

# Navigation-by-Music for Pedestrians: an Initial Prototype and Evaluation

**Matt Jones**

FIT Lab, Computer Science Department  
University of Wales, Swansea, UK  
always@acm.org

**Gareth Bradley, Steve Jones & Geoff Holmes**

Department of Computer Science  
University of Waikato, NZ  
{gb20,stevej,geoff}@cs.waikato.ac.nz

## ABSTRACT

Digital mobile music devices are phenomenally popular. The devices are becoming increasingly powerful with sophisticated interaction controls, powerful processors, vast onboard storage and network connectivity. While there are 'obvious' ways to exploit these advanced capabilities (such as wireless music download), here we consider a rather different application—pedestrian navigation. We report on a system (ONTRACK) that aims to guide listeners to their destinations by continuously adapting the spatial qualities of the music they are enjoying. Our field-trials indicate that even with a low-fidelity realisation of the concept, users can quite effectively navigate complicated routes.

## Author Keywords

Mobile navigation, digital music devices, audio adaptation.

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): audio, ambient.

## INTRODUCTION

People like to listen to music while on the move. Since the inception of the first tape-based Walkman, Sony has sold more than 150 million personal music players. Five and a half million Apple iPod devices, the current leader in the market, were sold in the first months of 2005 alone. Millions of songs are purchased and downloaded from digital music stores daily, even whilst on the move, providing a musical background to everyday activities.

Some of the time that people are on the move is spent traveling well-known routes with a rehearsed path to a destination; navigation is easy and requires little thought. In unfamiliar locations, however, a great deal of attention may be afforded to navigation, with reference to maps, instructions, street signs, landmarks and other cues.

In an earlier paper, we introduced the novel approach of continuously modifying the playback of music tracks on a portable player to indicate both direction and distance from target locations to a listener [14]. The aim is to introduce a minimally intrusive mechanism into a common activity such that it provides effective support with little attentional overhead for the user.

The concept is illustrated in this use-scenario:

*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 concept of an audio navigation system is attractive because most situations where we need a navigation aid require the use of our eyes. A non-speech based system is perhaps even more desirable because the user can process sounds in the background. This way they can be enjoying their surroundings or concentrating on some other task while they navigate intuitively.

This is one of the most important aspects of the system: rather than making users learn a new behaviour, it extends an existing behaviour. Because it builds on an enjoyable activity, there is an increased likelihood of user adoption. Other ambient audio systems may suffer from the fact that they are neither natural nor fit with the everyday use of audio [4]. The benefit of this concept, then, is in exploiting a widely used and enjoyable activity—listening to music whilst walking—to provide intelligence about the user's environment.

Previously, we evaluated the concept in a laboratory-based virtual world prototype. That is, participants performed route navigation tasks using a mouse while listening to audio tracks and viewing a 3D block world on a desktop computer. The results indicated that the scheme had interesting potential [14].

This paper builds on our prior proof-of-concept laboratory work. We describe the mobile version of our ONTRACK system, and present the results of a field evaluation, which reinforce the potential of our minimally intrusive approach to the provision of audio navigation cues.

### NAVIGATION AND AUDIO

Conveying navigational information to users can take many different forms. GPS technology has been used in various systems, like ship and plane navigation, for years. Commercial car navigation systems routinely use speech to provide route guidance. Such systems—such as the Garmin StreetPilot c320<sup>1</sup>, Magellan RoadMate 700<sup>2</sup>, TomTom Go<sup>3</sup>, among many others—display instructions on a screen, and, when approaching a turn, vocalise the instructions, for instance “turn left” or “veer right”.

Pedestrian navigation can require different sorts of information as May *et al* discuss. They found that, “landmarks were by far the most frequently used category of navigation information identified” [7, p.336]. Information relating to the visual appearance of bridges, shops, restaurants, and post boxes, among other items, was the category of information participants found most useful. Distance information and street names were among the types of information participants found least useful.

Work has been done to consider various methods of providing this information on handheld computers, covering text and spoken instructions, 2D route sketches, 2D maps, and 3D maps [6].

