
Paper:

Robinson, S., Eslambolchilar, P. & Jones, M. (2010). Exploring casual point-and-tilt interactions for mobile geo-blogging. *Personal and Ubiquitous Computing, 14*(4), 363

http://dx.doi.org/10.1007/s00779-009-0236-5

Exploring Casual Point-and-Tilt Interactions for Mobile Geo-Blogging

Simon Robinson, Parisa Eslambolchilar, Matt Jones

25th March 2009

Future Interaction Technology Lab, Computer Science Department,
Swansea University, Swansea, SA2 8PP, UK
{s.n.w.robinson, p.eslambolchilar, matt.jones}@swansea.ac.uk

Abstract

People record and share their experiences through text, audio and video. Increasingly they do this blogging from mobile devices. We illustrate a novel, mobile, low interaction cost approach to supporting the creation of a rich record of journeys made and places encountered: by pointing and tilting a mobile, users indicate their interests in a location. We built three mobile prototypes to explore the approach – one combines gestures and visual map feedback; the second provides a simpler visual interface; the third supports eyes-free interaction, allowing the user to simply point-and-tilt, with no visual display required. We describe two field studies undertaken to understand the value of the interaction styles afforded, then continue with a further user study to assess the interaction speed and accuracy between these interaction methods. We present the results of these studies and raise issues relevant to their design and to the wider class of devices and services concerned with mobile spatial information access.

1 Introduction

Increasingly, people are documenting their lives – reporting on their emotions, thoughts, plans and actions – to remember, make sense of or share their experiences [10]. While mobile, this can be an arduous task, involving fiddly interaction with an interface that consumes most of the user's attention. Our motivation is to improve and simplify this interaction, allowing users to casually mark places of interest while mobile. In our approach, sensor data is used both to collect and provide content, which is later integrated into an interactive map visualisation to help the user re-trace their journeys and discover related geo-tagged content around each interesting location.

Consider this example interaction, supported by our systems, to indicate location interest by pointing a mobile device:

> *Sam is in Singapore. Just across the road he notices some colourful, old houses, an interesting contrast to the shining newness of everything else around him. He takes his mobile phone out of his pocket and points at the area; he holds the phone almost vertically as the houses are so close by. Later he's downtown. Across the river he sees a statue – a cross between a lion and a mermaid. Bringing his phone in front of him, he points, tilting it nearly horizontally as the statue is far away. When Sam returns to his hotel room, he enjoys re-tracing his journey and viewing the photos and web links associated with Arab Street and the Merlion on the automatically generated map.*

This map generating, journey reminiscing scenario is the one we focus on in this article. In this case, an interaction performed at an earlier point in time is used to provide information later. While mobile blogging is often characterised as an activity where content creation and posting occur in short order, often at the scene of the experience, our aim is to support a more reflective combination of a user's mobile and, later, non-mobile interactions. The very large number of users already creating personalised content using web map services is a strong motivation to consider additional ways of enabling creativity.

Previous authors have investigated the use of point-to-select location selection using complex geo-spatial models to predict the user's desired target. In this work, however, we remove the need for a location model by combining pointing with tilting to allow both direction and distance to be specified in a single interaction. This lightweight gesture helps to support a more casual and interactive experience whilst still allowing users to pinpoint locations for later information perusal.

We built two experimental sets of apparatus to explore options for supporting this interaction, and tested them in separate field studies. The first prototype, described in section 3, combines pointing and tilting with visual feedback to help mark areas of interest. In the second, detailed in section 4, only pointing and tilting is required, removing the need for visual display interaction, and allowing a heads-up approach to place marking.

These two systems were first reported in [13]. Here, however, we extend their discussion and compare their performance with a third prototype inspired by the results of the earlier field studies. Our third prototype aims to address issues with our earlier designs by presenting a minimal visual display to attempt to improve the precision of the tilting interaction. To help understand the systems further, we carried out a lab-study that considered the speed and accuracy of location marking over each of the three systems. We present this experiment and its results in section 5, then discuss our findings and conclusions, and suggest pointers to future work to complete the article.

2 Background

In this section, first we explore related literature concerning a need for tools for mobile geoblogging. Other researchers have focused on pointing to select locations: we describe selections of the related work in this area, then conclude by reviewing related delayed search systems.

