# Navigation via Continuously Adapted Music

**Nigel Warren, Matt Jones, Steve Jones**
**& David Bainbridge**
Department of Computer Science
University of Waikato, Hamilton, NZ
{nw53,mattj,stevej, davidb}@cs.waikato.ac.nz

## ABSTRACT
Listening to music on personal, digital devices while mobile is an enjoyable, everyday activity. We explore a scheme for exploiting this practice to immerse listeners in navigation cues. Our prototype, *Ontrack*, continuously adapts audio, modifying the spatial balance and volume to lead listeners to their target destination. An initial lab-based evaluation has demonstrated the approach's efficacy: users were able to complete tasks within a reasonable time and their subjective feedback was positive. Encouraged by these findings, we are building a pocket-sized prototype for further testing.

### Author Keywords
Mobile systems, navigation, audio, ambience, GPS.

### ACM Classification Keywords
H5.5. Sound and music computing.

## INTRODUCTION
Personal portable music players are essential pieces of street equipment. While such devices have been available for over a quarter of a century, recently digital versions have emerged. These provide powerful onboard processing and integration with other computing resources via fixed and wireless connections. Commercial services are exploiting these innovations providing, for instance, mobile online music purchase and graphical information to complement the audio stream.

Our interest, however, is the types of information that can be presented to a listener using only the music they are immersed in. Here, we introduce and evaluate the notion of adapting the acoustic signal to provide location and context awareness while mobile.

We have built a prototype system – *Ontrack* – to help a pedestrian navigate to a target destination via calm feedback. The concept is illustrated in this use-scenario:


*CHI 2005,* April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

*Ben is going to meet a friend in a restaurant across town. After alighting from the tram station closest to the restaurant, he puts on his headphones and turns on his portable music player to hear a new song he's just downloaded. At the same time, he turns on the navigate-by-music feature; the restaurant location was copied from the email his friend sent him.*

*He's fairly certain of the direction to head and as he moves off, the music playback is clear and strong, at a normal volume and through both headphones. As he approaches a cross-roads, Ben perceives the music shifting slightly to his left, with the volume decreasing. He crosses the road and heads in that direction and the music is now balanced and the volume returns to normal. As he walks further along the street, he notices the music beginning to fade, the volume is decreasing; he looks around and notices he's just walked past his destination*.

The remainder of the paper is structured as follows. First we consider alternative approaches to assisting listeners as they navigate on foot, and place our work in the wider context of ambient interfaces. Next, we describe the implementation of the *Ontrack* lab prototype used to evaluate the scheme. The experimental method and result are then discussed. We conclude by outlining our ongoing work aimed at further understanding the role of the approach and the implementation on a pocket-sized device using readily available technology.

## RELATED WORK
Commercial navigation systems routinely use speech to provide route guidance. In the research literature, there has also been interest in using discrete, meaningful clips of non-speech audio to convey directional information. Spatial audio – where the listener's perception of the location of the audio source is manipulated – has been used in conjunction with such notifications: e.g., a harpsichord tone emanating from the listener's left means they should move onwards in that direction [5]. In these systems, the user often walks around "listening" to silence, interrupted with audio cues when they need to take notice. For everyday users, this sort of scheme does not seem attractive: who wants to wear headphones to listen to nothingness interspersed with

ambient noise? Rather, our approach is a symbiotic one, exploiting a desired, natural use of a personal player, it continuously modifies the characteristics of the listener's own choice of music.

*Ontrack* provides a type of minimal attention user interface [6] - the listener does not have to use their eyes or hands to make use of the system. When the user does choose to attend to the navigational information, they can do so while still being immersed in the music: additionally, then it is minimally disruptive.

The location and context information is provided in an ambient way. Conventionally, an ambient display adds information to move from the peripheral to the centre of a user's attention: a dangling wire twitches as data activity increases; a surface of water increasingly ripples as a web server dishes out more pages [3]. In contrast, our approach sees information being subtracted when the significant events occur – the full, normal stereo audio presentation only occurs when the user remains on track to the target. If they wander off the path, audio output is reduced in either one of the channels or in terms of volume.