While speech seems an obvious solution to navigation, there has also been interest in using discrete, meaningful clips of non-speech audio to convey directional information. Nemirovsky observes that such alternative methods of information delivery can provide people:

*“... in situations of information overload, with a background information channel, leaving our foreground concentrated on the more thought-demanding tasks.” [8, p.2].*

To explore this notion, he developed GuideShoes, using GPS navigation and “emons”, short audio clips, to help users navigate. The system uses a GPS unit attached to the user’s leg that transmits coordinates to a remote base station. The base station calculates the navigational route and chooses which emon to play, sending it back to the user

over FM radio. The project was developed in the context of investigating whether it is possible to construct a language vocabulary using short pieces of non-speech audio that convey meaning.

Spatial audio – where the listener’s perception of the location of the audio source is manipulated – has also been explored in conjunction with such discrete navigational notifications. So, in AudioGPS the user is presented with audio cues emanating appropriately from the left or right channel to indicate the direction they should head [4]. As the system could only represent 180 degrees of directional information, two audio cues were used to inform the user whether the target is directly in front of them or behind them. Additionally, target distance was mapped to a series of audible clicks—the more rapid the clicks the closer the user is to the target.

In systems such as AudioGPS and GuideShoes, 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.

Using a continuous non-speech audio source such as music has potential in light of a study by the US Army on various types of acoustic beacons for navigational purposes [13]. The study concluded that listeners preferred continuous rather than pulsed beacons, and also that they preferred non-speech over speech audio. Additionally, in discussing ESPACE 2, an extensively featured speech and non-speech audio environment, the authors note that:

*“... auditory artifacts such as data changing over time [...] are better represented with continuous patterns of sound.” [10]*

The NavBelt [11] system demonstrates the use of both discrete and continuous audio cues to assist visually impaired users. The system is able to give the listener an awareness of obstacles in the path as well as providing guidance to a target. ‘Clicking’ noises played using a form of spatial audio indicate locations. In the system’s ‘image mode’, complex audio processing is used to present a soundscape to convey a feel of the environment in order that the user may navigate independently, in effect using their ears as eyes.

<sup>1</sup> Garmin International. Garmin StreetPilot c320 car navigation system. <http://www.garmin.com/products/sp320/> (last accessed 13/01/2006).

<sup>2</sup> Thales Navigation, Inc. Magellan RoadMat 700 car system. <http://www.magellangps.com/en/products/product.asp?PRODID=955> (last accessed 13/01/2006).

<sup>3</sup> TomTom BV. TomTom Go car navigation system. <http://www.tomtom.com/products/product.php?ID=74&Language=1> (last accessed 13/01/2006).



**Figure 1. ONTRACK on the move. User is holding iPaq in their right hand, the GPS device in their left and listening to the adapted music via headphones.**

#### **ONTRACK PROTOTYPE**

We developed the ONTRACK system on a HP iPaq hx4700 Pocket PC, connected to a Garmin eTrex Summit—a Global Positioning System (GPS) device that includes a magnetic compass and a serial connection through which GPS and heading data is output at 2 second intervals (see Figure 1).

The prototype utilizes predefined routes along which users are guided. A route is defined by a sequence of audio beacons. From the user's perspective each beacon is a sound source from which music emanates, and the music that they hear is modified to provide direction and distance cues to indicate the location of the next beacon in a sequence.

Directional cues are presented by panning the music between a pair of stereo headphones. When the user is directly facing a beacon the music is centered within the stereo sound field.

As the user rotates, the music is panned within the stereo sound field to indicate where the current beacon lies in relation to the user's new orientation. To do this, we compute the angle between the user's heading and the target heading (the direction towards the beacon). Spatial adaptation thresholds were set every 30 degrees. That is, if the user was off-track by between 15-44 degrees, the audio would be panned 30 degrees in the opposite direction; 45-74, by 60 degrees and so on. This approach was used instead of one that panned the music in a fine-grained way, both to take account of the jitter present in the compass reading due to gait, and given our previous lab experience

that showed that users could cope better with such macro rather than micro guidance.