2.1 Mobile blogging

Espinoza *et al.* [4] created GeoNotes, an early location annotation system that is very similar to the location blogging tools available today. They list several problems with location-based place annotation, and highlighted that creating and placing a virtual placemark needs to be simple and fast. Later, Bamford *et al.* [1] describe LocoBlog, a system for creating travel blogs while on the move. Users are able to take photos with a mobile phone then submit these, coupled with a GPS location fix, to a central server where their photos and associated text entries are added to a public journey blog. They found that users appreciated the ability to share their location and surroundings with family and friends in other places, and that, even when short of time, users often quickly submitted an entry so that they could share their experiences.

Beale [2] comments on the need for non-intrusive tools for mobile blogging, and designed SmartBlog, a blog authoring tool for users on the move. Users are able to share photos and associated comments on a public blog from a simple mobile phone interface, but no location information is included. Publicly available applications have also been produced (e.g. Nokia LifeBlog[1]), but these focus on annotating and uploading user-created content rather than automatic collection of related geo-tagged information.

---

[1]

1. http://www.nokia.com/lifeblog

2.2 Pointing to select

Gesture-based control of mobile device and service functionality has received much attention, with several authors combining gestures and other sensor data to create smart spatial appliances. Before the necessary technologies were integrated into mobile devices, Egenhofer [3] proposed several *Spatial Information Appliance*s. These included a *Smart Compass*, providing turn-based GPS guidance toward a location, and *Geo-Wands*, that help users to identify geographic targets by pointing toward them. Increasingly, prototypes combining gestures and sensor data to create devices such as these are being made. Wasinger *et al.* [19] created a pointing-based location interaction system, combining GPS and compass data with speech recognition to allow a user to say a query (e.g. 'what is that?') whilst facing a location. Their system processes this data, recognising the information request, but does not yet present the user with the requested information.

Several authors discuss implementations of 'point-to-select' methods of interaction to retrieve information. This interaction technique allows a user to simply point at an object to indicate an interest in it. These implementations, though, differ to our approach on two counts: firstly, they use more complex geo-spatial models to interpret gestures and retrieve content; and, their focus is real-time retrieval *in situ*. In contrast, in our approach, the mobile interaction is far more casual and speculative in intention and is integrated with a later interactive experience.

Peter Fröhlich *et al.* [5] conducted a Wizard-of-Oz style user study to assess the viability of point-to-select against several other methods of interaction, concluding that pointing gestures were 'highly attractive and efficient' forms of location selection. Building upon this work, Simon *et al.* [16] describe the spatially-aware mobile phone, a conceptual device to connect the physical and digital worlds. Their framework uses a three-dimensional model of a location in conjunction with knowledge of a user's position in order to create a line-of-sight visualisation from the user's position. Continuing this concept, Simon *et al.* [17] create a point-to-discover application using this framework. Their application prototype uses location and heading information to, at the push of a button, calculate the visible points from the user's location and display relevant information about them. A further paper by Simon and Fröhlich [15] discusses a similar concept that presents the user with Wikipedia articles about locations near to them based upon their location and the direction they are facing.

These systems use gestural point and tilt data from a mobile device to determine a line of sight from the user's current position. Our designs, however, use pointing (heading) data to determine the direction the user is facing, and tilting to allow the user to specify the distance of an information point. Similarly, Strachan *et al.* [18] use location and heading data in conjunction with real-time trajectory prediction to guide a user along a path to a desired target location. By pointing and tilting a device around the environment, the user can browse the features around them, with both audio and haptic feedback directing them toward the specified target. Their system intentionally presents the uncertainty in the system to the user, and allows them to probe possible future routes in the available space, sensing the feedback from routes up to 20m ahead of their current location.

Each of these approaches demonstrates active, focused mobile spatial interaction, with the user conjuring up data by actively pointing the device and pulling in content. In contrast, the RelateGateways project [6] uses less complex spatial contextual information to push directional information about pervasive services available to the user, including the heading and distance of these objects. This information is presented on the screen of a mobile device, and users can select each of these objects by tapping on the screen, which expands the information about the selected object. Pering *et al.* [11] describe a similar system in which users can connect to and control physical objects using electronic tags and simple gestures.

