

Large-Scale Mobile Audio Environments for Collaborative Musical Interaction

Mike Wozniowski &
Nicolas Bouillot
Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada
{mikewoz,nicolas}@cim.mcgill.ca

Zack Settel
Université de Montréal
Montréal, Québec, Canada
zs@sympatico.ca

Jeremy R. Cooperstock
Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada
jer@cim.mcgill.ca

ABSTRACT

New application spaces and artistic forms can emerge when users are freed from constraints. In the general case of human-computer interfaces, users are often confined to a fixed location, severely limiting mobility. To overcome this constraint in the context of musical interaction, we present a system to manage large-scale collaborative mobile audio environments, driven by user movement. Multiple participants navigate through physical space while sharing overlaid virtual elements. Each user is equipped with a mobile computing device, GPS receiver, orientation sensor, microphone, headphones, or various combinations of these technologies. We investigate methods of location tracking, wireless audio streaming, and state management between mobile devices and centralized servers. The result is a system that allows mobile users, with subjective 3-D audio rendering, to share virtual scenes. The audio elements of these scenes can be organized into large-scale spatial audio interfaces, thus allowing for immersive mobile performance, locative audio installations, and many new forms of collaborative sonic activity.

Keywords

sonic navigation, mobile music, spatial interaction, wireless audio streaming, locative media, collaborative interfaces

1. INTRODUCTION & BACKGROUND

With the design of new interfaces for musical expression, it is often argued that control paradigms should capitalize on natural human skills and activities. As a result, a wide range of tracking solutions and sensing platforms have been explored, which translate human action into signals that can be used for the control of music and other forms of media. The physical organization of interface components plays an important role in the usability of the system, since user motion naturally provides kinesthetic feedback, allowing a user to better remember the style of interaction and gestures required to trigger certain events. Also, as digital devices become increasingly mobile and ubiquitous, we expect interactive applications to become more distributed and integrated

within our physical environment. This prospect yields a new domain for musical interaction employing augmented-reality interfaces and large multi-user environments.

We present a system where multiple participants can navigate about a university campus, several city blocks, or an even larger space. Equipped with position-tracking and orientation-sensing technology, their locations are relayed to other participants and to any servers that are managing the current state. With a mobile device for communication, users are able to interact with an overlaid virtual audio environment, containing a number of processing elements. The physical space thus becomes a collaborative augmented-reality environment where immersive musical interfaces can be explored. Musicians can input audio at their locations, while virtual effects processors can be scattered through the scene to transform those signals. All users, performers and audience alike, receive subjectively rendered spatial audio corresponding to their particular locations, allowing for unique experiences that are not possible in traditional music performance venues.



Figure 1: A mobile performer

1.1 Background

In earlier work, we have spent significant time exploring how virtual worlds can be used as musical interfaces. The result of this investigation has led to the development of the *Audioscape* engine [30]¹, which allows for the spatial organization of sound processing, and provides an audiovisual rendering of the scene for feedback. Audio elements can be arranged in a 3-D world and precise control over the directivity of propagating audio is provided to the user. For example, an audio signal emitted by a sound generator may be *steered* toward a sound processor that exists at some 3-D location. The processed signal may again be steered

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME08, Genova, Italy

Copyright 2008 Copyright remains with the author(s).

¹Available at www.audioscape.org

towards a virtual microphone that captures and busses the sound to a loudspeaker where it is heard. The result is a technique of spatial signal bussing, which lends particularly well to many common mixing operations. Gain control, for instance, is accomplished by adjusting the distance between two nodes in the scene, while filter parameters can be controlled by changing orientations.

