

Alter Audio: Mobile and Locative Sound Experiences

Geoffrey Shea

Ontario College of Art and Design
100 McCaul Street
Toronto ON M5V 1W1
+1 519 369 3025
gshea [at] faculty.ocad.ca

Paula Gardner

Ontario College of Art and Design
100 McCaul Street
Toronto ON M5V 1W1
+1 416 977 6000
pgardner [at] ocad.on.ca

ABSTRACT

In this paper, we describe the software applications and audio content that have been created for cell phones as part of the Mobile Digital Commons Network research project being led by Principal Investigators Sara Diamond (Ontario College of Art and Design) and Michael Longford (Concordia University). Four of ten planned iterative, mobile sound experiences have been built and tested. We look at the different strategies sound designers took to creating content for these applications.

Keywords

Mobile, audio, music, cell phone, locative, iterative

INTRODUCTION

As part of the MDCN research project at OCAD, Banff NMI, and Concordia/Hexagram, OCAD faculty Paula Gardner and Geoffrey Shea have constructed a series of interactive user experiences for presentation on cell phones. These were designed as a series of iterative development exercises, each exploring different technical possibilities and a range of content types and user engagements. At a charette (or workshop) in Toronto in June, 2006 three interns worked with Gardner and Shea and MDCN engineers to implement the first four of ten planned iterations. Here we review the technologies and strategies adopted in preparation for continuing work on the further iterations.

The MDCN project has, in part, led to the development of the Mobile Experience Engine (MEE), a code-generating tool that allows designers and engineers to input the parameters of a game or interactive experience in XML and then generates the appropriate underlying software application (in C++ for example) for installation on a mid-range cell phone. For this research the Toronto MDCN team has adopted a phone manufactured by i-Mate, running the Windows Mobile operating system. These phones can communicate via Bluetooth with each other and ancillary devices such as GPS locators.

The first four applications that we wanted to implement, during an intensive four-day charette at OCAD, were:

1) Orchestra

Four pre-recorded, looped sounds are installed on one phone and a user can turn sounds on and off with buttons and an on-screen display.

2) Choreography

One pre-recorded, looped sound is associated with each of four phones and when the phones are within Bluetooth range of each other, each user can hear the other's sound.

3) Choreography & Control

Two pre-recorded, looped sounds are associated with each of four phones and when the phones are within Bluetooth range of each other, each user can hear the other's sounds. Each user has the option to turn either or both of their sounds on or off; other users within range will hear these sounds turn on or off.

4) Environmentally Located Sounds

Up to five sounds are associated with five specific GPS regions; a phone within a region plays that sound.

Six additional applications are planned and will be built taking into consideration the results of these initial productions.

The three interns who worked with us during the charette were students or recent graduates from both the Art and Design Faculties at OCAD: Mark Poon, John Pavicic and Nigel Craig. They were presented with the capabilities of the iterations outlined above, shown the Orchestra application and the plans for future development. They were asked to imagine what approaches they would take to preparing content for these applications and what sort of additional capabilities they would like see. Then they went into brainstorming and production, creating actual audio content that could be deployed on the applications of their choosing.

The third iteration, Choreography & Control ran into an engineering bottleneck and was not available for testing, although Nigel did create a suite of audio content intended for it. The fourth iteration, Environmentally Located

Sounds, only became available on the last day and a production by John could not be tested during the charette. Other than that, all of the productions were installed on the applications and tested in the field.

APPLICATIONS IN USE AT THE CHARETTE

Prior to the charette the researchers created two content applications. The first was for Iteration 1: Orchestra. This included four tracks, two seconds long each. Three tracks contained different one-bar, musical phrases in 3/4 time: a bass, spacey guitar and a snare drum; the fourth track was a short, spoken phrase. When each loop was started by the user, it would not likely start in synchronization with the other loops. But because these were so short and there was no harmonic progression (i.e. movement from chord to chord), the effect was a vague sense of rhythmic complexity: the guitar, for example, playing a fraction of a beat behind the bass. Users very quickly acclimatized to each new 'dub' version that starting and stopping sounds created.