Rukzio *et al.* [14] studied three techniques (touching, pointing and scanning) for locating smart objects, finding touching and pointing to be the preferred interaction techniques if the user had a line of sight to or was close to the target device. Pointing was seen as a quick technique that required some cognitive effort but a low amount of physical effort, especially when objects were not within touching distance. Results from their study also showed both pointing and touching to be intuitive techniques, particularly among older participants who wanted to be able to avoid mobile device input as much as possible.

Each of these papers demonstrates that point-to-select is a viable method of interaction, and can provide users with valuable location-specific information. However, they all require virtual location models in order to be able to pinpoint the user's targets. Whilst these authors provide valuable insights into possible methods and uses of location-based interaction, the aim of our project is to provide users with similar data but without the need for complex location models and visibility calculations. We imagine situations where such models are unlikely to be created: landscapes, seascapes and very rural areas, for instance, are unlikely to be mapped in great detail in the near future. In addition, our work could allow users to mark objects that are evolving and may never be modelled – live events, shows or funfairs for example – by allowing them to mark any item in the space around their position.

2.3 Delayed interaction

Our work is not focused on the real-time delivery of location information but rather on providing later access to content related to places the user has visited. This delayed approach to information retrieval resembles the concept of 'slow technology' (for example, in [7]). Jones *et al.* [9] present another approach to delayed search. Textual notes jotted onto a handheld device are later used to provide packaged web information via a search engine. More recently, major commercial search engines have provided means for users to read and reflect on their own search histories[2].

Rekimoto *et al.* [12] created LifeTag, a location recording tool using WiFi to find and store a person's position as they travel around a city. As the user moves, their location is tracked automatically, but it is also possible for them to manually 'bookmark' a notable location if desired. Later, when analysing the resulting data, the authors are able to create sculpted views of the visited areas, highlighting in more detail those places where lots of visits occur and shrinking those which are less popular. In addition, their system is able to detect location events that are out of the ordinary, and the authors suggest that these could be used as memory cues or as a basis for blog entries, a similar application to that of our prototypes. Currently developed applications for their system include an automatic photo geotagging tool, and playback of a user's journey augmented with related geo-tagged pictures, sourced from Flickr using the locations recorded along the route.

3 Prototype 1: Gestures and Visual Feedback

The first prototype segregates the interaction experience into two phases. First, when mobile, the user can mark any number of points of interest with a simple point-and-tilt gesture (as illustrated in Figures 1 and 2); then, when they return to their computer, a map is generated showing the routes taken and information about the areas selected (see Figure 3). In this section, we outline how these facilities are provided.

3.1 Mobile hardware

We use SHAKE (Sensing Hardware Accessory for Kinesthetic Expression, see [20]) sensor packs for real-time recording of tilt and heading (compass) data. The SHAKE SK6 is a small wireless device incorporating three-axis accelerometers, magnetometers and angular rate sensors, dual-channel capacitive input sensors

---

[2] http://www.google.com/psearch

and a navigation switch. Signals from these sensors are communicated wirelessly using Bluetooth serial port emulation. The SK6 also includes a programmable vibrating motor that can be used to simulate a range of vibrotactile effects and provide feedback to users.

We use a standard Bluetooth GPS receiver to determine a user's location, and a Dell Axim x51v PDA to record the data. Each SHAKE is attached to the back of a PDA, so that any movements made by the user whilst holding the device are recorded by the SHAKE (see Figure 1). The equipment is designed to be as simple as possible to use, allowing the users to just leave the devices powered on at all times, and then mark places whenever they like.

3.2 Marking points of interest

<< Figure 1 about here >>

<< Figure 2 about here >>

The system displays an aerial photo of the user's current location, overlaid with an arrow showing the direction they are currently facing, pinpointing the location that can be marked. To mark areas of interest, users point at the area of interest with the PDA in their hand. They then tilt the device toward or away from their bodies to refine the targeting. When the marker is positioned over their desired target location, the user presses a button to mark that place (Figure 2).

Our motivation for using tilting and button presses instead of simply allowing a user to tap the touch screen is two fold. Firstly, our interest is in understanding techniques that will be relevant to a range of future devices, many of which will not have full screen touch input. Secondly, we wished to avoid the user being too focused on the digital map at the expense of the physical world around them.

