Urban Sound Gardens: Supporting Overlapping Audio Landmarks in Exploratory Environments

Yolanda Vazquez-Alvarez[1†], Ian Oakley[2] and Stephen Brewster[1]

[1] Glasgow Interactive Systems Group, Computing Science
University of Glasgow, Glasgow G12 8QQ, UK
www.gaime-project.com
{yolanda, stephen}@dcs.gla.ac.uk

[2] Madeira-ITI, University of Madeira
Funchal 9000-390, Portugal
www.m-iti.org
ian@uma.pt

ABSTRACT
Sound gardens rely on location tracking technology to situate audio landmarks in public spaces, typically urban parks. These landmarks are surrounded by a proximity zone, and content relating to them is displayed to users within this region. In densely populated sound gardens, overlap between zones is inevitable and will result in simultaneous playback of multiple audio sources – a potentially confusing situation. We propose an urban sound garden design featuring both overlapping proximity zones and spatialized animal sounds to attract the user's attention to particular landmarks in a non-guided exploratory environment. We first present a detailed description of our sound garden design and then report results from a user study in which our proposed sound garden design provided a much greater sense of discovery and immersion when compared to three less complex designs.

Keywords
Sound garden, proximity zone, spatial audio, audio landmarks.

1. INTRODUCTION
A sound garden [1] is a virtual audio environment composed of a set of precisely situated sounds spatially overlaid on an urban park. Using GPS or WIFI technology, the position of users in relation to specific audio landmarks is tracked and information related to their proximity to these landmarks is presented. A sound garden is usually intended for users to casually explore and experience rather than navigate via predefined paths. The unstructured nature of this activity presents unique challenges for the design of audio feedback to support exploration. Fundamentally, individual landmarks need to advertise themselves both to attract the user's attention and support subsequent targeting. This is typically achieved using an acoustic beacon– sounds which activate when a user is within a specific distance from a landmark [2]. Two concentric levels of beaconing feedback are beneficial, the first in a wide proximity zone and the second in a narrower activation zone. The goal of audio cues in the proximity zone is to provide unobtrusive audio guidance, which enables a user to move towards the activation zone. Once this inner zone is successfully reached, additional content is made available to the user, either to indicate that a landmark has been found or to provide structured information describing it. Beyond other design considerations, the size of these zones is influenced by the accuracy of the tracking technology: larger errors (e.g. conventional GPS) require larger zones for stable results, while improved tracking systems (e.g. differential GPS) are able to effectively operate with smaller zones.

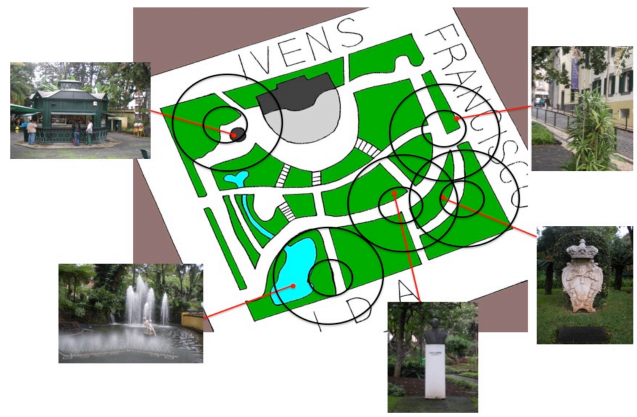

Figure 1. Municipal Gardens in Funchal, Madeira. Still images of the landmarks and illustration of proximity and activation zone per landmark.

This paper considers a sound garden using conventional GPS technology (widely available and with an accuracy of approximately 10m) and featuring a dense population of landmarks in order to provide a rich user experience. In such a system, it is inevitable that the audio activation zones around multiple landmarks will overlap, as illustrated in Figure 1. This is especially likely in sound gardens where users are not expected to follow a predefined route. Managing the resultant simultaneous playback of sounds requires careful design of cues and interaction – for example, playing two or more conflicting speech samples may overload users and may not be desirable. One way to manage the presentation of this information is via content minimization and representing sounds with either Earcons or Auditory Icons [e.g. [3]]. This paper explores and extends these techniques. It proposes the use of symbolic earcons [4], environmental audio sounds which rely on abstract mappings loosely associated to landmarks (providing a low impact cue to presence and location), and spatial audio to make audio landmarks appear to originate from different locations. Ultimately, we aimed to answer the following questions: 1) how useful were the symbolic earcons? Did users find them appropriate for disambiguating the audio landmarks? 2) To what extent did the use of spatialization and spatial audio aid users when they were faced with overlapping audio landmarks? Did these techniques help augment the experience?

† Thanks go to the Ken Browning Scholarship, University of Glasgow.

