An Investigation into the use of Spatialised Sound in Locative Games

Dr. Kirsten Cater
Department of Computer Science,
University of Bristol,
Bristol, BS8 1UB, UK
cater@cs.bris.ac.uk

Dr. Richard Hull
Hewlett-Packard Laboratories,
Bristol, BS34 8QZ, UK
richard.hull@hp.com

Tom Melamed
Hewlett-Packard Laboratories,
Bristol, BS34 8QZ, UK
tom.melamed@hp.com

Robin Hutchings
Department of Computer Science,
University of Bristol,
Bristol, BS8 1UB, UK

CHI 2007, April 28 – May 3, 2007, San Jose, USA
ACM 1-xxxxxxxxxxxxxxxxxx.

Abstract
This paper shows how spatialised sound can be used to guide users around a located gaming environment. Thus far, despite growing interest in delivering location-relevant media information to users, accurate delivery of virtual spatialised sound using limited-processing portable devices, such as Personal Digital Assistants (PDAs), has not yet been explored. The use of spatialised sound allows users to judge accurately which direction a virtual sound is coming from through a pair of stereo headphones. The initial findings of this research demonstrates that spatialised sound can be used to navigate users in a locative game running on a PDA.

Keywords
Spatial Audio, Location Based Media, GPS, Immersion.

ACM Classification Keywords
H1.2 User/Machine Systems; H5.1 Multimedia Information systems; H5.2 User Interfaces.

Introduction
With the recent introduction of portable gaming devices with integrated GPS and the launching of a variety of Location Based Services (LBS) along with the popularity of Geocaching, it will not be long for located media games to be the next fad in the public domain. There are already located games on the market, for example *'The Shroud'* created by Your World Games [13],

'RealReplay' created by Mopius [11], both for GPS enabled mobile phones, and *'Colors'* created by Tiger Telematics for the Gizmondo [2], who stated on their website that *'Basing a mobile game on your actual location is more than the thinking of a great mind: it will change gaming for good. GPS will be for the games of tomorrow what 3D was for the games of yesterday'*. There are many more examples of one-off experimental installations due to this revolutionary technology having been in the research domain for slightly longer than the commercial one, for example '*Can you see me now?*' [1], *Savannah* [4] and *Cititag* [14]. It is important to note that these applications, both commercial and experimental, typically only use GPS-derived maps for visual content navigation rather than providing cues within the located media itself.

Researchers looking into locative gaming are already stating that these games can change the perception that a person has of a particular area through merging the virtual with the physical reality. This gives users of mobile games a chance to withdraw from their 'serious life' in a familiar space to travel to a new, unpredictable, unexplored space, thus blurring the known space with the un-known [3]. To achieve a better, more immersive, experience for the user in a game, or any virtual experience, the author should try to match the real world scenario that it is trying to imitate. Therefore it is hypothesised that an even greater level of immersion in mobile gaming could be achievable by using spatialised sound, thus morphing the virtual reality with the real reality!

Navigation for Locative Gaming

Most navigation cues in a game are delivered via a graphical interface; however this is inappropriate with use of pervasive computing as the user would have to try to keep their eyes on both the screen on the handheld device and the surrounding environment to prevent themselves walking into a physical obstacle such as a tree [9]. This can lead to an uncomfortable experience for the user; therefore an auditory navigation system would seem to be more appropriate as the user can receive directions without being visually distracted from the physical environment. However, most auditory systems guide the participant by using precise spoken instructions which require high attention levels. The major disadvantage of this method in a pervasive game is that without use of compass there is no way of knowing which direction the user is facing even though we known their location. Thus if the spoken guide told the user to *'turn left'*, the direction in which they would turn would depend on the way that the participant was facing, i.e. they could end up going in completely the wrong direction. By using a compass in a pervasive game we are not only able to deliver the correct spoken navigation but furthermore guide the users by providing them with an awareness of where the target destinations are located through spatialised sound as will be shown in this paper.

