarina* is one of the earliest music instrument mobile apps that mimics an actual acoustic instrument by realization of musical expressivity mapped to gestures, such as pitch, scale, dynamics, timbre effects like vibrato(See Figure 3). Another interesting feature of *Ocarina* is the World Listener view, where one can see the locations of other *Ocarina* players and hear them playing. Wang suggests that mobile phone could be a point of social interaction in music making.

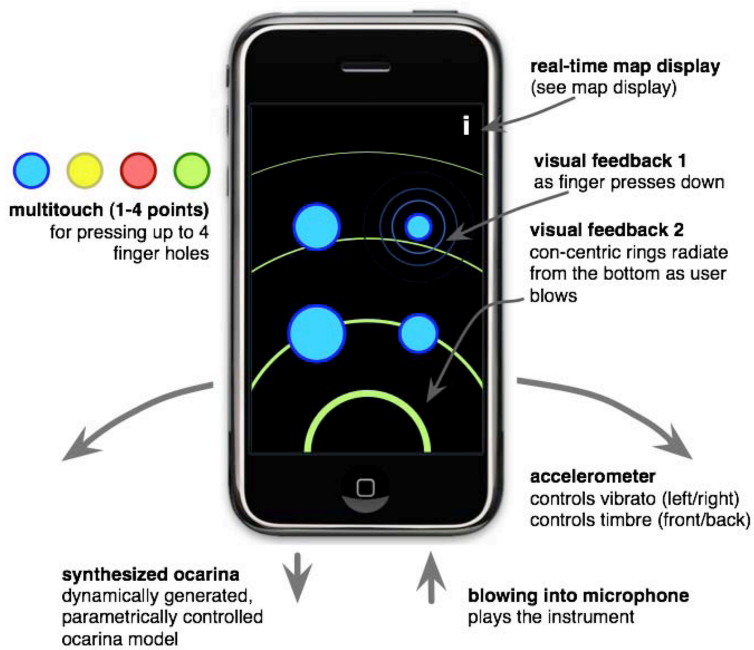

Figure 3 Design schematic for the Ocarina interface

While *Ocarina* focuses on simulating an existing acoustic instrument, there are apps, which furnish a mobile compositional environment. *ZoozBeat* by Weinberg et al. provides a simplified music studio environment on a mobile phone [18]. The application offers a full-length, multi track composition environment based on loop and music sequencer type graphical interface. It focuses on the mobility in user interaction so that *ZoozBeat* users can shake the phone or tap the screen to enter notes. Another app, *beatmaker* crafts a mobile music workstation close to a traditional digital audio workstation[19].

Another novel approach in mobile music applications is to utilize machine listening techniques to create intelligent music-making environment for casual smart phone users. There are a number of music apps that utilize machine listening techniques for music recognition , i.e. Shazam[20] and SoundHound[21]. A recent application from Khush Inc, *LaDiDa*, enables users to sing into an iPhone, refines the recorded voice with auto-tuning and generates back-up tracks in multiple music styles[22]. A sister application, *Songify*, turns speech into music by more intensive usage of auto-tuning[23]. Both applications utilize social networks and let users share their music creation online in and out of the application.

One of the recent and noteworthy movements in mobile music is a mobile phone ensemble. With the concept of turning mobile phones into musical instruments, parallel to the notion of laptop orchestras, Wang, Essl and Penttinen founded Mobile Phone orchestra (MoPho) at CCRMA, Stanford University Fall 2007 and had their debut concert in 2008 [24]. The project involved tasks ranging from crafting a musical instrument inside a mobile phone to writing dedicated music for the mobile ensemble. In the beginning of their first piece called *Drone In*, with a human conductor standing alone on the stage, the members of the ensemble began to play their instruments in hand, sitting disguised among the audience, providing spatial sounds to demonstrate the benefit of mobility. After the successful launch of MoPho, a number of mobile phone orchestras were founded in various academic institutions[25-27].

Figure 4 Mobile Phone Orchestra (MoPho)

Due to continued research in Mobile Phone Orchestra, the researchers naturally needed a development environment for mobile phone to expedite mobile music application prototyping and development. MoMu Toolkit by Bryan et al. is an open-source software development toolkit focusing on musical interaction design for iOS devices[28]. The API can be easily embedded with iOS development environment and covers common functionality required for developing mobile music applications, from sound synthesis to sensor access. Essl developed an environment named UrMus to support interface design, interaction design and patching for synthesis, on both iOS and Android devices[29].

Throughout the history of mobile music, researchers utilize the portability of devices, wireless communication capabilities and built-in sensors to create novel music interfaces and instruments. At the same time, however, the small size of these devices often presents challenges. The small screen size of mobile phone makes the interface somewhat limited in terms of the resolution of control in musical expressivity or virtuosity. In addition, most mobile music performances must face the poor-quality internal speakers on mobile devices; researchers either separate sound generation from the mobile phone or use supplementary external speakers to increase the level of sound for a music performance. Today, with the emergence of new type of mobile devices such as tablets and with more advanced mobile technology, there appear alternatives to solve these problems and the boundary among mobile devices, other electronic devices, and a computer is getting blurred.

Evolution of Mobile Music

The era of mobile music opened new possibilities in computer music in various aspects. As I listed many of the significant works in mobile music, I characterized them based on which quality of mobile phones each piece utilized. Just to recap, Table 1 shows the reviewed works in mobile music and its pertinence in seven distinct characteristics of mobile music - ubiquity, sound synthesis, mobility, built-in sensors, social interactivity, mobile composition and machine listening.

Table 1 Evolution of mobile music

	Ubiquity	Sound Synthesis	Mobility	Built-in Sensors	Social Interactivity	Mobile Composition	Machine Listening
Dialtones	X						
PDa/ MobileSTK		X					
Pocket Gamelan		X	X				
Malleable Mobile Music			X	X			
Urban Remix			X	X			
CaMus			X	X			
ShaMus		X	X	X			
Ocarina		X	X	X	X		
ZoozBeat/Beatmaker		X	X	X	X	X	
LaDiDa/Songify		X	X		X		X
MoPho		X	X	X	X	X	

Even though *Dialtones(a Telesymphony)* could not explore any other dimension of mobile music making, it is the only piece from the table which actively used mobile phones that casual phone users brought to the concert in the performance. Most of the works in mobile music are somewhat motivated and benefited by its ubiquity, but ubiquity of mobile phones is not a necessity. I believe there is relatively uncharted space left where one can combine ubiquity and current mobile technology to seek a novel musical opportunity, which naturally led me to pursue a research in audience participation using mobile phones. The audience participation music pieces that deploy mobile phones will be discussed later in this chapter.

In addition, audience participation music piece using mobile phones is also motivated by the intent of mobile music researchers to expand mobile phone ensembles to large-scale orchestra[30]. Even though previous works proved that mobile phones could be easily turned into musical instruments in a live concert, the mobile phone ensemble is mostly limited to the computer music researchers who are experts in mobile application development and computer music. By transferring the ownership of instruments to audience regardless of their musical backgrounds, I wish to push the boundary of mobile phone orchestra by having audience as a member of the ensemble. This work will demonstrate the possibility of large-scale mobile phone orchestra from the perspective of accessibility.

Audience Participation at a Music Performance

There is a definite separation between audience and performers in a traditional music concert. However, musicians and composers try to blur this separation by making audience involved in a music performance. Audience participation in a music concert is not a new concept and there have been numerous attempts to make audience participate both in musical and non-musical way, in a wide variety of genres from contemporary music to popular music.

In popular music, many musicians make use of audience participation to engage audience at a live concert, where the audience is often expected to sing a part of the lyrics. For instance, the band first sings and then they signal the audience to repeat it. The artists often hold their microphone toward the audience, which we usually term as call and response song. On the other hand, there is a case where audience already knows the music and how to participate in the music. *We Will Rock You* by British Rock Band, Queen, which is famous for repetition of simple participatory rhythm “stomp, stomp, clap”, leads to a very successful participation. Brian May, the guitarist of Queen recalls:

> *“We will rock you” was a response to a particular phase in our career when the audience was almost becoming a bigger part of the show than we were. They would sing all the songs. In a place like Birmingham, they'd be so vociferous that we'd have to stop the show and let them sing to us. So both Freddie and I thought it would be an interesting experiment to write songs with audience participation specifically in mind. My feeling was that everyone can stomp and clap and sing a simple motif* [31].

It can be seen that the audience participation was intended right from its composition and it was taken into account even in the studio recording by making clapping and stamping sounds from various distances to emulate a live concert[32].

Another type of audience participation conducted by Bobby McFerrin is noteworthy. He is a renowned musician for his unique vocal techniques and singing styles. In McFerrin’s concert appearance, he often performs a unique vocal improvisation with audience’s chorus. In one of his live concert films, *Spontaneous Invention*, McFerrin tries out various kinds of a vocal improvisation with audience involvement. In one case, he split the whole audience into two sections, assigned each group a simple two note melody, led the audience to sing the melody with his hand signs and sang a solo with audience's accompaniment synchronized with his conducting[33]. In his talk at the World Science Festival 2009, McFerrin demonstrated audience participation, having the audience sing in a scale depending on his standing position[34]. In both examples, audience participated in a musical improvisation context. Even though participation was very simple, audience acted as a choir and each individual performed as a singer with the limited musical space that McFerrin controlled; the first one involved synchronizing with McFerrin's hand signs with little expectation on the music, his solo, and the latter included audience's musical interpretation on his movements to sing the expected notes. Audience took more active roles than the one in Queen's example of *We will rock you*, in which audience knew the music and anticipated what would happen with their participation.

An audience participation piece from contemporary music incorporates deeper commitment of the audience compared to previous examples from popular music. In Jean Hasse's *Moths*[4], audience performed as musicians, reading a graphical music score and interpreting the score expressively. However, no musical background was necessary for the musicians. The audience was instructed to whistle along to conductor's gestures and a graphical score (Figure 5). The performance included seven minutes of practice and three minutes of performance. The only sound of music is whistle by the audience and there is no technology to augment sound of audience participation. In the recording of *Moths*, the applause at the end shows that the audience was engaged with the music. It involves two important factors for audience to get connected to music; synchronization with conductor's gestures and musical expression of audience by reading non-standard music notation.

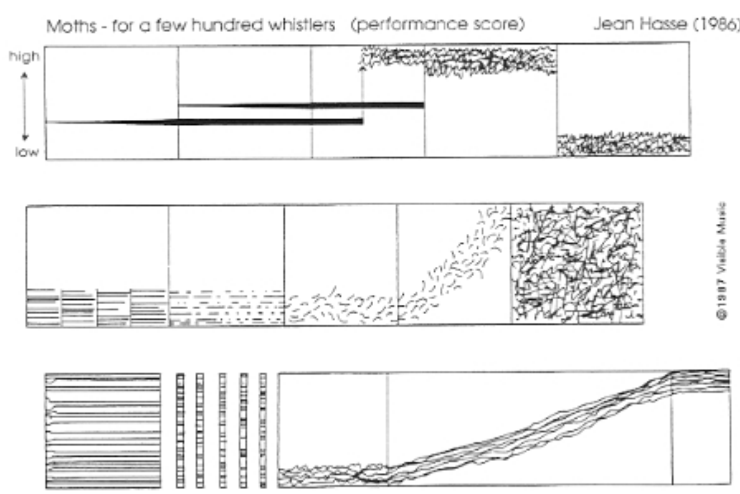

Figure 5 Graphical score of Moths

One of the earliest participatory music pieces in computer music is the collaborative work of The Hub and Ramon Sender, *HubRenga*[35]. It was a music/poetry piece where members of a poetry conference could submit poems under a common theme and the moderator, Sender, browsed the submissions, reading them as part of the music. Participants were limited to poets from the poetry conference of The Well, the Bay Area's online computer network, not open to audience in general and it was far from audience participation in a traditional performance setting.

A common approach to make audience participate in the music is to provide limited choices for them to vote on so that they can influence music without being responsible for making sounds. In Thomas C. Duffy's *the Critic's Choice*, audience participate in the composition by casting a vote on the possible ending of music[36]. In Kevin Baird's *No Clergy*, audience submitted values of parameters for algorithmic composition software via a web browser so that music notation could be generated based on submitted parameters for each musician in a small ensemble[37]. Even though audience had a continuous range for parameters so they have infinite number of choices, it is similar to a voting system in a way that the submissions are collected and averaged.

We can find similar approaches outside the musical context. There have been efforts to use technology to engage large-scale audience in a game. Cinematrix developed a computer vision system and a reflective paddle[38]. The paddles were distributed to audience and used to collectively navigate game objects on a projection by holding up either red or green side of the paddle. Another work took a similar approach to control a mass-audience game by computer vision tracking of audience movement [39].

These types of interactions make large-scale participation effective and let organizers easily simulate the possible scenarios. On the other hand, the nature of audience engagement in these cases is very different from that of previous examples (e.g. McFerrin, *Moths*, Queen). While, in previous cases, audience followed musician's instructions (or gestures) to actually contribute to music and became a part of the performance like musicians usually do, here the audience took a governing role to decide a direction of the performance similar to a conductor. In addition, in the voting scenario, as audience can choose from multiple choices, the averaged (or aggregated) result of whole audience interactions may be different from each individual's choice. In contrast, in the singing audience scenario, the goal of participation is to make individual voices united as a choir.

On the other hand, there were efforts to involve audience in music composition such as to represent audience's gestures into music notation and to have acoustic instrument musicians play music based on their interpretation of the real-time notation. In Freeman's piece for chamber orchestra, *Glimmer*, he built a feedback loop model of interaction so that audience can shape the composition during the performance[40]. In the concert, audiences of 200-600 members were divided into groups and used battery-operated light sticks during the piece to influence the music. With computer vision technology, light stick motion and collective intensity of light of each group was detected, analyzed and compared with that of other groups (Figure 6). Based on the aggregated movements of a group, the system generated music notations of a musician who was associated with the group. The following work by Freeman and Godfrey, *Flock*, is a full-evening music for sax quartet and audience participation[41]. *Flock* also included a computer vision system to track the participants' location in an open space venue, and the system produced real-time notation for sax-quartets along with electronic sounds. The author increased the length of the piece to a full evening so that audience will have enough time to learn the mapping between participation and music to have a more meaningful participation[3].

Figure 6 Audience waving light sticks in Glimmer

Audience Participation Using Mobile Phones

Early Works of Audience Participation Using Mobile Phones

Not surprisingly, using mobile phone as an interface of audience participation is not new. The first project, which has audience involvement using mobile phones, is *Dialtones (A Telesymphony)*[6]. However, in this example, the audio being generated through speakers of mobile phones was not the result of audience's interaction but just ringtones triggered by musicians on stage. Even though *Dialtones* is an early piece that engages large-scale audience using mobile technologies, many music researchers who have pursued audience participation point out that *Dialtones* is not an audience participation piece due to the passive role of the audience[2, 5, 30].

McAllister et al. utilized PDA to have audience participate in a music improvisation[5]. The application on PDA was implemented to capture and transmit the graphic gestures of each participant, sampled from the audience. With the transmitted data, a server computer generated a graphical notation on a digital screen for the corresponding musician on stage to read. The system created a closed loop between the audience performer and musician linked by the graphic gestures and the sonic result.

