Audio Bubbles: Employing Non-Speech Audio to Support Tourist Wayfinding

David McGookin, Stephen Brewster and Pablo Priego

Department of Computing Science
University of Glasgow, Glasgow G12 8QQ,
{mcgookdk, stephen}@dcs.gla.ac.uk, www.dcs.gla.ac.uk/~mcgookdk

Abstract. We introduce the concept of Audio Bubbles - virtual spheres filled with audio that are geocentered on physical landmarks, providing navigational homing information for a user to more easily locate the landmark. We argue that the way in which tourists navigate is not well supported by traditional visual maps, and that Audio Bubbles better support the serendipitous discovery and homing behaviours exhibited in such tourist activities. We present a study comparing Audio Bubbles to a visual map in a real world navigation task. Navigation with Audio Bubbles appeared to be faster and was preferred by most of the participants. We discuss the findings and outline our future development plans.

Keywords: Non-visual Navigation, Wayfinding, Auditory Display

1 Introduction

The evolution of hand-held mobile computers into generic, powerful, location-aware connected devices, has dramatically increased their utility as mobile navigation and tourism guides. This has allowed dynamic contextual information to be easily provided to the user. However, most of these applications, whilst allowing dynamic overlay of tourist related information, such as restaurants, and other points of interest, present this information on a visual display. Most commonly this is achieved by overlaying the information onto a visual map. This requires the user to look at the map to locate, and then navigate towards, points of interest. The user is required to make an explicit decision to find a feature on the map and then determine how to navigate towards it. Whilst this is useful when the user has a pre-determined goal in mind, there are many situations when being a tourist where there are no concrete destination goals.

Brown and Chalmers [1] carried out an ethnographic study of tourist wayfinding and navigation in Edinburgh in the UK. Visiting activity was found to be planned, but only up to a point. They found that the user would "wander" along the streets and serendipitously discover interesting sights and sounds. In picking a restaurant for example, tourists would select an area with a number of restaurants and wander about until they found one that met their needs. The user would not find a particular restaurant in a guidebook and then constantly refer to a map to head towards that restaurant. Similar requirements, that mobile

guides should not be over controlling or constraining, were found by Kassinen [2] and Cheverst *et al.* [3]. However, Brown and Laurier [4] identified situations where a tourist would study a map out of context in an effort to learn an unfamiliar area and its points of interest. An obvious conclusion to draw is that the user does not want to be explicitly burdened with constant referral to a visual map, but at the same time does not want to miss important (to the tourist) points of interest in an area that is being explored. This is something that existing visual maps fail to adequately support, primarily as the user must make an explicit decision to refer to the map, which is unlikely in such a "wandering" scenario.

Another situation where traditional visual maps fail to properly support wayfinding is when the user must relate a position of a point of interest in a printed map or guidebook, to his or her current position in the environment [1]. This context change is most common at the very end of a route. For example, identifying which building on a street was the birthplace of a famous author, when the user is standing on the same street. At this point, traditional street level maps provide poor support for the user to "home in" on the desired point of interest. Brown and Chalmers [1] describe an example of finding a particular building in amongst other buildings, but many other points of interest may be equally unobtrusive, and considerably smaller in the environment (e.g. the zero kilometer point in Place du Parvis in Paris - represented by a small brass disc on the ground). Again, visual maps are not particularly suitable to support this activity.

Both of these situations, wandering and homing, are related - the user may not have an explicit goal in mind, preferring to find interesting features as he or she walks. However, when the user identifies an interesting feature or landmark, perhaps from a street sign, a handed paper flyer, or the pre-study of guidebooks [4], it can be difficult to use a visual map to home in on the point of interest. In both cases the visual nature of maps (both paper and electronic) means that they are not the most appropriate way to support these types of activity, either due to context changes, or that when wandering there would be little need for constant referral to maps.

2   Audio Bubbles

To better support the kinds of activities previously mentioned, we have developed a technique called Audio Bubbles. These provide serendipitous, non-visual awareness and "homing" information onto nearby points of interest, without the explicit user intervention that would be required with a visual map. An Audio Bubble is a virtual sphere geolocated on a real world point of interest. Within the bubble audio information is played. This audio is manipulated to indicate the user's distance from the point of interest that the bubble is centred upon. As a user walks around the environment, he or she will come into contact with the boundary of the Audio Bubble, and become aware of a nearby point of interest, or that a pre-determined point of interest is being approached. Using

Fig. 1. Right: A screenshot of the visual display showing the current user location (A) and the location (B) and name of the next landmark). **Left:** A picture of the landmark.

the distance information encoded in the audio, the user is guided the last few meters onto the point of interest. In this way, the Audio Bubble acts as a form of augmented reality, overlaying virtual audio in the soundscape of the environment, and increasing awareness of nearby interesting things that can either be attended to or ignored, without the need to make the explicit decision to look for points of interest that is required with a visual map.

