gpsTunes – controlling navigation via audio feedback

Steven Strachan[1]
[1]Hamilton Institute
NUI Maynooth
Maynooth, Ireland
steven.strachan@nuim.ie

Parisa Eslambolchilar[1]
[1]Hamilton Institute
NUI Maynooth
Maynooth, Ireland
parisa.eslambolchilar@nuim.ie

Roderick Murray-Smith[1,2]
[2]Dept. of Computer Science
University Of Glasgow
Glasgow, Scotland
rod@dcs.gla.ac.uk

ABSTRACT

We combine the functionality of a mobile Global Positioning System (GPS) with that of an MP3 player, implemented on a PocketPC, to produce a handheld system capable of guiding a user to their desired target location via continuously adapted music feedback. We illustrate how the approach to presentation of the audio display can benefit from insights from control theory, such as predictive 'browsing' elements to the display, and the appropriate representation of uncertainty or ambiguity in the display. The probabilistic interpretation of the navigation task can be generalised to other context-dependent mobile applications. This is the first example of a completely handheld location- aware music player. We discuss scenarios for use of such systems.

1. INTRODUCTION

We describe the first prototypical implementation of a system, *gpsTunes* which adapts the music currently playing, in order to guide a user to a desired physical location. Put simply, this a working example of a pocket-sized location aware system, based on GPS (Global Positioning System) signals, which presents different audio cues to the user according to location and heading. For the illustrative example we show that with something as simple as volume adaptation and panning, with fixed targets, people are able to navigate successfully.

The merging of functionalities of different types of mobile device has become more common in recent times. Portable music players, personal digital assistants and smart phones have become extremely popular in their own right, with huge reductions in size achieved along side massive increases in capacity. There is now a drive to merge the different functionalities of these devices. Examples of this include Nokia N91 and Sony-Ericsson W800 'walkman' phones which combine the functionality of a high-capacity mp3 player with a standard mobile phone, and number of other phones now on the market which offer all of the functionality of a standard PDA. Mobile GPS, in the past, was largely confined to maritime use, but many new cars already come with built-in GPS, hikers and hill walkers can commonly be found with a GPS and some GPS-based games have been developed, e.g. [1]. Here, we incorporate GPS with an mp3 player to provide a handheld system capable of providing a user with the ability to navigate themselves to a desired location, through continuously adapted music, whilst 'on the move'.

1.1 Location-aware audio

Problems associated with the usability of mobile devices have been extensively documented [4]. There are significant problems associated with user attention and cognitive load, when using a mobile device whilst actually 'on the move'. Will the navigation cues actually work in real-world scenarios? Recent work in this area includes that of Holland *et al.* [3] who describes a prototype spatial audio user interface for a mobile GPS. Their interface is designed to allow mobile users to perform location tasks while their eyes, hands or general attention are otherwise engaged. They found with the use of a spatial, non-speech geiger counter style sound and a prototype 'back-pack' based system that very simple and computationally inexpensive spatial mappings are surprisingly usable for helping users to find locations. Warren *et al.* [8] have more recently shown, with their VRML simulation of a similar system, that embedding navigation cues within music is a promising new approach to the problem of supporting mobile users in route tracking tasks. Their initial lab-based experimental work suggested that users would be able to follow routes by simply keeping track of the volume and perceived direction of a music source.

2. EQUIPMENT

The equipment used for our system consists of an HP iPAQ 5550 equipped with a MESH [5] inertial navigation system (INS) backpack consisting of 3 Analog Devices ±2g dual-axis ADXL202JE accelerometers, 3 Analog Devices ±300deg/s Single chip gyroscopes, 3 Honeywell devices HMC1053 magnetometers and a vibrotactile device, used for feedback purposes. A standard orthogonal inertial sensor arrangement is used with the sensitive axis of the respective inertial sensors mounted coincident with the principle device axes providing us with direct measures of lateral accelerations, turn rates and magnetic field strength as well as the current GPS latitude and longitude.

The GPS system we use is a Trimble Lassen Sq module, produced for mobile devices, and is also built-in as part of MESH (see figure 1). This module boasts a 9m resolution

Figure 1: Left: Mesh device alone and attached to an HP5550 Pocket Pc. Right: The MESH circuit board showing the main components related to the navigation task

with up to 6m resolution around 50% of the time it is used. It also provides us with velocity resolution of 0.06m/s and an 18m altitude resolution. This module suffers the same problems that most GPS modules suffer, in that there are occasional problems with resolution, latency, signal shadowing and noise in the signal which can be detrimental to a system such as ours, as we have found with initial testing. It is for this reason that systems like this need further support from other inertial sensors such as accelerometers, gyroscopes and magnetometers, which we have at our disposal with MESH. This equipment provides us with an unprecedented level of information about the current inertial state of our mobile device and sensor fusion algorithms are currently being developed to further increase the general usability of this hardware and to integrate the INS/GPS components of the system, eliminating as much uncertainty as possible.

