Audvert: Using Spatial Audio to Gain a Sense of Place

Liam Betsworth[1], Nitendra Rajput[2], Saurabh Srivastava[2], and Matt Jones[1]

[1] Future Interaction Technology Lab, Swansea University, SA2 8PP, UK
{csliamb,matt.jones}@swansea.ac.uk
[2] IBM Research, 4, Block C, Vasant Kunj, New Delhi – 110070, India
{rnitendra,sausriva}@in.ibm.com

Abstract. We introduce Audvert – a system that facilitates serendipitous discovery and navigation through spatial audio; used to navigate and discover points of interest in large, unfamiliar indoor environments. Our main aim was to create a lightweight spatial audio display that can convey a sense of a place without complex point and select interactions. We conducted a preliminary study comparing two audio types to see which best suited sound localization and a study of Audvert used in a real world scenario. Our findings suggest that long continuous audio performs better than short intermittent audio for sound localisation. We also discover a change in behaviour when using the system, with a large percentage of users wanting to visit newly discovered shops after using the system. We discuss the findings and draw research conclusions.

Keywords: Non-visual Navigation, Wayfinding, Auditory Display.

1   Introduction

When a visitor enters a shopping mall for the first time, it is likely that they do not know where they are going. To maximize the available space in a mall, shops are often packed closely together, and depending on the specific layout of the shopping mall and the location of the visitor, it is also likely that not all shops are visible. All of these factors can make a large, unfamiliar indoor location a confusing place to be.

Previous research on users' use of technology in unfamiliar places has shown that tourists often exhibit a wandering behaviour and then home in on serendipitous discoveries [1]. To aid such activities, maps (physical and digital) are sometimes used, although as mentioned by McGookin et. al [8], maps are not best suited to these kinds of activities due to changing contexts and the need to constantly refer to a map. We argue that tourist behaviour can be linked quite closely to visitor behaviour when in large and unfamiliar places such as shopping malls. It is easy to imagine a visitor to a shopping mall, wandering, homing in on anything of interest to them.

In an effort to aid such visitors, we have developed a system that attempts to facilitate navigation (homing in) and serendipitous discovery (wandering) using only spatial audio. Audvert is a mobile application that attempts to give a user a sense of place. Conventional maps are good at summarising what points of interest are available to a user in the nearby environment, but Audvert attempts to give the user a sense

of what is actually around them by delivering spatial audio feedback. When a user hears a point of interest (POI) played back through their headphones, the audio sounds as though it originates from the real physical location of that POI. A visitor may not always be able to see a shop, but with Audvert they can infer the direction and proximity from the audio feedback, enabling them to then look around to see possible pathways of getting there. Where existing audio navigation and discovery systems have simply focused on alerting a user to the proximity or direction of a POI [4, 5, 7, 8], Audvert is unique in that it is used indoors, layering useful, spoken information about points of interest with directional and proximity feedback. Audvert is a lightweight system that offers the user continual feedback, allowing the user to engage when they want. The following scenario, along with Fig. 1, illustrates how a visitor might use Audvert in a shopping mall.

Scenario. *Fred is at a large shopping mall but is unsure about which shops to visit. Putting on his headphones and opening the Audvert application, he holds his phone out in front of him (1). The Audvert app randomly picks a store from the mall and begins playing an audio clip containing the name and description of the goods and services available there. Fred can hear this information on his right side indicating that this particular store is in that direction (2). Deciding he is not interested in this shop, Fred waits a few seconds for the next random selection to be played. After listening to the name and short description of the second store, which is now being played in his left ear, Fred decides he wants to know more and signals this via a shake gesture with his phone (3). Fred now hears additional information and uses the direction the sound is coming from to guide him to the store (4). As he approaches the store, the amplitude increases.*

Fig. 1. User interaction in Audvert (explained in scenario above)

2   Related Work

As inexpensive, powerful smart phones have become more popular, soundscape and auditory display technologies have reached the consumer domain. A large amount of research on these topics has focused on their deployment for museum exhibits [3] and outdoor, augmented-reality soundscape experiments [2]. In terms of its applications in way finding, earlier work has tended to focus on the audio modality as a basic means of A to B navigation [4, 5], but more recently, researchers have focused on the specific problem of tourist way finding. McGookin et. al developed a system named Audio Bubbles [8], "employing non-speech audio to support tourist way finding." The researchers argue that the way in which tourists navigate is not so well supported by traditional visual maps. We argue that this is also the case for unfamiliar *indoor*

environments such as shopping malls. Our system similarly attempts to facilitate serendipitous discovery – something that visual maps do not focus on.