Previous studies have encoded distance to the next beacon in the volume of the audio, however we chose to rely on encoding our directional guide with spatial audio. Warren's [14] paper and prior research with our laboratory prototype has gained user feedback that distance to the next beacon was difficult to notice and interpret

GPS, of course, has inherent inaccuracies. Even with a clear line of sight to GPS satellites, our GPS device would provide accuracy ranges of 5–15 metres the majority of the time. We attempted to accommodate this in two ways. First, the GPS positions of beacons were defined by taking multiple readings and finding the average. Second, each beacon was defined as having a radius of 5-15 metres. As soon as the user walks into the boundary of a beacon, their arrival is recognized by ONTRACK, and the audio music cues are modified to indicate the next beacon on the route.

#### **METHOD**

To evaluate the usefulness of the approach in guiding people around routes, we carried out an outdoor, mobile trial. Ten participants were recruited. All were university students between the ages of 19 and 45. Although all had experience with mobile phones, and a few with handheld computers, none had previously used a personal, mobile navigation system. Each participant took part in an evaluation session that lasted approximately 40-50 minutes and which consisted of three parts: a training session; three test tasks; and, a post-use interview and questionnaire.

During the training session, the purpose and operation of the system was described. Then, the participant used ONTRACK to follow a short training route that was designed to demonstrate the audio cues, the update delay present in the system, and the feedback given when a route is completed.

Participants then performed three evaluation tasks. Each task involved walking from a start location to a target location along a route consisting of predefined beacons. The only directional information provided was the continuously adapted music. The music used was a set playlist of three songs for each route, playing in the same order for each participant, looping if necessary. All songs in the playlist were reduced to mono quality, to eliminate the possibility of their stereo mix affecting the study

The routes were all of a similar length and complexity as illustrated in Figures 2.1–2.3. The start and end points are marked with "S" and "E", respectively. The intermediate audio beacons are shown as circles. The participant pressed a button on the handheld computer when they started the task.

Participants were encouraged to walk at a regular pace. The task was completed either when the target had been reached or the participant decided to terminate the task. The former was indicated by a completion sound being played and the

music halted, the latter by the participant notifying the observer. As each task progressed, the observer followed the participant, noting their reactions and demeanor. In addition, the GPS and heading readings were logged automatically. At the end of task, the time to complete the route (or to quit) was recorded.

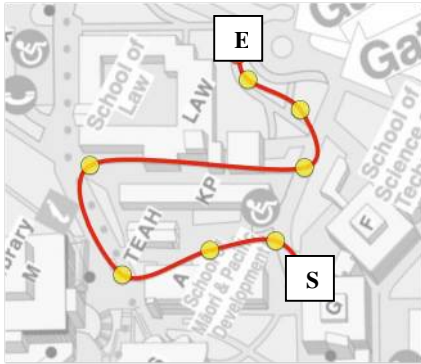


Figure 2.1. Route A. “S” and “E” are the start and end points respectively.