The paradigm of organizing sound processing in three-dimensional space has been explored in some of our previous publications [27, 28, 26]. We have seen that users easily understand how to interact with these scenes, especially when actions are related to every-day activity. For instance, it is instantly understood that increasing the distance between two virtual sound elements will decrease the intensity of the transmitted signal, and that pointing a sound source in a particular direction will result in a stronger signal at the target location. We have designed and prototyped several applications using these types of interaction techniques, including 3-D mixing, active listening, and using virtual effects racks [27, 29]. Furthermore, we began to share virtual scenes between multiple participants, each with subjective audio rendering and steerable audio input, allowing for the creation of virtual performance venues and support for virtual reality video conferencing [31].

While performers appreciated the functionality of these earlier systems, they were nevertheless hampered by constraints on physical mobility. These applications operated mainly with game-like techniques, where users stood in front of screens, and navigated through the scene using controllers such as joysticks or gamepads. The fact that the gestures for moving and steering sound were abstracted through these intermediate devices resulted in a lack of immersive feeling and made the interfaces more complicated to learn.

We thus decided to incorporate more physical movement, for example, sensing the user’s head movement with an orientation sensor attached to headphones, and applying this to affect changes to the apparent spatial audio rendering. To further extend this degree of physical involvement we began to add real-world location awareness to the system, allowing users to move around the space physically instead of virtually. For example, our *4Dmix*³ installation [4] tracked up to six users in an 80m² gallery space. The motion of each user controlled the position of a recording buffer, which could travel among a number of virtual sound generators in the scene. The result was a type of remixing application, where users controlled the mix by moving through space.

In the remainder of this paper, we explore the use of larger scale position tracking, such as that of a Global Positioning System (GPS), and the resulting challenges and opportunities that such technology presents. We evolve our framework to support a more distributed and mobile-capable architecture, which results in the need for wireless audio streaming and the distribution of information about the mobile participants. Sections 2 and 3 describe the additional technical elements that need to be introduced into the system to support wireless and mobile applications, while Section 4 demonstrates a prototypical musical application using this new architecture. Musicians in the *Mobile Audioscape* are able to navigate through an outdoor environment containing a superimposed set of virtual audio elements. Real physical gestures can be used to steer and move sound through the space, providing an easily understood paradigm of interaction in what can now be thought of as a mobile music venue.

1.2 Mobile Music Venues

By freeing users from the confines of computer terminals and interfaces that severely limit mobility, application

spaces emerge that can operate in a potentially unbounded physical space. These offer many novel possibilities that can lead to new artistic approaches; or they can re-contextualize existing concepts that can then be revisited and expanded upon. An excellent example is parade music, where sound emission is spatially dynamic or mobile; a passive listener remains in one place while different music is coming *and* going. One hundred years ago, Charles Ives integrated this concept into symphonic works, where different musical material *flowed* through the score, extending our notions of counterpoint to include those based on *proximity* of musical material. The example of parade music listening expands to include two other cases: a mobile listener can walk with or against the parade, yielding additional relationships to the music. Our work also integrates the concept of active listening; material may be organized topographically in space, produced by mobile performers and encountered non-linearly by mobile listeners. From this approach come several rich musical forms, which like sculpture, integrate *point of view*; listeners/observers create their own unique rendering. Thus, artists may create works that explore the spatial dynamics of musical experience, where *flowing* music content is put in counterpoint by navigation. Musical scores begin to resemble maps, and listeners play a larger role in authoring their experiences.

1.3 Related Work

With respect to collaborative musical interfaces, Blaine and Fels provide an overview of many systems, classifying them according to attributes such as scale, type of media, amount of directed interaction, learning curve, and level of physicality, among others [7]. However, most of these systems rely on users to be in a relatively fixed location in front of a computer. The move to augmented- or mixed-reality spaces seems like a natural evolution, offering users a greater level of immersion in the collaboration, and their respective locations can be used for additional control.