We explicitly make two decisions with Audio Bubbles and our future implementations. Firstly, that the sound we use within the bubble is unspecified. There are several types of sound that could be used to fill an Audio Bubble, however what sound is used will depend on the use to which the bubble is being put, and the point of interest it represents. In some cases a sound representative of the point of interest would be most appropriate, but as with many other areas of auditory display, such sounds may not be possible and a different kind of sound may need to be chosen. We discuss this issue later in the paper. Another decision with the design of Audio Bubbles is that they should augment the existing soundscape in an environment. Both Brown and Chalmers [1] and De Certeau [5] point out that the sights, sounds and smells - the ambiance of an area - is one of the main aspects of a location a tourist wishes to experience.

3   Related Work

Whilst there has been a great deal of work on visual navigation (e.g. [3] [4]), less has been done on auditory navigation. The bulk of this work on auditory navigation has investigated the presentation of route and nearby environmental information to people with visual impairments (e.g. [6]). One of the first systems for sighted users was Holland *et al's.* [7] AudioGPS system. This used spatial panning, with the user wearing headphones to communicate accurate direction of a specified point of interest. To communicate distance, a Geiger counter metaphor was used, with a sound repeating at an increasing rate as

the user became closer to the point of interest. Whilst informal comments on AudioGPS proved positive, no formal evaluation was carried out. Strachan *et al's.* [8] gpsTunes worked in a similar way, but modified the perceived spatial location of music that the user listened to to provide bearing information and modified volume to provide distance information. Again, the user had to wear headphones to support the spatialisation. An informal field evaluation was carried out which showed promise that the audio manipulation may be useful in navigation. Reid *et al's. Riot! 1831* [9] study used similar techniques to present an auditory soundscape of the Bristol riots of 1831 in the streets of modern day Bristol. As users walked around a city square wearing headphones and a GPS equipped mobile device, their current locations were used to trigger sound effects and vignettes of real world events from the riots. Reid *et al.* found that their system provided a deep level of immersion within the experience. Stahl's [10] Roaring Navigator used a similar technique, with spatialised sounds of animals being used to indicate the directions of, and guide visitors to, various enclosures in a zoo. Software running on HP iPaq mobile computers was used to determine the user's position and to adjust the sounds so that they appeared to be coming from the direction of the appropriate enclosure. Participants could either browse, hearing the nearest four enclosures, or could navigate by hearing only the destination sound source. Again, Stahl's system required the user to wear headphones, but shared the audio being played so that users could interact with each other and use the sound as a shared reference point. Stahl notes that in audio museum guides, visitors often do not interact with each other due to the wearing of headphones causing social isolation. An evaluation of the system was performed with encouraging user feedback, however Stahl did not compare performance with a standard visual map of the zoo.

Whilst these previous systems show that there are uses of auditory wayfinding, they also have issues when being applied as solutions to the issues of tourist navigation discussed in the introduction. Firstly, they all use spatialised audio, which works by varying the waveforms that reach the left and right ears, allowing the user to determine the direction of the sound source. However, such spatialisation requires the user to wear headphones. Headphones can have the effect of blocking out the existing auditory soundscape. In some cases this may be desirable, such as on a crowded commuter train; in other cases it may be less desirable, such as wandering through the streets of an historic Italian town. It is very much a preference of the user whether to wear headphones or not. Additionally, as noted by Stahl, wearing headphones can cause the wearer to be more socially isolated from his or her fellow travellers. Given that travel is a social activity [1], this social isolation would not be desirable. Whilst there are other types of headphone such as bone conductance [11], which do not occlude the ears and other sounds reaching them, they are not common or widely available. As we would like Audio Bubbles to augment the existing environment (and certainly not block it) and that some users may not wish to wear headphones, we consider that monaural presentation (without headphones) via the loudspeakers on the mobile device presents a "worst case" scenario, and worthy of investiga-

tion. Secondly, whilst users can perform navigation using the audio, none of the studies discussed have compared performance between visual only and auditory navigation. Does the addition of sound lead to superior performance in near field navigation over navigation with a visual map? Finally, the sounds used in these systems always present some sort of audio no matter how far away the user is from the point of interest. As already described, this is unlike Audio Bubbles, which have a defined outer exterior beyond which no sound is heard. This is important as if the user has no defined route, such as when wandering, the audio that is presented must be controlled in some way so that too much is not presented concurrently, and the overall effect is not cacophonous [12].