2. FUNCHAL'S URBAN SOUND GARDEN

To explore these issues we created a virtual sound garden in the Municipal Gardens in Funchal, Madeira. The sound garden ran on a Nokia N95 8GB using software adapted from the Mobile Trail Explorer[1] application together with the Head Related Transfer Functions (HRTFs) [5] and the JAVA JSR-234 Advanced Multimedia Supplements API to position the audio sources. The user location was determined using an external Qstarz BT-Q1000X Bluetooth GPS receiver, and head orientation (compass heading) was determined using a JAKE[3] sensor pack also connected via Bluetooth. They listened to the sounds planted in the garden using a pair of headphones. The GPS receiver was placed on the headphone's left ear-cup and the JAKE on the crown of the head, in the middle of the headphone's headband. Both sensors were mounted using Velcro tape (see Figure 2). No pre-determined route or visual aids such as maps were provided to the user.

2.1 Audio Landmarks: Content and Configuration

Five different symbolic earcons in the form of recordings of animal sounds, i.e. owl, goose, cricket, nightingale and frog, were created to alert the user of the presence of five physical landmarks, i.e. the Rua Sao Francisco; a Coat of arms of Saint Francis convent; the Statue of Joao Reis Gomes; the café and the pond (see Figure 1). Animal sounds were used to identify landmarks because they seemed a good fit to the natural environment. For each landmark, a brief audio clip containing factual information about the site was synthesized using Cereproc's (www.cereproc.com) British English male RP voice. Both the animal sounds and the audio clips were mono, 16-bit, sampled at 16 kHz and normalized to a conversational volume level (approx. 60-70dB).

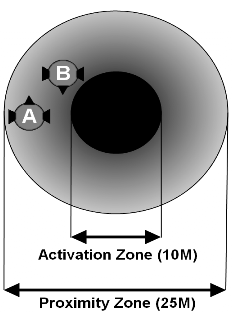

Figure 3. Audio landmark, gradient indicates volume. User A (looking up in figure) hears a quiet sound to the right; User B (looking down) hears a louder sound front left.

Two circular zones surrounded each landmark: activation (radius 10m) and proximity (radius 25m) zones. Due to the size of the garden (82m x 109m), only three landmarks had overlapping proximity zones while the other two were isolated. Figure 1 shows the audio landmark configuration.

When the user entered the proximity zone, the symbolic earcon, i.e. animal sound, corresponding to that landmark was triggered (see Figure 3). The animal sound increased in loudness and updated its spatial orientation as the user walked towards the landmark. Once in the activation zone, the audio clip could be played (and the animal sound stopped) by pressing the central navigation button on the mobile phone.

3. USER STUDY

Eight users (6 male, 2 female, all right-handed) participated in a study comparing a fully spatialized audio landmark configuration, as described in the previous section, to three other less complex alternatives: 1) Symbolic earcons and limited audio spatialization (distance only), 2) Symbolic earcons but no audio spatialization, and 3) No symbolic earcons or audio spatialization, only audio clips with landmark information. Two different participants tested each configuration. When wandering around the park, the use of earcons made the discovery of landmarks less "abrupt" and more "lively". In situations with multiple sound sources, although users did find overlapping sounds harder to deal with, they reported that "overall the localization was easy" and heading information helped. Hearing overlapping sounds at a distance that the user had already heard offered the added benefit of "familiarization with the surroundings". Overall, when spatialized audio was used, participants reported a greater sense of "discovery" and "immersion", and therefore spent longer in the garden (on average 21 mins when spatialized compared to 11.49 when not spatialized). Users also reported the experience to be playful and engaging.

Figure 2. Experimental setup. 1) JAKE sensor, 2) GPS receiver, and 3) mobile device.

4. CONCLUSIONS

This paper presented an urban sound garden design supporting overlapping audio landmarks (with a mix of earcons, spatial audio and speech content). Results from an initial user study showed that our sound garden design provided a greater sense of immersion, discovery and playfulness even in a densely populated audio space. It also allowed users to discover the content and their surroundings at their leisure with no guidance or help from visual aids like maps. We believe these results show that an exploratory sound garden design like the one described in this paper can create a rich and compelling locative mobile audio environment.

5. REFERENCES

[1] Shepard, M. Tactical Sound Garden [TSG] Toolkit. *3rd International Workshop on Mobile Music Technology*, Brighton, UK (2006).

[2] Walker, B. N. and Lindsay, J. Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice. *Human Factors*, 48, 2 (2006), 265-278.

[3] Stahl, C. The roaring navigator: A group guide for the zoo with shared auditory landmark display. In *MobileHCI 2007* (Singapore, 2007). ACM Press, New York, NY, 282–386.

[4] McGookin, D. K. *Understanding and improving the identification of concurrently presented earcons.* PhD, University of Glasgow, Glasgow, 2004.

[5] Begault, D. R. *3D Sound for Virtual Reality and Multimedia*. Academic Press, Boston, MA, 1994.

[1] http://code.google.com/p/mobile-trail-explorer/

[3] http://code.google.com/p/jake-drivers/