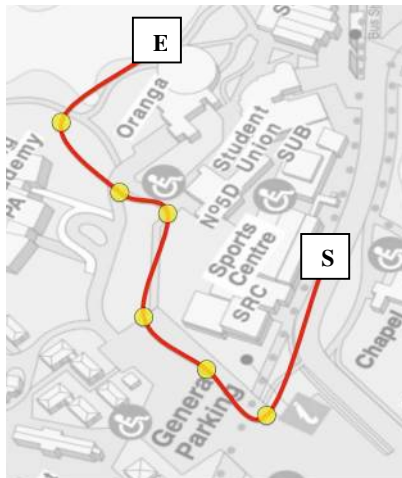


Figure 2.2. Route B

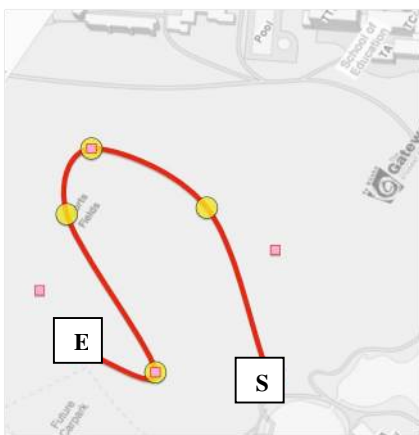


Figure 2.3 Route C. Small squares are goalposts on sports field

After the three tasks had been completed, participants were asked a series of qualitative questions relating to the usefulness and usability of the system. Finally, they completed a questionnaire based on the NASA Task Load Index (TLX) [3]. This questionnaire required participants to rate a series of statements on a ten point Lickert scale.

**RESULTS**

Table 1 shows how successfully participants were able to complete the navigation tasks using ONTRACK. Of the twenty-eight tasks, 86% were completed. Two tasks were not logged due to software failure.

The table also shows the mean time to complete each route. To put these figures in perspective, we asked an ‘expert’ who knew the exact route of each path to complete the routes as a control. They did so in 257s, 239s and 302s, for routes A, B and C, respectively. The individual paths taken by each participant over the three routes are shown in Figure 3.

Table 2 shows the mean rating scores for the NASA TLX questionnaire. A higher number indicates a more positive user experience.

**DISCUSSION**

In most cases, participants could follow a route from the starting point to a target location, via the predefined sequence of beacons. This is encouraging, given that the only knowledge they had of the route was provided by the adapted music. The average route completion times are in the order of 75% slower than the time taken by the control ‘expert’.

Route	Completion Rate	Mean time to complete in seconds (std. dev)
A	8/9 (89%)	443.50 (279.57)
B	9/10 (90%)	418.78 (173.32)
C	7/9 (78%)	534.00 (210.60)
All	24/28 (86%)	

Table 1. Route completion rates and times.

Factor	Mean rating (std.dev)
Mental load	6.85 (1.25)
Success	6.40 (2.31)
Satisfaction	6.30 (1.87)
Confidence	5.55 (2.01)
Enjoyability	7.20 (2.20)

Table 2. NASA TLX questionnaire results



#### Session 4: Why wear?

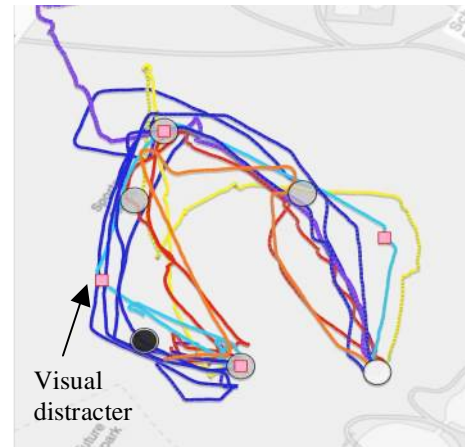
The large deviations in time to complete may be explained partly by the range of walking paces in our participants; however from the observations we made by following the users during the trials, we do not feel this was a major factor. Rather, firstly we note the impact of near outliers – for each route, one to two participants took nearly two standard deviations longer than the mean time to complete.



**Figure 3.1 – Route A Participant Paths. Each line represents a participant.**



**Figure 3.2 – Route B Participant Paths. Note outlier meandering.**



**Figure 3.3 – Route C Participant Paths. Some users are led off path by a visual distracter.**

These outlier routes were very much more off-track than others (see Figure 3.2, for instance).

Secondly there were two quite different strategies when using the system that impacted on completion time. A few participants developed an efficient technique of moving the handheld device around in front of them while walking, in order to ‘probe’ the sound source and to reassure themselves they were heading in the right direction. This approach allowed for faster completion times, and increased confidence.

More commonly, though, when people became unsure of where to head next, they would stop walking, re-orientate themselves, listen to how the music had changed and if necessary, repeat this process until their confidence had increased.