In terms of locative media, some projects have considered the task of tagging geographical locations with sound. The *[murmur]* project [2] is one simple example, where users tag interesting locations with phone numbers. Others can call the numbers using their mobile phones and listen to audio recordings related to the locations. Similarly, the *Hear&There* project [20] allows recording audio at a given GPS coordinate, while providing a spatial rendering of other recordings as users walk around. Unfortunately, this is limited to a single-person experience, where the state of the augmented reality scene is only available on one computer. Tanaka proposed an ad-hoc (peer-to-peer) wireless networking strategy to allow multiple musicians to share sound simultaneously using hand-held computers [22]. Later work by Tanaka and Gemeinboeck [23] capitalized on location-based services available on 3G cellular networks to acquire coarse locations of mobile devices. They proposed the creation of *locative media instruments*, where geographic localization is used as a musical interface.

Large areas can also be used for musical interaction in other ways. *Sonic City* [16] proposed mobility, rather than location, alone, for interaction. As a user walks around a city, urban sounds are processed in real time as a result of readings from devices such as accelerometers, light sensors, temperature sensors, and metal detectors. Similarly, the *Sound Mapping* [19] project included gyroscopes along with GPS sensors in a suitcase that users could push around a small area. Both position changes and subtle movements could be used to manipulate the sound that was transmitted between multiple cases in the area via radio signal.

Orientation or heading can also provide useful feedback,

since spatial sound conveys a great deal of information about directions of objects and the acoustics of an environment. Projects including *GpsTunes* [21] and *Melodious Walkabout* [15] use this type of information to provide audio cues that guide individuals in specific directions.

We take inspiration from the the projects mentioned above, and incorporate many of these ideas into our work. However, real-time high-fidelity audio support for multiple individuals has not been well addressed. Tanaka’s work [22], as well as some of our past experiences [8], demonstrate how we can deal with the latencies associated with distributed audio performance, but minimizing latency remains a major focus of our work. The ability to create virtual audio scenes will be supported with some additions to our existing Audioscape engine. To address the need of distributed mobile interaction, we are adding large-scale location sensing and the ability to distribute state, signals, and computation among mobile clients effectively. These challenges are addressed in the following sections.

2. LOCATIVE TECHNOLOGY

In order to support interaction in large-scale spaces, we require methods of tracking users and communicating between them. A variety of mobile devices are available for this purpose, potentially equipped with powerful processors, wireless transmission, and sensing technologies. For our initial prototypes, we chose to develop on Gumstix (verdex XM4-bt) processors with expansion boards for audio I/O, GPS, storage, and WiFi communication [17]. These devices have the benefit of being full-function miniature computers (FFMC) with a large development community, and as a result, most libraries and drivers can be supported easily.

2.1 Wireless Standards

Given that the most generally available wireless technologies on mobile devices are Bluetooth and WiFi, we consider the benefits and drawbacks for each of these standards. For transmission of data between sensors located on the body and the main processing device, Bluetooth is a viable solution. However, even with Bluetooth 2.0, a practical transfer rate is typically limited to approximately 2.1 Mbps. If we want to send or receive audio (16 bit samples at 44kHz), approximately 700 kbps of bandwidth is needed for each stream. In theory, this allows for interaction between up to three individuals, where each user sends one stream and receives two. Given the need to support a greater number of participants, we are forced to use WiFi.² Furthermore, the range of Bluetooth is limiting, whereas WiFi can relay signals through access points. Furthermore, we can make use of higher-level protocols such as Optimized Link State Routing protocol (OLSR) [18], which computes optimal forwarding paths for ad-hoc nodes. This is a viable way to reconfigure wireless networks if individuals are moving.

2.2 GPS

GPS has seen widespread integration into a variety of commodity hardware such as cell phones and PDAs. These provide position tracking in outdoor environments, typically associated with the 3-D geospatial coordinates of users. However, accuracy in consumer-grade devices is quite poor, ranging between approximately 5m in the best case (high-quality receiver with open skies) [25] to 100 metres or more [6]. Several methods exist to reduce error, for example, differential GPS (DGPS) uses carefully calibrated base sta-

tions that transmit error corrections over radio frequencies. The idea is that mobile GPS units in the area will have similar positional drift, and correcting this can yield accuracies of under 1m. Another technique, known as assisted GPS (AGPS), takes advantage of common wireless networks (cellular, bluetooth, WiFi) in urban environments to access reference stations with a clear view of the sky (e.g., on the roofs of buildings). Although accuracy is still in the order of 15m, the interesting benefit of this system is that localization can be attained indoors (with an accuracy of approximately 50m) [6].