3. FEEDBACK FOR CONTROL

The behaviour exhibited by any system can be interpreted as trying to control its perceptions [7]. We view the task of a mobile navigation system as being that it should provide appropriate feedback to allow the user to control their location, compared to their goals. In the case of GPS-based navigation, this will typically be longitude, latitude and bearing. On a mobile device, the user perceives the feedback via audio, vibrotactile and visual displays. In this paper we are focussing on the audio element, in order to reduce the need for screen-based interaction with the device while mobile.

In many situations we have to make decisions, and control a system based on uncertain evidence. Representing ambiguity is especially significant in closed-loop continuous-control situations, with sensor uncertainty, and where the user is constantly interacting with the system to achieve some goal, which may not be known to the machine. Formulating the ambiguity in a probabilistic framework, we consider the conditional probability density functions of sequences of actions associated with each potential goal in the system. [9] introduced the use of granular synthesis methods to sonify probability distributions, and this general approach can be applied to the use of location-aware audio presentation. The concept of 'location' is viewed as being described by a probability density, conditioned on the current sensed position. This is a general technique which can be applied to other notions of context than spatial context, i.e. if we view a context as the probability of a particular random variable, given the evidence available. If the display has the 'appropriate' amount of ambiguity, this should lead to smoother control actions, as they can judge how much to 'trust' the feedback. The opposite case of an uncertain variable always being presented as the complete truth would lead to much more variable behaviour. For a location-aware system, we need to consider how we can represent the uncertainty of goals and states in an appropriate manner.

We also need to give the user some predictive ability – to know, if they keep going in that direction they will get to the goal. 'Quickening' (or 'predictive displays') is a method for reducing the difficulty of controlling second-order or higher order systems, by changing the display to include predictions of future states, reviewed in [2]. E.g., a quickened display for an acceleration control system shows the user a weighted combination of position and velocity. This weighted summation effectively anticipates the future position of the system. It can greatly improve human performance in controlling these systems. Quickening in general is a prediction of the future state of the system based on the current state vector (for example position, velocity, acceleration) and a model of system behaviour and expected user action. An example of this is based on the Doppler effect, which highlights the user's approach to a target, or a target's movement from the current state, in the same way we hear the pitch of an ambulance siren change as it passes us at speed. It can also be applied to browsing tasks – the user could wave the device in a particular direction and get predictive feedback indicating what would happen if they continued in that direction. Utilizing other display modalities, such as vibrotactile feedback, could help the user disambiguate predictive from current state information. This was used in a PocketPC context in [6] – applying such techniques to a larger scale system such as this may prove to be effective. Below we discuss the 'browsing' mode in our system which helped users find their targets in a more direct fashion than having to run back and forth to estimate gradients.

4. GPS-TUNES

To our knowledge, our initial prototype implementation of the *gpsTunes* system is the first of it's kind implemented in a truly hand-held, real-world situation.

Figure 2: User holding the PocketPC and MESH in hand.

Our current system is designed to guide a user to a desired target by varying the volume and 'bearing' or direction of the currently playing song. So, for example, if a user enters an area with which they are not familiar and they wish to locate their desired building, they may inform the system of where they wish to go with a click of a map, which will then alter the volume level and bearing of the music played. They may then attempt to move towards the sound source keeping the music in front. As they move closer to the target, the volume of the music will increase, reaching maximum (user preferred) volume at the point where the target has been reached. At this point they will be notified of their arrival by an additional pulsing sound played over the current track.

When building a system such as this the two most important pieces of information to convey to the user are the *distance from their desired target* and the *current direction, relative to targets* [3]. In our system the distance is conveyed by a change in volume. A Gaussian density is placed around the chosen target, and this is mapped to volume of the sound source. The music switches to the lowest audible volume on the edge (a threshold value) of this distribution. As the distance to the target is decreased the volume increases back towards the users preferred level. The direction of the current target is conveyed to the user by panning the sound 'source' around their head. When the user clicks their desired target, the bearing to the target is calculated using the current GPS estimate of latitude and longitude. Using the heading calculated from the calibrated magnetometers in MESH allows the system to pan the sound source to the correct position, from the user's perspective. The user can rotate on the spot and hear the sound source effectively moving round their head.

