

Audio GPS: spatial audio in a minimal attention interface

SIMON HOLLAND & DAVID R. MORSE

Computing Department, The Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom.

Email: S.Holland@open.ac.uk, D.R.Morse@open.ac.uk

Phone: +44 (0) 1908 653148, +44 (0) 1908 858463 Fax: +44 (0) 1908 652140

Abstract

In this paper we consider a prototype audio user interface for a Global Positioning System (GPS) that is designed to allow mobile computer users to carry out a location task while their eyes, hands and attention are often otherwise engaged. Audio user interfaces for GPS have typically been designed to meet the needs of visually handicapped users, and generally (though not exclusively) employ speech-audio. In this paper, we consider a prototype audio GPS user interface designed primarily for sighted mobile computer users who may have to attend simultaneously to other tasks, and who may be holding conversations at the same time. The system is considered in the context of being one component of a user interface for mobile computer users. The prototype system uses a simple form of spatial audio. Various candidate audio mappings of location and distance information are analysed. A variety of tasks, design considerations, technological opportunities and design trade-offs are considered. Preliminary findings are reported. Opportunities for improvements to the system, and future empirical testing are explored.

Keywords: *Audio, Global Positioning System, GPS, Sound, Navigation, Minimal Attention User Interface*

Introduction

Recent empirical work on mobile computer usage has argued that the GUI-based interaction style currently embodied in most mobile devices is inappropriate in many situations in which mobile computers are used [1]. Such situations are often characterised by:

- interactions with the real world being more important than interactions with the computer so the user has a strictly limited attention capacity for the computer;
- user s hands are often used to manipulate physical objects;
- users being involved in tasks that demand a high level of attention (to avoid danger to the individual as well as to engage with the task at hand);
- users being highly mobile, and adopt a variety of positions and postures during the task;
- the user s interactions with the environment being context dependent;
- interactions with the computer can be high speed and driven by the external environment [1, 3].

In order to overcome some of the perceived drawbacks of the Graphical User Interface, both Kristoffersen [4]& Pascoe [5] proposed alternative user interaction styles that aimed to reduce the amount of user attention, but not necessarily the amount of user interaction, that is required to perform a particular task. The intention of both interaction styles being to transfer interaction tasks to interaction modes that take less of the user s attention away from their current task. Kristoffersen [4] proposed a technique called MOTILE that is based on the three principles of (i) little or no visual attention, (ii) structured, tactile input and (iii) the use of audio feedback. In contrast, Pascoe [5] proposed a Minimal Attention

User Interface (MAUI), linked to context-awareness that successfully reduced the amount of attention that a fieldworker had to pay to their mobile computing device [5].

Both MOTIVE and MAUI were developed to support fieldwork applications, where a principal requirement is data capture through making observations on the local environment, be it on a boat or a game reserve in Kenya. A second situation in which a MAUI would be useful is in navigation. Whatever mode of transport is used, frequent reference must be made to the navigation aid usually a map or written instructions. This takes one's attention away from the primary task (moving through the local environment), which can result in accident, particularly when driving a vehicle. Therefore, one of the primary requirements for navigation aids is that they be low-attention, which has driven a long tradition of research into electronic navigation aids particularly for the visually impaired [6]

In practical terms, a MAUI navigation aid would generally use the audio channel as a primary means of communication with the user. Since the deployment of satellite-based navigation networks (the American Global Positioning System (GPS) and the Russian GLONASS network), and with recent advances in speech technology, the majority of navigation aids have been based on these two technologies. Such navigation aids include the Personal Guidance System [6,7], Personal Navigation System [8] Mobility of Blind and Elderly People Interacting with Computers [9] and a navigation system for the blind that uses a GPS and a mobile phone connection [10]. These devices have the dual function of supporting navigation whilst also informing the user about relevant landmarks of interest to them. Similar applications that have been developed with sighted people in mind electronic tourist guides have become the canonical context-aware demonstrator application [11, 12].

In recent years, handheld GPS systems (such as those manufactured by Magellan, Trimble and Garmin) have ceased to be the preserve of amateur mariners and are now being used in all walks of life. For most people, GPS systems are used to (i) find out where you are now, and (ii) to find out how far it is, and in which direction your destination (by inference on the part of the user, or after being programmed by said user). The small screen of most handheld GPS units present users with the same usability problems that were found in the fieldwork applications quoted earlier that led to the development of the Minimal Attention User Interface concept.