As a default, the user is presented with a view that shows an area up to 175m from their current positions. They are, though, able to zoom out through three levels from this view to allow them to select targets up to 350m, 700m and 1400m from their location. The lowest zoom levels let the user select close places with high precision; while the highest zoom levels allow selection of distant places with less accuracy (see Figure 2).

To select places furthest from the user, the device is held horizontally; for places close to the user, though, the device is tilted back towards their body. That is, when held horizontally (at 0 degrees relative to the horizontal plane), the distance recorded is the maximum range of the current zoom level; when it is held vertically (at 90 degrees relative to the horizontal plane) the distance is set to 0 metres. Between these two extremes the distance is a continuous function of the degree of tilt.

There is no requirement for the user to have a line of sight to objects of interest as in some mobile location aware proposals (e.g. [17]). The user can browse the map and place markers at any location they are interested in from the aerial view alone.

All of the map visual data is pre-loaded on the PDA – no online access is necessary. The system can be used in any location but if visual map data is not available, the display shows only the arrow overlay.

3.3 Making maps

When the user has completed their journey and docked the PDA with the computer, we generate an interactive map of their route, presenting them with information about each of their points of interest. The first stage of this process is to determine the user's route from logged GPS readings, and to retrieve the latitude and longitude coordinates of each of the user's points of interest. When this process is complete, at each marked position an area of interest is defined.

3.3.1 Finding content

The coordinates of each area of interest generated in the first phase are used to retrieve postal codes (similar to zip codes but at a much finer level of granularity) and district names from a purpose-built gazetteer of the area. Each of these descriptors specifies a small area at street level, allowing for detailed location-specific results. We select location descriptors from a 100m^2 area around each point of interest in order to try to ensure that at least one postal code is included for the majority of the locations searched. These location surrogates are then used as query terms in a series of searches, ranking closer objects higher. We use Google for text and image results and further images are garnered from social photo publishing sites (Flickr, Panoramio, Geograph).

3.3.2 Map visualisation

<< Figure 3 about here >>

The map visualisation shows the route taken during each journey – as coloured lines, a different colour for each route – and the areas-of-interest marked by the user as overlays. All of the standard Google Maps functions – such as different views and panning and zooming – are available.

Markers for each area-of-interest show the number of images retrieved for that location. This meta-datum was chosen over others – such as the total number of all items retrieved or the sequence count of the marker along the route – after an earlier focus-group study suggested the higher value of image items. Locations with no retrieved content are shown as grey markers.

When the user browses the map they can hover over markers. When they do this, two additional pieces of content are shown to the user: the number of search results and the number of images; and, where the location was marked during the journey – a line is drawn from the marker to the point on the user's journey. When the user clicks on a marker, the search result data is combined into a simple information pop-up that is overlaid on the map. Items in this pop-up can be selected to see the full document or image context. Figure 3 illustrates the possible interactions and information retrieved.

3.4 Evaluating the System

A field study was performed to help understand aspects of the use and general value of the approach and the efficacy of specific interaction methods including the visual point-tilt method and the post-use interactive map visualisation.

3.4.1 Method

Fifteen participants aged from 18 to 45 were recruited for a multi-day study. Five participants were University staff members, ten were students; six participants were male and nine female. Three participants were from disciplines connected to computer science, the remainder worked in unrelated areas. Six participants had previously used accelerometers with the Nintendo Wii games console and two used PDAs regularly; the remainder had no prior experience of this method of interaction.

At the start of the study each participant was met individually and introduced to the equipment and its purpose, followed by a short demonstration of its use. As a form of training, participants were then asked to use the system from the lab to mark a number of points from a window in our laboratory.

Participants used the system for a 4-day period. Seven used it during Tuesday-Friday; eight others Friday-Monday. They were asked to leave the devices switched on at all times during any journeys they made during that period, and then use the system to mark any places they were interested in at any time.

At the end of the study period the data logs were collected from each participant's PDA. These were then analysed to identify the marked places for each journey, and a personalised map of their routes was produced.