The second content application was for Iteration 2: Choreography. The aural experience of Iterations 1 and 2 are the same: four audio tracks turn on and off in different combinations, and loop with no particular synchronization, but in Iteration 2 the control is less specific. Rather than direct control by buttons, sounds turn on and off when other users move in and out of Bluetooth range. Bluetooth range during our testing was somewhere between 3 - 10 meters, depending on the electronic 'noisiness' of the environment, the number of phones involved, physical structures like walls, etc. By design, a single user could not affect a particular combination of sounds by themselves. If, for instance, they wanted to hear the sound from User B's phone and moved towards User B, they might inadvertently move out of the range of User C. During testing all the users would first tend to understand how close to each other they needed to be. Since this distance tended to vary, and the other users tended to be continually moving, it was common for any user to hear only one or two sounds at a time. All four users would eventually bunch together, though, to hear the possible combination of all available sounds and to verbally compare experiences.

When this iteration was first deployed a user never heard their own sound; this added some confusion about which sound was associated with which other user's phone. There also seemed to be some lag in the Bluetooth connectivity, so even if a user moved into range their sound might not turn on for 5 - 10 seconds, sometimes creating the impression that the sound was triggered by some other event such as another user approaching or a user walking away from the associated cell phone.

It is important to note that all testing to date has been conducted by people who are working on the project in some capacity: either designers and engineers who were creating the experience, or research assistants, or members of one of the other parallel research teams that are working

under the general MDCN banner. Once we have achieved a level of technical reliability we are committed to testing with more diverse and less familiar user groups.

Also, we have been working with one model of cell phone, the i-Mate SP3i, running Windows Mobile, and primarily one model of Bluetooth connected GPS device. Some technical issues, outline below, may be in part due to built-in limitations of these particular devices. We are looking into testing the same software and content applications on other hardware platforms for comparison.

The i-Mate phone has a 176 by 220 pixel screen and supports stereo sound. Sound quality, clarity and volume are greatly enhanced by the use of headphones. This can be beneficial, creating an enhanced immersive experience, but also can create a sense of disassociation from the immediate physical environment.

In these iterations the basic software design works on the premise that each phone involved has been specially prepared: they have the software application and the sound files pre-loaded, and they have each been assigned identifies that will allow them to be properly recognized by the other phones running the application (for example, this phone is 'guitar', that one is 'bass', that other is 'voice', etc.). So when another user approaches the effect is that you hear 'their' sound, but in reality a sound associated with their identity and which is already installed on your phone, begins to play.

CHARETTE GUIDELINES

The interns were given minimal guidance regarding the design process that would ensue during the charette. We asked that they create a variety of types of content, pay attention to the relationship between form and content, and the relationships between technology, content and space as it relates to the user experience. We gave them some examples of how we have imagined these relationships, for example: choreography, metaphor and orchestra. And finally we asked them to consider what changes in the iterations would push the limits of the currently imagined and designed sound experiences. As a goal, we asked each intern to produce 2-3 audio productions per iteration, in rapid prototype fashion.

With each day, interns became more familiar with the charette model, the software capabilities and with imagining sound experiences. Some interns tended to enjoy rapid prototyping while others took a more deliberate and thoughtful approach to content creation. Importantly, the interns immediately engaged in a critique of the perceived strengths and weaknesses of the software and imagined different user experiences based on more complex and capable software. Interns were guided to work within the existing limitations and to see how far creative content could push the limits of these iterations and still be engaging or enjoyable to users.

Participant Observations

Mark felt that iteration 3 seemed like part of an orchestra, whereby individuals could each have an instrument and try to create a symphonic arrangement from it. He imagined hundreds of people and lots of symphonies existing, with users joining the group that interested them.

Nigel was interested in the possibilities of community building through collaboration, and interactivity through the sound experience. He was also interested in addressing human scale and building to that scale. He strongly felt that the available iterations were constraining and wouldn't allow sufficient interactivity and so brainstormed new future iterations. For example, he tried to think of interactivity in other ways such as outdoor games or user content feeding into the experience. Nigel wanted to turn the user's expectation upside down through their experience.

John's initial response was to think ahead, imagining future iterations because he felt early ones seemed like prototypes. He adopted a 'marketing perspective' that users would want to 'buy in' down the line. John was interested in leaving virtual items and artifacts in the space, thinking of games that had strategies and got people out into the park, developing more mysterious and fun experiences, and creating interactive experiences that made people want to play with the phone instead of simply creating an experience as a 'tacky add-on' to existing phone use. Finally, he thought about creating an application that would allow each user to have a unique experience. This led to an excellent group conversation about how the game might change over time, as users played, so that one would never have the same experience twice. We thought about programming the experience so that it changed based on events that previous players engaged in.