Audio Bubbles uses a Geiger counter metaphor to signify proximity to a POI, similar to the early audio navigation system, AudioGPS [4]. In our design, we have chosen against using this common geiger counter metaphor. Instead, we use amplitude to give feedback of proximity. We believe this is a more natural mapping, as sounds naturally get quieter as they are further away. Unlike Audio Bubbles, we also incorporate directional feedback. We encode directional and proximity feedback in our audio recordings at runtime, similar to ONTRACK [5], although using the full 360° for direction instead of simple left and right panning. By not using the Geiger counter metaphor, we are also able to use audio recordings such as shop descriptions.

Different kinds of audio feedback for auditory displays have been explored in depth by Vazquez-Alverez et. al [13]. One auditory display system that uses spoken word is PULSE [7]. PULSE turned local Twitter tweets into spoken word using speech synthesis. When users approached the physical location where a tweet was composed, it would play to the user. Audvert is again, different from PULSE as we use spatial audio that takes distance into account. PULSE was designed to provide the user with a social vibe – "an intrinsic understanding of the people, places and activities around their current location." We strive to achieve a similar sense of place in Audvert, but through using pre-recorded audio content relevant to the shopping mall.

One problem that can occur when too many spatial audio streams are playing concurrently is increased cognitive load [12]. To combat this problem, we have chosen only to play one spatial audio stream at a time. In doing this, we also make the process of stream selection easier for the user. Most audio displays make use of a bearing-based selection technique (e.g. [6]), but this technique has its problems. When the number of selectable entities (audio sources in this case) increase, the bearing gap between each decreases. According to Strachan and Murray-Smith [11], effective selection was maintained down to around a 20° separation with a 10° target width. As Audvert only ever uses a single audio stream, we use a simpler interaction technique – a shake gesture to select or deselect the audio stream. We have chosen to use a gesture over a button, as previous research on gestures as a means of control for mobile devices has shown that they are a reliable, eyes-free interaction technique [9]. Further research has also been carried out on the practical applications and social acceptability of mobile gestures in the real world [10]. We believe the shake gesture is a simple and discreet gesture that most users will be able to use and will be comfortable with.

In summary, Audvert is novel in that it is an attempt at indoor audio navigation and discovery, not only providing proximity and directional feedback like previous systems, but also giving useful spoken information. It also tries to simplify the task of spatial display selection by making only one choice available at a time.

3   Audvert

Audvert is a prototype system designed to give users a better sense of the place that they are in using only spatial audio. To use the system, a pair of headphones and the

Audvert application is required. When using Audvert, sounds appear to originate from their actual physical locations. A user should be able to infer proximity through amplitude (closer is louder) and direction through the panning of sounds. Spatial audio feedback in the system is constantly changing due to lightweight interactions such as the user walking around (changing their location in the system) or the user changing direction (changing their listening orientation). Audvert uses a digital embedded compass, so as the user rotates, sounds continue to appear as though originating from the same physical location in space. With regards to location tracking, we decided upon a Wizard of Oz study methodology for Audvert. Users are not told that their location is entered manually elsewhere, so it is fair to say that users believe that their location is being automatically tracked. We use a separate purpose built system where a researcher may tap a map to update the location of a user. This Wizard of Oz approach was chosen as it offers a high accuracy with little complication. GPS was not a viable option as the system was used indoors. In reality, this kind of system would most likely work indoors with WiFi localisation.