4 Evaluation

In our initial evaluation we chose to investigate only the "homing" scenario as discussed in the introduction. Primarily, this is because we wanted to gain feedback on the size and utility of the Audio Bubbles. Once we have investigated this scenario we can consider more fully the utility of the Audio Bubbles and investigate the wandering scenario, where the user has no fixed destination in mind. Given the discussion of the previous work, we sought to answer two related research questions about Audio Bubbles:

RQ1: Can Audio Bubbles presented using monaural (without headphones) sound be effective navigation aids to help users locate points of interest?
RQ2: Do Audio Bubbles improve navigation performance over a visual only map?

Our initial Audio Bubbles implementation was developed as an application on a Nokia N810 Internet Tablet. We used the built-in GPS system on the tablet to determine the position of the user. The visual display of the device showed a scrollable overview map of the area (obtained from `www.openstreetmap.org`), the current user position and one or more landmarks (see Fig.1). The Audio Bubbles were implemented along the Geiger counter principle of Holland *et al.* [7]. Distance was mapped to the repetition rate and volume of a short "click" sound. The closer the user was to the point of interest, the louder and more frequently the click was played. We set the initial boundary to be 250 meters from the point of interest, and no sound was played out with this. To evaluate our initial Audio Bubble implementation, we carried out a within groups study on a route navigation task. Eight participants (4 women and 4 men) aged 24-27 took part in the experiment. All of the participants reported normal hearing and normal, or corrected to normal sight. None were paid for participation. Each participant walked a route using a visual map representation (Visual Condition), or the same map representation but with each landmark encased in an Audio Bubble (Audio Bubble Condition). The order in which participants carried out each condition was counterbalanced. Two different routes (of length 940m and 810m respectively) around different landmarks (4 and 3 respectively) in a local park were used. Routes were counterbalanced between conditions to avoid
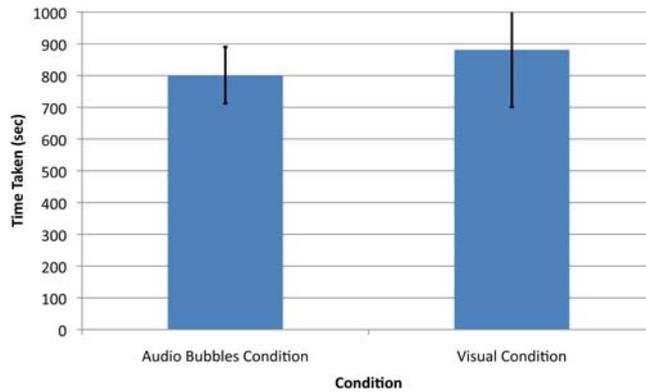
Fig. 2. Graph showing the mean time taken in both the Visual and Audio Bubbles Conditions. Shown with one standard deviation.

difference effects. We chose the park as it has a number of small points of interest, such as statues, that are easy to miss even when close by. These are the sorts of points of interest that we consider the Audio Bubbles to be particularly useful to find, as they are not obvious in the environment. This is similar to the situations where a user is trying to home onto a very old pub, or a niche museum in a city centre, but keeps walking past as these are unobtrusive landmarks in the environment. None of the participants were regular visitors to the park. Participants were shown how to use the system for the current condition and told that they should navigate to the landmarks given. The position of the first landmark and its name was displayed on the map, as was the participant's current location (see Fig. 1). In the Audio Bubbles condition only the Audio Bubble for the current point of interest was active. As the participant entered a zone approx 10m around the landmark, a trumpet sound was played, signalling that the landmark had been correctly identified. The next landmark was then shown on the display. When all landmarks had been found, the participant was given a message that the experiment was over. At the end of the experiment the participant was verbally debriefed and asked for any comments.

4.1 Results

As participants were required to reach each point of interest, accuracy was uniform. In addition, we recorded the time taken to complete each route. One participant's results were found to be over two standard deviations from the mean for each of the conditions. We judge this participant as an outlier and exclude those results from the analysis. The mean time taken for the other seven participants is summarised in Fig. 2. As the groups were not fully counterbalanced we elected not to perform a formal statistical analysis of the time taken. However the results do seem to indicate that there was some influence of the Audio Bubbles in reducing the time taken to find the points of interest.

4.2 Qualitative Results

The interviews with participants after the experiment yielded valuable insights into the usefulness of the Audio Bubbles, and how they compared to the Visual condition. Participants were asked a number of open-ended questions about their experience with both of the systems. These included the usefulness of the Audio Bubbles, as well as how the sounds changed as the participant moved around the environment. The participants were also asked for any further issues that they felt relevant.

