Hear&There: An Augmented Reality System of Linked Audio

Joseph Rozier, Karrie Karahalios, and Judith Donath
Sociable Media Group
MIT Media Lab
20 Ames St.
Cambridge, MA 02139 USA
+1 617 253 5098
{jrozier, kkarahal, judith}@media.mit.edu

ABSTRACT

This paper presents an augmented reality system using audio as the primary interface. Using the authoring component of this system, individuals can leave "audio imprints," consisting of several layers of music, sound effects, or recorded voice, at a location outdoors. Using the navigation component, individuals can hear imprints by walking into the area that the imprint occupies. Furthermore, imprints can be linked together, whereby an individual is directed from one imprint to related imprints in the area.

Keywords

Augmented reality, audio imprint, navigation, authoring environment, audio links

Hear&There: An Augmented Reality System of Linked Audio

Joseph Rozier, Karrie Karahalios, and Judith Donath
Sociable Media Group
MIT Media Lab
20 Ames St.
Cambridge, MA 02139 USA
+1 617 253 5098
{jrozier, kkarahal, judith}@media.mit.edu

ABSTRACT

This paper presents an augmented reality system using audio as the primary interface. Using the authoring component of this system, individuals can leave "audio imprints," consisting of several layers of music, sound effects, or recorded voice, at a location outdoors. Using the navigation component, individuals can hear imprints by walking into the area that the imprint occupies. Furthermore, imprints can be linked together, whereby an individual is directed from one imprint to related imprints in the area.

Keywords

Augmented reality, audio imprint, navigation, authoring environment, audio links

INTRODUCTION

Augmented reality (AR) systems are designed to enhance the known physical environment. Typically, this is done by aligning an overlay of a virtual world onto the physical world. Most overlays used in augmented reality systems consist of video, audio, or a combination of the two. AR differs from the completely immersive world of virtual reality (VR) in that users are able to maintain the context of their surrounding environment, while still obtaining the benefit of additional sensory input and information.

Hear&There uses sound overlays to allow users to leave personal imprints that are associated with a particular location in physical space. An author creates a sound and embeds it at a location using an authoring toolkit. An individual can then hear that sound while he or she is traversing the designated location. *Hear&There* approaches augmenting the space with an emphasis on social interactivity. Unlike most systems, the users of the system create the augmented content of the space. A simple, yet flexible authoring system is thus an integral part of *Hear&There*. Furthermore, *Hear&There* does not require adding any physical objects (such as tags) to the environment. This makes it particularly useful in the augmentation of public spaces.

There are several fundamental issues we are addressing with *Hear&There*. How do we define an audio imprint? How do we embed it in a space? How are we made aware that it exists? How can we link sounds to compose a coherent path for users to navigate freely?

PREVIOUS WORK

Enhancing a space with audio is not new. Many museums today offer cassette tapes that direct a user through an exhibit while providing additional information about the various pieces of art. This approach is linear and somewhat inflexible: the user follows the guide on the tape and cannot stray to a different path. Bederson [1] built a prototype augmented reality system for museums that went beyond the traditional cassette-based systems by allowing for non-linear browsing. Using Bederson's system, the curators of a museum could place electronic transmitters above pieces of art, and then store audio information about each piece of art on a minidisk. As patrons using the system walked near one of the tagged pieces of art, the appropriate audio was recalled from the minidisk and played back through the patron's headphones.

Both of these approaches have some advantages. With the cassette tape, presumably a curator has taken the time to arrange a tour in a meaningful manner, making the order of the exhibit, as well as the content, an integral part of the presentation. With Bederson's system, the material is also created beforehand. However, the user can move back and forth, linger, and create his or her own tour. The visitor to a museum can now make the museum visit a very personal experience.

Another AR system that influenced *Hear&There* is the *Touring Machine* [2]. This is a visual AR system designed to explore an urban environment, specifically, the Columbia campus. Using this system, individuals are provided with information about the buildings he or she encounters on a walk through Columbia. Context sensitive web pages are accessible via a hand-held display. An important contribution of *The Touring Machine* is its reliance on the global positioning system (GPS),

which removes the need for placing physical tags in an area. GPS therefore allows greater flexibility in placement of augmented information.

One issue with the *Touring Machine* that we investigated in *Hear&There* is how much attention an augmented reality system should grab from the user. The goal of AR is to provide a layer of information over the real world. Determining how much this layer should distract from the real world is an important question in any AR project. In *The Touring Machine*, the textual information overlaying buildings and the hand-held device may become overly distracting. Because *The Touring Machine* was created as a way to provide information to newcomers, this approach may be appropriate. However, in *Hear&There* we attempted to make the AR layer a more subtle and integral part of the environment.

Most previous work in augmented reality has focused on the experience of the end-user. Content creation has been left to the designers of the system (as in *The Touring Machine*) or to professionals (as in Bederson's system). In *Hear&There*, allowing the users of the system to create content for the system was an important goal from the beginning. As a result, we have viewed the system as a tool that individuals can use to enhance an environment, rather than as a tool to provide particular information to end-users.