Before viewing their map, participants completed a questionnaire based on the NASA Task Load instrument [8]. This questioned their perception of the costs in using the mobile element of the system. They were asked to rate the mental, physical and temporal demand, their success in performing the marking task, overall effort needed and frustration with the system. Each of these dimensions was rated on a scale of 7 (positive, e.g. low mental demand, high performance) to 1 (negative, e.g. high frustration, low performance). In addition, each participant was asked to rate on the same scale of 1-7, specific aspects of the mobile prototype's use and usability. The features rated were: the overall ease-of-use of the system; the identification of a location from the visual display; the overall approach to marking locations; the support for accurate marking; and, the time taken to mark locations.

Participants were also asked to give feedback about any problems they had experienced with the system, and discuss any notable observations that had come to light.

Participants were then asked to browse their map, exploring their routes and the search results of each marked location in turn, whilst thinking-aloud to explain their interaction with and impressions of the map. Participants also rated the content retrieved for each marked location on a scale of 1-7; where 1 indicated the content was not at all relevant and 7 that it was very relevant.

Each participant was rewarded with a bookstore gift voucher at the end of the study.

3.4.2 Findings

<< Figure 4 about here >>

We discuss our findings below in relation to the automatically logged data during the mobile system use; and, the subjective rating, interview and think-aloud feedback provided by participants.

Journeys and marked points of interest All participants successfully used the system for at least one journey over the course of the study. In total 57 separate journeys were recorded, with a maximum of nine per participant. The mean journey time was two hours, with four locations marked per journey on average, and 241 marked in total over all participants.

The journeys recorded represented a diverse set of routes spread over a large area of the region. Twelve participants explored an area covering approximately 14 square miles, enclosing the main populous areas of the region, but three participants ventured further afield, up to 56 miles from the starting point. The routes produced covered a variety of places such as town and city centres, parks, residential areas and rural countryside. Figure 4 shows all of the routes and locations marked within the main common journey area.

The majority of the places marked were public buildings, such as leisure centres, shops and museums; others were ones of historical interest (e.g. castles, old houses) and participants also used the system to point at landscape features, such as beaches and headlands, rather than man-made areas. 63 (26%) of the locations marked were landscape features, 45 (18%) were points of particular historical interest, and the remainder (56%) were public buildings. Approximately 160 of the 241 places marked were unique – that is, in 66% of cases, a location was marked by only one participant.

The aerial view zooming functions were not often used; indeed, six participants did not zoom at all, leaving the display in its default view, and the remainder made use of the zoom only four times on average. This behaviour limited the maximum range of participants' location marking to 175m, the maximum range at the default zoom level.

The average distance from a participant's position to a marked location was 230m (std. dev.: 319m). 85% of the marked locations over all participants were between 0 and 175m from the participant's location; 36 of the 241 marked locations were more than this distance away. The mean time taken to mark each location was $4\frac{1}{2}$ seconds (std. dev.: 6.7 seconds). Three participants spent a large amount of time making their gesture very specific. For example, one participant took 28 seconds to mark an individual room in a building, rather than just the building as a whole. Figure 5 shows the distribution of times taken to mark locations.

<< Figure 5 about here >>

Mobile Task Load and Usability Ratings Figure 6 (top) shows the average rating of the task load dimensions; the overall assessment of the system's demand and impact was positive although not overwhelmingly so; in particular, their assessment of confidence in performing marking tasks was low.

Figure 6 (bottom) presents the average ratings of aspects of the mobile system's usability. The overall 'ease-of-use' of the system was rated positively but again participants found elements of the mobile use less than optimal.

<< Figure 6 about here >>

Interview and think-aloud findings relating to the mobile system Most participants found pointing while looking at a location to be a natural way to request information. All except one participant were able to identify places on the aerial photo of their location with little effort. However, many found that accurate marking of each location was too time consuming and commented on the mental impact of having to match the map view with their actual view.

Participants were positive about the facility to 'see beyond' their field-of-view; that their line-of-sight did not restrict their ability to browse places they were situated within.

Two participants travelled to an area that was not covered by our map images, so were unable to see the actual aerial photo; instead they could see only the arrow showing their target direction and, more importantly, an imprecise indication of distance. One of these participants commented that they had found marking locations easier when the map was *not* visible; with the map they had felt they always had to mark the exact position of the target, but without the map they were able to be more imprecise yet still retrieve relevant results about their target location.