In GPS systems, the core information to be communicated to the user typically takes the form of time-varying spatial information and relatively simple quantitative information. Given this, GPS systems are good candidates for investigating the design issues arising in minimal attention user interfaces. As noted above, GPS navigation systems for the visually impaired have been studied by several authors, but in this paper we are considering instead a minimal attention user interface for the sighted person who is simultaneously involved in other demanding tasks. The two situations have somewhat different needs. For example, an audio GPS system for visually impaired users needs to be able to communicate a wide range of information through the auditory channel to allow the GPS to be programmed, and to communicate many kinds of information to the user about the surroundings. Given the potential complexity of this information, speech audio is more or less essential. However for the sighted but minimally attentive user, a simpler interface which is less prone to interference from conversational speech can be considered, namely a simple non-speech spatial audio interface. Such a system could be a useful component of a more general mobile user interface albeit for sighted users in hazardous or demanding situations, for example, when it is dark, raining, when hands are full of shopping or when there are children to be supervised. In this paper we consider some of the issues involved and report on a prototype system.

The AudioGPS

The basic conception that has guided us relies on a virtual acoustic display as the user interface. This display takes an audio signal and transforms it into a binaural signal that the user listens to through headphones. The result is a sound that appears to emanate from a given environmental location. Such a system would be most useful in assisting a user to navigate from their current position towards their destination. Of course, the assumption is that the path (road, waterway etc.) is navigable via a shortest distance, straight-line path from the current location to the destination. In those situations (probably the majority) in which this is not the case, then intermediate waypoints have to be inserted by the user either dynamically during use or before undertaking the journey. Yet more sophisticated ideas have been considered by Moon [9].

We used 'open' stereo headphones that do not exclude background environmental noise because such noise provides important navigational cues and warnings to all people, not just the visually impaired. Furthermore, we used non-speech audio rather than speech because we do not wish to overload the user through using a communications medium that places a large processing and attention burden on the user. In addition, we do not want to compete with the voice channel, which might be used for other communication tasks. For example, the user might be engaged in conversation with companions, or via mobile phone, while walking.

Audio representations of direction and distance

When using a GPS system for navigation there are at least two pieces of information that have to be communicated to the user: the distance to the destination (or intermediate waypoint) and its bearing, relative to the current direction of movement or direction in which the user is facing. On a graphical user interface, the distance typically would be displayed numerically and the direction would be displayed by a compass bearing. In our prototype AudioGPS, we have used the following audio representations of direction and distance.

Direction

To a first approximation, the bearing to the next waypoint (from the current position, relative to the current direction of motion) is indicated by the simple means of panning a sound source representing the destination across the stereo sound stage. Later prototypes are planned to incorporate a more capable 3D audio rendering system, but the current prototype has the positive benefit of permitting an investigation of the usefulness of a simple, computationally inexpensive audio rendering resource. The sound source used is a briefly repeated tone, which we will describe in more detail in a moment. With simple panning, it is not hard to distinguish, informally speaking, between sound sources at the extreme left, extreme right, straight ahead or at some intermediate points. However, simple panning does not help distinguish between a source placed in front vs behind the listener (indeed, even relatively capable 3D audio systems employing more precise generalised Head Related Transfer Functions can suffer from similar confusions). As a pragmatic way around this perceptual problem, we use a sharp tone (in fact a harpsichord sound) when the destination is in the semicircle in front of the user, and a relatively muffled tone (in fact a muted trumpet sound) when the destination is located in the semicircle behind the user.

Ideally, the bearing of a destination should be computable even when the user is stationary, and should be heard to pan around the head if the user revolves on the spot. This is easily realisable using a flux magnetometer (electronic compass) in addition to a GPS. However currently the AudioGPS does not incorporate an electronic compass. Hence the audio GPS, just like most visual GPSs, has no way

of knowing which way the user is facing. However, again like most GPSs, provided the user is moving, or has been moving, the direction of motion can be calculated. Hence, just as with its visual equivalent, the audio GPS uses the convention that if the user faces in the direction of travel, the destination will be rendered in the correct direction.