HEAR&THERE

Hear&There is being developed within a larger exploration of audio augmented reality. Many of the technical and design challenges, such as creating user-location triggered sounds or evaluating the importance of directional audio, are common to all of our applications. *Hear&There* focuses on the *in situ* placement of oral histories and reflections; this scenario emphasizes the development of a navigational system (for exploring a series of related stories) and an easy to use authoring environment.

The space currently used for developing *Hear&There* is a courtyard adjacent to the Media Laboratory. It is a locale that is frequently used by members of the community and it allows for casual encounters for it is also a throughway of the campus. Finally, it is an area that is constantly evolving: seasonally and architecturally. It is ripe for historical perspectives as well as reflection to those who visit it frequently.

Figure 1. *On the left is a GUI representation of the augmented environment. (E15 is the Media Laboratory.) The semi-transparent circles represent sound imprints. On the right is a photograph of the area we are augmenting, from the perspective of the Media Lab.*

The Sound Imprint

The basic unit in *Hear&There* is the sound imprint. The sound imprint is a multi-layered, customizable collection of sounds that can be placed in a space. An important design consideration for the sound imprint was utilizing the precision afforded us by technology that can provide the location of the user within tens of centimeters of his or her actual location. Audio AR systems such as [1, 3] could determine if the user was directly in front of an object or in a given room. The higher resolution available in *Hear&There* allows one to experiment with nuances of sound that were not previously possible.

The imprint consists of a single primary sound, with other audio braided into the periphery. These braids of audio overlap the imprint, with each braid of audio shifting into and out of prominence. The use of braided audio in addition to the primary sound allows for the creation of highly individualized imprints. For example, the primary sound may be the reflections of an individual on the nearby architecture. The author's favorite classical music could then be braided around the edge of the reflections.

It is important to note that most properties of these imprints are adjustable. The amount of braiding that occurs, whether the audio loops or plays once, and even whether a braid exists or not is left up to the author. The intention of our project is not to determine what is or is not an ideal imprint, but rather to provide the underlying structure and leave the form up to the person creating the imprints. This is similar to HTML (Hyper-Text Markup Language), in which the tools are provided for individuals to make a web page, but it is up to the author to make the web page interesting.

Navigating

To browse the sound imprints, the user carries the navigation toolkit as he or she travels through the augmented space. The toolkit, shown in Figure 2, consists of a pair of headphones with a digital compass attached, a laptop, a PalmPilot, a GPS receiver, a battery, and a microphone. Using the information from the GPS and the digital compass, the Hear&There system is able to determine the location and head position of the user with a high degree of accuracy. As the navigator encounters a sound imprint, he or she hears the imprint's audio. Because the head position is known, the sound can also be spatialized, which makes the sound "appear" to be coming from a particular location in space.

Figure 2. *The navigation toolkit, including a GPS receiver, a laptop computer, headphones with an attached digital compass, and a PalmPilot. Because the high-accuracy GPS receiver is rather bulky, we have placed it on a luggage cart that the user can roll around the environment.*

To aid the user in movement through a space, a PalmPilot is provided. The PalmPilot can show the distance and direction to imprints in an area. It is important to note that the information supplied on the PalmPilot is not necessary to use the system. Rather, it provides information that can aid the traveler through the augmented space. In future implementations, we will provide other, audio-only cues to the user about the presence of imprints in an area (such as "sound beacons" emanating from the center of the imprints). Both of these systems may be unnecessary for the bulk of users, as many people may wish to simply wander around an area and seek out the imprints without any help from the system.

Perhaps the biggest technical challenge in the navigation system we use is accurately sensing the position of an individual. We are currently using a GPS receiver to determine positioning. Using a system such as GPS is convenient for it allows for arbitrary positioning of sounds. The space does not have to be specially prepared with IR tags or other beacons. A sound can be placed on a busy street, by a statue, or at an arbitrary location in a field. Some of the drawbacks to GPS are that it can only receive signals outdoors and that it does not provide a consistent unique position. Upgrading to a real-time kinematic (RTK) GPS receiver dramatically increases the precision from stand-alone GPS. Further inconsistencies in position are accounted for in software. The emphasis of navigating with *Hear&There* is on the perceptual interface and not the computing and hardware behind it.

Authoring

Many AR systems have been developed for navigating spaces. It is rare to find one that also allows the user to easily create or modify a space [4]. With most systems, one must view or listen to material created by others. It would require programming experience and extensive knowledge of the internal structure of the AR system to create a new environment.

The goal with the authoring component of *Hear&There* was to create a simple tool that would allow an individual to build his or her own audio overlay on a physical space. That is, to allow one to record sound, place sound in the space, and then modify it or rearrange it quickly and intuitively.