Although it was not designed for traditional music performance, *Net_Dérive* by Tanaka is one of the earliest participatory mobile music pieces[42]. It was an audiovisual installation in a gallery and its surrounding urban environment connected to mobile phones. Participants were given a scarf with two mobile phones and were told to wander around in the neighborhood of the gallery. The sounds and pictures collected in the streets and participants' location were deployed as materials for visualization and sonification of the installation, which was streamed to participants as well.

As pioneering works that used mobile phones for audience participation, two research works described above utilized mobile devices to involve listners in the music making process. However, in both cases, participation was small-scale and limited to a few people selected. Also, both works were not at the stage where researchers took advantage of ubiquity since mobile phones that contain the developed application were not the ones that audience owned but the ones that researchers provided to audience.

Ubiquity of Mobile Phones and Audience Participation

One of the earliest examples which are built upon audience's mobile phones is Sello's work, *isms* [43]. It is a network music piece using text messages that audience members send to a designated phone number. These textual messages, which may contain comments, questions or urban poetry, are transcribed into score data for the performers and transformed into video-projections. The performers respond to the audience's contributions and generate an interactive communication between audience and musicians.

Recently, Stanford Mobile Phone Orchestra (MoPho) held their annual MoPho Concert in November 2010 and the theme of the concert was audience participation[30]. They performed five pieces that incorporated audience participation in various ways. One of the novel approaches they introduced was to utilize social aspects of mobile computing technologies to collect sound sources for music making, often called crowdsourcing. In Oh's piece, *Converge 2.0*, she encouraged the public to submit audio-visual material prior to the concert. The submitted materials were then used as sonic and visual objects in the actual performance. The audience participation took place in asynchronous and ubiquitous manner, similar to social usage of mobile phones, such as sending emails and sharing pictures with friends any time from any place. Similar approach of collecting audiovisual objects from audience was used for the piece *Madder Libs* by Kruge, using video snippets that people submitted to trigger rhythmic audiovisual events in a grid controller inspired from MPC controller[44].

In the concert, real-time audience participation was also conducted in a music piece, *TweetDreams*. In *TweetDreams*, a certain set of hash tags were selected by performers and tweets containing those hashtags were sonified and visualized with its textual contents[1]. Audience members used their personal mobile devices to tweet using keywords and were able to trigger audiovisual events in the performance. Furthermore, audience members could interact with anyone in and out of the performance space by tweeting textual messages.

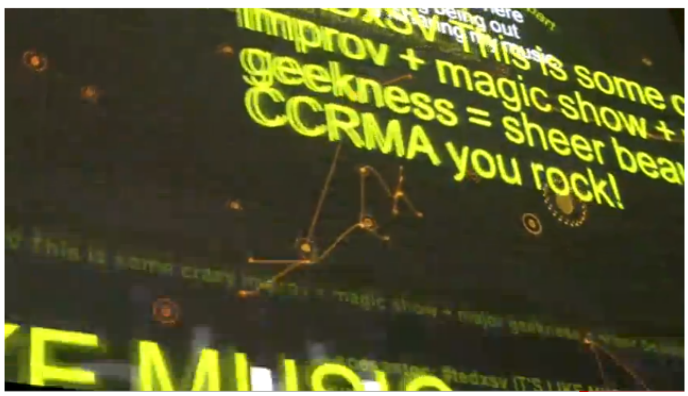

Figure 7 TweetDreams - tweets projected on screen

Another audiovisual performance, *Moori* by Kim*,* took a similar approach as *TweetDreams* to engage audience through multiple channels[45]. In *Moori*, participants could send a text message using their mobile phones or portable smart devices (such as iPod Touch or iPad), in response to guided questions by the performer. As users sent text messages, they appeared on the screen and words were spoken with text to speech (TTS) software. The messages could be sent via SMS or an OSC controller application, *MrMr*[46]. *MrMr* also let participants to control visual elements on the projection screen. Prior to the performance, audience members were asked to configure their devices with an instruction handout and a verbal explanation.

massMobile by Weitzner et al. facilitates rapid development and enables plug-and-play setup on mobile platform, especially for audience participation[47]. Instead of implementing specific participation model, massMobile offers a development framework for audience participation. The system includes a Max/MSP API for configurable user interface on a mobile platform and for audience interaction data query from server. In addition, it utilizes web standards and runs on any mobile web browser. Hence any smart phone user can participate regardless of operating system, and the configuration for audience is simplified to loading a URL.

Researchers and artists deployed mobile phones as an interface of audience participation music piece, appreciating its mobility and ubiquity. All the works proved that audience could effectively participate in an interactive music piece only if they had mobile phones. One commonly found aspect throughout the pieces discussed in this section, which is not easily apparent in previous audience participation works, is that mobile phones operate to trigger (or provide) audiovisual objects which become source material for music, accompanied by visualization projected on stage screen. I believe it is an effective way in which performers can shape the music as they desire, with reduced risks and let audience instantly identify their participation visually from the projection. On the other hand, audience would have participatory experience with little musical intention. They would not have an insight on how their participation would contribute to the music since participation is limited to triggering (or providing) audiovisual events because performers musically organize their participation in selective manner. As audience participated with limited musical expressivity in the performance, the visual and textual feedback of their participation seems to be a natural choice to foster engaging experiences for audience.

Even though these pieces needed the ubiquity of mobile phones, the audience's usage of mobile phones in these pieces is different from the usage of the devices in mobile phone orchestra where they tried to craft a complete musical instrument inside the device (See Table 2). Hence I believe an audience participation using mobile phones as musical instruments can be realized and audience can participate as a musician similar to singing and clapping in Queen's united voice or McFerrin's improvisation cases. In the following two sections, revisiting audience participation music works, qualities that make successful audience participation and novel criteria to characterize audience participation music piece are proposed. A novel form of audience participation is explored to have more musically engaging participatory experience with mobile phones.

Table 2 Evolution of mobile music with audience participation pieces

	Ubiquity	Sound Synthesis	Mobility	Built-in Sensors	Social Interactivity	Mobile Composition	Machine Listening
Dialtones	X						
MoPho		X	X	X	X	X	
isms	X		X				
Moori/TweetDreams	X		X		X		
Converge 2.0	X		X				
massMobile	X		X				

Key Success Factors for Shaping Large Scale Audience Participation

Winkler conducted an insightful research on audience participation on a interactive installation in a gallery setup[48]. In the paper, he emphasized that non-digital aspects (such as physical, social, and personal factors) of a digital installation have a significant impact on audience perception. He pointed out that one big challenge to audience participation for an interactive installation artist is "to know the full range of audience members and provide a multi-layered work that will be engaging on many cognitive, physical, and emotional levels". I strongly believe this challenge also applies to the one in a music performance setup and this requires that the participatory performance should satisfy a set of qualities to be successfully delivered.

First of all, an interaction must be designed to be easy and straightforward for the audience to understand within a very short amount of time so that any individual can easily participate. Not only the interaction should be easy to participate, but also the relationship between their participation and its musical consequence should be clear. There is a variety of strategies to make a participatory performance accessible to spectators. Sometimes it is inherently easy since audience members exactly know what they have to do at the time of participation like the case of *We Will Rock You*. Often times, performers prepare multiple choices that each audience member can vote on. Or one can have a simple and expressive gesture that may not necessarily be a musical activity but interpreted and mapped into musical events. There are a number of attempts to utilize visualization to help audience understand their interaction.

Revisiting McFerrin's improvisation with audience participation, another challenge of audience participation can be inferred from the following event. In one excerpt from his live 1986 concert, *Spontaneous Invention*, McFerrin initiated audience's singing a well-known nursery rhyme, itsy bitsy spider, with hand gestures[49]. After the first iteration of large-scale singing, he expressed his disappointment with low participation of one particular section from the audience. Given that it's reasonable to think singing nursery rhyme with simple gestures is easy, audience would not feel like participating actively because they were not intellectually, emotionally or musically interested with the participation. Would it have been different if the target audience were kindergarten class? Probably, yes. It indicates that the participation should have some attraction and should be engaging enough to sustain interest during the piece.

In the video, there was another interesting moment when McFerrin urged the group of audience to stand up and sing the song but found that no one wanted to go first. This event tells us that there's a social barrier of participation where people in public want to avoid the situation that will draw attention of others, which is also noted by Winkler[48]. A person usually does not want to be the first one to participate so that he or she would have a significant influence on the music and would draw the attention of the other audience. This is not because the participant thinks it is difficult or boring but because he or she will not feel comfortable in being alone and observed. Therefore, any audience participation requires a good initiating procedure to boost a large-scale participation. However, once audience feels a relief of "I am not the only one" about their participation, the barrier rapidly disappears. At the very end of the video, it shows that McFerrin needed only one audience member with courage to make the whole group stand up and sing the song in front of the rest of the audience. Most of researchers and artists had a strategy (such as gestures, short rehearsal, anonymity or tutorial session prior to the participation), which dramatically reduces this psychological obstacle. Especially, in cases where audience participates through their mobile phones, such as *TweetDreams*, *Moori or massMobile*, initiation of participation becomes easy with their private screen, interface that is not shared with others, and its anonymous influence to the music.

Another type of security is necessary both for audience and performers. Typically, except the cases where all the audience knows the music in advance, musicians do not truly know how audience will participate until the moment of the performance. A possible scenario of audience participation may risk the whole performance to unwanted musical result, such as a complete silence or a complete chaos. The performers need measures to connect a variety of participation into an organized music piece so it will not exceed musicians' acceptable range of musical aesthetic and coherence. Audience would also want their music to be satisfying even with the understanding of the interactive nature and would not want to participate if each individual is afraid of "Will I do something wrong?" For this matter, practitioners mostly set inherent constraints in the participation process to deliver the music with chance involved. For instance, providing binary choices for alternate ending of a music piece is a great way to secure the music with the open-endedness with little risks.

Lastly, people like to receive clear feedbacks on their participation. When audience's participation is not linked to direct sound generation or transparent change in music, people lose their interest in their interaction since it does not seem like their action influences music at all. Like in the cases of McFerrin's improvisation or Hasse's *Moths*, sometimes participation (audience's voice) itself is direct (and sometimes only) sonic contribution to the music so the clarity of interaction is easily achieved. Nonetheless, it is not very clear all the time. In Freeman's *Flock*, the author tried to build transparency between interaction and algorithm by preparing video animation, which visualizes position of audience members and musicians and abstraction of music notation with respect to their location. The author found that some audience would have even liked to read the actual music notation that musicians read to satisfy their curiosity[40]. Outside the musical context, there has been visualization work with audience participation for the purpose of science education[50]. In that work, it was observed that audience enjoyed participation with the individual visual object that they can interact with, even though the nature of participation is not far different from mass polling. Both examples suggest that audience like to have a transparent mapping between participation and feedback.

To sum up, reviewing research and performances of audience participation gives me a set of key factors in designing interaction for a successful audience participation; i) accessibility, ii) attraction, iii) initiation, iv) musical security and v) transparency. Most of the times, these five qualities are interdependent, dynamic and sometimes even conflicting. For example, if the interaction is too simple, it may be monotonous so that audience will not be intrigued about the participation. Or people would not want to stand out among the audience in the beginning of the participation but want the exact opposite towards the end; hoping their musical actions to be distinguished from others. Hence we need a comprehensive knowledge of these qualities in designing a participation system. In the following section, revisiting audience participation music works, four ways to classify audience participation music piece will be proposed to relate each type with five qualities and to find a white space for audience to have more musically engaging participatory experience.

Classification of Large Scale Audience Participation

Any audience participation performance has a unique strategy to motivate audience and to secure the sonic outcome. I propose five methods to classify audience participation based on similarity and differences of strategy in each performance. There can be multiple types of audience participation in various setups but the domain of audience participation in this section is limited to the one in a traditional music performance setting.

Role of Audience: Composer vs. Performer vs. Audiovisual Objects

For most of audience participation works that do not involve any technology, participants sang, clapped or stamped. Even though those are simple tasks, each one incorporates some level of cognitive load and gestures similar to a musician playing an instrument. In these cases, the role of audience is to play music as a performer. On the other hand, some of the works we reviewed involve audience in shaping the music rather than playing the music, which is what a composer usually does. From the case where audience chooses the ending of the music to the cases where audience's input affect music notation generation in various levels, participants involve in completing the composition as one of the composers. Freeman's works fall into this category where the result of participation is music notation. In *Glimmer*, the author chose this design framework of audience influencing composition rather than directly generating sounds in order to make the audience as comfortable in participating as possible assuming that they have no musical background (accessibility). The separation of sound generation from audience participation needed one or more intermediate steps, for example, abstraction, algorithmic composition, score generation or human's interpretation and those steps help the composer to convert audience's participation into music in limited ways (musical security).

More recently, especially alongside the usage of mobile phones, there have been cases where the role of audience is neither a performer nor a composer. For example, in *TweetDreams* or *Moori*, even though audience has a chance to trigger an audiovisual event on screen and in the music, the will of their actions lean more on the textual communication with the musicians and the rest of audience rather than musical activities. In his paper, Dahl also states that "audience members do not play the instrument in the sense of directly controlling what sounds are made, however their actions trigger musical and graphical events whose details are determined by their actions"[1]. Another approach that musicians often have taken for audience participation is to request sound so that participants can submit audiovisual objects and musicians utilize those as compositional resources. In *Williams Re[mix]ed* by Larry Austin, he invited public to upload sound files to a website and those submissions were used in a generative sound installation[51]. In another piece, *Converge 2.0*, participants submitted audio snippets or pictures prior to the concert so that the performer can make music based upon those materials [52]. In both cases, audience might or might not experience music-making process at the time of submission, and musical aesthetic decisions are made by the performer who would select, place, and process those materials in the music piece. A new category to cover these types of participation is necessary, which can be called audiovisual objects. While this approach will lower the barrier, audience participation is limited in influencing the musical aesthetic of the piece.

Application of Participation: Collective vs. Individual vs. Selective

In singing audience case, every participant can sing individually and participate in large-scale at the same time due to the nature of participation; to make one united voice of a chant. The pieces keep individual contribution and one to one relation between audience participation and voice. Hence the result of individual participation is transparent. In *TweetDreams*, and *Moori*, their participation was also linked to a single event in music, which was easy to identify with textual contents on screen(tweets, messages, username).

However, most of the times, it is hard to link a participatory gesture to one musical event, especially when the music is performed in a smaller number of voices than the number of audience. In those cases, participation affects music in a collective manner and individual interaction is hard to identify from the music. In Thomas C. Duffy's *the Critic's Choice*, audience members participate in the composition by casting a vote on the possible ending of music and the decision is made upon the choice that they chose the most[36].

Collective participation is less transparent for audience than the case of individual participation, especially when decisions of individual audience members differ from the aggregated result of whole audience. In Freeman's *Glimmer*, audience was divided into several sections and the collective sum of each section's participation was mapped to the music notation generation of each musician. The author noted that aggregating group interaction made less change in music than there could have been because averaging all participation seemed to cancel out audience's various interactions. Some audiences were frustrated since their actions were not linked to noticeable changes in music[2]. In the author's original thought, he expected that organic group behavior would emerge during the piece but the piece was not self-explanatory enough to achieve the goal within a short amount of time.