Managing discrete event notifications so that the level of disruption is acceptable to the user can be highly challenging. Filtering rules, past interactions and current contextual information – such as the user's level of activity – have been proposed [7]. In *Ontrack*, as notification is continuous and embedded within the audio stream, there is a simple and direct way for the user to finely control the prominence of the navigation information. Just as they might increase or decrease the volume or bass/treble mix of the music playback, they can similarly increase and decrease the "volume" or degree of the adaptations. So, a user may amplify the level of guidance in circumstances when they are more uncertain of the route; and turn it down when they have greater confidence.

## LAB PROTOTYPE

Before committing to a mobile implementation, we wanted to evaluate the efficacy of the approach. To this end, we built a lab-prototype, simulating the navigation task on a desktop PC.

The prototype places users in a 3D virtual world, in the grounds of a local large public hospital. The user can navigate through this world freely, using cursor controls, negotiating block-form buildings. While wandering around the environment, music plays through their headset.

Routes through the world are created by placing audio beacons at appropriate locations along the path to the destination. The user's relative positioning to a beacon affects the quality of the stereo playback. When they are facing a beacon, music is played in both left and right ears. If they turn their gaze away from the beacon towards the left, the music will be shifted more to the right ear than the left, and vice versa if they turn to the right. As the user approaches a beacon, the music volume gradually increases. Conversely, if the user heads away from a beacon the volume decreases to indicate they are going in the wrong direction.

Navigation proceeds, then, by the user heading towards successive beacons along the path. Once the user reaches a beacon, the music continues to play but the spatial balance and volume are adjusted to signal a new intermediary target. That is, the perceived location of the music changes to point to the direction of the new beacon by shifting to the right or left channel, and the volume is decreased to indicate the need to progress to a new source. In this way the user is able to navigate by following an audio route from the starting location to the destination.

The prototype is implemented using a third-party open source 3D engine, to manage the on-screen virtual world, and an open source audio engine, for music playback. On each program cycle, the user's distance from the current beacon as well as orientation relative to the beacon are computed and used to determine the volume level and stereo panning of the audio. Thus, as the user moves throughout the virtual world, the location of the sound relative to the user is constantly updated to reflect their current position and give the illusion of sound emanating from a specific point in the world.

## EXPERIMENTAL EVALUATION

### Method

A between groups design was employed with 25 subjects. One group used the *Ontrack* system while the other used a control prototype, referred to from here on as *Map*. The *Map* users were presented with the same 3D world and also listened to identical music tracks as they navigated. However, the music stream was not adapted in any way. This group was also given a birds-eye view paper map of the world which they could consult at any time.

Each user had to complete three route navigation tasks. In each case the starting and destination points were different. Two task routes (A & C) were of a similar length and complexity, while one (B) was shorter and simpler. The order of tasks was randomized for each participant. Routes were not dynamically recalculated to accommodate user errors.

Trials proceeded in the following way. After completing a consent form, the user was given instructions explaining how to move through the 3D world. In the *Ontrack* condition, a brief explanation of the how the music would be adapted was also provided. Then, the users were asked to perform a training task, navigating to a location.

For all of the tasks, including the training one, a birds-eye view map was presented on screen. The starting and target locations were marked on this display. After 20 seconds, the user was placed within the 3D world at the initial location and had to navigate to the destination. While moving, the users could not vary the speed of their progress. All targets were represented as lit doorways and when reached a "Success" message was displayed. The user could quit a task at any time by hitting the "escape" key.

The *Ontrack* condition simulated more closely the experience of a visitor to an unfamiliar site, such as a hospital. They encounter a plan layout initially and then have to find their way to their destination.

The program tracked the users' routes through the map and automatically calculated times to completion. After a user had performed all three tasks, they were asked to complete a questionnaire based on the NASA TLX [4]; this was designed to assess the physical and cognitive load of the activities.

All participants were drawn from the staff and students of a local university. Participants were randomly assigned to a group – 13 were placed in the *Ontrack* group (8 male, 5 female); and the remaining 12 in the control (7 male, 5 female).