Spatialised Located Sound

Spatial or 3D audio, is sound as we hear it in everyday life. Sounds come at us from all directions and distances, and individual sounds can be distinguished by pitch, tone, loudness, and by their location in space. The spatial location of a sound is what gives the sound its three-dimensional aspect. A sound generated in space creates a sound wave that travels to the ears of the listener. When the sound is to the left of the listener, the sound reaches the left ear before the right ear, and thus the right ear signal is delayed with respect to the left ear signal. In addition, shadowing from the head causes the right ear signal to be attenuated. Both ear signals are also subject to a complicated filtering process caused by acoustical interaction with the torso, head, and in particular, the external ear [7]. Thus we unconsciously use the time delay, amplitude difference, and tonal information at each ear to determine the location of the sound. To process conventional stereo signals to give an enhanced audio field is relatively uncomplicated, for example multi-speaker 'surround sound' systems have been around in various incarnations for several years and have now been produced cheap enough that most people will have a 3D audio system setup in their home. However, these systems do not create a truly 3D audio environment as the listener cannot perceive audio coming from an arbitrary point in space. This is harder to achieve but with present advances in both of our knowledge of the human auditory system and the progression of processor power, it is now possible to realistically create spatialised sound by using several specialised filter functions. These are called Head-Related Transfer Functions (HRTFs) [6]. Therefore, by creating genuine 3D spatialised audio a truly compelling experience is possible, which can leave the user questioning whether they were hearing a real sound from a particular location or whether it was a virtually located sound.

The present challenge however is transposing these functions onto a limited processor device such as a Personal Digital Assistant (PDA) as digital signal processing needs a high volume of computing power along with the fact that the audio filters have to be applied in real-time to the audio. In order to implement this system on PDAs efficient, quick and high-quality audio filters need to be developed.

Implementation

This paper will present the findings from a trial organised to look at the possibility of creating rudimentary spatialised sound by using a basic 2-axis compass, a Bluetooth GPS, an HP iPAQ and programming the base spatial filters needed for a single user. It will then go on to describe the future work which will be carried out in the next few months to expand this work further.

As mentioned above the most accurate method at present for realistically creating spatialised sound is by using HRTFs however this method had to be optimized due to the limited processing power of the HP iPAQ being used. Thus a more simplified approach was taken implementing the following functionality:

- *Time Delay* – When a listener turns their head away from a sound source the time taken for the sound waves to reach their left and right ear changes. There will therefore be a time delay between when the sound reaches the left ear and when it reaches the right ear. This is known as the Interaural Time Delay (ITD). The ITD is at maximum when the angle between the listener and the sound source is 90 degrees. The head is thus modeled as a sphere and the time delay between the left and right ear is calculated, using this time value the sound source is delayed to the necessary ear.
- *Panning* – As well as a time delay the listener experiences a drop in intensity of the perceived sound in the ear facing away from the sound source. This is caused by a head shadow effect (the head obstructs the sound waves before they reach the

ear). To model this FMOD panning [5] was used, and in a similar way to the time delay it is at its maximum at 90 degrees.

- *Scaling* – To simulate front and back directional information a volume queue is added. As the listener turns away from the source the volume of the sound in both ears is decreased. This is done in a linear fashion, where the decrease is proportional to the angle between the user and source. The volume will therefore be a minimum at 180 degrees, i.e. when the user is facing directly away from the sound source.