Five of the eight participants stated that they found the Audio Bubbles helpful in trying to locate the points of interest. One participant (the participant whose results were excluded from those in Section 4.1) did not find the Audio Bubbles helpful, and two participants were unsure if the Audio Bubbles were helpful or not. All of the participants were able to determine the changes in sound as they walked through the Audio Bubbles. Comments primarily addressed the "range" of the Audio Bubbles, that is how far away from a point of interest the user should be before entering the bubble. Two of the participants found that the sound started too far away from the point of interest and was initially annoying, but with decreasing distance from the point of interest the sound became more useful and valuable. One participant saying *"I would use it, yeah. ... But the beeping would be annoying if it would beep for every little sight. I would turn off the sound until I got closer to the points "*, the same participant went on to say that the sound cues were useful *"especially when you were closer, when you were closer it really gave you hints I think"*. In the experiment a user entered an Audio Bubble when around 250 meters from the point of interest. However, in this situation where the user already has a destination determined, the Audio Bubble would only be useful (as discussed in the introduction) when close to that destination. Whereas, in a wandering scenario, where the Audio Bubble notifies the user of a nearby point of interest that he or she may not be aware of, a larger radius may be more appropriate. Further investigation is needed to provide concrete guidance in this area.

In conclusion, we can consider to have provided some answer to RQ1, that Audio Bubbles were a useful way to navigate and locate points of interest. However, we cannot make claim to have answered RQ2, that Audio Bubbles are superior to simple visual maps. The results we obtained appear to show that the Audio Bubbles allowed some reduction in time taken for the fine grained homing behaviour participants needed at the end of a route, and that the Audio Bubbles were certainly not poorer than only using a visual map. However, the results are not definitive and further investigation, given the qualitative feedback from users, is required to refine Audio Bubbles to be truly useful navigational aids.

5 Future Work

From the experimental results there are several refinements that should be carried out on the Audio Bubbles, and several future avenues for investigation.

5.1 Audio Bubble Size

Many of the participants stated that they found the audio started to play too early. In other words the Audio Bubble was too big. As sound can be annoying, determining the appropriate size of the Audio Bubble, so that it is neither too large or too small is important. In our initial implementation we chose 250m as the boundary of the bubble, primarily down to informal testing by the experimenter. While reducing the size of the bubble would be beneficial based on the comments by participants, we also need to consider the role the Audio Bubble plays at a particular time, and if its size should be adjusted accordingly. In our experimental situation, the Audio Bubbles were centred on locations that the participant was trying to reach and were felt to be useful in "homing" onto the point of interest when nearby. However when wandering, the purpose of the Audio Bubble is primarily to inform the user of a nearby, non-obvious point of interest, that he or she might be interested in. In such a situation, an Audio Bubble with a radius extending to a "reasonable walking distance" between the user and the point of interest, may be more appropriate. Further research into how the size of the Audio Bubble should change according to user context is therefore required.

5.2 Investigating Different Types of Audio for Points of Interest

In our experiment we chose to follow the same auditory design as that of Holland et al. [7], using a Geiger counter metaphor, where audio pulses repeat at a faster rate the closer the user is to the point of interest. We also increased the volume level of the pulses as the user came closer to the point of interest. We chose to investigate this kind of mapping as it provided more information in the sound as to how far the user was from the point of interest. As we did not use spatial presentation (see Section 3), we wanted to ensure that the distance information was communicated as clearly as possible. In addition, unlike the work of Stahl [10], many of our points of interest had no obvious sound that could be used to represent them, meaning a more abstract sound was required. The issue being that although the sound communicates distance information, it contains no information as to the point of interest being navigated towards. In our initial study this is not so important, but it would be useful to communicate some information about the point of interest in the sound, thereby assisting the user in the decision to walk towards that point of interest or not. In practice, it is likely that there will be a mixture of points of interest which can have appropriate natural sounds associated with them, and points of interest for which there is no obvious sound. Therefore, it would be necessary to use different auditory representations to represent different points of interest, and further investigation is required to identify how well these different auditory representations work with our monaural system. Dingler,

5.3   In-Field Evaluations

Our evaluation work described in this paper has sought to investigate both the utility of Audio Bubbles as an idea, and consider how they can be used in the homing scenario discussed in the introduction. However, another goal of our work is to allow discovery of points of interest in a wandering scenario - where the user has no particular goal or destination in mind, and would tend not to use a visual map as a means of navigation. Our intention is that Audio Bubbles would allow him or her to "stumble" upon interesting places that may not be obvious in the built environment. However, our current implementation is not advanced enough to be used as a full tourist application. Therefore, we are building an Apple iPhone application so we can evaluate the wandering scenario with real tourists.