### Results

Quantitative task performance is shown in Tables 1 and 2. Successful completion was defined as the user reaching the target destination. Table 1 presents the results for all such success cases. Performance of a Mann-Whitney test on the task times indicates there is no statistically significant difference between the two groups ($p=0.66$). Table 2 shows the performance in terms of the three routes.

For both conditions, there was the occasional extremely untypical performance: a subject would wander for long periods within the world. Defining such outliers as being at least two standard deviations from the mean completion time, we found two cases in the *Ontrack* condition (both in Route A tasks – where the subjects took 555s and 480s respectively); and, 4 in the *Map* data set. With these points removed, though, there is also no statistical significance between the conditions ($p=0.65$).

| Condition | Task success rate. | Mean successful task time (std. dev) *secs.* |
|---|---|---|
| *Ontrack* | 32/38 (84%) | 117.8 (110.8) |
| *Map* | 28/32 (87%) | 105.8 (57.1) |

**Table 1: Successful completion performance. 5 tasks were not logged due a software failure.**

| Condition | Route A | Route B | Route C |
|---|---|---|---|
| *Ontrack* | 184.7(164.1) | 57.1(13.9) | 124.6(39.6) |
| *Map* | 124.5 (51.2) | 63.9(22.9) | 120.2(69.1) |

**Table 2: Successful mean completion times in seconds (and standard deviations) for the three routes. Route presentation orders were randomised for each subject.**

| Task-load dimension | *Ontrack* mean (std. dev) | *Map* mean (std. dev) |
|---|---|---|
| Mental effort | 7.3 (2.3) | 5.9 (2.4) |
| Perceived success | 8.2 (1.5) | 7.9 (1.5) |
| Performance satisfaction | 8.1 (1.3) | 7.6 (2.6) |
| Confidence in ability to complete tasks | 8.4 (1.6) | 8.0 (1.9) |
| Frustration level | 8.7 (1.5) | 7.1 (2.4) |
| Overall task load rating | 8.1 (1.6) | 7.3 (2.3) |

**Table 3: Subjective task load ratings for *Ontrack* and control. Normalised rating scale: 1 (least positive) to 10 (most positive rating).**

Subjective task load ratings for the two groups are given in Table 3. Ratings were made on a 10 point scale, where 1 was the worst and 10 the best score: higher scores are more favourable, so, for example, the degree of frustration experienced with *Ontrack* was low. The overall task load index ($p=0.09$) and individual components were examined using the Mann-Whitney test and none were found to show statistically significant results.

### Discussion

In the *Map* condition, the subjects could refer to the printed plan at will and often did, glancing from the screen to the map. Meanwhile, the *Ontrack* group used only the music-based cues and their memory of the plan shown at the start of each task. Even so, somewhat to our surprise, this group was able to complete the tasks, doing so within a reasonable period of time.

*Ontrack's* unconventional way of presenting navigation information, and the paucity of training provided, did not seem to be a major obstacle. To put the completion times in perspective, an expert with good knowledge of an optimal route can complete the A, B and C routes in 70, 44, and 83 seconds, respectively.

There is evidence, then, that the approach is intrinsically helpful – users can reach their destinations using navigational information encoded in the music; and, the subjective ratings were favourable. Looking at *Ontrack* in comparison with the *Map* case suggests a similar level of performance and subjective response (however, no statistically robust conclusions are able to be drawn from this initial study).

All routes in both conditions led to a high degree of completion time variability. Increasing the number of test subjects or further constraining the experimental conditions may lower this. Plots of the paths taken by each participant during the tasks, however, indicate that while task success time varied somewhat, the actual number of different gross routes was lower (see Figure 1 for an example).