A musician who leads the audience participation can take a different approach. While, in most of the examples we reviewed, it is audience's individual decision whether to participate in the performance, but sometimes performers selectively sample a subset of audience to participate. As an extreme, we can find cases where one audience member participates in the music performance. McFerrin often sampled one audience member in his performance[53]. The selective process prior to actual participation provides a few advantages to musicians who organize audience participation. For instance, volunteers have more motivation and less hesitation to initiate their participation than audience in general and they are willing to be involved even if the participation is more difficult, such as singing alone, in front of public. In another example such as *Converge 2.0*, the musician picked out a subset of submission to use those as a source material of music and to build a musical narrative of the performance [52]. Same approach is taken in *Piano Etudes* by letting the pianist select one piece from audience's submissions[54]. In general, although participation of selected audience can influence music either in individual or collective manner, reduced number of participants will have benefits of having more transparency in their contributions in either way since each individual will have more impact on music compared to the large-scale participation. In addition, the rest of the audience who do not participate in the piece will have a different kind of engagement from the participants; they usually empathize with the selected audience members and are amazed with the music of the participants.

Sound Source of Participation: Audience vs. Stage

Another distinction can be made based on where the sonic result of participation is generated. The sound source can be either audience themselves, instrument musicians or computer music system which takes input from the audience. When sound source is individual and right next to the participant, their role becomes close to the notion of traditional acoustic instrument player where the sound is coming from where they stand. The location of sound source matters if it comes from audience. It means that they have direct control over their sound and individual participation can be easily differentiated from overall music. This approach can naturally achieve the clarity of interaction for audience (transparency) due to the proximity (or sameness) between the sound sources and the participant. Singing audience in McFerrin's improvisation and whistling audience in *Moths* are examples where the sound is coming from each individual.

The notion of a "Stage" can include human musicians, electronic sound from computers, or any sound from main speakers. The distinction between human musicians and other computer systems is not important in this context since it is hard to achieve clarity of interaction by listening to sound from stage, regardless of its type.. For all the reviewed examples except *Moths*, *McFerrin* and *Queen*, the sound comes from a stage, either from musicians or main speakers on stage and it requires an intermediate process (physical and virtual) between audience participation and music.

Time of participation: Synchronous vs. Asynchronous

One of the criteria that fundamentally changes listening experience and participation process is whether audience participate at the time of performance (Synchronous) or prior to the performance (Asynchronous). Researchers have already considered temporal distinction to categorize network music and mobile music in time dimension[52, 55]. Most of the examples we reviewed chose real-time participation, except a few pieces - *Madder Libs* and *Converge 2.0*. There are a few more noteworthy works that allow non real-time participation prior to the actual performance. Jordà developed an online collaborative music composition system called *F@ust Music On-Line* (*FMOL*)[56]. Through the networked instrument, the system allows anyone - from novice users to professional musicians- to participate in composing music for a live concert, using computer mouse as a controller. Combining the pervasive social web medium with traditional live musical performance, Freeman composed two interactive pieces – *Graph Theory* (2006) and *Piano Etudes* (2009)[57]. *Graph Theory* provides a unique online compositional interface on the web where users can build a path in a graph where each node in the graph is a short fragment of pre-composed music and eventually the path generates a fixed score for the live concert. A final piece for the live concert was generated from the graph by searching an optimal path based on weights collected from participants' composition. *Piano Etude* takes a similar approach with more focus on individual creation of music rather than collective composition in *Graph Theory*.

Non real-time participation seems to have a great advantage in a way that it surpasses most of the physical constraints- time, tool and location. Users can participate in the music piece remotely with dedicated tools, typically through web or mobile platform and they can spend as much time as they want. However, due to the absence of live music while participating, it seems that their interactions are limited to indirect participation in music making, composition or recording submission. Also, as they cannot have instant feedback on their participation and are not exposed to interaction of the others, it would be difficult for a majority of the audience to find connection between their individual participation and the live music generated, due to the lack of instantaneousness.

Space of Participation

As Oh already mentioned in the decentralized models of audience participation[52], the participants do not need to be in the concert hall to participate in the performance. For the asynchronous examples we just discussed in the previous section, they naturally require remote interaction prior to the concert. *TweetDreams* is a different example where participants are mostly at the music performance but people can possibly participate in the performance remotely as well, even without knowing if they are participating[1]. This research focuses solely on audience who are physically in the performance. Given the limited definition of audience, space of participation criteria is not necessary.

Technology, Audience Participation and Innovation

Four ways of classifying audience participation are suggested to characterize audience participation piece based on its participation nature; Role, Sound Source, Application and Time. I believe these will be insufficient to cover all types of audience participation in the future and to account for differences between the pieces. Nonetheless, it gives me a good starting point to find an unexplored space in the field. In order to investigate previous works, I listed up the audience participation works reviewed and assign the appropriate attribute for each criterion in Table 3.

Table 3 Audience participation pieces and its attribute on each criteria.

	Role	Sound Source	Application	Time
We Will Rock You	Performer	Audience	Individual	Synchronous
McFerrin's improv.	Performer	Audience	Individual	Synchronous
Moths	Performer	Audience	Individual	Synchronous
No Clergy	Composer	Stage	Collective	Synchronous
The Critic's Choice	Composer	Stage	Collective	Synchronous
Glimmer	Composer	Stage	Collective	Synchronous
Flock	Composer	Stage	Collective	Synchronous
Graph Theory	Composer	Stage	Collective	Asynchronous
Piano Etude	Composer	Stage	Selective	Asynchronous
isms	Audiovisual Obj.	Stage	Individual	Synchronous
TweetDreams	Audiovisual Obj.	Stage	Individual	Synchronous
Moori	Audiovisual Obj.	Stage	Individual	Synchronous
Converge 2.0	Audiovisual Obj.	Stage	Selective	Asynchronous
Madder Libs	Audiovisual Obj.	Stage	Selective	Asynchronous
massMobile	Composer	Stage	Collective	Synchronous

As you may easily notice, there is a clear separation between audience participation works without technologies (the first three rows) and the rest, especially on the first two columns; Role and Sound Source. This research finds its contribution by following the paradigm of audience participation as an individual instrument player and introduces an audience participation in mobile music with four attributes in the suggested classification- Performer, Audience, Individual and Synchronous.

In those non-technological audience participation pieces, people act as musicians with musical activities like singing, clapping, or whistling. I am particularly interested in this traditional model of audience participation where each audience member has an individual musical space of his/her own and plays a musical instrument (e.g. own voice, whistle) with the ownership of the instrument. As each audience member has control over their own instruments, one can become more engaged by removing the intermediate process between their interaction and musical change from the stage. As they play note by note by their own gestures, it is natural for them to understand how their interactions contribute to music. I believe this approach naturally accomplishes transparency.

In fact, this research will not be able to accomplish "Individual" type of participation completely among the suggested types of application; Collective, Individual and Selective. As the instrument is built upon a platform of smart phone, the opportunity of participation is limited to people who own a certain type of mobile devices. Although it is different from the participation examples in which musician selectively limits the number of participants with the intention to avoid issues of large-scale participation, the nature of audience participation is a selective procedure depending on ownership of certain devices. On the contrary, it has been impossible for audience participation researchers to provide individual digital instruments unless it is a selective small-scale participation, due to the scalability of the participation and the cost associated with large-scale expansion. However, many people today use smart phones and developers can distribute a free mobile music application through an online application store in less than a minute. Therefore, using mobile phones of audience is a natural choice for me to realize audience participation in computer music context. An increasing number of smart phone users make large-scale audience participation possible, with no cost of preparation.

In terms of musicality, providing a musical instrument to audience is a challenging task for both audience and me. Unless I perform a well-known piece with very simple and repetitive gestures, e.g. *We Will Rock You*, or unless audience was guided to repeat the same vocal phrase as McFerrin does, giving them a musical instrument, even if it is as simple as one-octave piano app, will increase the risks of music being chaotic or completely silent, which is not the musical aesthetic that I envision. The closest example from the related works will be Hasse's *Moths* where audience freely improvises by whistling. In *Moths*, having a conductor, a visual score and a practice session helped audience improvise in a synchronized fashion. Likewise, to make participation easy and interesting at the same time, and to establish the music piece within the musical aesthetic, equivalent security measures should be planted in the mobile instrument. To accomplish accessibility and musicality of the performance, the concept of networked collaborative instrument is applied in developing the digital instrument. In the following section, I briefly explain multi-user collaborative musical instrument and find an appropriate model for this particular setting.

Collaborative Networked Music Instruments

As mentioned in the previous section, most of audience participation needs a component that can guide audience to know when and how to participate, unless audience participates voluntarily and organically (i.e. waving lighters). Gestural signs by a conductor or a musician were widely used[4, 32, 33, 58] and there was an attempt to have dancers among the audience to guide participation[3]. When audience participation involved technologies, most of the pieces had a visual projection on screen so that it would help audience members to receive feedback from the visualization [1, 3, 30]. In addition, providing multiple choices to audience to choose from is another strategy[36].

For this particular research where the audience actually generates sound, it is much more challenging to guide audience to perform as if a composer usually does for professional musicians with musical notation or conducting. To resolve the issue, I would like to place a constraint on audience's application so that audience can improvise on the instrument and contribute to the music simultaneously, while another musician controls the constraint in a hierarchical manner. The idea originates from traditions of collaborative musical instruments in and out of computer music history.

An early example of a collaborative musical instrument can be found in traditional Indonesian music. Gamelan is an Indonesian musical ensemble where a number of people play a variety of musical instruments [59]. However, Gamelan music is not for a soloist's art. No melody can be singled out and played by one member of the ensemble. Because of the interrelated structure of Gamelan music, when one learns to play, one needs to know almost all components of Gamelan ensemble, which also implies that the whole ensemble is one collaborative instrument.

While there is no mutual interaction beyond music in Gamelan's example, in John Cage's piece, *Imaginary Landscape No.4* for twelve radios, 24 performers were paired to control one radio - one of them controlled the frequency dial to traverse the radio station as indicated in the notation while the other controlled the volume level. The sonic result is a product of two musicians' individual control (contents and dynamics) and the inherent randomness coming from the radio. The piece can be considered as the first multi-user instrument in Western music history[60, 61].

Later, computer music researchers proposed taxonomy for classifying collaborative and networked music-making environments or multi-user musical instruments. Barbosa proposed a thorough classification of computer supported collaborative music based on synchronisms and tele-presence[55]. Weinberg classifies network music environments depending on the roles of computer and the level of interconnectivity among users[61]. He also describes a number of topologies depending on relations among musicians (centralized / decentralized) and the nature of interconnectivity (synchronous / sequential). Jordà describes the paradigm of multi-user instruments and explored some examples in distinct paradigms[60].

With the goal to provide constrained musical instrument for novice users, the dynamics of musician and audience here is very different from the typical collaborative instrument examples, such as *DaisyPhone*[62], *reacTable* or *FMOL*[63], where often the goal of the system is to facilitate group creativity. Each musician has a balanced control over music and takes a flexible role in the collaboration. In contrast, the target of this study is analogous to the examples where each musician has a specific role in a global sequence of generating sound. *The Squeezables* exemplifies this type of approach in crafting networked instrument[64]. There are three performers and five squeezable balls for accompaniment and one soloist ball with a specific function for each ball. While three balls determined the low-level aspects, mostly timbral parameters of a voice, the other two balls controlled the higher-level contents of the voice such as rhythm and arpeggio. The five accompaniment balls also affected the melody ball, which controls the pitch contour.

More recently, *Beatscape* by Albin et al.[65] took a similar approach by introducing two different types of performers; one using tangible objects on a tabletop interface and the other type of musician using Wii Remote. The former type of performer place a sound object, which is a physical object, on a tabletop interface and adjust the pitch through a rotating gesture, while the latter type of performers actually generate sound by triggering a playbar using Wii Remote. As I personally had a chance to perform in the *Beatscape* ensemble, I am inspired from the concept of *Beatscape*, which separates the roles of selecting sound source and generating sound.

The most direct predecessor of this study is *Pazellian* by Pazel et al., a distributed interactive music application using harmonic constraint[66]. The application is designed specifically for a wide audience to enjoy a musical experience with the computer application. In the application, there are three different types of roles to control music - Performer, Conductor, or Maestro. The conductor-role user can control the high level aspects of the music such as volume, tempo and, most importantly, the harmony of the current music. The performers can choose a voice (instrumentation) of the music and perform in real-time by moving indicator in a two dimensional space (pitch-volume) using mouse. The system automatically corrects the pitch selected by performer user according to the harmony of the music, which is concurrently controlled by the conductor user. The user study indicates that participants from a diverse age group could enjoy the performer's role while the conductor's role was not appreciated in general. This approach that a conductor set a harmonic constraint on other performers interaction is considered as a proper model for non-musicians to involve in playing instruments without training, while keeping the harmonic progression of music reasonable.

Recently, d'Alessandro et al. applied this concept of distributed roles in music making to mobile space. Combining two different interfaces- *Vuzik* and *ChoirMob*, they explore the traditional choir and turn it into a digital choir for mobile devices[67]. Following a hierarchical relationship between a choir and their conductor/music, the system utilizes *Vuzik*(conductor), an interactive display for composition and *ChoirMob*, a singing synthesis mobile application (each member of choir). In performance, a pre-composed music piece is drawn in *Vuzik* and it distributes musical phrases to *ChoirMob.* Each mobile phone choir member can perform in a musical space (note onset, dynamics, pitch deviation, and vowel of voice) within the broadcasted reference pitch. With the automated conducting system, the mobile application accommodates novices to play music with expressivity.

In this study, the mobile application needs to be intuitive to use so that audience can interact with it immediately after they have their instruments. At the same time, non-musicians' performance with the musical instrument has to deliver satisfying musical result to audience regardless. The harmonic constraint model of *Pazellian* is applied to the study so that the application will delegate a high-level musical decision to a musician and let the audience improvise in a relatively safe zone under the constraint.

CHAPTER 3

DESIGN PRINCIPLES AND IMPLEMENTATION OF ECHOBO

The goal of this research is to perform an audience participation music piece and to make audience connect to the performance by letting them play musical instruments on their mobile phones. My hypotheses for this research are as follows; i) The use of a multi-user interactive mobile music application at a live musical performance will offer an accessible audience participation environment and help audiences feel more connected to the music performed. ii) Incorporating elements of social interaction among audience members in the mobile instrument will enable an individual member to have a greater influence on the music and give a sense of collaboration among the audience.

Through an investigation of the related works, I chose to develop a collaborative and networked music instrument on mobile platform. There are two challenges in developing the mobile instrument for successful audience participation. First of all, the application needs to be easy to use in various perspectives since audience will have no prior exposure to this particular music and the instrument (Accessibility). It will be challenging for audience to get familiarized with the instrument promptly enough to make music. Second, it is also challenging to keep audience members from being lost during the performance and losing interest in their participation since a major part of music comes from their mobile phones (Attraction). While there are multiple possible reasons that can discourage active and consistent participation, I would like to incorporate social interaction element in the application to motivate audience to keep musicianship during the piece.

With these considerations, a mobile music application, named *echobo*, was developed. The rest of this section will cover design principles and implementation details of *echobo*. The application is free and available in the Apple App Store since January 2012.