This system was then trialed using a simple locative game with 15 users in College Green in Bristol, UK. The goal for the users in the located game was to navigate their way around the physical environment by listening to the virtually spatialised sound files located around them. When the user came across a sound file for the first time they had to navigate their way to the centre of where that sound source was located. Each sound file covered a 55m radius area, with the centre region being set to a 5m radius area. There was nothing physical in the space to tell the users where the regions were, they had to rely on their spatialised sound to guide them. Once the user correctly found the centre location of that sound they received a cheering sound file to tell them they had correctly navigated to the centre and then the next spatialised audio file would start. There were five located sound files in total for the users to find and only one sound file would be played at a time. Once the participant had found all five virtual sound files the user would receive the final cheering sound file and an image on the PDA to tell them that the trial was over. The equipment used to trigger the sound files was a 2-axis compass, a Bluetooth GPS and a HP iPAQ as can be seen in figure 1. The located spatialised audio game was created using the Mobile Bristol Authoring Tool [10]. The layout of the source sound files can be seen in figure 2, the black lined circles show where the 5 sound files are located and the red circles are the centre regions for each of those sound files. The blue regions are paths and physical objects which helped to locate the regions when authoring the locative game.

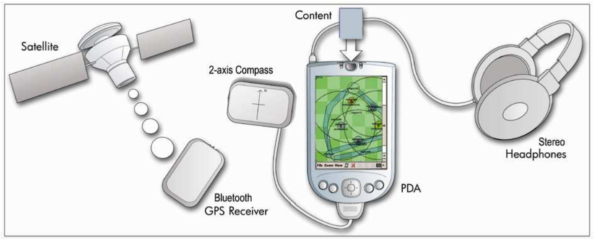

figure 1. Image to show the equipment used in the trial.

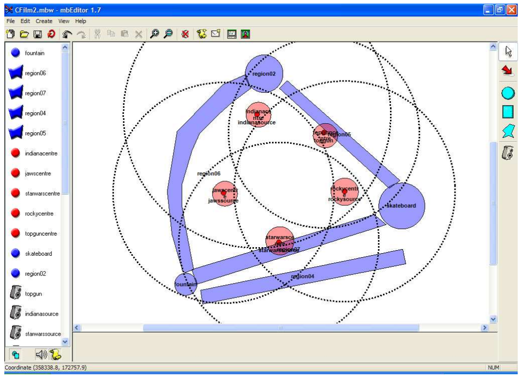

figure 2. Layout of the Located Spatialised Sound Game in College Green, Bristol, UK.

Results

The results of the trial proved very promising, all 15 participants of the trial managed to correctly navigate their way round the physical environment to find the centre of the virtual sounds simply by relying on the spatialised sound to guide them. This has several implications for future HCI research, mostly because it demonstrates that there is no longer a need for complicated spoken navigational cues in locative gaming, instead the user can be guided to their target destinations using the much simpler method of spatialised sound which should also improve their level of immersion as seen in indoor virtual reality applications [8]. However, there was a major weakness that became apparent with the first users in this trial and that was the type of compass used. The compasses being used were only 2-axis and therefore they were very sensitive to tilt, thus any movement by the user from the vertical, which it transpired occurred quite often, meant that the calculated direction that the user was facing would be incorrect. Once we had realized this was occurring all participants were asked to keep their head as straight as possible to minimise any tilt. Once we had told the users of this problem they managed to find the centre of the sound regions much quicker as they weren't being distracted by incorrect directional readings.

Improvements

The initial findings using the 2-axis compass were very promising however due to errors occurring when the user tilted their head there is still room for improvement. We were surprised about how easy it was for the users to find the location of the virtual sound simply by using base filters to mimic spatialised sound rather than trying to implement efficient, high quality audio filters similar to HRTFs to run on the PDAs. Thus the next stage for this research is to focus on fixing the errors from the head tilt which is a higher priority than improving the audio filters which create the spatialised sounds which seemed to work well from the results of this trial.