When considering the user interaction with the prototype system, we decided that it should facilitate three separate functions, all relative to tourist way finding [1] - to aid serendipitous discovery of POI (wandering), to deliver useful information about them and to offer users the ability to navigate towards them if they wished to do so (homing). To facilitate serendipitous discovery, Audvert randomly chooses points of interest in the local environment, playing them back one at a time to the user. Useful information about the points of interest is spoken in the audio recordings (sales, offers and general information about the shop). Once a user has selected a POI, Audvert facilitates navigation through continually playing spatial audio for that POI.

4   User Evaluation

Two studies were conducted to test Audvert. The first was used to determine audio preference and performance. The second study gave participants the opportunity to use the system fully as they would do in a real life scenario. Both studies were conducted in a shopping mall in New Delhi, India.

Prior to the evaluation, early testing with users alerted us to the issue of users pointing the device in a different direction to their orientation. Because of this, a static arrow was added to the GUI as an affordance, so that users would remember to point the device in front of them at all times. There were also instances where users would forget the name of the shop they were currently listening to. For this reason, the name of the current shop the user is listening to was also added to the interface.

4.1   Preliminary Study: Cube Placement

Before embarking on a usability/user experience study of Audvert, we wanted to ensure we picked the optimal type of audio for use within the system. With this in mind, we recruited 24 participants (18M, 6F, 16-37 years) for a between-subjects study testing the effectiveness of two different types of audio; short intermittent and long

continuous. We define short intermittent as audio clips that do not have continuous background sound and are short in length (i.e., purely spoken word, 5-6 seconds long). Long continuous clips contain continuous background sound and are longer in length (i.e., spoken word with a musical backing track, ~25 seconds long).

Our goal in this preliminary evaluation was to determine which of these two types of audio provided users with the most accurate path to the audio source location. This was achieved by playing five audio sources and asking each participant to place five paper cubes on a map of the shopping mall that best matched the location of where sounds were originating from. The map was A3 in size and each cube was $2cm^3$. Direction scores were calculated depending on the cube placement angle of error. Scores from 10 to 0 were given for 22.5° segments. Proximity scores where calculated depending on the cube placement error in centimeters (distance from user). None of the participants we recruited had any prior knowledge of the shopping mall.

The results of this evaluation showed that on average, participants using long continuous audio performed better in terms of perceiving both proximity ($p < 0.0068$; Unpaired t-test; 1.8cm average difference) and direction ($p < 0.0018$; Mann-Whitney; 11.4/50 average score difference). We therefore decided that short intermittent audio should be used as a brief alert of a shop's existence (when a shop is unselected), and long continuous audio should be used to give a person more information about a shop and allow a user to navigate towards it if they wish (when a shop is selected).

4.2   Usability and User Experience Study

The purpose of this study was to gain a better understanding of the usability and user experience offered by the system. We wanted to discover if Audvert could be used to navigate towards a shop that a user does not know the location of, and more generally, how a user feels about their surroundings when using the system. In search of answers, we attempted to evaluate the system in a real usage scenario, such as when a visitor is walking in a shopping mall and wants to know more about the shops there.

Participants. Participants of this study were again, real visitors to the shopping mall. We recruited 19 participants (10M, 9F, 17-47 years), none of which had taken part in the preliminary study. Approximately 3 were regular visitors to the mall. All participants could speak and understand English. No participant reported hearing problems. All participants had used a mobile phone before, but not all had used a smartphone.

Task and Measurements. Participants of this study were expected to explore the shopping mall using the system. Each participant was personally allowed to choose what he or she used the system for. Qualitative data on user experience with the system was gathered through a post-study interview. Questions asked in this interview were: Has the system given you a better or worse understanding of the shop locations? How would you describe your sense of place after using the system? *(Level of comfort and familiarity with the place)* Did you hear any shops that you have not visited before? *(Would you consider visiting one of them now?)*

Procedure. Participants were recruited individually as they walked through the shopping mall and were given a short introduction to the Audvert application. Participants were asked to explore the shopping mall using Audvert. As participants moved around the shopping mall, a researcher stood around 20 meters away and updated the participant location every few seconds. While exploring, participants were free to interact with the system as they wished. At the end of each study session, the participant was asked a series of questions in a post-study interview regarding their experience with the system. No incentive was given for completing the study.