Networked Performance Concept and Target Performance Setting of Echobo

As a major design decision, the mobile instrument has collaborative aspects and network constraints as a multi-user instrument. The main concept is to separate the level of control in music and to assign simpler and flexible tasks for audience. More specifically, I would like to have a master musician control the high level music structure, connect the master musician's application with audience musicians' applications and let the audience play the instrument at low level, note by note, under the constraint. That way, audience is restricted to play notes in a specific measure of the piece.

As described, the responsibility of master musician is to control musical structure of the piece at a high level (Figure 8). More precisely, the master musician controls chord progression of the music. The master musician can set a sequence of chord progression on the master interface and progress the music by selecting a chord at the right timing. However, the master musician cannot generate any sound from the application but the audience can. On audience's mobile phone screen, a very simple keyboard is implemented that can be played using touch screen. The keyboard has only eight keys in major/minor scale that are associated with the chord that the master musician on stage selects at the moment. The keys only appear when the master musician selects a chord.

Whenever the master musician selects a different chord from the previous one, keys on the audience interface will be transposed. For example, suppose the master musician selects C Major chord, eight keys in major scale with root note C will appear on audience's application. If the master musician then changes the chord to E minor, existing keys in C Major scale will fade out and a new set of keys in E minor scale will fade in. Projection of the master musician's interface is desirable so that audience will anticipate the timing of keys change.

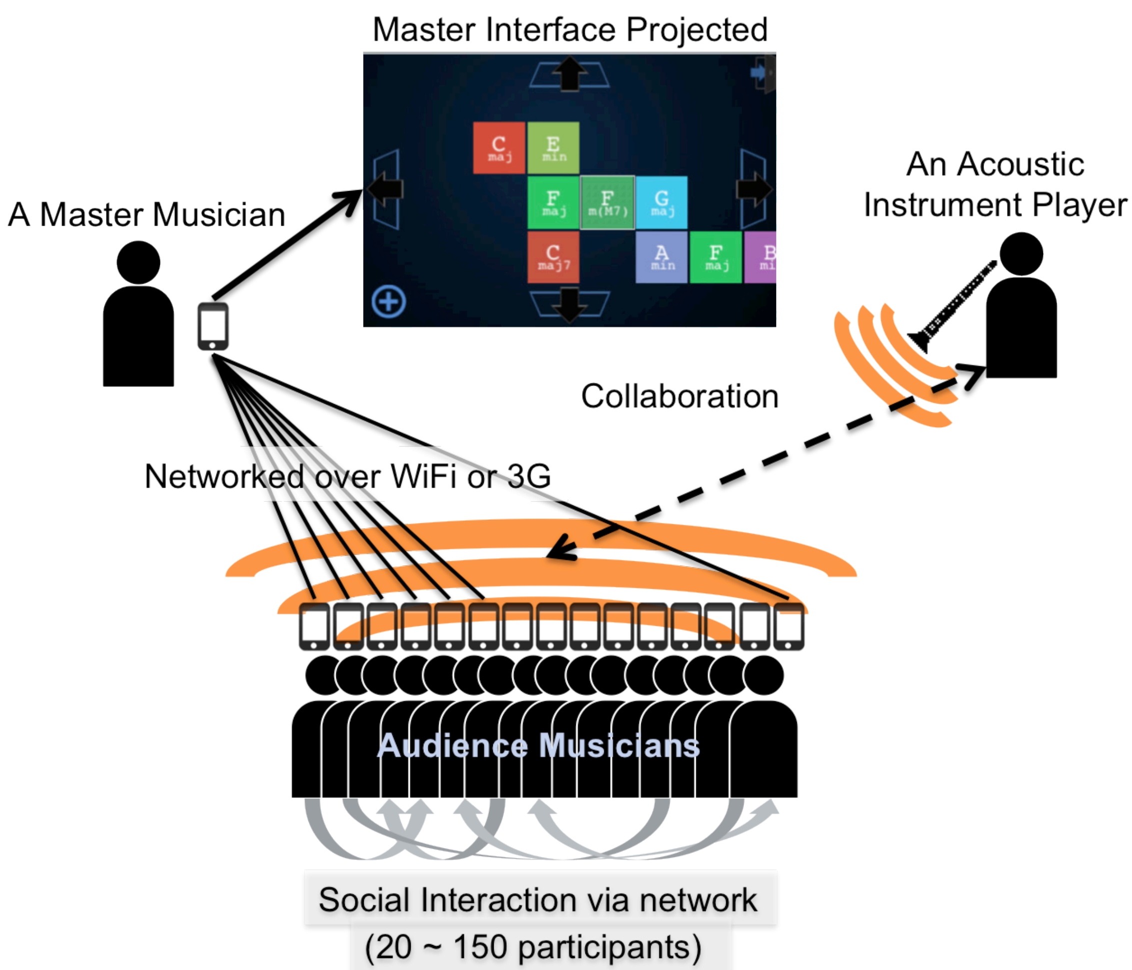

Figure 8 Performance concept diagram of echobo

By having limited keys based on the chord progression sequence that the master musician plays, any possible sound generated by mobile phones would be in a correct scale with the musical structure that the master musician changes over time. An analogy can be made for the musical relationship between the master musician and the audience musicians as the relationship between the left hand and the right hand while playing an acoustic guitar. Imagine a professional musician having a grip of chord on fret board and another person (possibly non-musician) plucking strings to make sounds. Hence the master musician's grip will eliminate the risk that the music will be out of key throughout the performance.

The dual constrained application structure, in which the master musician controls the pitch space while the audience generates sound, will help the musicians on stage take the responsibility to build the music while leaving the ownership of sound generation to audience within the controlled space. This constraint will work effectively especially under condition that the mass-audience will participate in the piece with the instrument they get right before the performance. While there can be other accessible strategies that makes large-scale audience participation possible, such as singing repetitive chorus or clapping to the beats, this constrained model is more applicable to the notion of musical instrument in mobile phones. Additionally, it can create a novel sonic experience where each audience member can perform as an individual member of an ensemble and the whole audience produces an accompaniment, rhythmically disorganized but harmonically organized.

In the audience interface of *echobo*, pressing the keys makes an electric piano sound. It could have been mapped to any other instrument sound but I wished the sound of instrument to have definite attack and decay envelope so that it is easy to distinguish from other sounds. Collective sound is generated when entire audience plays the instrument. Since no rhythmic guideline is given, the individual sound is defined by each member's musical decisions and it is hard to expect how each individual will play these keys. At its extreme, audience may play not even a single note or play very densely. Recall the analogy of left hand (the master musician) and right hand (audience musicians) in an acoustic guitar and imagine that there are 100 people plucking 100 guitar strings in their own individual way while they have the same grip on C major chord. Eventually, when all the sounds from audience are layered, the collective result will sound like continuous guitar strumming without any particular rhythmic pattern.

The aggregated sound results in a dense and arbitrary series of notes in a chord and can be employed for background harmony in the same key. The instrument will produce a unique pileup of distinct electronic piano sounds. Additionally, the texture of the sounds can vary based on the number of participants and the variety of audience's improvisation in note duration and note density. The nature of collective improvisation in this limited and controlled musical space forces the musicality of the instrument to be limited; a harmonic accompaniment of chord progression. This will also limit the music style into which the instrument can fit, for example, minimal music with static chord change or a simple harmonic accompaniment similar to a synth pad sound. Considering the whole set of mobile phones connected to master's mobile phone as one instrument, the master musician can coordinate a group collaboration of audience and, in some level, control the texture by giving a textual instruction or gestural instruction to audience musicians. Those methods to control overall texture with a group of audience are analogous to the notion of a high-level control with randomness used in stochastic music, where randomness is factored in the improvisation of each audience member. This is similar to the way Xenakis manipulated note density with single parameter, the average number of events, in the algorithm of his piece, ST/10-1 080262[68]. More detail about the textual instruction will be discussed later in this chapter.

Due to the nature of its sound similar to a harmonic accompaniment, having an acoustic instrument is one of the appropriate music styles that requires an accompaniment of chord progression for the solo instrument. Therefore I decided to have an acoustic instrument player who plays a solo along with the sound that audience generates (Figure 8). It is also similar to the model of McFerrin's improvisation example where audience sang repetitive phrases as accompaniment, and he sang as a solo instrument. Having an acoustic instrument player will not only improve the musical aesthetic, but also help the audience connect to the music in collaboration with the musician.

Target number of audience for this particular performance is relatively small-scale (from 20 to 150 people) just to avoid the issue of large-scale network traffic (either wireless, 3G or 4G) in the performance space and the server application. In addition, the acoustics of the concert space is important due to the limited volume of a mobile phone. For example, an auditorium or a traditional concert hall in which acoustic instruments are not amplified are more appropriate than an open-air theatre where a set of speakers is desirable.

Stage layout is flexible. As there is no cable or amplification required in the piece except the projection of the master musician's interface, a master musician, an acoustic instrument player and audience musicians can stand in any layout; i) a traditional layout where performers (a master musician and an instrumentalist) on stage while audience is seated, or ii) a circular format where performers are in the center of a performance space, surrounded by audience with audience being seated, standing or mobile! The latter has an advantage to performers since they can listen to audience playing better than the former, but it will not be appropriate when the acoustic instrument has certain directionality in sound generation, such as a trumpet.

Both the applications for the master musician and audience musicians are built inside one mobile application. While the reason behind using mobile phones of audience is already explained, I also wanted to build a master's instrument on the mobile platform as well. The motivation is to follow the spirit of mobile phone orchestra to craft a music instrument inside the mobile phone. As the instrument stays within the mobile platform, it will be easy to distribute this instrument to anyone without any technical requirement for the target performance setting only if there is a master musician with understanding of chord progression.

Room Creation, Configuration and Joining

Two completely different interfaces for the master musician(or "meta musician" inside the application) and audience musicians (or "echo musicians") reside inside *echobo*. After a user launches the app, he or she can choose one of two interfaces. I added a notion of "a room" for each master musician. A master musician can create a room and the identity of the room is the username of the master musician. A master musician can just type his/her own username to create a room in the green box on the room creation view (Figure 9). As far as the audience musicians know the username of the master musician, which will be instructed in the performance space, they can join the room by typing the username of the master musician in the purple textbox, pressing search button and join button. There is a minimal description about the application on the bottom of the view.

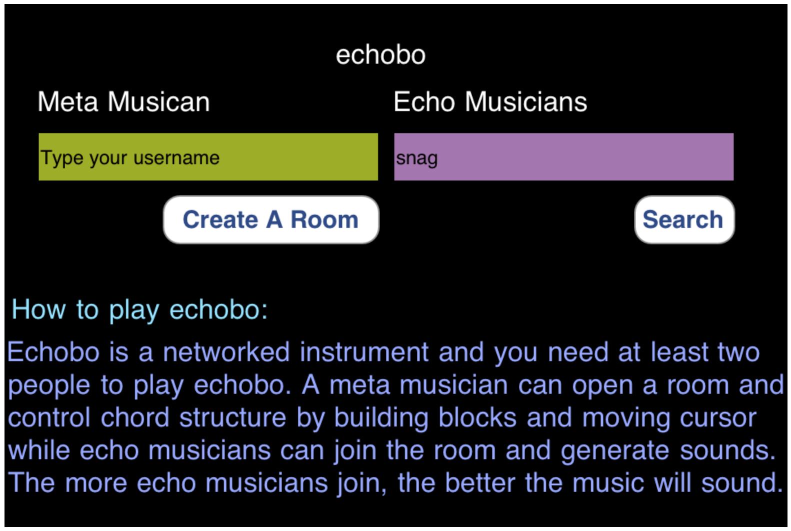

Figure 9 Room creation, search and joining view

Once the master musician presses the create button, the view is replaced with the room configuration view (Figure 10). In the configuration view, there are two forms that the master can specify. The slider on the top is for setting tempo of the music. Once the tempo is set, it will later be used to give a visual cue (a horizontal playhead) of tempo for the participants and musicians. In addition, the tempo will be used as a reference of one cycle for pattern broadcasting function, which will be discussed later in this chapter.

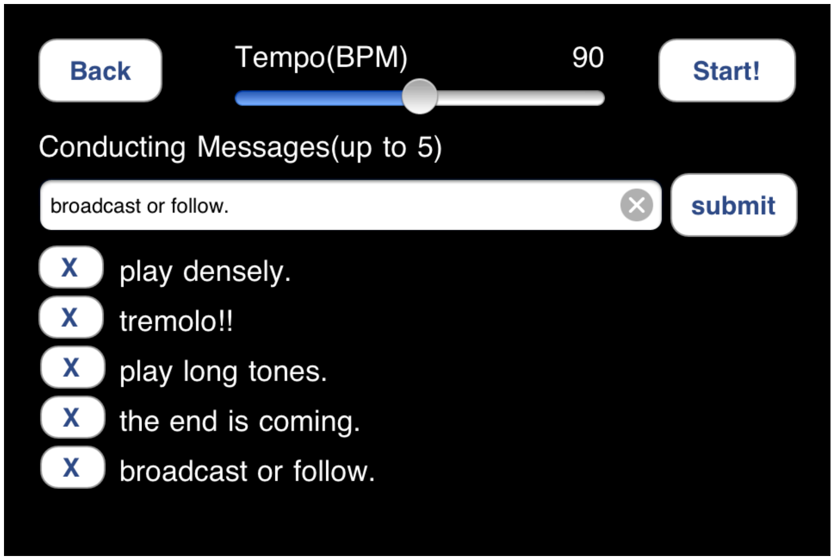

Figure 10 Configuration view -tempo slider and conducting message input form

At the bottom of configuration view, a master musician can register multiple textual messages. As a group of both master musician and whole audience forms one networked instrument, textual messages are used for the master musician to coordinate collaboration in audience's performance like conductor's gestures. Once the master musician submits multiple messages in this view, he or she can broadcast this text later during the performance to all audience screens so that they can follow the instruction to shape the sound of the instrument as the master musician wants. For example, the conducting message may contain a variety of instructions such as, "Play densely.", "Play long tones.", "The end is coming.", "Tremolo!", so on.

Master Musician Interface

The design of master musician interface in *echobo* directly comes from one of my past group projects, named *Crossole.* Meaning crossword of sound, *Crossole* is a musical meta-instrument where the music is visualized as a set of virtual blocks that resemble a crossword puzzle. In *Crossole*, the chord progressions are visually presented as a set of virtual blocks. With the aid of the Kinect sensing technology, a performer controls music by manipulating the crossword blocks using hand movements. The performer can build chords at a high level, traverse over the blocks, and step into a low level to control the chord arpeggiations note by note. With a similar role of the master musician controlling chord progression, I applied the same interface of chord progression part in *Crossole*(Figure 11) into the master's interface of *echobo*. For detail, see [69].