The authoring component consists of a graphical user interface (GUI) and a mobile authoring device. The authoring GUI allows for far more options than a user might want on a brisk walk. The intention here is that the user will record the audio clips in the spots he or she chooses in the courtyard, and will later refine the spots using the graphical interface.

Figure 1 depicts the graphical user interface for the authoring component. It consists of a map of the Media Lab courtyard with superimposed sound imprints. This tool has two modes: a draft mode and a demo mode. The draft mode allows the author to create a new sound imprint, place it on the map, adjust the various audio properties, and record audio. This is done using a drag and drop interface written in Java.

The demo mode of the authoring interface allows the author to simulate the experience as if they were outdoors in the courtyard. By using the mouse to designate the navigator's position, the author can hear what the navigator would hear if he or she were following that path. In this manner, the author can experiment with different scenarios, audio braiding, and audio clips with the aid of visuals.

Allowing a person to record an imprint in the field is an important goal of this project. Because the audio imprints are contextualized to the space, it is reasonable for a user to create an imprint while standing or sitting in that space. To make adding imprints to a space as easy as possible, the interface for recording a sound in the field will be very simple. The user will hold a small microphone, and a very simple interface will be provided on the PalmPilot (with "record," "stop" and "cancel" buttons). The imprint will have default characteristics that can be adjusted later using the more complex GUI environment.

Linking

In order to enhance the sociable and informational aspects of the augmented space, the Hear&There system allows authors to link imprints together. On the World Wide Web (WWW), a user can create a page containing links to other sites he or she finds interesting or useful. Using this as inspiration, we added the ability for an author to link his or her imprints to any other imprints in the system.

As an example of how links could be useful in this environment, consider a hypothetical user from Brazil. This user decides to create an imprint near some trees in the Media Lab courtyard reflecting on the colorful New England autumn, and comparing it to the lack of season changes in her native country. She can then link this imprint to an imprint she created about the cold New England winter, and an imprint one of her Brazilian friends made.

Although the WWW was an inspiration for this project, the notion of linking in an augmented space poses a challenge. This is the fact that an augmented space is a true three-dimensional world, whereas the WWW is, for all intents and purposes, represented by a two-dimensional monitor. That is, when an individual viewing a web page follows a link, the linked page is loaded onto the same screen the individual was already using. In a system such as *Hear&There* where the imprints are contextualized to a particular area, the user must move to a new location in space to follow a link.

This poses a challenge in the design of the system, as mechanisms must be put in place to direct the user to the location of the linked imprints. In our current implementation, this guidance is provided by the PalmPilot interface. On the screen, only those imprints that have been linked to from the imprint the user was in recently are shown. Because we are exploring the use of audio only interfaces as well, the audio beacon idea described above would also work in this situation.

FUTURE WORK

One of the most important questions we are attempting to address with this system is the feasibility of navigating an augmented space using sound exclusively. Current AR technology could be used to provide a visual indication of imprint location. However, because the focus of *Hear&There* is on creating audio imprints in a space, providing such rich visual indicators would distract from the content. As the project develops, we will explore different ways of orienting a user in a space with audio. As a result, the use of this system by the visually impaired will become an interesting area to explore.

We also intend to more fully explore various notions of audio augmentation. In the current incarnation, imprints exist only as spots of sound. It will be interesting to explore other ideas of augmentation. One example we have considered is the concept of an "audio stream" in which a particular location consists of ever-moving snippets of sound. While standing in this stream, the user would hear sounds rushing past (using the spatialized audio system described above).

One of the limitations we have encountered in this project is the bulk of the system. To obtain the accuracy that we need, the GPS requires two large antennas and a bulky receiving unit. As a result, most of the system is contained on a luggage cart

that the user can move from spot to spot. (We chose not to use a backpack to avoid the issue of user fatigue from the weight.) In the future, we hope that more accurate GPS systems will approach the size of current off-the-shelf, but less accurate, implementations. Once this goal is realized, the entire Hear&There system could conceivably be enclosed in a Walkman sized device, allowing more natural use of the system.

ACKNOWLEDGMENTS
Many thanks to the members of the Sociable Media Group, the Digital Life (DL) and Things That Think (TTT) consortia of the MIT Media Lab, and many colleagues in the Media Lab for their suggestions and support.

REFERENCES

1. Bederson, B. Audio Augmented Reality: a prototype automated tour guide. *Conference companion on Human factors in computing systems* , 1995, Page 210.
2. Feiner, S., MacIntyre, B., Höllerer, T., Webster, A. A Touring Machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. *Proceedings of the First Annual International Symposium on Wearable Computing*, Cambridge, MA, October 13-14, 1997. pp. 74-81.
3. Mynatt, E., Back, M., Want, R., Baer, M., and Ellis, J. Designing Audio Aura. *Conference proceedings on Human factors in computing systems* , 1998, Pages 566 - 573.
4. Rekimoto, J., Ayatsuka, Y., and Hayashi, K. Augment-able Reality: situated communication through physical and digital spaces. *ISWC'98*, 1998.