Several participants noted that they couldn’t navigate during quiet parts or fade ins/outs of some songs. One participant commented that as he was unfamiliar with the music in the playlist, there would be a period of approximately 30 seconds to become acquainted with that particular track to have a comfortable understanding of the changes made to it.

While the only cues we gave people were audio ones, the behaviour we observed during the test, and comments given in the post-study discussions, served as a good reminder that people are ecological rather than technological. That is, participants made some use of the physical environment—the apparent routes available, landmarks etc—to help them make decisions.

Many times this additional information was useful. For example, at points with multiple turnings they would often slow down, as they didn’t want to miss a turn. However in Route C, which was on a sports field with very few landmarks, we saw that people would be falsely drawn towards the visible landmarks. The route uses two of the

goal posts on the field of the four immediately available. The four goal posts are marked in Figure 1.3 as squares. After the top-most beacon was reached, we can see that some users overstepped the beacon between the two posts and headed for the post to their right, rather than moving downwards to the fourth beacon. In this case a misleading visual landmark – the non-route goalpost – was a stronger attractor than the audio cues.

Overall the system scored well in terms of the participants' subjective satisfaction, as measured by the task-load questionnaire. However they gave the system a lower rating with respect to the confidence they had in it to support their task. In the post-study interviews, an important factor that reduced their confidence was the lower than required responsiveness given by the implementation. Because the GPS/compass apparatus we used only provided updates every 2 seconds, the music adaptations sometimes failed to cope when the participant rapidly changed position and heading.

#### RELATED WORK

Since our initial lab proof of concept, two independent implementations of similar approaches have been reported. The Melodious Walkabout system [2] implemented a non-speech audio navigation system with the inclusion of a low-pass filter to provide spatial feedback when a target was behind the user.

A study of 24 users was undertaken, investigating the system's ability. The users were asked to provide one track of their choice that they enjoyed and were familiar with for one of the routes, and the researcher provided music for the other. Two routes were used, with the order of the routes and the assignment of music randomly assigned in each case.

The study showed that users adapted to the system with experience, completing the second route in an average of 27% faster, and covering 25% less ground. The study also found that the selection of music did not have a significant influence on completion time, though user feedback suggested that the cues were significantly clearer and satisfaction with the navigation was greater when using their own music.

The gpsTunes system [12] employed a handheld computer complemented with a 'backpack' containing many devices – gps, accelerometers, gyroscopes, fluxometer (electronic compass) and a vibrotactile device. Only the gps and compass were used directly, with readings from the other devices stored and later used for analysis.

A test of 4 users over 2 routes was undertaken on a flat sports field, finding that the system was able to successfully guide the users along the defined paths, but the routes were seen as simple, so complex routes remain unevaluated.

#### FUTURE WORK

There are some obvious issues to be addressed in the next phase of our research. In this study we converted the music

tracks to mono, so that any stereo panning embedded in the music could not be interpreted as a directional cue. Next, we will retain the stereo characteristics of the music and determine the extent to which users can discriminate between ONTRACK cues and similar effects within the music tracks themselves.

While our participants found the experience of using the prototype enjoyable, we have not studied the impact of the approach on the enjoyment of listening to the music itself. It will be interesting to consider such impacts and how to reduce any negative effects in trials involving the use of the system over an extended period of time, with participants' own music collections.

Additional extensions to the work include experiments to compare the efficacy of the system with a range of other cues – such as spoken and discrete audio ones – in a controlled way (a pilot of such a study is outlined in [5]). Further, we are continuing to investigate the performance impacts of more sophisticated models of the user's location and movement as well as additional audio adaptations.

While we have focused on navigation, here, we are also interested in the use of music cues to give the user an awareness of resources in their environment. This notion of providing extra information when the user is near a point of interest is prompted by the NAVICAM [9] and Audio Augmented Reality [1] projects.