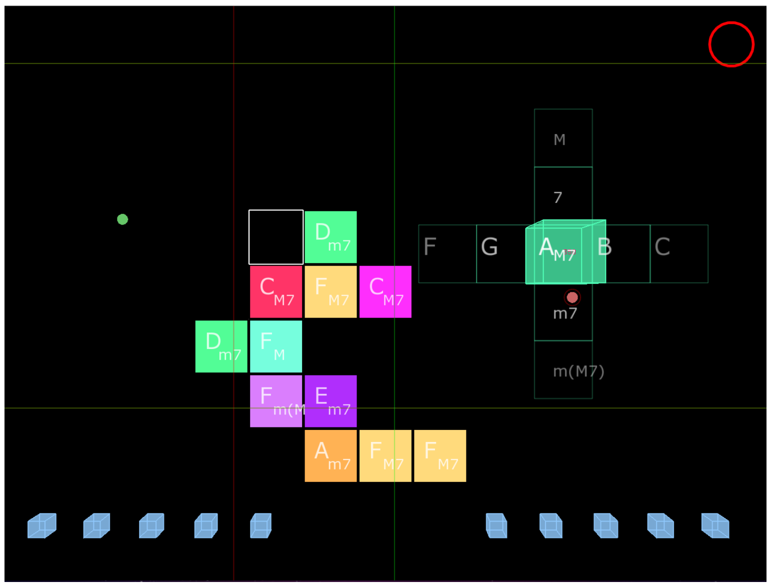

Figure 11 Screenshot of Crossole

The visualization of master interface in *echobo* employs a novel representation of musical structure; chord is symbolized by a colored square (or a "Block"') and a piece of music can be presented as a set of blocks, which resembles a crossword puzzle at the end. As the first step, the player instantiates a block with the plus sign button on the left bottom (Figure 12). Then, the player associates the block with a base note of the chord (e.g. C, D, E) and a chord type (e.g. major, minor, major 7, etc) by traversing the scrolling menu and selecting a right menu of choice. A combination of base note and chord type determines color-coding with which the associated block is drawn. Finally, the chord block can be placed anywhere on the screen with a drag and drop gesture using the grab buttons (two blue hand icons). The screen only shows one fourth of the space available virtually and a user can change the viewpoint by finger dragging.

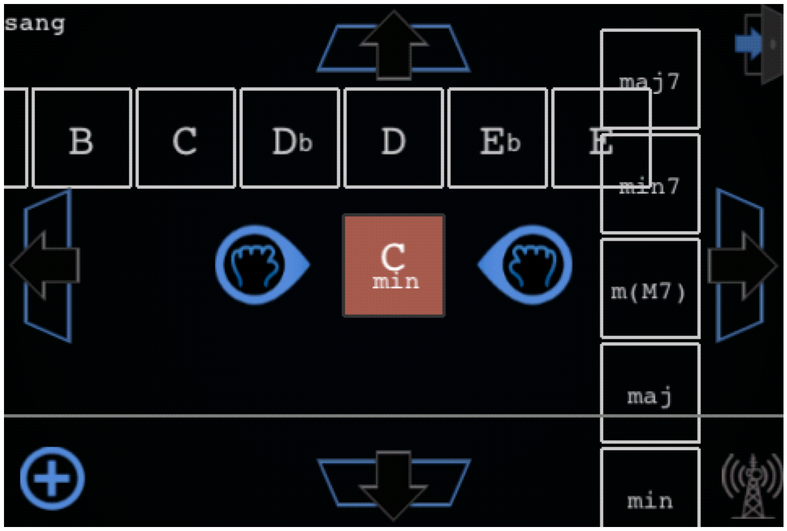

Figure 12 Selecting a chord menu when "plus" button is pressed

Another important element that completes chord progression control is the play cursor, which is depicted as a white empty square (Figure 13). While adding blocks provides a base for the music structure, the cursor control determines the temporal progression of chords. The cursor can move only one unit in four directions (up, down, left and right) on each step. The buttons for moving the cursor in four directions are placed on the four sides of the screen. A master musician can select which chord to play on the fly by moving the cursor upon the blocks specifying the chords. Therefore, the set of blocks can be turned into a unique sequence of chords depending on when and how the cursor travels on the screen. Naturally, a new block needs to be placed adjacent to a block that it follows in the chord sequence.

Figure 13 A music structure represented by a set of virtual blocks. The cursor is on F major block

Building the block structure is again essential in a way that the position of each chord determines a possible chord sequence with the constraint of cursor move. On the other hand, in order to achieve a correct progression, the cursor control needs to be precise in time. Four arrow buttons for cursor control are always available, even while selecting chord and broadcasting conducting messages. In case of pre-composed piece, although creating the chord blocks can be prepared prior to the performance, it is recommended to build the structure in real time to engage audience through the projection. In contrast, in improvisational setup for all three musicians - a master musician an instrumental player, and audience, the master musician would build a chord structure on the fly, responding to what the other two types of musicians play and giving them anticipation on the improvisation, especially to the instrument player.

Other than two main elements, block and cursor, there is an exit button on the top right and the name of the master's musician on the top left (Figure 13). The grey horizontal playhead is a visual cue for temporal synchronization of master's phone with others (both audience and an acoustic instrument player). One sweep of the playhead line from top to bottom is one measure long, based on the BPM (beat per minute) specified by the master musician and time signature (4/4) hard-coded inside the application. Presently, the master musician can specify BPM in the configuration view and time signature selection feature will be added in the near future. The antenna button on the bottom right is for broadcasting conducting messages, which will be discussed later in this chapter.

The visualization of master interface in *echobo* employs a novel representation of musical structure (Figure 13). While the player can simply place blocks in a long line according to the order of chord sequence from left to right, the two dimensional depiction of blocks without a specific directionality encourages the master musician to understand the musical structure and find an efficient way of constructing chord clusters. For example, if a sequence is five iterations of four different chords, it is obviously much easier to place four blocks in a circular manner than to place twenty blocks from left to right. The repetitions, variations, and musical form (such as A-A-B-B, or verse, chorus, bridge) can be interpreted by the player in an efficient representation. As there are many ways of building a block structure for one chord sequence, finding an effective representation of musical structure and building it on the fly are the virtuosity of the master musician for this instrument.

The motivation behind this visual interface is effective audience engagement. It is recommended to have visualization projected on a screen in order to accomplish transparency of the relationship between the master musician and audience musicians, to build audience's expectation of chord change and to facilitate synchronous engagement. More specifically, the visualized block structure lies between a musical notation and a musical interface; it has general information about what the current chord is and gives audience some anticipation on the next chord based on observation of cursor moves. However, it is not a music notation but a graphical interface because it has no temporal information about the music and the timing of cursor move is determined by the master musician. With its incompleteness, the block structure and the cursor move give a rough idea of what the musical structure is like. At the same time, a real-time gesture of moving cursor remains in the visualization similar to the way musicians in a traditional ensemble watch each other to synchronize their musical gestures. In summary, the visual interface helps 1) to reveal all musical mappings of the master's instrument rather than hide it behind the screen, 2) to have a real-time involvement of playing a musical instrument and 3) to make audience comprehend and anticipate the musical structure.

Audience Musicians Interface and its Sonic Space

Even though the audience side of *echobo* is dependent on the master's side, it is only the audience's mobile phones that generate sound. Each audience's application is close to the notion of a traditional musical instrument. In general, it is a virtue in developing a new musical instrument to offer *low entry fee with no ceiling on virtuosity*[70]. In this particular case of the whole networked instrument, the virtuosity mostly applies only to the master musician; developing the block structure quickly on the fly, controlling the cursor precisely, organizing active audience participation before and during the performance with verbal introduction, conducting messages, and possibly physical gestures. In contrast, the goal of audience application is not to help a user develop virtuosity in this one-time playing instrument but to help them enjoy the participation in the performance. Therefore, audience members need not have a full freedom of expressivity like a traditional instrument in this disposable instrument that will be used for less than half an hour and probably never used again. I wished that the interface is accessible and playable just enough to make audience interested during the performance period. Therefore, wide range of expressivity and room for virtuosity are not provided, which might possibly make music out of control and produce low participation.

For the audience musician interface, a keyboard-like interface is selected (Figure 14). It was natural choice because the keyboard-like instruments (e.g. piano, kalimba) have a better learning curves at initial stage than other types of instruments[71], considering the majority of audience need to play this instrument within 10 minutes from the time they encounter it. I believe that the current design has a limited but adequate musical space for audience to participate consistently during the piece without losing interest and being overwhelmed.

Figure 14 Keys on audience's interface. For the upper case, the C major chord is selected and eight keys are one octave of C major scale. The lower case is F minor 7.

As described in the beginning of this chapter, audience can join a live room through the room-creation/joining view. Once he or she joins a room, if the master musician selects any chord, the key interface will appear with a transition animation (Figure 14). It always has eight keys and the leftmost and the rightmost notes on the screen will be the root-note pitch class of the selected chord with one octave difference. Eight keys will be either in major or minor scale depending on which type of chord is selected. Each key has its own color based on its pitch class and the same color mapping is used to set the color of a virtual block on the master's interface. If the master musician changes a chord, keys on audience's interface will be replaced with a new set of corresponding keys with fade-in transition animation. Other than keys, there is a grey horizontal playhead, which is synchronized with master's phone and the rest of audience, the name of the room on the bottom left, the exit button on the bottom right and the broadcast button on the top right. The broadcasting feature will be discussed later in this chapter.

The initial design had only four keys in the interface, restricting the keys strictly to the notes that belong to the chord that a master musician selected. For example, when a master musician selected C Major, the four keys provided would be C5, E5, G5, and C6. Hence audience could play any key on screen, which would be one of the four notes in the selected chord. Through a preliminary evaluation and discussion, the design was changed to have an octave of a scale so they can play a melody within the scale to give more musical space that audience can explore while being still limited compared to a general instrument. However, incorporating the pitch class other than chord notes has the risk of sound being dissonant. To resolve this issue, a key of the pitch class of the selected chord is marked with a black arrow and audience is informed about the meaning of the arrows and instructed to play keys with arrows more than others so that they can make their own decision to play the instrument understanding the associated risks.

There are a few more features that *echobo* supports for expressivity. Duration of note is determined based on how long the user holds a finger pressed on a key; it works just like an electric piano key without dynamics so that the sound sustains while it is pressed down and slowly fades out. Or if the user lifts his/her finger right after he or she presses it, it sounds like a *staccato* in electric piano. In addition, *tremolo* is implemented by applying a modulating envelope. If one moves a finger inside the key while it's pressed down, *tremolo* will be applied to the sound. Another natural feature I could consider was to add dynamics to the instrument. However, as the limited volume of mobile phone was one of key constraints in using a mobile phone as a sound source, I chose to apply same level for each key. Audience members at the performance are instructed to maximize volume.

Finally, as mentioned earlier, the master musician can broadcast a textual message to shape the sound coming from audience. Once the master musician presses the antenna button on the bottom right, the message list registered during configuration(Figure 10) will reappear(Figure 15 top). If the master musician clicks a message, the message will slowly flash both in the master's screen and at the bottom of the audience's screen(Figure 15 bottom).

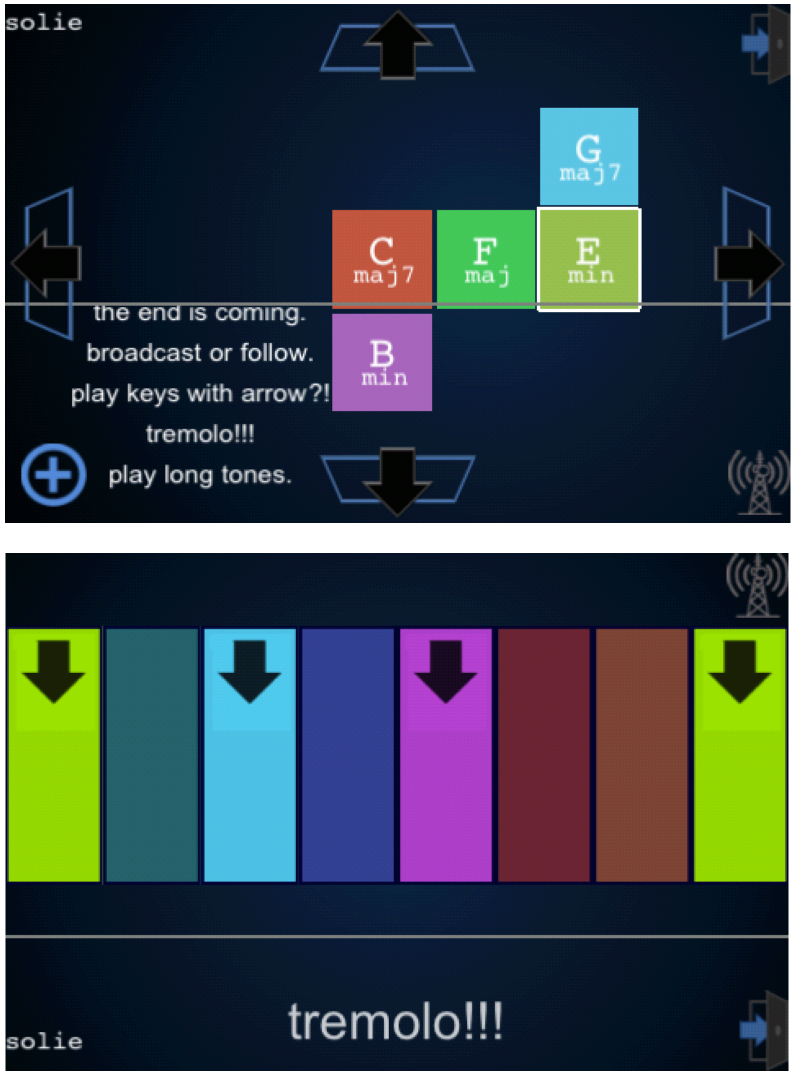

Figure 15 Broadcasting conducting message by a master musician(upper) and flashing message on audience's side(lower)

Broadcasting a Musical Pattern in Audience Interface

As one of the goals is to achieve large-scale musicianship for non-musicians, it is required to build the instrument as accessible as possible. From a musical perspective, even though the sound generated can be chaotic in terms of its rhythmic nature, all the notes (with the tendency of audience playing keys with arrows) are in harmony. Therefore the aggregated result of the instrument is somewhat predictable and audience performance will produce desirable accompaniment – harmonious tones, unless audience does nothing. However, from usability perspective, the interface of audience-side application is very limited compared to traditional instruments so that audience may soon find the interactions monotonous and boring to keep playing. During the preliminary experiment conducted in a classroom setup, participants expressed concerns of the limited interaction on a mobile music app, especially with constraints, towards the end of the music.

In general, expanding the degree of freedom is a solution to this problem. However it is not applicable to this particular scenario. First, providing a general music instrument close to a piano or violin naturally requires users to spend time and effort to be musically virtuosic. Not only it will be difficult to do so within the time window of the performance length, but also a majority of people is not willing to do so. Furthermore, offering a high ceiling for virtuosity even might discourage participants with too many features that they can never understand, making them feel lost, and lead to low participation. In the worst case, bored musicians might stop playing the instruments and music will come to an abrupt halt. Therefore, a different approach is required to keep participants motivated and to design the interaction interesting while keeping simplicity of the interface. More specifically, to sustain active participation during the piece, it seems that the instrument should be easy to play and, at the same time, entertaining to play beyond just a musical reason.