2.3 Orientation & Localization

While GPS devices provide location information, it is also important to capture a listener’s head orientation so that spatial cues can be provided, the resulting sound appearing to propagate from a particular direction. Most automotive GPS receivers report heading information by tracking the vehicle trajectory over time. This is a viable strategy for inferring the orientation of a vehicle, but a listener’s head can change orientation independently of body motion. Moreover, the types of applications we are targeting will likely involve periods of time where a user does not change position, but stays in one place and orients his or her body in various directions. Therefore, additional orientation sensing seems to be a requirement.

In human psychoacoustic perception, accuracy and responsiveness of orientation information are important, since a listener’s ability to localize sound is highly dependent on changes in phase, amplitude, and spectral content with respect to head motion. Responsiveness, in particular, is a significant challenge, considering the wireless nature of the system. Listeners will be moving their heads continuously to help localize sounds, and a delay of more than 70ms in spatial cues can hinder this process [10]. Furthermore, it has been demonstrated that head-tracker latency is most noticeable in augmented reality applications, as a listener can compare virtual sounds to reference sounds in the real environment. In these cases, latencies as low as 25ms can be detected, and begin to impair performance in localization tasks at slightly greater values [11]. It is therefore suggested that latency be maintained below 30ms.

To track head orientation, we attach an inertial measurement unit (IMU) to the headphones of each participant, capable of sensing instantaneous 3-D orientation with an error of less than 1 degree. It should be mentioned that not all applications will require this degree of precision, and some deployments could potentially make use of trajectory-based orientation information. For instance, the *Melodious Walkabout* [15] uses aggregated GPS data to determine the direction of travel, and provides auditory cues to guide individuals in specific directions. Users hear music to their left if they are meant to take a left turn, whereas a low-pass filtered version of their audio is heard if they are traveling in the wrong direction. We can conceive of other types of applications, where instantaneous head orientation is not needed, and users could adjust to the paradigm of hearing audio spatialization according to trajectory rather than line of sight. Of particular interest, are high-velocity applications such as skiing or cycling, where users are generally looking forward, in the direction of travel. Such constraints can help with predictions of possible orientations, while the faster speed helps to overcome the coarse resolution of current GPS technology.

3. WIRELESS AUDIO STREAMING

The move to mobile technology presents significant design challenges in the domain of audio transmission, largely

²We note viable alternatives on the horizon, such as the newly announced SOUNDabout Lossless codec, which allows even smaller audio streams to be sent over Bluetooth.

related to scalability and the effects of latency on user experience. More precisely, a certain level of quality needs to be maintained to ensure that mobile performers and audience members experience audio fidelity that is comparable to traditional venues. The design of effective solutions should take into account that WiFi networks provide variable performance based on the environment, and that small and lightweight mobile devices are, at present, limited in terms of computation capabilities.

3.1 Scalability

Reliance on unicast communication between users in a group suffers a potential n^2 effect of audio interactions between them and in turn, to bandwidth explosion. We have investigated a number of solutions to deal with this problem.

Multicast technology, for instance, allows devices to send UDP packets to an IP multicast address that virtualizes a group of receivers. Interested clients are able to subscribe to the streams of relevance, drastically reducing the overall required bandwidth. However, IP multicast over IEEE 802.11 wireless LAN is known to exhibit unacceptable performance [14] due to unsupported collision avoidance and acknowledgement at the MAC layer. Our benchmark tests confirm that multicast transmission experienced higher jitter than unicast, mandating a larger receiver buffer to maintain quality. Furthermore, packet loss for the multicast tests was in the order of 10-15%, resulting in a distorted audio stream, while unicast had almost negligible losses of 0.3%. Based on these results, we decided to rely for now on a point-to-point streaming methodology while experimenting with emerging non-standard multicast protocols, in anticipation of future improvements.