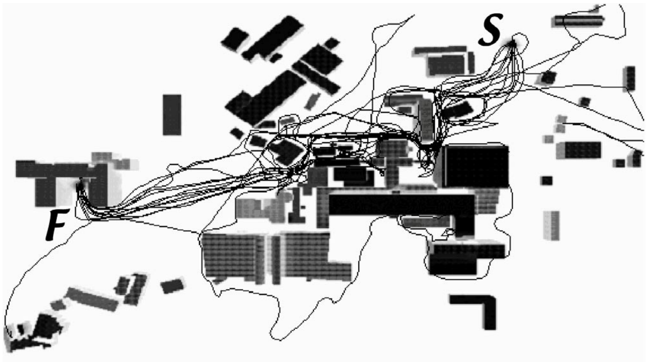


**Figure 1. Each line shows the route taken by a participant as they navigate from the start (S) to finish (F) point in the *Ontrack* condition. Most people followed the route suggested by the audio.**

## ONGOING WORK

At a theoretical level, there are two interesting areas for investigation. Firstly, the model currently used to adapt the music is very simple. While appropriate for the controlled 3D world, in cases where there is more uncertainty (e.g. a moving target or inaccurate location readings), sophisticated, probabilistic models seem more suitable [8]. Secondly, we are investigating a vocabulary of music adaptations, in the tradition of work on auditory earcons [1]. What, for example, would be an effective way to notify a listener of an interesting location or person in their vicinity? There is much scope for a richer set of modifications, the simplest involving use of a true 3D spatial audio system [5].

While we have considered navigation from and to a fixed point, the approach is applicable in a wide range of route following and target location tasks. For instance, a *tourist service* might guide a listener around an interesting route while they listen to either their own music or a commentary. A *find-a-friend service* could lead the listener to rendezvous with a nearby acquaintance. In a gaming context, these people would be targets as in [2]; again the audio content – a gaming soundtrack, or communications with other players – could be seamlessly integrated with navigation information. Meanwhile, a *child monitoring system* could act as a reassurance aid: a parent sits in the park, reading their magazine, listening to music as their young children play. The music grows softer as the children move further away; when the volume is at uncomfortably *low* level, the parent looks up and calls them back.

While users were able to perform effectively with the lab-prototype, finding the experience pleasant and not onerous, field conditions may impact on the performance and acceptance. For example, our lab participants rated the *Ontrack* experience as not frustrating; but, how would they feel after prolonged use in conjunction with normal music listening?

To further explore the approach we are implementing a Pocket PC based prototype. Location information is received (via Bluetooth) from a small, commercial GPS unit. While there are resolution accuracy and latency problems with GPS, prior prototypes suggest these can be accommodated [5]. In particular, we are incorporating an inertial sensor which will enable very accurate detection of orientation.

## CONCLUSIONS

Embedding navigation cues within music is a promising new approach to the problem of supporting mobile users in route tracking tasks. Our initial experimental work suggests that users will be able to follow routes by simply keeping track of the volume and perceived direction of the music source. Further, the work indicates that the process of doing so will not be cognitively demanding and may become an enjoyable activity in itself. These results were encouraging given the relatively simple nature of the music adaptations involved.

## ACKNOWLEDGMENTS

Ben Bederson & Carl Gutwin for their insights.

## REFERENCES

1. Brewster, S. (1993). An evaluation of earcons for use in auditory human-computer interfaces. *Proc. SIGCHI conference on Human factors in computing systems (CHI'93)*, 222-227.
2. Crabtree, A., et al. (2004). Orchestrating a Mixed Reality Game 'On the Ground'. *ACM CHI Letters, 2004* **6**(1): 391-398.
3. Dahley, A., Wisneski, C. & Ishii, H. (1998). Water lamp and pinwheels: ambient projection of digital information into architectural space. *Proc. SIGCHI conference summary on Human factors in computing systems (CHI'98)*, 269-270.
4. Hart, S.G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, in *Human mental workload*, P.A. Hancock and N. Meshkati,(*Eds.*). 1988, North Hollland: Amsterdam. p. 139-183.
5. Holland, S., Morse, D. & Gedenryd, H. (2002). AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal and Ubiquitous Computing,* **6**(4): 253-259.
6. Pascoe, J., Ryan, N. & Morse, D. (2000). Using while moving: HCI issues in fieldwork environments. *ACM Transactions on Computer Human Interaction*, **7**(3): 417-437.
7. Sawhney, N. & Schmant, C. (2000). Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interaction*, **7**(3): 353-383.
8. Williams, J. & Murray-Smith, R. (2004). Granular Synthesis for Display of Time-Varying Probability Densities. *Proc. International Workshop on Interactive Sonification*, Bielefeld, Germany.