Another participant noted that if the distance in metres they were marking was overlaid on the map then they would quickly get used to the tilt required, and would not need to refer to the map display all the time.

Some participants wanted to precisely mark locations and indicated frustration when the system or circumstances made this difficult: for example, one had trouble keeping up with the map updates when using the system from a moving bus. Others demonstrated a less exact use of the system: for example, one participant reported often holding the device flat, pointing and marking, as they wanted to find out what was generally in the area and did not have specific targets in mind.

Interview, think-aloud and content rating findings relating to the map visualisations Think-aloud sessions with each participant using their personalised map generated largely positive comments, with all participants noting the potential applications and benefits of the system. All participants commented on the use of the device for tourism purposes around a new location, and two remarked that seeing their route would help them find their way back to new or interesting places they had visited during the day.

Several participants found information about places they had marked on the mobile map but had not actually been able to see from their location, and two were able to find contact details for local businesses they had marked for later use.

It was noticeable that most participants tended to skim over the textual information about the marked places, instead preferring to skip directly to pictures of the location. When participants did concentrate on textual information, it was only briefly glanced over, rather than analysed in detail.

Two participants highlighted privacy concerns from their usage of the system, and were worried about their location being tracked continuously, despite the option to turn off and disable GPS positioning. Conversely, several participants remarked that it was interesting to be able to see where they had been throughout the day, and that simply viewing this journey information (regardless of the markers) would remind them about things they would otherwise have forgotten.

Of the 241 locations marked, content was retrieved for 158. On average 6 images and 8 other forms of content were found for these places. For this content, the overall mean relevance rating for markers was 5.4 on a scale of 1 to 7.

Several participants commented on the need for more control over the sorts of information presented. For example, two participants indicated they would have liked to be able to select categories of results before viewing the information retrieved. Similarly, two others commented that had the text results been split up into categories (for example, what's at a location and the events that occur there) then they would have found the results easier to sort and filter according to their interest. Several participants noted that they had meant to mark a public building, such as a museum or leisure centre, but had been given results crowded by "yellow pages" type results, such as house prices and restaurant directories from that location.

3.4.3 Discussion

It was encouraging that the participants were able to mark locations and find value in the resulting maps with very little training and exposure. Participants were familiar with the areas they used the system in; even so, they found unexpected, interesting information. The approach may have further benefits for new-comers to or tourists in a location.

The locations participants marked but that had no content associated with them could be considered as opportunities rather than disappointments. That is, they could act as spurs to further user-generated content: noting content 'barren' places people are interested in could prompt contributions from others.

Many of the routes overlapped but, despite this, the majority of areas of interest were unique. Many experiments in conventional information retrieval show that there is a Zipfian distribution in query terms – that is, there is a 'long tail' in user requests. Our data is not conclusive but might indicate a similar diversity of future physically initiated 'queries'.

If there is a sparsity of location specific content, or content that is geotagged with a low degree of precision, it could be argued that approaches such ours are overly involved. Rather, one could envisage a simple, single button push that indicates that a desire to know about any content within, say, a 300m radius of the current position. This would be akin to the ‘blog this’ button-push available via some web browsers to capture pages viewed, or the ‘bookmark this location’ method used in [12]. However, in built-up, highly-populated or visited areas it is likely that such a blunt tool will lead to users being overwhelmed with unhelpful content. As more content becomes geo-locatable to increasing degrees of precision, we would expect more refined pointing mechanisms to become increasingly important. Filtering mechanisms such as the ones suggested by our participants will also be needed to ensure that the content is not only relevant but also useful.

The use of visual feedback to help the user more accurately indicate their areas of interest appears to have had both positive and negative impacts on the user experience. Firstly, it is worth noting the value reported by users in just being able to see the satellite map view of their surroundings. People have long enjoyed browsing topographical maps with limited textual content, first on paper and more recently in services such as MSN Live and Google Earth. While it is possible to now add many more sophisticated location-based information features, we should, perhaps, be careful to maintain the elegance of the simpler, less interactive views.