We also include an additional *browsing* facility in our current version. The user may switch from the general navigation mode into a browsing mode which allows them to gauge the current direction of all targets by rotating around, effectively pointing the device at each target.

5. USER FIELD TRIALS

Two scenarios were devised to test this system, walking in a sports field. In the first case, four users were asked to walk to four hidden targets, in order, using only the sound presented to them by the system. First they move to target 1 and when this is found they select target 2 and so on. In the second scenario each user is also allowed to use the browsing facility described previously. First they attempt to locate target 1 then browse for target 2 and so on until all 4 targets have been located.

Figure 3 shows the different paths taken by each user in the first scenario. It was found that at the beginning all users first rotated before realising the general direction of the first target and moving off in that direction. All users then perceived the sound source was moving to their left and adjusted their heading appropriately to move onto target 1. A similar procedure was used to traverse to the remaining targets. Most confusion with the system was displayed on the approach to target 1, as can be observed by the paths in figure 3, but after this each user located then moved towards the next target with no significant problems.

Figure 4 shows the paths taken to four new target locations with the introduction of the additional browsing facility. The paths in this figure indicate much less initial confusion by the users, although this may be due to the fact that the users were now more familiar with the system. All but one user used the browsing facility at the beginning of their journey to locate the direction of the first target and then moved off in that direction. Two of the users then used the browsing facility again to update their path to target 1. The introduction of the browsing facility reduced significantly the variable nature of the paths shown in figure 3, making them much more direct once the direction of the next target had been pinpointed during browsing.

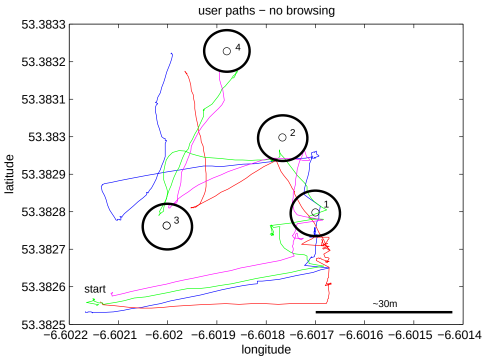

Figure 3: Paths taken by four users to four targets.

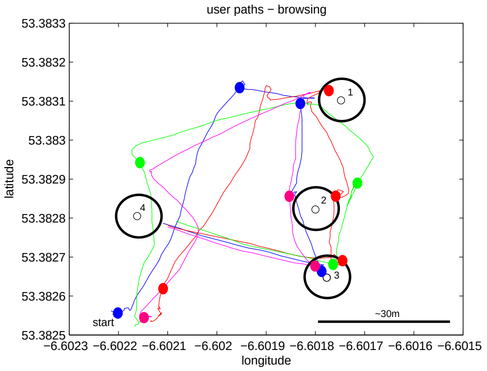

Figure 4: The paths taken by four different users to four targets with an additional browsing facility (activity highlighted by filled circles).

It is also possible to gain insight from our other sensor data. Figure 5 shows a time-series plot of one user's z-axis acceleration trace and their heading, plotted against time. Flatter areas of the acceleration trace show times when the user was not walking and more variable areas show where the user was walking and it is observed that the browsing by this user coincides with times when they have stopped walking.

6. FUTURE WORK AND OUTLOOK

6.1 Current Status

For this initial prototype we have implemented the most basic 'skeleton' system. Initial testing has shown that this concept does allow users to navigate in the real world, using a combined Audio/GPS player to aid navigation, but obviously are too preliminary to to say much about the usability of the system for more complex tasks. Our prototype, which

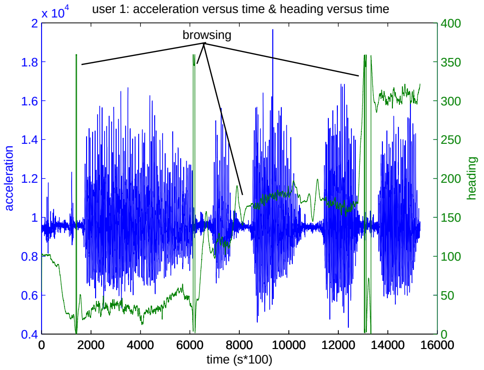

Figure 5: z-acceleration & heading time-series.

combines GPS latitude and longitude with magnetometer-derived bearing data, has shown that it is possible to test this in a handheld device. It is hoped that in future prototypes we may take fuller advantage of the other sensors at our disposal in MESH to produce a more robust and flexible system.