The team discussed the benefits and drawbacks of using headphones: that they allow the use of left and right stereo channels that can produce the experience of being unsure which space is generating which sound, and can pan 360 degrees. We discussed whether we were less engaged in the physical environment if we used headphones and what would happen if we could loop ambient sound through the headset.

We looked at issues of simulation versus actuality—for example, the different experiences of simulated versus actual cell phone conversations and the potential attraction of each to users. We discussed various ways to motivate users to keep using the limited and constrained iterations, such as providing rewards, pleasure, and playing on the "negative space" of the functions. For example, if users were allowed to turn off certain functions, how would this affect the experience?

We brainstormed about general approaches to creating appropriate content for the existing Iterations. Iteration I possibilities included: jazz music in same key; music; sound effects; hip hop; sampled sounds and music of

different lengths. Regarding GPS, we wondered if users themselves could deposit information at a virtual geographic location that would link data to a space that other users could access. We imagined this would play into the nature of the cell phone, making it distinctive from other locative AV media, since the cell phone inherently implies movement or mobility. We imagined possibilities for site memory, such as a message board, which creates a cache-like record, so that everyone can leave 'scar tissue.' By connecting to a cache, users would be able to hear the messages such as, for example, electrical connections to a netherworld. We users themselves being sampled by the technology as they passed a node, and the counter argument that this could be a form of surveillance. Finally, we imagined using the phone as a never-ending blog to capture stuff the designers wouldn't think to capture.

PRODUCTIONS

Since synchronization of the looped sounds was not available (although we are considering revisiting these iterations and adding synchronization as an engineering requirement) the musical compositions that were created sought to work within that given reality.

Jam

John created content for Iteration 2: Choreography that allowed four users to share sounds based on their Bluetooth proximity. Because sounds couldn't sync up, John focused on rhythm, using triplet notes in his beat. The tracks included one instrument track that played to a 12/4 melody resolving briefly on the fifth and second of B, one playing long, sustained tones (aii, bii and bi)¹, one playing syncopated percussion sounds and one featuring two distinct bird sounds in sequence. Each loop was 5 seconds long. Because the sounds were so different in texture and tempo, he projected that they would work equally well together regardless of the lack of synchronization.

Vocal Chords

Nigel took a very different approach to Iteration 2: Choreography. He created four 60 second loops that included a sequence of sung tones in the key of C, two using a male voice and two using a female voice. He reasoned that the human voice was a common musical element in many different cultures and that the C scale was, if not exactly universal, at least common. Each loop had 5 to 6 notes of about 10 to 12 second durations. These were:

Female: ci	a	g	di	f	b
Male:	B	f	e	g	c
Female: di	b	ci	ei	g	f
Male:	g	d	f	e	c

Because there were a different number of sounds combining, depending on the position of the four users and their phones, each user would hear from as little as nothing, up to a three tone chord. (Remember, in this build of the iteration a user would not hear their own sound.) Because the sound loops would start at different times there was an

immense variety of possible chord combinations and the chords would tend to evolve into other chords one note at a time, at times suggesting other modes based on the C scale. The reaction among testers was that the constant and unexpected variation of the tones and chords created a pleasant sense of anticipation that was seldom stifled by recognition of repetitious or predictable patterns.

(Nigel prepared an eight track version of this piece for implementation of Iteration 3: Choreography & Control, in which each of four participants has two associated sounds that they can optionally turn on or off. He projected that having all sounds on would be unsatisfying because six or eight note chords are difficult to discern, and this would encourage users to start stripping away sounds until a satisfying combination was achieved. This remains to be tested.)

Orchestras East/West

Mark created a suite of sounds for Iteration 2: Choreography which specifically created a contrast between two different music contexts.

Soundtrack to Your Life

Mark created an experience on Iteration 4: Environmentally Located Sounds, which allows designers to specify five GPS locations that would each be associated with a different sound, to be experienced by a single user. The locations would typically have a centre and a region defined by a radius (i.e. anywhere within 10 meters of this rock is one region). One problem that we anticipated with GPS is that the regions tend to 'drift', that is, they would move around erratically as the GPS device connected with different satellites. We are still measuring this drift but it seems that it may move scores of meters every few seconds.