3.2 Low Latency Streaming

Mobile applications tend to rely on compression algorithms to respect bandwidth constraints. As a result they often incur signal delays that challenge musical interaction and performer synchronization. Acceptable latency tolerance depends on the style of music, with figures as low as 10ms [12] for fast pieces. More typically, musicians have difficulty synchronizing with latencies above 50ms [13]. Most audio codecs require greater than this amount of encoding time.³ Due in part to limited computational resources available on our mobile devices, we instead transmit uncompressed audio, thus fully avoiding codec delays in the system.

Other sources of latency include packetization delay, corresponding to the time required to fill a packet with data samples for transmission, and network delay, which varies according to network load and results in jitter at the receiver. Soundcard latencies also play a role, but we consider this to be outside of our control. The most effective method for managing these remaining delays may be to minimize the size of transmitted packets. By sending a smaller number of audio samples in each network packet, we also decrease the amount of time that we must wait for those samples to arrive from the soundcard.

In this context, we have developed an dynamically reconfigurable transmission protocol for low-latency, high-fidelity audio streaming. Our protocol, *nstream*, supports dynamic adjustment of sender throughput and receiver buffer size. This is accomplished by switching between different levels of PCM quantization (8, 16 and 32 bit), packet size, and receiver buffer size. The protocol is developed as an external

³Possible exceptions are the Fraunhofer Ultra-Low Delay Codec (offering a 6ms algorithmic delay) [24] and the SOUNDabout Lossless codec (claiming under 10ms).

for Pure Data [3], and can be deployed on both a central server and a mobile device.

In benchmark tests, we have successfully transmitted uncompressed streams with an outgoing packet size of 64 samples. The receiver buffer holds two packets in the queue before decoding, meaning that a delay of three packets is encountered before the result can be heard. With a sampling rate of 44.1kHz, this translates to a packetization and receiving latency of $3 \times (64/44.1) = 4.35\text{ms}$. In addition, the network delay can be as low as 2ms, provided that the users are relatively close to each other, and typically does not exceed 10ms for standard wireless applications. The sum of these latencies is in the order of 7-15ms.

Practical performance will, of course, depend on the wireless network being used and the number of streams transmitted. Our experiments show that high packet rate results in network instability and high jitter. In such situations it is necessary to increase packet size to help maintain an acceptable packet rate. This motivates us, as future work, to investigate algorithms for autonomous adaptation of low-latency protocols that deal both with quality and scalability.

4. MOBILE AUDIOSCAPE

Our initial prototyping devices, Gumstix, were chosen to provide: 1) wireless networking for bidirectional high-quality, low-latency audio and data streams, 2) local audio processing, 3) on-board device hosting for navigation and other types of USB or Bluetooth sensors, 4) minimal size/weight, and 5) Linux support. A more detailed explanation of our hardware infrastructure can be found in another publication [9], in particular, the method of Bluetooth communication between Gumstix and sensors.

To develop on these devices, a cross-compilation toolchain was needed that could produce binaries for the ARM-based 400MHz Gumstix processors (Marvell's PXA270). The first library that we needed to build was a version of Pure Data (Pd), which is used extensively for audio processing and control signal management by our Audioscape engine. Particularly, we used *Pure Data anywhere* (PDA), a special fixed-point version of Pd for use with the processors typically found on mobile devices [5]. Several externals needed to be built for PDA, including a customized version of the Open Sound Control (OSC) objects, where multicast support was added, and the *nstream* object, mentioned in Section 3.2. The latter was also specially designed to support both regular Pd and PDA, using sample conversion for interoperability between an Apple laptop, PC and Gumstix units.