6.2 Future Work

Our prototype implementation of *gpsTunes* on a pocket pc with a MESH combined IMS and GPS provides us with a strong basis for any future expansion of the system. A future implementation would involve the use of our accelerometers and gyroscopes to decrease uncertainty, given the GPS and magnetometers alone. Use of the accelerometers and gyroscopes would allow us to incorporate tilt into the system which could, for example, be used as a form of range finding from the predictive browsing component. As the user tilts the device, the system could effectively change its predictive horizon to different distances in the scene ahead, and in changing uncertainty about range, control the number of sources of music heard.

Although in this paper we are not attempting to maximise the music listening pleasure of the user, this should become more of a priority in future versions of the system. E.g., using a more sophisticated audio feedback transformation than a simple stereo pan, used in the present system, to convey direction. We can be creative in how we implement our ideas involving quickening, ambiguous interfaces and probabilistic interfaces in a multimodal context. So, for example, instead of simply varying the volume of a song to convey distance, we may attempt to convey this distance by parameterising general signal processing filters (e.g. emulating tuning in and out of a radio station), complimented by vibrotactile feedback. True spatial signal processing based on the relevant psychoacoustic phenomenon, to convey the direction would eliminate any ambiguity as to whether the desired direction to progress is in front or behind. We may also provide alternative distance cues such as frequency attenuation at high distances, akin to muffled voices in the distance with less frequency attenuation at lower distances. For added realism, reverberation could be added in decreasing amounts as the user approaches the target, further reinforcing their sense of distance.

There is a host of possible research questions we can now address with the current system. While a person is walking down a street immersed in their own personal choice of music they may be detached from the outside world. How may we adapt what they are listening to, in order to provide some form of location awareness without necessarily disrupting the user's listening pleasure? Could users continuously shape and evolve, in a stigmergic fashion (the way termites build their nests, where previous work directs and triggers new building actions), a musical landscape in parallel with their physical space? Another interesting possibility with this system is to present the audio source to the user in a 'carrot and stick' fashion. So, instead of having the user traverse to a set target, they may follow an audio source around a set path with out any volume level alteration. This path may, for example, represent the best route around a town, which doesn't involve passing through buildings or may be used as a virtual guide to guide users around outdoor tourist locations. Use of dynamic features such as velocity would allow it to be used for sports training purposes, you could, for example, always compete against your best run for a familiar course.

7. ADDITIONAL AUTHORS

Additional authors: Stephen Hughes (Stephen Hughes Engineering Services, 1 Leopardstown Drive, Blackrock, Ireland, email: `stephenahughes@gmail.com`) and Sile O'Modhrain (Lecturer in Haptics, Sonic Arts Research Centre, Queens University, Belfast, email: `s.omodhrain@qub.ac.uk`).

8. REFERENCES

[1] M. Chalmers, M. Bell, B. Brown, M. Hall, S. Sherwood, and P. Tennent. Gaming on the edge: Using seams in ubicomp games. In *to appear in Proc. ACM Advances in Computer Entertainment (ACE'05)*, pages 306–309, 2005.

[2] R. J. Jagacinski. J. M. Flach. *Control Theory For Humans - Quantiative Approaches to Modeling Human Performance.* Lawrence Erlbaum Associates, London, first edition, 2003.

[3] S. Holland, D. R. Morse, and H. Gedenryd. AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal and Ubiquitous Computing*, 6(4):253–259, 2002.

[4] S. Kristoffersen and F. Ljungberg. Making place: to make IT work: Empirical explorations of HCI for mobile CSCW. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work 1999*, pages 276–285. ACM Press, 1999.

[5] I. Oakley, J. Ängeslevä, S. Hughes, and S. O'Modhrain. Tilt and feel: Scrolling with vibrotactile display. In *EuroHaptics 2004*, 2004.

[6] R. Murray-Smith P. Eslambolchilar, A. Crossan. Model-based target sonification on mobile devices. In T.Hermann and A.Hunt, editors, *Proceedings of Interactive Sonification Workshop, Bielefeld*, January 2004.

[7] W. T. Powers. *Behavior: The control of Perception.* Aldine de Gruyter, 1973.

[8] N. Warren, M. Jones, Jones S., and D Bainbridge. Navigation via continuously adapted music. In *Extended Abstracts ACM CHI*, 2005.

[9] J. Williamson and R. Murray-Smith. Sonification of probabilistic feedback through granular synthesis. *IEEE Multimedia*, 12(4):45–52, April/June 2005.

Funding Acknowledgements

We are grateful for support from: HEA project *Body Space*, SFI, EPSRC project *Audioclouds* and EC-funded *PASCAL* network.