Visual feedback appeared to allow participants to position points of interest at the level of precision they required. However, it is possible that level of visual detail provided increased the effort required. The demand levels reported in the TLX questionnaire along with the comments made by participants seem to provide evidence of this effect. Trying to match up the aerial map view with the physical surroundings is a potentially fiddly task.

One of the design objectives of the system was to provide fast capture of things of interest. The logged average time to mark a location was low (4.5s). It is interesting, though, to note that participants subjectively rated the amount of time to carry out the action as too high.

4 Prototype 2: Gesture Only

Given the potential added burden of the visual feedback and the possibility, as noted by participants, of users learning to mark distances without looking at the screen, it is worth considering a lighter-weight, gesture only approach. We created a prototype that illustrates such a minimal attention user interface method and deployed it in a field-study.

The map creating and visualisation elements were essentially the same as those described in Section 3. However the way of marking points of interest simply involved holding and tilting the SHAKE; study participants carried the PDA and GPS as in Section 3.1 during the studies but no visual feedback was provided.

4.1 Marking points of interest

<< Figure 7 about here >>

<< Figure 8 about here >>

Users mark areas of interest by pointing with the SHAKE held horizontally (flat) in their hand. They then tilt the device back toward their bodies to give an indication of the distance of the point-of-interest from their current location (see Figure 7). The gesture is like casting out a net and then drawing it back in to the correct position .The action can be completed very quickly (see Figure 8).

For places in the distance, the user tilts back only a small amount; for places close to the user they tilt back to a greater degree. When held horizontally (at 0 degrees relative to the horizontal plane), then, the distance recorded is approximately 1000 metres; when it is held vertically (at 90 degrees relative to the horizontal plane) the distance is set to 0. Between these two extremes the distance is a continuous function of the degree of tilt. This interaction allows approximately the same range of targeting as the zoom level interface used in prototype 1, but in a non-visual interface.

As in prototype 1, there is no requirement for the user to be able to *see* objects of interest. The user can then 'throw' their net over the visible horizon to see, for example, what lies behind the immediate cityscape.

The sensor readings recorded during a user's journey are analysed using a standard back-propagation neural net to identify pointing gestures. The recogniser looks for patterns where the sensor data first becomes relatively stable (e.g. the user has taken the device out of their pocket while walking and holds it flat to begin the gesture) and then shows characteristics contingent with the tilt-back action.

When a gesture is recognised, the compass reading (for orientation) and GPS (positional) readings at that point are combined. The tilt angle is then used to project this point into the distance. At this position, we define an area of interest 100 metres squared in size.

4.2 Exploring the approach

Seven participants were recruited for a 7-day field study. The study period included both weekdays and one weekend. Five participants were University staff members, two were students; three participants were male and four female; none were from computer science or related disciplines. One of the participants had previous experience with the use of accelerometers in an environmental health context; the remaining participants had no prior knowledge of this type of interaction. A similar study protocol to the one described in Section 3.4 was deployed. In addition, at the end of each day of the trial, participants circled places on a paper map that they thought they had gestured at earlier. None of these participants were involved in the visual-gesture field-study described earlier.

Each participant was met individually at the start of the study and introduced to the equipment, then given a short demonstration of its use. Once confident using the equipment, participants were asked to leave the devices powered on at all times during any journeys they made, and then make gestures at any places they were interested in.

Each participant was rewarded with a bookstore gift voucher at the end of the study.

4.2.1 Findings

We focus here on reporting the findings related to the marking by gesture process. Twenty-three separate journeys were identified and analysed with each journey lasting, on average, around two and a half hours. Most marking gestures were successfully recognised. Participants enjoyed the freedom of casually being able to target a location with nothing but a gesture; one participant remarked that using the system was like '*Googling the real world*'.

On average these locations were 127 metres from the participant's position. Locations far away from participants were marked less frequently; the furthest correctly identified area was some 500m away from the origin. All participants said they felt less confident in judging distances far away from themselves.

For all participants there were false-positives – that is gestures were recognised when none occurred. In most cases, these unintended markers were clustered around a place of interest – that is, they were identified in a period shortly before or after a true gesture.

4.2.2 Discussion

It was encouraging that the participants were able to mark locations with a simple gesture. The false-positives issue is a problem that can be mitigated by refinements to the gesture segmentation algorithm and by using training data from each user to calibrate the algorithm.