We also supplied each user with an HP iPAQ, loaded with a customized application that could graphically represent their location on a map. This program was authored with HP Mediascape software [1], which supports the playback of audio, video, and even Flash based on user position. The most useful aspect of this software was the fact that we could use Flash XML Sockets to receive GPS locations of other participants and update the display accordingly. Although we used the Compact Flash GPS receivers with the iPAQs for sending GPS data, the interface between Mediascape software and the Flash program running within it only allowed for updates at 2Hz, corresponding to a latency of at least 500ms before position-based audio changes were heard. The use of the GPSstix receiver, directly attached to the Gumstix processor, is highly recommended to anyone attempting to reproduce this work.

The resulting architecture is illustrated in Figure 2. Input audio streams are sent as mono signals to an Audioscape

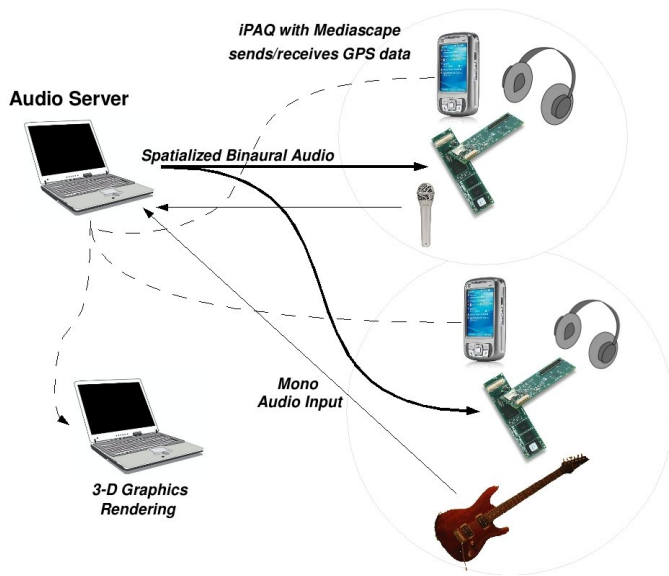


Figure 2: Mobile Audioscape architecture. Solid lines indicate audio streaming while dotted lines show transmission of control signals.

server on a nearby laptop. The server also receives all control data via OSC from the iPAQ devices and stores location information for each user. A spatialized rendering is computed, and stereo audio signals are sent back to the users. For all streams, we send audio with a sampling rate of 44.1 kHz and 16-bit samples.

In terms of network topology, wireless ad-hoc connections are used, allowing users to venture far away from buildings with access points (provided that the laptop server is moved as well). Due to the number of streams being transmitted, audio is sent with 256 samples per packet, which ensures an acceptable packet rate and reduces jitter on the network. The result is a latency of $3 \times (256/44.1) = 17.4\text{ms}$ for packetization and a minimal network delay of about 2ms. However, since audio is sent to a central server for processing before being heard, these delays are actually encountered twice, for a total latency of approximately 40ms. This is well within the acceptable limit for typical musical performance, and was not noticed by users of the system.

The artistic application we designed allows users to navigate through an overlaid virtual audio scene. Various sound loops exist at fixed locations, where users may congregate and jam with accompanying material. Several virtual volumetric regions are also located in the environment, allowing some users to escape within a sonically isolated area of the scene. Furthermore, each of these enclosed regions serves as a resonator, providing musical audio processing (e.g., delay, harmonization or reverb) to signals played within. As soon as players enter such a space, their sounds are modified, and a new musical experience is encountered. Figure 3 shows two such performers, who have chosen to jam in a harmonized echo chamber. They are equipped with Gumstix and iPAQs, with both unobtrusively in their pockets.

5. DISCUSSION

Approaching mobile music applications from the perspective of virtual overlaid environments, allows novel paradigms of artistic practice to be realized. The virtualization of performer and audience movement allows for interaction with sound and audio processing in a spatial fashion that leads to new types of interfaces and thus, new musical experiences.