As I already stated that I wish to follow the paradigm of singing audience, one good solution can be easily found from the previous examples; *We Will Rock You*, McFerrin's Improvisation and *Moths*. One common factor we can find in these three examples, or most of audience participation in popular context, is synchronization. The piece, *We Will Rock You*, is all about synchronization. Audience's rhythmic gestures are synchronized with the music and they feel engaged when all audience sing and clap in one united voice. McFerrin's example includes being synchronized with his hand gestures, which brings tension of being ready to follow his signs and offers release when all audience sing in one voice, following his signs. *Moths* also has a synchronization aspect. They are engaged with the performance not only by being expressive based on interpretation but also by listening to the sonic result of whole audience as synchronized with progression through the the music score. I believe synchronization in audience participation is an effective way for both musicians and audience to feel connected with each other.

Even though *echobo* provides a musical space for each individual to claim the ownership of a musical instrument, I wished to provide a ground in which audience can communicate musically and be synchronized with others. To attain the purpose, broadcasting function is implemented in the audience side of *echobo* for an audience member to share musical patterns with the other audience members. Whenever an audience musician creates an interesting pattern that he or she wants to share with other people, they can press the antenna button on the top right of the screen. Once it's pressed, eight empty squares appear right below eight keys and the countdown is shown on the top left corner (Figure 16 Top). Once the countdown changes into red "REC" sign(Figure 16 Bottom), whatever the player plays with the key interface will be recorded as a pattern for one cycle. The length of one cycle will be a measure length and it will be visualized by the grey playhead moving from top to bottom.

Figure 16 Recording a pattern. After countdown starts (top), there appears REC sign on the top left (bottom), anything you play with key interface will be recorded.

Once it is recorded, the pattern is visualized as a series of falling squares (notes) from the top of the screen over the key interface at the next cycle (Figure 17). This visualized score is displayed not only in the application of the user who created it but also in the applications of two random audience musicians in synchronized manner. Any of those three can play that pattern similar to a rhythm following game such as guitar-hero. More specifically, once each of the falling notes approaches close to one of the empty squares on the bottom, the note is activated (color of the note changes to white) and the player can play the note by pressing the white solid square inside the empty square. It makes an electric piano tone of the same key above so that the pitch contour and the rhythm of original recorded pattern are kept. Therefore at most three audience musicians, only if they want to, can play one musical pattern at the same time for one cycle. If they continue to follow the pattern, the falling notes will keep coming. Once they stop playing the notes coming down and resume playing on the key interface, the pattern will disappear at the next cycle. The interface for following pattern is different from the main key interface so that people will be able to separate two functions and play key interface and pattern at the same time. Also people who are familiar with the rhythmic following games will easily pick up the concept of using the interface.

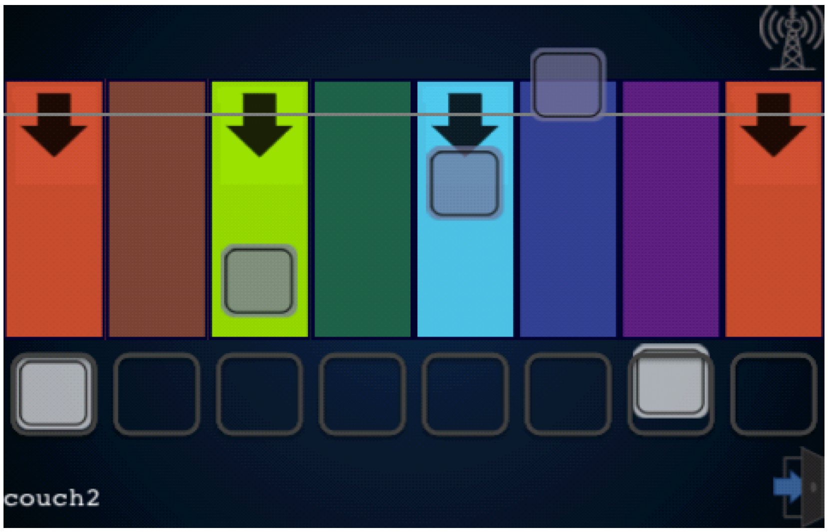

Figure 17 A broadcasted pattern (falling white squares)

In addition, broadcasting feature is designed to produce a viral action among audience. As mentioned, if an audience musician broadcasts a pattern, a visualized score will be displayed to oneself and two other people from the audience. If anyone of three people follows the pattern, the visual pattern will be broadcasted to one more audience member so that there can be at most three more people who receive the displayed notes at the next cycle. As one pattern can be spread virally, theoretically, it is possible that the whole audience plays one pattern synchronously, which will make a significant difference in music compared to the situation where everyone improvises freely.

When a master musician switches the chord, while the broadcasted pattern is followed, the pattern will be transposed with a new root note and new scale so that the pitch contour of the pattern is kept and the followed pattern is in harmony with current chord selected.

The highlight of broadcasting feature is that it gives two different types of motivation beyond improvisation on the key instruments, which may be soon boring with the musically limited space. The one who creates (broadcasts) has a motivation because, in that way, he or she can listen to someone else playing his/her pattern in synchronized manner. Furthermore, with its infectious nature, one broadcasted pattern may make a huge difference in music if a significant number of people follow it. Therefore one has the motivation to develop a pattern that is easy and interesting at the same time for the rest of audience to follow. On the other hand, those who receive and follow a pattern also have the pleasure of music making by shadowing someone else's performance through the pattern, feeling a sense of belonging and contributing in making difference to the music. This social dynamics of being a leader(pattern creation) or a follower will have rewards in both directions and help audience musically synchronized with each other.

Implementation

Echobo is built on the iOS platform, coded in Objective-C and only available in Apple iOS devices such as iPod Touch, iPhone and iPad. The whole visual interface is coded upon COCOS 2D API, which is the open source API for developing iOS games[72]. The API is efficient to build an interactive visualization with objected oriented programming. For sound generation, MoMu STK API[28] is deployed for playback of wave files, for envelope generation depending on note on and off, and for tremolo modulation.

To make the master musician and audience musicians communicate, server scripts are coded in PHP and MySQL. MySQL database is used mainly for room creation and room search for audience. Other than that, all data transfer is in JSON format using HTTP/POST method and the transmitted data are stored in server files generated for the specific room. Instead of storing data in a database, file write/read is used because of the small amount of data and frequent monitoring required by audience application. Each created room will have one main file that contains information such as status of the room, current chord selected, timestamp for synchronization, the number of people in the room, and broadcasted textual message from a master musician.

Every mobile phone is synchronized with the clock of the server so that the grey playhead in application moves in sync. A master mobile phone pings the server every few hundred milliseconds and updates latency and time difference recursively. It also submits chord progression change and the timestamp of next cycle. Audience's mobile phones then retrieve the information to change the keys and to synchronize clock with the timestamp for playhead movement. As the playhead is synchronized with master musician and other audience, audience will have same clock during the performance.

Broadcasting pattern by audience members also is stored in a set of files. Once a pattern is recorded, it creates a file named with a unique pattern index and store all the keys pressed in sequence and their timestamp. To reproduce the visual notes, it downloads detailed data of the pattern from the corresponding file. The server also maintains a finite length queue (twice the number of audience) for each room in a file. Whenever an audience musician creates a pattern, it adds two occurrences of pattern index in the end of the queue. Whenever an audience musician successfully follows a pattern, it adds one occurrence of that pattern in the end of the queue of the room. The audience side of *echobo* will consistently monitor the queue of the room and withdraw a pattern from the queue if there is any.

CHAPTER 4

ECHOBO PERFORMANCE

Preliminary Tests and Performances

Before organizing a performance to evaluate the application, I have conducted four experiments at different stages of development. Initially, before I started to develop *echobo*, I had a chance to present the idea and to experiment the simulated performance in a classroom setup, with 15 to 20 university-wide graduate students. Hence, with various music backgrounds, participants were close to target audience. At that time, I used the existing iOS music application called *Soundprism*[73]. The app has a similarity with *echobo* in a way that user can play the synth-pad sound of a chord instead of a note. Since *Soundprism* has no networked constraint like *echobo* would have, I prepared slides that indicate which chord the audience should play at a specific moment. I played a MIDI keyboard and progressed the slides, simulating a master musician and an acoustic instrument player at the same time.

Overall, it seemed that audience enjoyed the experiment. Based on the music that was generated, they seemed to be very good at following the visual score and my page-turning gesture, regardless of prior knowledge about music. The most significant issue was the limited volume of mobile phones. With the constraint of playing a chord, it is observed that people tend to choose chord tones in higher pitch range. It is well known that humans perceive high-pitch sound louder than low-pitch sound in the audible range[74]. I found that it would not be very effective to support a wide pitch range, especially the lower part, considering increased complexity (i.e. scrolling keys) of interface on limited size of the screen. A few more experiments and discussion with colleagues in Music Technology Program using *echobo* helped me push the range of expressivity from four binary interactions to the current design, one octave of scale, note duration and tremolo. Those experiments also gave me concrete ideas on how to lead the pre-performance session to describe the instrument.

In one of the experiments near the performance date, I had a chance to rehearse the piece and play *echobo* as a master musician, sitting right next to an acoustic instrument player. During the experiment, I was very disappointed with the balance between an acoustic instrument and the collective volume of mobile phones since I could barely hear the audience's accompaniment. Nonetheless, participants seemed surprisingly fine with the limited volume compared to my frustration. While it seemed that the overall volume of the audience performance is very soft, I realized each individual would experience a personalized and spatialized experience depending on what he or she plays and what the neighbors play. In addition, the acoustic instrument player was asked to play the instrument soft comparable to mobile phones level at the actual performance.

Instrumentation, Music Composition and Rehearsal

As mentioned earlier, the target performance setting includes a master musician, audience and one acoustic instrument player. Ideally, to demonstrate and to evaluate, it was desirable to have an acoustic instrument player who can improvise given chord progression accompaniment from audience. Improvisation of an acoustic instrument would create a collaborative setup where the acoustic instrument player can respond to audience if the musician hears a group behavior from the audience with a broadcasted pattern. The purpose was to help audience feel connected to the music and the acoustic instrument player by building a call and response relationship.

For the instrumentation, I believe it is beneficial to have a contrast between the sound that audience generates and the sound of acoustic instrument, for a clear distinction in music. The sound of *echobo*, as a whole, is polyphonic, discrete, limited in expressivity, and soft in low pitch register. Hence a monophonic, expressive and relatively low-pitched instrument would be desirable, for example, saxophone, bass clarinet, or cello. Unfortunately, I could not find an appropriate musician with improvisation background who was willing to collaborate with the piece within the time period. Instead, I collaborated with one of my colleagues, Yonatan Sasson, who is a clarinet player in the university Orchestra.

The most challenging part for this project and most of audience participation in general is to write music without a chance to play music with audience until the moment of performance. It is more challenging that I could never listen to the actual sonic result of the piece until the performance. The composition for audience will just be a chord progression structure while low-level information about music like rhythm, notes and density of sound is completely left to audiences' hand. Personally, I have limited music composition background and hence I had to pick an existing music composition for evaluation.

I chose one instrumental music piece from an album of the independent electronic ensemble named *byul*[75], and arranged the piece for a clarinet. Applying the same chord progression, the electronic piece became a 6 minute long music piece for solo clarinet. The music is in a constant tempo and has a simple musical form (A-B-A-B) of repetitive chord progression so that audience can anticipate the musical structure. Additionally, 18 measures were added, which were not in the original piece, especially for audience to broadcast or follow patterns so that a group behavior can emerge from the piece. The tempo of the music is set to 54 bpm, which is slower than the original piece.

With the prior experience in *crossole*, I performed as the master musician in the performance conducted for this study. As the master musician, practicing the performance comprises of three different processes; moving the cursor based on the composition, exploring various textures of sound using simulator for textual messages and collaborating with the clarinet player. For the first part, there is a slight network delay between the transition of keys and the cursor move; the goal of practice is to move cursor accurately so that audience interface transition can be accurate in music. I practiced with multiple phones, one for master's interface and others to observe the key transition timing. Although I could not hear any sound from those mobile phones due to absence of audience musicians, I could observe transition in audience's interface and get used to the network delay.

The second part was to explore different textures of sound. To simulate sonic result of true performance at the initial stage of development, the simulator of audience participation, which generates sound, was developed in Max/MSP separately from the application itself. The simulator was used in the composition process as well; to listen to the texture, to try different sounds, to listen to the piled end result and eventually to make appropriate choices. Using the simulator, a computer can simulate large-scale audience with two parameters; note density for each individual, and the number of audience members. The note density without particular rhythm for each individual can be easily implemented with random number generation using an exponential distribution, which well describes random time interval between two consecutive events with a parameter, average interval. Once I get the sense of texture from simulator, I used the simulator to try out different textures so that I would be able to coordinate audience's collaboration via broadcasting textual messages whenever I wanted denser texture or sparser texture.

Final process was to rehearse with the instrument player. To help the clarinet player practice music without audience, I let him play with the recorded track of background chord progression. At the same time, I could also rehearse cursor moves on the block structure in the master's side of *echobo* with the clarinet solo.

Quasi Rehearsal with Audience Prior to Performance

Regardless of the assumption that an eight key interface is easy and transparent to play even for a novice user, audience in public will have a fear of playing the instrument they have not used before for a music performance. All three qualities – accessibility, musical security, and initiation - are required for them to participate without hesitation. While those three qualities are taken into account in developing the application, it again needs to be well delivered to audience so that they can easily pick up the instrument and participate.

One of most common methods is to have a verbal introduction and explanation prior to the performance[5, 45, 76]. From the evaluation of previous works without an explanation, audience needed to be introduced to understand the participation system prior to the performance[3, 50]. Reinforcing explanatory procedure by having a verbal introduction and a short participation time will give audience a chance to learn the application along with musical and social interaction of the piece. At both performances, a verbal introduction was given to the audience and they had a chance to play the instrument prior to the performance.

The before-performance session was composed of i) configuration ii) introducing a relationship between the master musician and audience interface iii) explaining musical expressivity of the instrument iv) introducing broadcasting feature v) questions and answers vi) rules of interaction; I stated two basic rules for them to follow - a) Play the instrument as you like whenever there is a key interface and b) Play keys with arrows more than the others. During the verbal introduction, audience had a chance to try the instrument and use the features that had been introduced. This procedure is important to assure audience that playing the instrument in a group does not harm the music. At the last step, I also emphasized that they don't really have to use all the features implemented if they are not confident about it. Revisiting the basic rules and encouraging the simplified interaction was to reinforce their initiating procedure, to have them in a comfort zone with their participations and to accommodate audience at different levels of musical backgrounds in the piece. Even though some people may think some features are difficult to use, it is natural that people will want to use more features, as they get familiar with the instrument.

Right before we started the performance, I asked them to be professional. It is to distinguish the introductory procedure from the actual performance and to help them take on a serious attitude towards the performance. Eventually, I wished the performance and the rehearsal process would give them a sense of musicianship and make them feel proud of their performance, which will lead to satisfaction from their performance.

CHAPTER 5

EVALUATION

For evaluation, two performances were conducted in two graduate classes at Georgia Tech on March 7^{th}(Figure 18) and March 13^{th}. I wished to evaluate the application rather with general audience than colleagues inside the music technology department. There were a total of 37 participants from both classes. About half of the participants either had no experience (12/37) or are at beginner level (6/37) in playing a musical instrument. I visited both classes a week before actual performance in order to know how many iOS devices are available. Although small portion of people had iPhone, more than half of the class had iOS devices, including iPod Touch and iPad. For those who did not have any iOS devices, iPod Touch(the 2^{nd} generation) were provided, although the maximum volume of the old IPod device is softer than the volume of iPhone, iPad and newer generation of iPod Touch.