We have recently gained access to some state-of-the-art 3-axis compasses, which have been developed as a result of collaboration between the University of Bristol and Hewlett Packard Laboratories, thus this will prevent the errors of head tilt from occurring and thus give the users a more accurate experience. Once we have the software working to interpret the users head-orientation from the new 3-axis compasses, along with the users location from the Bluetooth GPS receivers it will then be possible to play the audio to the participant through the headphones in order for them to receive accurately spatialised sound. We hope to devise a more complex located game that will test the user's ability to be guided around their physical environment by using spatialised sound and thus will test the full potential of the system. This will be trialed to a group of invited guests who are new to the technology, but not inexperienced in using computer games. The game will be run with and without the use of spatialised sound to understand its potential. Participants will be asked to fill out questionnaires and interviews will be conducted to ascertain their evaluations of the experiences. The users' movements throughout the game will also be traced in order to determine directional effectiveness and speed, which weren't recorded for the first trial.

Immersion

The level of immersion felt by a user in an experience is deemed a very important factor to the success of a virtual application - this is where the user's full attention has been captured by the experience i.e. they have become immersed in the virtual world. Research has proven that by accurately synthesizing spatial sound, the immersiveness of a virtual reality environment can be increased [8]. Thus by using spatialised sound on a PDA for a located game it is hypothesised that the users would then become more immersed in the experience. For the next stage of this research the trial will include an immersion questionnaire to backup this hypothesis, i.e. to see if spatialised sound really does heighten the user's sense of immersion as it does with other virtual reality environments proven by the research carried out by Hendrix and Barfield [8].

Conclusion

We hope to have shown in this paper the potential of using spatialised sound for locative gaming, as well as showing promise of advancing the system further to enhance the experience of the user. This project has demonstrated that users can indeed be guided around a virtual environment layered onto a physical one with the use of spatialised sound, allowing users to judge accurately which direction a virtual sound is coming from through a pair of stereo headphones. However, this trial has only scratched the surface of a very interesting and relevant topic for HCI research and may produce some interesting results that will inform future applications to investigate the possibility of using spatialised sound to guide visually impaired people safely around indoor and outdoor physical environments whilst preserving their natural auditory perception [12].

Acknowledgements

We would like to acknowledge the rest of the Mobile Bristol Research team for their support in this work. Dr. Cater would also like to thank The Nuffield Foundation Newly Appointed Lectureship Scheme for help to fund this research.

References

[1] Benford, S., Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., Tandavanitj, N., Adams, M., & Row-Farr, A. Can you see me now? In *Proc. CHI 2005*, ACM Press (2005).

[2] Colors - Gizmondo's GPS game http://www.gizmondo.com/news/item.asp?id=457

[3] De Souza e Silva, A. Location Based Mobile Games: Blurring the borders between physical and virtual spaces, ISEA (2004).

[4] Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., & Kirk, D. Savannah: mobile gaming and learning? Blackwell Publishing Ltd, Journal of Computer Assisted Learning 20, 399–409, (2004).

[5] FMOD - http://www.fmod.org/

[6] Gardner, W. G., & Martin. K. D. HRTF measurements of a KEMAR. Journal of the Acoustical Society of America, 97 (6), 3907-3908, (1995).

[7] Gardner, W.G. Spatial Audio Reproduction: Toward Individualized Binaural Sound. The Bridge, V34 – 4, (2004).

[8] Hendrix, C. & Barfield, W. The sense of prescence within auditory virtual environments. Prescence: Teleoperators and Virtual Environments, 5, 274-289, (1996).

[9] Kristoffersen, S. & Ljungberg, F. "Making place" to make IT work: empirical explorations of HCI for mobile CSCW. In Proc. GROUP'99, conference on Supporting group work, ACM Press (1999).

[10] Mobile Bristol - http://www.mobilebristol.com

[11] RealReplay by Mopius http://realreplay.mopius.com/

[12] Strothotte, T., Petrie, H., & Reichert L. Development of dialogue systems for a mobility aid for blind people, In Proc. ACM conference on Assistive technologies, 139-144, (1996).

[13] The Shroud by Your World Games http://www.shroudgame.com/

[14] Vogiazou, Y., & Raijmakers, B. Urban space as a large-scale group playground. Ubicomp, 54-55, (2004).