By these means, the location of a destination at the four cardinal points left, right, ahead and behind (at 90... intervals) are relatively straightforward to communicate. In addition, it is generally possible to perceive, informally speaking, if a sound source is at some intermediate point, giving more or less eight approximate cardinal points. By altering one's track, it may be possible to "steer" a destination into an easily identifiable angular location, such as straight ahead. However, if a better angular precision is required than approximately eight cardinal points, clearly more affordances are needed.

As an experimental way of addressing this problem, we have employed a 'chase tone', which works as follows. Leaving aside the chase tone for a moment, a repeated tone at a fixed pitch, but appropriately panned in space, always indicates the destination. A second tone, the chase tone, coincides exactly in pitch with the first tone if the destination is straight ahead, but progressively diverges in pitch to a maximum pitch difference of an octave as the angle of the destination reaches a maximum of 90 degrees of pan. Both tones are always spatially located together. Rather than using a chase tone whose pitch varies continuously, we have used a chase tone that varies its pitch discretely and chromatically. The pitch variation is scaled so that by the time the bearing is 90..., the chase tone is exactly one octave below the primary tone. Therefore, if the bearing is 90... from ahead then the user hears two tones which are one octave apart. We considered using two tones on a continuous scale but in preliminary pilot experiments, we found that the chromatic scale was easier to interpret than the continuous scale. Experiments are needed to find out the usefulness of this mapping taking into account individual differences among users.

Distance

We have been experimenting with several representations of distance. Our primary representation is a Geiger counter or "warm/cold" metaphor, where the number of pulses of sound, together with their rapidity, gives an indication of how far it is to the next waypoint. If the user is a long way from the waypoint then few, widely spaced sound pulses are emitted. If the user is close to the waypoint, then the user hears many, closely spaced, sound pulses. This metaphor is easy to communicate using the familiar "geiger counter" idea. Typically, on arrival, a distinctive arrival tone is used. This is the mapping we have used most to date.

An alternative representation communicates the distance to a destination using an discrete analog mapping to an appropriate scale, e.g. five clicks in each run of clicks indicate fifty metres, whereas one click indicates ten metres. Martyn Cooper has referred to this as an "inverse Geiger counter" mapping. When distant, clicks are emitted more rapidly. Different scales are appropriate to different situations - pedestrians vs car users, for example, and a different scale might be needed for the last part of a very long journey. Heirarchical versions of such a mapping are possible to permit greater precision over a wider range of distances

With any of these mappings, at some predetermined distance from the waypoint, the system should emit an arrival tone to indicate that the user has reached the waypoint within the limits of the precision of the GPS system that is being used to support navigation. If a differential GPS system was being used (which gives a more precise estimate of location than an uncorrected GPS system) then the arrival zone would be smaller than if an uncorrected GPS system is being used.

Implementation

A primary requirement for any navigation aid is portability. Notwithstanding that requirement, our implementation platform was chosen for its development facilities and flexibility rather than the convenience with which it can be carried. As we explain in the Further Work section of this paper, we intend to experiment with a number of peripheral devices. It is more convenient to work with such devices in standard development environments (for which device drivers are easily available) than it is to work with more appropriate, but more specialised, mobile computers. Therefore, we are using VisualWorks Smalltalk and Java, running on Apple Macintosh and Microsoft Windows laptop computers. We are equipping specialised rucksacks with forced electrical cooling to allow both hands to be kept free. Sound synthesis is carried out using software sound synthesis controlled via MIDI, which is available via a common Quicktime API for both software platforms.

User trials

We have performed three informal pilot user trials of the AudioGPS which were designed to evaluate it in the following areas.

- i. **Audio representations of direction and distance** We have carried out an informal pilot formative evaluation of the audio representations of direction and distance both with musicians (who might be expected to have a sound grasp of the principles upon which the audio representations are based) and non-musicians. We found that the distinction between being in front, behind, to the right and to the left of the user (represented by changes in the tone and placement of the sound on the stereo stage) was easy to discern, as were the intermediate four compass points. In other words, it is possible to perceive the direction of the sound source.
- ii. **Does the system work under field conditions?** Informal tests of whether the system works under field conditions were carried out on foot at night and in a car. In the former experiment we were primarily concerned with finding out how long it took for the AudioGPS to provide usable navigational information after having been switched on. (It takes time for any GPS system to acquire fixes from the orbiting satellites and use these to calculate its location, velocity and direction of movement.) Secondly, we were interested in how responsive the system was to directional changes in the path that the user was taking. Typically, it took approximately twenty seconds (or twenty metres while walking) for the system to initialise and start reporting usable bearing information. Subsequent changes in direction were reported more quickly than that.