Figure 18 Performance at a classroom on Mar. 7th.

There were no technical glitches in both performances and *echobo* worked reliably with the small scale. As I performed as the master musician, I could not really have the same listening experience that each audience would have. From the position where I stood, the accompaniment by audience was soft as anticipated. To listen to more realistic result, both performances were videotaped and one of the performances was recorded from the audience's angle.

In addition, a web-based survey was conducted for evaluation purpose. All participants filled an anonymous survey. The survey comprised 20 questions. The survey contained questions to collect the musical backgrounds of participants, subjective rating questions to evaluate the performance quantitatively, a few questions regarding broadcasting feature and one open-ended suggestion/comment question. The subjective rating questions use 5 level Likert-Scale (5 – Strongly Agree, 4- Agree, 3 – Neutral, 2- Disagree, 1- Strongly Disagree). The survey result is listed in the Table 4.

Table 4 Survey questions and interpolated media of for each question (5 – Strongly agree, 4- Agree, 3- Neutral, 2- disagree, 1- Strongly disagree)

Question	Rating
The mobile application used in the performance was easy to use.	4.55
During the performance, I felt connected to the music.	3.59
During the performance, I felt connected to the master musician (the one who controlled a mobile phone on stage).	3.45
During the performance, I collaborated with other audience members.	3.35
During the performance, I collaborated with the acoustic instrument player.	3.80
Overall performance was musically satisfying.	3.98
The performance would have been different without me.	2.96
The description/rehearsal before the performance was effective in helping me understand how the technology works.	4.28
I can perform better next time.	4.05
I could have been more musically expressive but I felt limited by the application.	2.77
Towards the end, the participation got boring.	2.29

The networked instrument seemed most successful in terms of accessibility. The survey question that asks if the application was easy to use, most of the audience felt the application was accessible (interpolated median 4.55, where 5 indicates strongest agreement). Audience also felt that the rehearsal and introduction process before the performance was effective to understand how the technology works (4.28). Audience felt they could do better next time (4.05), which also proves their confidence in skill development with the instrument. Overall, the application developed was accessible and the audience felt the interaction was transparent.

In addition, audience seemed to enjoy the participation. According to the observation of the video, most of the audience actively participated in the performance. There were some unexpected moments for audience, especially when the key interface disappears in the middle of performance, which is due to the intentional silence in the composition. Survey results also show their positive remarks on the performance in the question if the performance was musically satisfying (3.98). There were positive responses reported for the last open-ended question asking for comments and suggestions. Here are a few examples:

“Amazing application! Kudos!(...) Wish we could do it again!”

“good job!”

“Overall the performance was a fun experience.”

“It's fun, thank you. Would like to play more, in different places. It could be good improv at public space or restaurant, etc.”

“It was a really good application”

“Great App..”

“I love the application. thank you.”

At the second performance on Mar. 13th, after the performance and the participants' survey, audience actually wanted to perform one more time and we did. Most of the participants said the 2nd performance was better than the first one.

As the main goal of this research, I wanted audience to feel connected to the music, the musicians and the other audience members, which I believe is the most effective way to accomplish audience engagement. The survey included a question asking if they felt connected to the music and some people agreed that they were connected (3.59). More than half of participants agreed (20/37) while the rest of them was neutral (10/37) or showed disagreement (7/37).

To investigate their connection to the music in detail, participants were asked if they felt connected to the master musician and collaborated with the acoustic instrument player and other audience members. Interestingly, connection with acoustic instrument player (3.8) was stronger than collaboration with master musician (3.45) and the rest of audience members (3.35). This was not an intuitive result considering two networked elements in the application; i) a constraint between the master musician and audience musicians and ii) viral pattern broadcasting among audience. In contrast, an acoustic player is connected to audience only through sound.

The reason that audience expressed hesitation on the connection with the master musician was hinted from a post-performance discussion. One of the participants said that he felt the master musician took away the key interface from him whenever key interface makes transition to a new chord. A written response from the survey result well express a direct concern on the necessary constraint of dependency between the master musician and audience:

> *"The master musician's controls of the key made me feel limited. However I realize that this is most likely necessary."*

Also from the video recording, it was observed that few people watched the projection screen of master interface. As audience watching the master interface will help anticipate key-interface transition, the pre-performance session should have included an explicit instruction for the audience to watch the screen during the performance, which would improve collaborative feeling of audience with the master musician. Understanding that the cognitive load of audience to improvise and focus on the projection is high, an improvement on audience application is left for the future work to find a way to share the state of the master musician in audience's interfaces.

Through the survey, it seems that connection within audience was not strong enough (3.35). It means that broadcasting feature was not as effective as expected to facilitate collaboration among audience. Log data shows how actively audience members used the broadcasting features. At the second performance on Mar. 13th, during the six minute long piece, 61 patterns were created and broadcasted to audience members and 23 patterns out of 61 patterns were followed more than once. One participant's pattern was followed 63 times, which would have made significant number of people play the pattern. The most followed pattern was relatively easy to follow; it was composed of five notes with three distinct pitches; the root note, the major/minor 3rd and the perfect 5th.

About two thirds of participants (25/37) replied that they broadcasted pattern at least once. However only 8 of them could recognize someone else playing the broadcasted pattern. On the other hand, most of people (36/37) actively followed broadcasted patterns at least once but a subset of people (14/37) could recognize someone else playing the same pattern. The survey result indicates that it is hard to listen to the response of other audience members. Considering the limited volume of mobile phone, a group behavior of playing a broadcasted pattern was not prominent enough. Similarly, from recording of performances, group play of broadcasted pattern was not audible enough.

One of the reasons why the broadcasting feature was not effective is the lack of proximity in broadcasting pattern. Currently, anyone can broadcast a pattern and the pattern appears in two randomly chosen audience members. Ideally, it will be natural to broadcast pattern to people right next to the creator so that the creator and the followers can listen to the synchronized pattern being played. It was technically challenging to measure the proximity between the mobile phones without GPS data, which is not supported indoors. I believe other approaches are possible to involve the physical closeness in broadcasting. One approach is to let people electronically transmit pattern by targeted phones gestures rather than broadcasting to random two people. Similar phone gesture has been applied in an iOS app, *Sonic Lighter*, to light a neighboring iPhone's virtual candle inside the same app[77].

One commonly suggested improvement for broadcasting feature was that it would be better if the one who broadcasted pattern knew how many people were actually following his/her pattern. Following are the responses:

> *"Application needs some feedback in terms of whether users followed or broadcasted a pattern. Think guitar hero."*
>
> *"Would like to see visualization of what everyone was paying*[sic]*, also be able to identify the broadcasting patterns and who all have followed them."*

Originally, I hoped the one who broadcasted would be rewarded by the sonic result if the large-group behavior emerged. People would have liked having credits on their broadcasted patterns through screen, when many people followed the pattern. However, being anonymous would also help audience initiate the participation in some sense. As I understand this can be another research topic in audience interaction, the topic is left for future work.

Another reason why broadcasting was not effective was because the verbal introduction of broadcasting feature was not sufficient, being not clear on relation among the broadcasting button, the horizontal playhead, and the fixed length of recording pattern. A few participants expressed confusion when they recorded pattern to broadcast because they did not understand the length of recording was fixed to a measure. They rather believed that the key change by master musician interrupted recording patterns or the pattern broadcasted is of no use due to the key change. Here are some responses about it:

> *"The broadcasted patterns reset as soon as the scale was changed. It was kind of frustrating to not finish a pattern."*
>
> *"it seemed to stop recording sometimes when the key changed."*
>
> *"The chord progression was moving too fast. By the time I finished broadcasting, the key already changed, making my pattern less relevant."*

The responses above are not true; Key change does not stop pattern recording. With chord transition every measure from the composition, audience felt that recording is terminated by the chord change. Also, since any broadcasted pattern is recorded in relative interval in the current scale, the pattern recording/playback will be transposed if the master musician switches the chord. The verbal introduction over broadcasting function should be improved. Even though it resolves all misunderstanding regarding broadcasting, participants may felt limited to broadcast with the fixed length of pattern recording.

The two videos recorded from a distance also showed that the volume of audience's mobile phone was soft compared to the volume of clarinet. Even though the limited volume of mobile phones was anticipated, it seems to be less of a problem for the audience. A video and an audio recording from one audience member's seat give an insight on the issue. From the recording, it was clear that the level of an audience member was equivalent to the sound of clarinet. Also, the sound from other audience sounded like background accompaniment. The proximity of the instrument made the individual performance close to the performance of the clarinet player for each audience member. It is also evident from the survey result which suggests that audience felt most connected to the instrument player than the master musician and other audience members. The audience engagement here was fundamentally different from the one of the synchronized singing cases; each audience performs as an individual musician rather than as a member of choir.

Another goal for the application was to sustain interest in audience with limited musicality. The solution was to provide musical variability in the application and social interaction among audience. There was no noticeable decrease in level of participation based on the video recordings. Also, from the survey result, a small portion of participants (5/37) agreed to the statement that the performance got boring towards the end (2.29). Most of audience did not agree (15/37) or expressed neutrality (13/37) on the proposition that the application limited the expressivity of their performance (2.77).

At the two performances on Mar.7th and Mar. 13th, it took significant time just for verbal introduction compared to its length of music; 18 minutes and 15 minutes, respectively. Even though, the rehearsal and verbal introduction actually helped audience understand and initiate their participation, the time spent was not realistic for actual performance. The verbal introduction and configuration needs to be drastically simplified.

Recently, there was one more performance of the same music in a larger-scale. *Echobo* was presented in “Listening Machines” Concert 2012, the Music Tech performance that presents students’ own works every year (Figure 19). The performance log shows that total number of people joined the room was 105. Considering the number of people who joined the room more than once by accidentally pressing the exit button, I believe there were about 70 participants, which gives us much more realistic number for a targeted setting. An instruction for application configuration was prepared on projection prior to concert so that they could join the room before the verbal introduction began. With lessons learnt from two prior performances, the verbal introduction was more organized and simplified; i) explaining the basic relationship between the master musician and the audience musicians, ii) giving them two simple rules – [Whenever you see keys on screen, play the instrument.] and [Play keys with arrow more than the others.] – iii) simple introduction on broadcasting feature and iv) asking them to be professional. The verbal introduction took 6 minutes.

Figure 19 echobo performance at Listening Machines 2012 concert

The performance was successful under the large-scale environment. There were no technology issues. Since more people participated, the level between clarinet and echobo was more balanced than two preceding performances. A wide range of people participated in the piece, from children to the aged and from musicians to novices. One of my colleagues observed that even a baby participated in the piece actively. Although a formal evaluation was not conducted for this particular performance, there was a lot of positive feedback after the concert from the audience; both participants and non-participants, and both novices and musicians.

All in all, for the first hypothesis of accomplishing accessibility and audience engagement by applying multi-user instrument concept, *echobo* successfully builds a reliable environment for audience participation with accessibility and transparency, incorporating the notion of a collaborative instrument. On the contrary, for the second hypothesis of achieving collaboration though social interaction among audience members, it is partially achieved; the system includes a broadcasting feature that facilitates social interaction in the application and each audience can have a greater influence on the music. However, the strong sense of collaboration among audience was observed neither through the survey result nor the music, which leaves a space for improvement. Rather, it seems that *echobo* motivates audience to perform as an individual musician in improvisation settings..

Personally, as a listener, I could only guess that the end result was good from audience feedback since I could not experience audience's role and listen from audience's seating. However, I was very happy with the result as a performer because I had a strong connection to audience during the performance for their completing the music.

CHAPTER 6

CONCLUSION, IMPROVEMENTS AND FUTURE WORK

The *echobo* performance is designed to accomplish four characteristics of audience participation in music performance; i) individual interaction ii) synchronous participation, iii) audience as a performer and iv) sound coming from audience. This research aims to fulfill five desirable qualities of audience participation- accessibility, attraction, musical security, initiation, and transparency. In conclusion, a series of *echobo* performances were successfully presented. First of all, the application was very reliable given the targeted setup. It seems the delay between the master musician's gesture and state change in audience application is easy to handle with some practice on the master application. No network traffic issues were found for the targeted number of audience.

Secondly, audience found the participation very easy to use given the limited time. Wide range of audience was able to actively participate in the piece and contributed to the music by playing the instrument. The mapping between audience interaction and generated sound is as simple as a traditional key instrument. Compared to previous audience participation works, especially the ones with technology, the *echobo* performance allows participants to act as an individual musician and the sonic result gives them a personalized musical experience with transparent individual output. In addition, the prior performance process (verbal introduction, demonstration and rehearsal) worked effectively to make participants comfortable with the participation.

Third, provided that there are large-scale participants, with a proper verbal introduction, the musical output is within the expectation of the composer. Although there exists a certain limit in musicality of the performance that can be pursued with *echobo*, audience can contribute to music by playing an accompaniment of chord progression in slow-tempo or loose tempo music. The end result is musically satisfying for me with the application, the music and the performance.

Lastly, the application with constraints, given the length of the composed music, successfully sustained interest of participants by offering a limited space for musical expressivity and facilitating social interaction among audience. Throughout the survey and informal conversation, audience had a novel musical experience and would like to participate in it again. Broadcasting feature motivated participants to be synchronized with others, although it is questionable if they actually felt so. The success is also influenced by the accessible composition as well; the music gives audience a predictable musical structure and reasonable variation for climax at the same time, which gives them a balanced experience as a musician.

There remains a definite room for improvement in the application. First of all, while the application is easy to use for novice users, the opportunity of participation is still limited because the application is supported only in iOS devices, as mentioned earlier. While supporting more platforms such as Android or Microsoft and with more people using smart phone, the constraint can be alleviated; yet the nature of audience participation will be a selective procedure depending on ownership of a smart phone, which may be relevant to one's socioeconomic status. While completely opening participation opportunity for any class of audience will be impossible regardless of technology type, it is desirable to accommodate more people in the performance. The immediate action is to provide application in multiple platforms.

Another common complaint about the application was the soft level of sound coming from the mobile phone. In addition, it is observed and reported that the sound was clipping or distorted when the volume is maximized or when the user played very densely. The study lacks a thorough research on the most effective (audible) sound type and its synthesis on the platform given the limited capacity of internal speaker of handheld devices.

As mentioned earlier, the connection between master musician and audience musicians was not strong enough. Rather it seemed that audience felt that the master musician limited their musical space by switching keys too often. Projecting visualization was expected to build a bond between two groups but the projection was not effectively utilized in any of the performances. The verbal instruction to watch projection was not clear on the first two performances and the projection was not close enough for audience to read at the Listening Machines concert. In general, it seemed that audience focused solely on their own interface rather than watching master's interface. I believe it will benefit the connection to share master musician's state on audience interface with some visual cues; sharing complete block structure on the background is an easy way. Also there is an alternative to make audience actually affect cursor traversal of the chord progression with some voting mechanism so they can actively participate in composition and performance at the same time.

Another problem found was the difficulty for audience to configure the application. It required some time for audience to figure out how to join a room. Even though it is not required to list up all open rooms available since the master musician and audience musicians will be co-located and they can communicate with each other in this particular setting, it will be helpful to have a list of open rooms near the location of audience so that audience can easily join the room. Furthermore, the application can make audience join the room automatically depending on the proximity.