Mark chose to have all five regions concentric, so they existed in circular bands around a single, central GPS coordinate. Each ring was defined as 15 meters wide. If a user traversed each ring, moving directly towards the centre of the GPS region they would encounter five, equal sized bands with accompanying sounds. If the user moved on a lightly tangential path the rings would not appear equal sized and they might not ever encounter the innermost rings. (Johnny will begin to address this eventuality below.)

Mark took an orchestral approach to his sound creation. The five tracks move subtly from a 16 second sequence of four appoggiated chords combined with bird sounds to a fuller score with a double bass, a violin and cello playing in double time and a languid piano. Instruments are added one at a time over the original sounds creating the impression that sound is growing. Although the sounds might skip as they restart when users transition from one region to another, the overall impression among testers was that the 'score' was building much like it might during the title sequence of a film. One user dubbed it 'starring in my own

movie.' Mark used a 16 second sequence because he thought this was the time it would likely take an average user to traverse one band of the region. The immersive experience is engaging, and heightens one's experience of walking the park by adding a soundtrack and removing most of the ambient park sound, especially when one wears headphones.

Moving Crowd

Nigel considered the instability of the GPS coordinates – their tendency to jump several meters every few seconds – as a feature to be incorporated. He created an ambient sound installation for Iteration 4: Environmentally Located Sounds, using fairly densely packed regions, each one associated with a different crowd sound. He intended a nine region grid with each grid being square and only a few meters wide. He wanted to create the impression that, as a user stood still in the middle square of the grid, the crowd would be jumping around him and letting the GPS drift be the active agent in the experience. This remains to be tested.

Homing Signal

John created several series of short, looped sounds for Iteration 4: Environmentally Located Sounds, that would alert users that they are moving in the right direction. In each the sounds sped up as users moved from GPS region to region, reminiscent of the familiar radar pings we have seen in so many movies.

Streetscape

Nigel and Paula sampled sounds in the environment of John Street and Grange Park to be implemented as part of Iteration 4: Environmentally Located Sounds. This experience is similar in structure to Mark's Soundtrack to Your Life experience, but as a texturally different kind of walk in the Grange park: GPS nodes trigger various sounds that may or may not be exactly where the user is walking. This could be a component of a game (finding the sound) or simply an ironic or mystical aural experience that does not necessarily coincide with one's visual experience.

NEXT STEPS

We are working with the engineers to resolve some outstanding issues on the existing iterations: getting Iteration 3: Choreography & Control operating, ensuring users in Iteration 2: Choreography, hear their own sounds.

The remaining, planned iterations include:

5: Metaphor

A combination of the features in Choreography and Environmentally Located Sounds

6: Sampling

Users contribute sounds into the system by recording with the phone's microphone

7: Effects

Participants affect sounds in ways other than turning them on and off, such as reverb, panning, volume, mixing, pitch shifting)

8: Synthesizer

Users generate music in MIDI using the keypad

9: Gesture

Users affect sounds through gestural movements such as raising their hands

10: Uploading

Users contribute sounds to the system that remain available for future users

We are now involved in extensive testing of the available iterations and their content applications with a range of participants. User will be asked to try the experiences, provide feedback and make suggestions for where they would like to see more development, in either the applications or the content productions.

Once we have completed a thorough testing stage, we will engage a number of composers and soundscape designers

to create content that can be deployed on these applications for presentation to a broader public.

ACKNOWLEDGMENTS

We thank our research team, engineers: Tom Donaldson, David Gauthier, Jagmit Singh; research coordinator: Patricio Davila; research assistants: Ken Leung, Nevena Nilogova; production interns: Mark Poon, John Pavicic, Nigel Craig

REFERENCES

1. Helmholtz Pitch Notation: Middle C and the scale above are notated as c^i d^i , etc. The scale below is notated: c , d , etc. The next scale below is notated: C , D , etc. The next scale below is notated: C_i , D_i , etc. The scale above is notated: c^{ii} , d^{ii} , etc.
2. Benford, S., et al., Coping with Uncertainty in a Location-Based Game, *Pervasive Computing*, July 2003. Available at [http://www.equator.ac.uk/var/uploads /IEEEpervasive-nearlyfinal.pdf](http://www.equator.ac.uk/var/uploads/IEEEpervasive-nearlyfinal.pdf)