5.4   Social Points of Interest

Our experiment, as well as all of the existing work in Section 3, assumes that points of interest are fixed and determined by some prior authority. However with the growth of the mobile Web and geotagging (associating a geographic location with data) of information on the Web, we can add and augment points of interest from other users of the system. Axup and Viller [13] have looked at how backpacking is a temporally shared activity amongst travellers, and Goel [14] has looked at how information on tourist guides can be shaped and filtered by social networks. In addition, websites such as `www.tripadvisor.com` allow users to rate and comment on their experiences of hotels and restaurants. How should this information be represented through Audio Bubbles? One approach would be to modify the size of the Audio Bubble based on recommendations. For example, a well recommended restaurant would have a larger Audio Bubble than one with a poorer recommendation. However this may, given the results of our study, be annoying, as the audio would begin to play too far away from the point of interest. An alternate approach would be to use speech as the sound in the Audio Bubble, reading out the review comments that previous visitors have left. Which is better is unclear, and further investigation is needed in this area to identify the best approach.

6   Conclusions

In this paper we have introduced the concept of Audio Bubbles - virtual spheres filled with audio that are geocentered on physical points of interest to provide serendipitous discovery and assist in fine grained navigation. We argue such a technique is both necessary and useful, as visual maps fail to adequately support the non-directed wandering and homing activities that tourists engage in. The outcome of our experiment does indicate that the use of the Audio Bubbles provided some advantage to users when trying to locate points of interest in a park. Future versions of the Audio Bubbles will seek to identify how flexible the

concept is, by evaluation in less directed navigational scenarios than those in our experiment, and as such will provide answers to many of the issues raised in our future work section. In conclusion, we believe that with further refinement to the auditory design, future versions of the Audio Bubbles will provide significant advantages to users engaged in loosely directed navigation, in a way that is useful, but not intrusive or annoying, and support the undirected wandering behaviours that tourists currently engage in.

References

1. Brown, B., Chalmers, M.: Tourism and mobile technology. In: Eighth European Conference on CSCW. Volume 1., Helsinki, Finland, Kluwer Academic (2003) 335–354
2. Kaasinen, E.: User needs for location-aware mobile services. Personal and Ubiquitous Computing **7** (2003) 70–73
3. Cheverst, K., Davies, N., Mitchell, K., Friday, A., Efstratiou, C.: Developing a context-aware electronic tourist guide: some issues and experiences. In: CHI 2000. Volume 1., The Hague, The Netherlands, ACM (2000) 17–24
4. Brown, B., Laurier, E.: Designing Electronic Maps: An Ethnographic Approach. Volume 1 of Mapbased Mobile Services - Theories, Methods and Implementations. Springer (2005)
5. De Certeau, M.: The Practice of Everyday Life. University of California Press (1984)
6. Wilson, J., Walker, B.N., Lindsay, J., Cambias, C., Dellaert, F.: Swan: System for wearable audio navigation. In: ISWC 2007. Volume 1., Boston, MA (2007)
7. Holland, S., Morse, D.R., Gedenryd, H.: Audiogps: Spatial audio in a minimal attention interface. Personal and Ubiquitous Computing **6**(4) (2001) 6
8. Strachan, S., Eslambolchilar, P., Murray-Smith, R.: gpstunes - controlling navigation via audio feedback. In: MobileHCI 2005. Volume 1., Salzburg, Austria, ACM (2005) 275–278
9. Reid, J., Geelhoed, E., Hull, R., Carter, K., Clayton, B.: Parallel worlds: Immersion in location-based experiences. In: CHI 2005. Volume 2., Portland, Oregon, ACM Press (2005) 1733 – 1736
10. Stahl, C.: The roaring navigator: A group guide for the zoo with shared auditory landmark display. In: MobileHCI 2007. Volume 1., Singapore, ACM (2007) 282–386
11. Marentakis, G.: Deictic Spatial Audio Target Acquisition in the Frontal Horizontal Plane. PhD thesis, University of Glasgow (2006)
12. McGookin, D.K., Brewster, S.A.: Dolphin: The design and initial evaluation of multimodal focus and context. In: ICAD 2002. Volume 1., Kyoto, Japan, ICAD (2002) 181–186
13. Axup, J., Viller, S.: Augmenting travelgossip: Design for mobile communities. In: OZCHI 2005. Volume 2., Canberra, Australia, ACM (2005)
14. Goel, A.: Urban pilot: A handheld city guide that maps personal and collective experiences through social networks. In: Digital Cities LNCS 2362. Volume 1., Springer (2002) 384–397