Results. The following section details the responses given to the questions asked in the post-study interview. There were some participants that could not decide upon an answer to some of the questions. These have been included as N/A in the results. Percentages are relevant to the number of participants that answered each question.

Question	Response			
	Better	**Unaffected**	**Worse**	**N/A**
Has the system given you a better or worse understanding of the shop locations?	79% *(15)*	5% *(1)*	16% *(3)*	0
How would you describe your sense of place after using the system?	80% *(12)*	7% *(1)*	13% *(2)*	*(3)*
	Yes	**Maybe**	**No**	
Would you consider visiting a new shop you discovered using the system?	63% *(10)*	31% *(5)*	6% *(1)*	*(3)*

Table 1. Quantitative results from questions asked in the study

Informal Observations & Feedback. It became apparent in the study that not every participant was selecting a shop by shaking the device. By avoiding this function, the participant was not able to gain extra information about a shop (a long description) – the system however, is still usable without this function. As Audvert provides no visual feedback of where the user is in the system, one participant was unsure whether a shop was selected or unselected at any point in time. This participant carried on for a further two minutes and after this period appeared to be much more comfortable with the system. One participant that successfully used the shake gesture to activate a shop was seen attempting to navigate towards the selected shop. This participant attempted to follow the direction of sounds, and did so successfully by finding the shop. The participant mentioned that they knew they were going in the correct direction because they could hear the sounds getting louder as they got closer to the shop. This person said that they had never visited the shop before, and that they had no prior knowledge of its location. There was also another situation where a participant used the varying volume of sounds to navigate towards a shop. This participant had selected a shop, walked in one direction and then realised that the volume was decreasing - inferring that they were heading in the wrong direction. The participant then turned around to face the correct direction and successfully found the shop that they had selected.

4.3   Discussion

Most participants enjoyed the experience, claiming that the system appeared to be 'very accurate' and that the direction sounds came from was the same direction that they could see the shops. A large portion (79% (15)) of participants said that they had a better understanding of the location of shops after using the system. One participant who claimed a worse knowledge of shop locations said that it was "difficult to differentiate between the distances of each shop." Another participant agreed, finding it "difficult to tell the distance to a shop", but claiming the "direction was good." Although 79% of participants said that their understanding of shop locations had improved, 1/3 of these (5 participants) claimed that they could either infer direction or distance, but not both. In most of these 5 cases, participants could only infer direction (4 participants). This is an interesting result, as without knowing both, one would imagine it being difficult to pinpoint an exact location. These participants obviously believed that knowing only direction or proximity was sufficient for knowing the location of a shop. This strengthens the argument for spatial audio as a navigation aid.

A similarly high percentage (80% (12)) claimed that they had a better sense of place after using Audvert. There were 5 participants who did not answer this question, unable to differentiate between knowing the locations of shops and a sense of place. One participant said that the system gives a "good sense of the area. I could tell what kinds of shops and products were available." Another participant argued that the system gave a bad sense of place as there were "lots of shops and I only heard a few." Although a similar percentage said they had a better sense of place and an improved understanding of shop locations, not every participant who had an improved understanding of shop locations had a better sense of place, showing that the two weren't regarded as the same question by participants.

80% of participants heard a shop they had not visited before, of which, 63% (10) said they would definitely consider visiting that shop and 31% (5) said they might visit now. There were 4 participants who had previously visited every shop that they heard when using the system. Only one participant decided that they would not visit a new shop they had discovered, saying that they are the "kind of person that comes to the mall for one thing." There does appear then, to be a compelling change in behaviour, where the system seems to encourage participants to visit new, serendipitous discoveries that are made. Generally, most participants explained that they would like to visit these shops because of sales and discounts mentioned in the audio recordings. Participants who said they might visit a new shop often said that they would do so if more detailed information about a shop was available. Judging by this feedback, it seems that the content is an integral factor of this change in behaviour.

