An Augmented Reality Framework for Wireless Mobile Performance

Mike Wozniewski &
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
We demonstrate that musical performance can take place in a large-scale augmented reality setting. With the use of mobile computers equipped with GPS receivers, we allow a performer to navigate through an outdoor space while interacting with an overlaid virtual audio environment. The scene is segregated into zones, with attractive forces that keep the virtual representation of the performer locked in place, thus overcoming the inaccuracies of GPS technology. Each zone is designed with particular musical potential, provided by a spatial arrangement of interactive audio elements that surround the user in that location. A subjective 3-D audio rendering is provided via headphones, and users are able to input audio at their locations, steering their sound towards sound effects of interest. An objective 3-D rendering of the entire scene can be provided to an audience in a concert hall or gallery space nearby.

1. INTRODUCTION
Large-scale outdoor spaces offer an interesting interaction space for musical performance, where all participants involved are free to explore sound in a random and nonlinear fashion. Whereas most traditional music is composed and arranged in time, mobile musical applications need to consider the spatial aspect as well. Rather than just focusing on *when* sonic events occur, the composer must also consider *where* they should be located in space. This is a difficult task, since sounds must evolve over time to be heard, while users, who trigger such events with their motions, may do so in a nondeterministic fashion. Composers must therefore arrange their sonic scores in a coherent spatial fashion. There need to be boundaries between sounds that do not mix well, and transitions between adjacent sounds that work in both directions. In a sense, the composer must lay out a mix in a topographical fashion, becoming a sort of sonic cartographer who maps out the score in both space and time.

If we then consider the ability to add live audio input into the scene, a rich venue for live performance emerges, where dynamic sound sources are driven by mobile performers. However, a significant challenge arises, since users will need to transmit audio wirelessly while maintaining synchronization with everything in the scene. There are few tools available to artists for accomplishing all of these tasks. As a result, we have expanded the Audioscape framework[1] to support this kind of mobile interaction. In related work, we have designed an adaptive audio streaming protocol that can transmit sound between multiple individuals on an adhoc network, with very low latency [9]. Furthermore, we have explored the use of Global Positioning Systems (GPS) to track multiple users in an outdoor environment and immerse them in an overlaid virtual audio scene. The initial prototype that we developed, seen in Figure 1, allowed two performers to navigate about a physical space and encounter various audio elements such as sound loops an virtual acoustic enclosures. Each participant had a subjectively rendered audio display that allowed for a unique experience as they travelled through the shared virtual scene.

Figure 1: Mobile performers

2. RELATED WORK
To our knowledge, the interaction we aim to achieve has not been supported by any single system, though researchers have explored various subtasks related to this challenge. The *Hear&There* project [4] allowed users to record audio at a given GPS coordinate, while providing a spatial rendering of

[1] www.audioscape.org

other recordings as they navigated. Unfortunately, this was limited to a single-person experience, where the state of the augmented reality scene was only maintained on one computer. Tanaka proposed a peer-to-peer wireless networking strategy to allow multiple musicians to simultaneously share sound using hand held computers [6]. The system did not incorporate position awareness, but other work by the author [7] capitalized on location-based services of 3G cellular networks to provide coarse locations of users' mobile devices.

Position is, however, not the only type of data that has been explored in mobile applications. Projects including *GpsTunes* [5] and *Melodious Walkabout* [3] have used heading information to provide audio cues that guide individuals in specific directions. In fact, spatial audio and simulated acoustics can provide users with a wealth of information about the superimposed virtual audio scene. However, very few projects have used orientation information to *steer* audio propagation through virtual space. In our work (see www.audioscape.org for an overview), we have provided users with the ability to precisely control the direction in which they may emit or capture sound. Thus, by walking and turning, performers can steer their instrumental sound (e.g., harmonica) towards specific virtual effects units for processing. Virtual space thus becomes the medium for musical interaction, and the organization of musical pieces becomes spatial in nature.

3. APPROACH

From our experiences with the initial prototype we created, we discovered that GPS accuracy is a significant problem for augmented audio scenes where users move slowly. Consumer-grade devices provide readings with errors of about 5m in the best case [8] and 100m [1] in the worst case. Furthermore, the heading information that is inferred from a user's trajectory of motion requires the averaging of several measurements over relatively large distances, and can thus exhibit large latency and inaccuracies in pedestrian applications. For spatial audio applications, these delays and errors can deteriorate the quality of the experience. In particular, head tracker latency is most noticeable in augmented reality applications, since 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 can then begin to impair performance in localization tasks [2].

As a result, we propose a new strategy for the organization of virtual sound elements in an augmented reality scene where users move slowly. Our approach divides physical space into *interaction zones*, where musical material is clustered around a single location. When GPS readings indicate that a user has entered within a threshold distance of a such a location, their virtual position in the scene will be 'pulled' to that location over a short period of time, where they remain fixed until they enter another zone. In a sense, we create a discrete number of attractor locations that help to minimize the effects of GPS errors, and thus lock users into positions where interesting sonic material is present.

4. DEMONSTRATION

Our demonstration allows a user to explore a sonically augmented space, such as a small park or playing field. While roaming in that space, the user can enter different kinds of virtual zones, each one with a specific potential for sonic interaction – like a musical instrument. For instance, when inside an echo zone, sounds emitted by the user (or a previous user) will be heard echoing for several minutes, and musical layers can be built up. In a remixing zone, the user is surrounded by a number of synchronized playback voices which can be selectively listened to using the orientation of the listener's head.

The user is outfitted with a headset (plus mounted orientation sensor) and GPS receiver, all connected to a tiny wearable computer. An WiFi connection is established with a local (laptop) server, that maintains bidirectional audio streams and receives control signals. The audio scene is updated based on a user's current location, and spatial audio is rendered specifically for that user. For onlookers, or people waiting their turn, an additional audiovisual representation of the current scene can be displayed.

5. ACKNOWLEDGEMENTS

The initial prototype was produced in co-production with The Banff New Media Institute in Canada. The authors also wish to acknowledge the generous support of the NSERC / Canada Council for the Arts: New Media Initiative.

6. REFERENCES

[1] R. Bajaj, S. L. Ranaweera, and D. P. Agrawal. GPS: Location-tracking technology. *Computer*, 35(4):92–94, 2002.

[2] 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.

[3] R. Etter. Implicit navigation with contextualized personal audio contents. In *Adjunct Proceedings of the Third International Conference on Pervasive Computing*, pages 43–49, 2005.

[4] 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.

[5] 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.

[6] A. Tanaka. Mobile music making. In *Proceedings of New Interfaces for Musical Interaction (NIME)*, 2004.

[7] 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.

[8] M. Wing, A. Eklund, and L. Kellogg. Consumer-grade global positioning system (GPS) accuracy and reliability. *Journal of Forestry*, 103(4):169–173, 2005.

[9] M. Wozniewski, N. Bouillot, Z. Settel, and J. R. Cooperstock. Large-scale mobile audio environments for collaborative musical interaction. In *International Conference on New Interfaces for Musical Expression*, Genova, Italy, 2008.