Finally, as mentioned earlier in the previous chapter, broadcasting feature can be improved and extended in multiple ways. First of all, the broadcasting interface is not intuitive enough and it requires lengthy explanation to make participants understand the feature. Second, it would be better if the connection between a leader and followers is more transparent either with the proximity of sound followed or the aid of visualization. Third, seeking a way to motivate a leader beyond just sound response from other audience members will benefit the level of interaction, such as giving a credit(username) on the projection or providing number of current followers. Lastly, the broadcasting feature could have been more flexible in terms of its length so that advanced users may like to broadcast a longer pattern.

For the future work, I am hoping to extend the application for more general musical styles within the audience participation context in terms of its mapping and sound. While keeping the key concept - master musician restricting the musical expressivity of each audience member - I would like to make sound synthesis and mapping parts of the application programmable by a musician and push the application closer to the concept of a meta-instrument, with which a user can personalize the musical instrument. For example, the current sound generation is composed of eight keys, which are mapped to electric piano sound. However, the sound generation can be programmable by a developer to support various musical styles. The whole sound generation can be one patch that can be downloaded from web space right before the performance so that composer can provide different sound sets for audience in the specific composition. Using existing audio API for iOS devices, such as libPD[78], the whole sound generation module can be detached from the main application implementation.

Likewise, the current two-dimensional space of chord selection (root note selection and chord type selection) can be replaced by a completely different mapping. For instance, one dimension can be mapped to each section of music (A-B-C-A') and the other dimension could be mapped to instrumentation (Piano – Synth – Sample). The control space can be customized by a composer or a master musician and can also be retrieved from the web for the specific master musician in a simple XML/JSON format. Therefore, customization on mapping should involve a new sound generation patch due to the new mapping. All those customizations shall be open to the public so that users in general can enjoy the diversity of musicality contributed by developers.

In conclusion, the research delivered an accessible audience participation environment, utilizing the research in mobile music, audience participation and networked instrument. With improvements and extensions, I expect the system can be accessible to more musicians. I hope *echobo* can provide a novel example of audience participation to computer music researchers and offer an interactive tool in the arsenal of musicians for audience engagement at a music performance.

REFERENCES

[1] L. Dahl, J. Herrera, and C. Wilkerson, "TweetDreams: Making music with the audience and the world using real-time Twitter data," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Oslo, Norway, 2011.

[2] J. Freeman, "Glimmer: Creating new connections," *Transdisciplinary Digital Art. Sound, Vision and the New Screen,* pp. 270-283, 2008.

[3] J. Freeman and M. Godfrey, "Creative collaboration between audiences and musicians in Flock," *Digital Creativity,* vol. 21, pp. 85-99, 2010.

[4] J. Hasse, "Moths," in *Visible Music, Euclid, OH*, ed, 1986.

[5] G. McAllister, M. Alcorn, and P. Strain, "Interactive performance with wireless PDAs," in *the International Computer Music Conference*, San Francisco, USA, 2004, pp. 702-705.

[6] G. Levin, "Dialtone: a Telesymphony premiere at Ars Electronica Lynz Austria," ed, 2001.

[7] G. Geiger, "PDa: Real time signal processing and sound generation on handheld devices," in *the International Computer Music Conference*, Singapore, 2003.

[8] G. Essl and M. Rohs, "Mobile STK for Symbian OS," in *the International Computer Music Conference*, New Orleans, USA, 2006, pp. 278-281.

[9] G. Schiemer and M. Havryliv, "Pocket Gamelan: Tuneable trajectories for flying sources in Mandala 3 and Mandala 4," in *the International Conferences on New Interfaces for Musical Expression (NIME)*, Paris, France, 2006, pp. 37-42.

[10] A. Tanaka, "Malleable mobile music," in *the International Conference on Ubiquitous Computing(UbiComp)*, Tokyo, Japan, 2004.

[11] G. Geiger, "Using the touch screen as a controller for portable computer music instruments," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Paris, France, 2006, pp. 61-64.

[12] G. Essl and M. Rohs, "Shamus-a sensor based integrated mobile phone instrument," in *the International Computer Music Conference*, Copenhagen, Denmark, 2007, p. 50.

[13] M. Rohs, G. Essl, and M. Roth, "CaMus: live music performance using camera phones and visual grid tracking," in *the International Conference on New Instruments for Musical Expression (NIME)*, Paris, France, 2006, pp. 31-36.

[14] M. Shepard, "Tactical Sound Garden [TSG] Toolkit," in *the 3rd International Workshop on Mobile Music Technology*, Brighton, UK, 2006.

[15] J. Freeman, C. DiSalvo, M. Nitsche, and S. Garrett, "Soundscape Composition and Field Recording as a Platform for Collaborative Creativity," *Organised Sound,* vol. 16, pp. 272-281, 2011.

[16] G. Essl and M. Rohs, "Interactivity for mobile music-making," *Organised Sound,* vol. 14, pp. 197-207, 2009.

[17] G. Wang, "Designing smule's iphone ocarina," in *International Conference on New Interfaces for Musical Expression (NIME)*, Pittsburgh, USA, 2009.

[18] G. Weinberg, M. Godfrey, and A. Beck, "ZOOZbeat: mobile music recreation," in *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* Atlanta, USA, 2010, pp. 4817-4822.

[19] (2008, Apr. 22th). *BeatMaker*. Available: www.intua.net

[20] (2012, May 6th). *Shazam*. Available: http://itunes.apple.com/us/app/shazam/id284993459?mt=8

[21] (2012, May 6th). *SoundHound*. Available: http://itunes.apple.com/us/app/soundhound/id355554941?mt=8

[22] (2009, Dec. 15th). *LaDiDa*. Available: http://itunes.apple.com/ca/app/ladida/id326533688?mt=8

[23] (2011, Dec. 15th). *Songify*. Available: http://itunes.apple.com/ca/app/ladida/id326533688?mt=8

[24] G. Wang, G. Essl, and H. Penttinen, "Do mobile phones dream of electric orchestras," in *the International Computer Music Conference*, Belfast, UK, 2008.

[25] (2008, Apr. 15th). *Helsinki Mobile Phone Orchestra*. Available: http://www.acoustics.hut.fi/projects/helsinkimopho/

[26] (2010, Apr. 10th). *The Michigan Mobile Phone Ensemble.* Available: http://mopho.eecs.umich.edu/

[27] (2010, Apr. 15th). *KAIST Mobile Phone Orchestra*. Available: http://kampo.kaist.ac.kr/

[28] N. J. Bryan, J. Herrera, J. Oh, and G. Wang, "Momu: A mobile music toolkit," in *the International Conference on New Interfaces for Musical Expression(NIME)*, Sydney, Australia, 2010.

[29] G. Essl, "UrMus - an environment for mobile instrument design and performance," in *the International Computer Music Conference*, New York, 2010, pp. 76-81.

[30] J. Oh and G. Wang, "Audience-participation techniques based on social mobile computing," in *the International Computer Music Conference*, Hudderseld, Kirklees UK, 2011.

[31] A. D. Perna. (2002) Interviews -Queens < Queen are Dead! - Long Live Queen! >. *Guitar World*. Available: http://queenonline.com/en/the-band/interviews/brian-may/reigning-champion/

[32] NPR, "Queen's Brian May Rocks Out To Physics, Photography," ed, 2010.

[33] B. McFerrin. (1986, Sep. 4th). *Spontanious inventions*. Available: http://www.youtube.com/watch?v=5VfyMPHzYdQ

[34] B. McFerrin. (2009, Sep. 4th). *World Science Festival 2009: Bobby McFerrin Demonstrates the Power of the Pentatonic Scale*. Available: http://vimeo.com/5732745

[35] C. Brown and J. Bischoff. (2003, Jan. 5th). *Indigenous to the Net: early network music bands in the San Francisco Bay area*. Available: http://crossfade.walkerart.org/brownbischoff/IndigenoustotheNetPrint.html

[36] T. Duffy. (1995). *The Critic's Choice (score)*.

[37] K. C. Baird, "Real-time generation of music notation via audience interaction using python and GNU lilypond," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Vancouver, 2005, pp. 240-241.

[38] L. Carpenter, "Cinematrix, Video Imaging Method and Apparatus for Audience Participation. ," United States Patent #5210604, 1993.

[39] D. Maynes-Aminzade, R. Pausch, and S. Seitz, "Techniques for interactive audience participation," in *IEEE International Conference on Multimodal Interfaces (ICMI)*, Pittsburgh, PA, 2002, p. 15.

[40] J. Freeman, "Extreme sight-reading, mediated expression, and audience participation: Real-time music notation in live performance," *Computer Music Journal,* vol. 32, pp. 25-41, 2008.

[41] J. Freeman and M. Godfrey, "Technology, real-time notation, and audience participation in flock," in *the International Computer Music Conference*, Belfast, UK, 2008.

[42] A. Tanaka and P. Gemeinboeck, "Net_Dérive: Conceiving and producing a locative media artwork," *Mobile Technologies: From Telecommunications to Media,* 2008.

[43] J. Sello, "isms : network performance using sms for audience participation," Master's Thesis, 2009.

[44] N. Kruge and G. Wang, "MadPad: A Crowdsourcing System for Audiovisual Sampling," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Oslo, Norway, 2011.

[45] H. Kim, "Moori: interactive audience participatory audio-visual performance," in *ACM, Creativity and Cognition*, Atlanta, USA, 2011, pp. 437-438.

[46] E. Redlinger, "Mrmr OSC controller," ed, 2008.

[47] N. Weitzner, J. Freeman, S. Garrett, and Y. Chen, "massMobile – an Audience Participation Framework," in *the International Conferences on New Interfaces for Musical Expression (NIME)*, Ann Arbor, USA, 2012.

[48] T. Winkler, "Audience participation and response in movement-sensing installations," in *the International Computer Music Conference*, Berlin, Germany, 2000.

[49] B. McFerrin. (1986, Sep. 4th). *Itsy Bitsy Spider from Spontenaoues Inventions*. Available: http://www.youtube.com/watch?v=1GCVDIpW-oo

[50] R. Fisher, P. Vanouse, R. Dannenberg, and J. Christensen, "Audience interactivity: A case study in three perspectives including remarks about a future production," in *the 6th Biennial Symposium for Arts and Technology*, Connecticut College, New London, USA, 1997.

[51] L. Austin. (2003). *Williams Re[mix]ed, from email of public call for sounds*. Available: http://www.propheticdesire.us/microsound/html/2003/2003-02/msg00503.html

[52] J. Oh and G. Wang, "Converge: An Omni-Biographical Composition," in *Korean Electro-Acoustic Music Society's 2011 Annual Conference(KEAMSAC2011)*, Seoul, Korea, 2011.

[53] B. McFerrin. (2001, Sep. 4th). *Try this at home - Singing with an audience member* Available: http://www.youtube.com/watch?v=mQu_c3YXVZA

[54] J. Freeman, "Graph theory: interfacing audiences into the compositional process," in *the International Conference on New Interfaces for Musical Expression (NIME)*, New York, USA, 2007, pp. 260-263.

[55] Á. Barbosa, "Displaced soundscapes: A survey of network systems for music and sonic art creation," *Leonardo Music Journal,* pp. 53-59, 2003.

[56] S. Jordà, "FMOL: Toward user-friendly, sophisticated new musical instruments," *Computer Music Journal,* vol. 26, pp. 23-39, 2002.

[57] J. Freeman, "Web-based collaboration, live musical performance and open-form scores," *International Journal of Performance Arts and Digital Media,* vol. 6, pp. 149-170, 2010.

[58] B. McFerrin. (1988, Sep. 4th). *audience participation jam*. Available: http://www.youtube.com/watch?v=kdVypcHaoi8

[59] J. Lindsay, *Javanese Gamelan*: Oxford University Press, USA, 1980.

[60] S. Jordà, "Multi-user instruments: models, examples and promises," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Vancouver, Canada, 2005, pp. 23-26.

[61] G. Weinberg, "Interconnected musical networks: Toward a theoretical framework," *Computer Music Journal,* vol. 29, pp. 23-39, 2005.

[62] N. Bryan-Kinns, "Daisyphone: the design and impact of a novel environment for remote group music improvisation," in *the Conference on Designing Interactive Systems (DIS)*, Cambridge USA, 2004, pp. 135-144.

[63] S. Jordà, "Interactive Music Systems For Everyone: Exploring Visual Feedback As a Way for Creating More Intuitive, Efficient And Learnable Instruments," in *Stockholm Music Acoustics Conference*, Sweden, 2003.

[64] G. Weinberg and S. L. Gan, "The squeezables: Toward an expressive and interdependent multi-player musical instrument," *Computer Music Journal,* vol. 25, pp. 37-45, 2001.

[65] A. Albin, S. Şentürk, A. Van Troyer, B. Blosser, O. Jan, and G. Weinberg, "Beatscape, a mixed virtual-physical environment for musical ensembles," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Oslo, Norway, 2011.

[66] D. Pazel, S. Abrams, R. Fuhrer, D. Oppenheim, and J. Wright, "A distributed interactive music application using harmonic constraint," in *the International Computer Music Conference*, Berlin, Germany, 2000, pp. 113-116.

[67] N. d'Alessandro, Pon, A., Wang, J., Eagle, D., Ehud, S., Fels, S., "A Digital Mobile Choir: Joining Two Interfaces towards Composing and Performing

Collaborative Mobile Music," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Ann Arbor, MI U.S.A., 2012.

[68] D. Keller and B. Ferneyhough, "Analysis by Modeling: Xenakis's ST/10-1 080262," *Journal of New Music Research,* vol. 33, pp. 161-171, 2004.

[69] S. Şentürk, S. W. Lee, A. Sastry, A. Daruwalla, and G. Weinberg, "Crossole: A Gestural Interface for Composition, Improvisation and Performance using Kinect," in *the International Conferences on New Interfaces for Musical Expression (NIME)*, Ann Arbor, USA, 2012.

[70] D. Wessel and M. Wright, "Problems and prospects for intimate musical control of computers," *Computer Music Journal,* vol. 26, pp. 11-22, 2002.

[71] S. Jordà, "Digital Instruments and Players: Part I – Efficiencyand Apprenticeship.," in *the International Conference on New Interfaces for Musical Expression (NIME)*, Hamamatsu, Japan, 2004, pp. 59-63.

[72] COCOS2D. (2011, June 17th.). *COCOS2D API* Available: http://www.cocos2d-iphone.org/

[73] Audanika. (2010, June 20th). *SoundPrism*. Available: http://itunes.apple.com/us/app/soundprism/id386833491?mt=8

[74] C. J. Plack, *The sense of hearing* Lawrence Erlbaum Associates Publishers, 2005.

[75] Byul. (2004). *Silk Road, from the album "prostrating on a deck"*. Available: www.byul.org

[76] J. Freeman, "Large audience participation, technology, and orchestral performance," in *the International Computer Music Conference*, Barcelona, Spain, 2005, pp. 757-760.

[77] Smule. (2008, Apr. 27th). *Sonic Lighter*. Available: www.smule.com/soniclighter

[78] (2011, Apr. 20th). *libpd*. Available: http://libpd.cc/