The majority of participants liked the shake gesture. Two participants explicitly mentioned that they would "prefer the shake over a button." In contrast, two participants said that they were not comfortable with the shake gesture, preferring a button. One participant's concern came from not wanting to shake their expensive mobile phone. To satisfy all users, a multimodal approach could be introduced, where the user chooses whether the interaction occurs onscreen or through a gesture.

5   Conclusions

We have introduced Audvert to illustrate the potential of a lightweight system that facilitates navigation and serendipitous discovery of large indoor spaces using spatial audio. While the idea of an audio display for navigation and discovery is not a new one, we argue that the novelty of our approach lies within its indoor application and the feedback and selection techniques employed. We use amplitude and directional feedback to manipulate useful, spoken audio information about a point of interest in real time. Audvert has also introduced a simplified selection technique for spatial displays through using a single active element and a simple shake gesture to select.

Our user evaluations have shown that Audvert can convey useful spatial audio feedback, and that it is capable of helping users navigate towards and discover new points of interest. We saw multiple examples of users successfully arriving at a shop and for the majority of participants, we witnessed a change in behaviour with participants claiming that hearing new shops made them want to go and visit there.

References

1. Brown, B., Chalmers, M.: Tourism and mobile technology. In: Eighth European Conference on CSCW 2003, Helsinki, Finland, vol. 1, pp. 335–354. Kluwer Academic, Dordrecht (2003)
2. Cater, K., Fleuriot, C., Hull, R., Reid, J.: Location aware interactive applications. In: ACM SIGGRAPH 2005 Sketches, SIGGRAPH 2005 (2005)
3. Heller, F., Knott, T., Weiss, M., Borchers, J.: Multi-user interaction in virtual audio spaces. In: CHI EA 2009, pp. 4489–4494 (2009)
4. Holland, S., Morse, D.R., Gedenryd, H.: Audiogps: Spatial audio navigation with a minimal attention interface. Personal Ubiquitous Comput. 6(4), 253–259 (2002)
5. Jones, M., Jones, S., Bradley, G., Warren, N., Bainbridge, D., Holmes, G.: Ontrack: Dynamically adapting music playback to support navigation. Personal Ubiquitous Comput. 12(7), 513–525 (2008)
6. Marentakis, G.N., Brewster, S.A.: Effects of feedback, mobility and index of difficulty on deictic spatial audio target acquisition in the horizontal plane. In: Proceedings of CHI 2006, pp. 359–368 (2006)
7. McGookin, D., Brewster, S.: Pulse: the design and evaluation of an auditory display to provide a social vibe. In: Proceedings of CHI 2012, pp. 1263–1272 (2012)
8. McGookin, D., Brewster, S., Priego, P.: Audio bubbles: Employing non-speech audio to support tourist wayfinding. In: Altinsoy, M.E., Jekosch, U., Brewster, S. (eds.) HAID 2009. LNCS, vol. 5763, pp. 41–50. Springer, Heidelberg (2009)
9. Pirhonen, A., Brewster, S., Holguin, C.: Gestural and audio metaphors as a means of control for mobile devices. In: Proceedings of CHI 2002, pp. 291–298 (2002)
10. Rico, J., Brewster, S.: Usable gestures for mobile interfaces: evaluating social acceptability. In: Proceedings of CHI 2010, pp. 887–896 (2010)
11. Strachan, S., Murray-Smith, R.: Bearing-based selection in mobile spatial interaction. Personal Ubiquitous Comput. 13(4), 265–280 (2009)
12. Vazquez-Alvarez, Y., Brewster, S.A.: Eyes-free multitasking: the effect of cognitive load on mobile spatial audio interfaces. In: Proceedings of CHI 2011, pp. 2173–2176 (2011)
13. Vazquez-Alvarez, Y., Oakley, I., Brewster, S.A.: Auditory display design for exploration in mobile audio-augmented reality. Personal Ubiquitous Comput. 16(8), 987–999 (2012)