In summary, the system worked useably for pedestrians. Our preliminary experiments with in-car use of the AudioGPS demonstrated that the system was not responsive enough to cope with direction changes faster than every 10 seconds or so (while driving around a car park). However, it worked adequately once the GPS system had settled down while driving along the open road.

Further work

With the system as currently implemented, we intend to find out how individual differences affect the effectiveness of the audio mapping we have adopted. We will also find out how well the system can be used by sighted users to find unknown locations on foot at varying distances. We will also consider other forms of locomotion. At the same time, we are upgrading the system to use a flux

magnetometer (electronic compass) and a rapidly reacting head tracker. The system will be augmented sparsely with speech audio. The head tracker is expected to improve front-rear discrimination and angular resolution - which we plan to test in field trials.

Summary

Audio GPS systems have a very important role for visually disabled users, but they have a slightly different role for mobile sighted users in situations of minimal attention, and where conversations may compete with speech audio information. In contrast with many (though not all) GPS systems, we have focussed on spatial audio rather than speech audio. Preliminary trials have suggested that very simple and computationally inexpensive spatial mapping is surprisingly adequate for finding locations, irrespective of musical training.

Acknowledgements

Thanks to Steve Brewster for hiking from Willen and many tips, Jim Ballas for some basics, Dave Haniff for drivers, Kristina H k for Elks in the Wood.

References

- [1] S. Kristoffersen and F. Ljungberg, "Making place" to make IT work: empirical explorations of HCI for mobile CSCW, presented at Proceedings of the International ACM SIGGROUP conference on Supporting group work (GROUP'99), Phoenix, AZ USA, 1999.
- [2] M. Weiser, The Computer For the 21st-Century, *Scientific American*, vol. 265, pp. 66-75, 1991.
- [3] J. Pascoe, N. Ryan, and D. Morse, Using While Moving: HCI Issues in Fieldwork Environments, *ACM Transactions on Computer Human Interaction*, vol. 7, pp. 417-437, 2000.
- [4] S. Kristoffersen and F. Ljungberg, Designing Interaction Styles for a Mobile Use Context, in *First International Symposium on Handheld and Ubiquitous Computing (HUC 99)*, vol. 1707, *Lecture Notes in Computer Science*, H.-W. Gellersen, Ed. Karlsruhe, Germany: Springer-Verlag, 1999, pp. 208-221.
- [5] J. Pascoe, N. Ryan, and D. Morse, Issues in developing context-aware computing, in *First International Symposium on Handheld and Ubiquitous Computing (HUC 99)*, vol. 1707, *Lecture Notes in Computer Science*, H.-W. Gellersen, Ed. Karlsruhe, Germany: Springer-Verlag, 1999, pp. 208-221.
- [6] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, GPS-based navigation systems for the visually impaired, in *Fundamentals of wearable computers and augmented reality*, W. Barfield and T. Caudell, Eds. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2001, pp. 429-446.
- [7] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, Navigation system for the blind: Auditory display modes and guidance, *Presence-Teleoperators and Virtual Environments*, vol. 7, pp. 193-203, 1998.
- [8] G. Moon, Personal Navigation Systems for the Blind, in *Institute of Engineering Surveying and Space Geodesy*. Nottingham: University of Nottingham, 1998, pp. 328.
- [9] H. Petrie, V. Johnson, T. Strothotte, A. Raab, S. Fritz, and R. Michel, MOBIC: Designing a travel aid for blind and elderly people, *Journal of Navigation*, vol. 49, pp. 45-52, 1996.
- [10] H. Makino, I. Ishii, and M. Nakashizuka, Development of navigation system for the blind using GPS and mobile phone combination, presented at Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 18, Pts 1-5, 1997.
- [11] K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou, Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences, presented at CHI 2000: The Future is Here, The Hague, The Netherlands, 2000.
- [12] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton, Cyberguide: A mobile context-aware tour guide, *Wireless Networks*, vol. 3, pp. 421-433, 1997.