This is an area that ONTRACK could expand into by providing feedback on nearby attractions based on the user's current location and walking direction. It could be implemented with a publish/subscribe system using an active database containing a profile of the user's preferences. When the system detects the user is near a point of interest based on GPS coordinates and the user's profile, rather than giving spoken information, the audio could be altered to draw the user's attention. We are currently developing and investigating an early prototype of this idea, where the audio is shifted near a point of interest. We hope to find out if it is possible to draw users' attention to specific items in their surroundings. Additionally, we are performing tests where the audio shifts at random. By finding out what users *think* we are drawing their attention to, we aim to gain insight into what type of things it is possible to notify users of, and what they would want to be notified of.

#### CONCLUSIONS

We have produced a low-cost, low-fidelity prototype to evaluate the ONTRACK concept in the field, following on from our laboratory-based studies. The results are encouraging. Even with low-resolution apparatus and minimal prior experience, most users successfully followed predefined routes to reach a target location. Although the targets were reached in far from optimal times, we expect this aspect of performance to be ameliorated though more experience with the system.

#### Session 4: Why wear?

Our current apparatus is low-cost and widely available yet suffers from accuracy limitations, which in some cases caused problems for users. We will extend our prototype by employing an inertial cube unit that will provide continuous and accurate orientation data.

This study has strengthened our belief that music-based cues can support user navigation whilst demanding minimal attention and integrating seamlessly with the common behaviour of listening to music whilst on the move. More widely, our experiences are opening up interesting research into the use of music as a mediator to give listeners intelligence about the environments they journey through.

#### REFERENCES

1. Bederson, B.B. (1995). Audio augmented reality: a prototype automated tour guide. Conference companion on Human factors in computing systems, 210–211
2. Etter, R. (2005) Melodious Walkabout - Implicit Navigation with Contextualized Personal Audio Contents, Adjunct Proceedings of the Third International Conference on Pervasive Computing, ISBN 3-85403-191-2.
3. Hart, S.G., Staveland, L.E. (1998). 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 Holland, Amsterdam, 139-183.
4. Holland, S., Morse, D., Gedenryd, H. (2002). AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal and Ubiquitous Computing*, 6(4), 253-259.
5. Hunt, R., Apperley, M., Cunningham, S-J, Jones, M. & Rogers, B. (2006). Minimal Attention Navigation via Adapted Music. Proceedings of the 3rd international workshop on Mobile Music Technology (*to appear*)
6. Kray, C., Elting, C., Laakso, K & Coors, V. (2003). Presenting route instructions on mobile devices. *Proceedings of the 8th international conference on Intelligent user interfaces*, 117–124.
7. May, A.J., Ross, T. Bayer, S. H. & Tarkiainen, M. (2003). Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing*, 7(6): 331–338.
8. Nemirovsky, P. (1999). *Aesthetic forms of expression as information delivery units*. Masters thesis, Massachusetts Institute of Technology, accessible via <http://web.media.mit.edu/~pauln/research/guideshoes/thesis/Thesis.pdf>.
9. Rekimoto, J. and Nagao, K. (1995). The world through the computer: computer augmented interaction with real world environments. Proceedings of the 8th annual ACM symposium on User interface and software technology, 29–36.
10. Sawhney, N. and Murphy, A. (1996). ESPACE 2: an experimental hyperaudio environment. *Conference companion on Human factors in computing systems: common ground*, 105–106.
11. Shoval, S., Borenstein, J., and Koren, Y. (1998). Auditory guidance with the NavBelt—a computerized travel aid for the blind. *IEEE Transactions on Systems, Man, and Cybernetics*, 28(3): 459–467.
12. Strachan, S., Eslambolchilar, P. & Murray-Smith, R. (2005). gpsTunes – controlling navigation via audio feedback. Proceedings of Mobile HCI 2005, 275-278. ACM Press.
13. Tran, T.V., Letowski, T., and Abouchacra, K.S. (2000). Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics*, 43(6): 807–827.
14. Warren, N. Jones, M., Jones, S., Bainbridge, D. (2005). Navigation via continuously adapted music. Extended abstracts, conference on Human factors in computing systems (CHI '05), (Portland, Oregon, USA, April 3-7, 2005), 1849-1852. ACM Press.