Figure 3: Two participants jamming in a virtual echo chamber, which has been arbitrarily placed on the balcony of a building at the Banff Centre.

We have presented the challenges associated with supporting multiple participants in such a system, including the need for accurate sensing technologies and network architectures that can support low latency communication in a scalable fashion. The prototype application that we developed was well-received by those who experimented with it, but many improvements still need to be made. The coarseness of resolution available in consumer-grade GPS technology is such that an application must span a wide area for it to be of any value. This is problematic, since the range of a WiFi network is much smaller, mandating redirection of signals through additional access points or OLSR peers. If all signals must first travel to a server for processing, then distant nodes will suffer from very large latency.

One solution is to distribute the state of the virtual scene to all client machines, and perform rendering locally on the mobile devices. For the prototype application that we developed, this would cut latency in half since audio signals would only need to travel from one device to another, without the need to return from a central processing server. Furthermore, this strategy would allow users to be completely free in terms of mobility, rather than in within contact with the server for basic functionality. However, for scenes of any moderate complexity, this demands much more processing power and memory than is currently available in consumer devices, and of course, the number of users will still be limited by the available network bandwidth required for peer-to-peer streaming.

A full investigation into distributing audio streams, state and computational load will be presented in future work, but for the moment we have provided a first step into the exploration of large-scale mobile audio environments. The multi-user nature of the system coupled with high-fidelity audio distribution provides a new domain for musical practice. We have already designed outdoor spaces for sonic investigation, and hope to perform and create novel musical interfaces in this new mobile context.

6. ACKNOWLEDGEMENTS

The authors wish to acknowledge the generous support of NSERC and Canada Council for the Arts, which have funded the research and artistic development described in this paper through their New Media Initiative. The prototype application described in Section 4 was produced in co-production with The Banff New Media Institute (Alberta, Canada). The authors would like to thank the participants

of the locative media residency for facilitating the work and in particular, Duncan Speakman, who provided valuable assistance with the HP Mediascape software.

7. REFERENCES

- [1] HP Mediascape website. www.mscape.com.
- [2] The [murmur] project. murmurtoronto.ca.
- [3] Pure Data. www.puredata.info.
- [4] Webpage: 4Dmix3. audioscape.org/twiki/bin/view/Audioscape/SAT4Dmix3.
- [5] PDA: Real time signal processing and sound generation on handheld devices. In *International Computer Music Conference (ICMC)*, 2003.
- [6] R. Bajaj, S. L. Ranaweera, and D. P. Agrawal. GPS: Location-tracking technology. *Computer*, 35(4):92–94, 2002.
- [7] T. Blaine and S. Fels. Contexts of collaborative musical experiences. In *Proceedings of the conference on New Interfaces for Musical Expression (NIME)*, pages 129–134, Montreal, 2003.
- [8] N. Bouillot. nJam user experiments: Enabling remote musical interaction from milliseconds to seconds. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pages 142–147, New York, NY, USA, 2007. ACM.
- [9] N. Bouillot, M. Wozniowski, Z. Settel, and J. R. Cooperstock. A mobile wireless platform for augmented instruments. In *International Conference on New Interfaces for Musical Expression*, Genova, Italy, 2008.
- [10] D. Brungart, B. Simpson, R. McKinley, A. Kordik, R. Dallman, and D. Ovenshire. The interaction between head-tracker latency, source duration, and response time in the localization of virtual sounds. In *Proceedings of the International Conference on Auditory Display (ICAD)*, 2004.
- [11] D. S. Brungart and A. J. Kordik. The detectability of headtracker latency in virtual audio displays. In *Proceedings of International conference on Auditory Display (ICAD)*, pages 37–42, 2005.
- [12] E. Chew, A. A. Sawchuk, R. Zimmerman, V. Stoyanova, I. Tosheff, C. Kyriakakis, C. Papadopoulos, A. R. J. François, and A. Volk. Distributed immersive performance. In *Proceedings of the Annual National Association of the Schools of Music (NASM)*, San Diego, CA, 2004.
- [13] E. Chew, R. Zimmermann, A. A. Sawchuk, C. Papadopoulos, C. Kyriakakis, C. Tanoue, D. Desai, M. Pawar, R. Sinha, and W. Meyer. A second report on the user experiments in the distributed immersive performance project. In *Proceedings of the 5th Open Workshop of MUSICNETWORK: Integration of Music in Multimedia Applications*, 2005.
- [14] D. Dujovne and T. Turletti. Multicast in 802.11 WLANs: an experimental study. In *MSWiM '06: Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems*, pages 130–138, New York, NY, USA, 2006. ACM.
- [15] R. Etter. Implicit navigation with contextualized personal audio contents. In *Adjunct Proceedings of the Third International Conference on Pervasive Computing*, pages 43–49, 2005.
- [16] L. Gaye, R. Mazé, and L. E. Holmquist. Sonic city: the urban environment as a musical interface. In *Proceedings of the Conference on New interfaces for musical expression (NIME)*, pages 109–115, Singapore, 2003.
- [17] Gumstix. www.gumstix.com.
- [18] P. Hipercom. RFC 3626, Optimized Link State Routing protocol (OLSR), 2003.
- [19] I. Mott and J. Sosnin. Sound mapping: an assertion of place. In *Proceedings of Interface*, 1997.
- [20] J. Rozier, K. Karahalios, and J. Donath. Hear & There: An augmented reality system of linked audio. In *Proceedings of International Conference on Auditory Display (ICAD)*, 2000.
- [21] S. Strachan, P. Eslambolchilar, R. Murray-Smith, S. Hughes, and S. O’Modhrain. GpsTunes: Controlling navigation via audio feedback. In *International Conference on Human Computer Interaction with Mobile devices & services (MobileHCI)*, pages 275–278, New York, 2005. ACM.
- [22] A. Tanaka. Mobile music making. In *Proceedings of New Interfaces for Musical Interaction (NIME)*, 2004.
- [23] A. Tanaka and P. Gemeinboeck. A framework for spatial interaction in locative media. In *Proceedings New Interfaces for Musical Expression (NIME)*, pages 26–30, Paris, France, 2006. IRCAM.
- [24] S. Wabnik, G. Schuller, J. Hirschfeld, and U. Krämer. Reduced bit rate ultra low delay audio coding. In *Proceedings of the 120th AES Convention*, May 2006.
- [25] M. Wing, A. Eklund, and L. Kellogg. Consumer-grade global positioning system (GPS) accuracy and reliability. *Journal of Forestry*, 103(4):169–173, 2005.
- [26] M. Wozniowski. A framework for interactive three-dimensional sound and spatial audio processing in a virtual environment. Master’s thesis, McGill University, 2006.
- [27] M. Wozniowski, Z. Settel, and J. R. Cooperstock. A framework for immersive spatial audio performance. In *New Interfaces for Musical Expression (NIME)*, Paris, pages 144–149, 2006.
- [28] M. Wozniowski, Z. Settel, and J. R. Cooperstock. A paradigm for physical interaction with sound in 3-D audio space. In *Proceedings of International Computer Music Conference (ICMC)*, 2006.
- [29] M. Wozniowski, Z. Settel, and J. R. Cooperstock. A spatial interface for audio and music production. In *Digital Audio Effects (DAFx)*, 2006.
- [30] M. Wozniowski, Z. Settel, and J. R. Cooperstock. Audioscape: A pure data library for management of virtual environments and spatial audio. In *Pure Data Convention*, Montreal, 2007.
- [31] M. Wozniowski, Z. Settel, and J. R. Cooperstock. User-specific audio rendering and steerable sound for distributed virtual environments. In *International conference on auditory displays (ICAD)*, 2007.