As in the visual-gesture system, many of the points marked were close to the participants. There are at least two possible explanations for this. As with the first study, over the limited trial time, the most obvious use of it was to mark things they saw directly around them. Secondly, there was no opportunity for users to learn the effect of different tilt actions as no feedback was provided during the mobile trial with results only available at the end of the period. With extended use participants would have the opportunity to calibrate the actions with results, possibly increasing the range of tilt gestures to reach a wider range of targets.

5 Exploring the effect of feedback on target marking

From these first two prototypes we saw that participants were able to use both visual and non-visual marking systems to target points in their environment. However, the extent to which the presence of visual feedback affected the marking ability and accuracy displayed by the participants was not clear. To address this, we conducted a further study to explore the impact of visual feedback on point-tilt accuracy in a more controlled manner. We modified our prototypes to standardise the interaction methods and created a third system with an interaction style that combines those of our earlier prototypes. As before, each system allows the user to mark interesting targets with a simple point-and-tilt gesture (as illustrated in Fig. 1). In the following section we outline how each system is used, and how its interaction differs from the others.

5.1 Marking Points of Interest

<< Figure 9 about here >>

In each of our three test systems, to mark target areas of interest users point at the location with the PDA in their hand. They then tilt the device toward or away from their bodies to refine the distance of their targeting gesture. When they feel they have indicated the approximate distance to their desired target location, the user presses a button to mark that place. Each system has a maximum range of 175m and does not allow zooming, a feature removed as a result of user feedback from our previous systems.

Our test systems all use the same point and tilt interaction, but each presents a different display to the user (see Fig. 9). System one, the *aerial view* uses the same interface as prototype 1 (see Section 3). System two, the *arrow view*, dispenses with the aerial photo display, and instead only shows the user an arrow, the length of which indicates the distance of the location they are targeting. Distance markers (in metres) are displayed alongside the arrow in 25m increments. Turning to face a new location does not adjust the displayed arrow;

instead the arrow always points toward the front of the PDA; i.e. the direction in which the user is facing. Our third prototype, the *non-visual* system, removes all indication of distance and heading, presenting the same display as our earlier prototype 2 (see Section 4).

5.2 Evaluating the systems

We performed a lab-based study to enable us to evaluate and compare each of our three design prototypes with regards to efficiency and accuracy in marking locations. We were particularly interested in the differences in time taken to mark locations between systems, and the targeting accuracy that is afforded by each interaction method.

5.2.1 Method

<< Figure 10 about here >>

Thirty-eight participants aged from 18 to 65 were recruited for a ten-minute study. 16 participants were University staff members, 22 were students; 15 participants were male and 23 female. Nine participants had used accelerometers with the Nintendo Wii games console; the remainder had no prior experience of this method of interaction.

At the start of the study each participant was met individually and introduced to the equipment and its purpose, followed by a short usage scenario example and a demonstration of its use. As a form of training, participants were then asked to familiarise themselves with the system by marking a number of points of their choosing from a window in our laboratory, until they felt they were comfortable using the system for the study.

After this initial training, participants were then asked to mark six preset points from a fixed location in our laboratory. These points were picked specifically to be at a range of distances between the minimum and maximum marking range of the systems, and also in several different viewing directions (see Fig. 10 for details). The points used were all clearly visible from the laboratory window and all participants stood at the same position in the lab during the study.

Thirteen participants were asked to use the aerial view for training and during the study. A further thirteen participants used the arrow view for training and the study.

Twelve participants were asked to use the arrow view for training and the non-visual system for the actual study. Our reasoning behind this method is that non-visual system gives no indication of the distance of the location being marked; had we asked these participants to use the non-visual system without prior knowledge of the distances they would be marking then the results from these participants would be of no use. As the method of marking locations is identical for each system – only the display is different – we did not expect any adverse effects on our results from this choice of training method.

For each participant we measured the time taken to mark each target. In addition, three accuracy-related measurements were recorded (see Fig. 11). These were:

- The absolute distance the participant's marker was from the intended target (e).
- The angular difference between the compass heading of the target and the direction the participant pointed toward (α).
- The distance of the participant's marker from the point they were standing in the lab (d).

<< Figure 11 about here >>

Measuring absolute distance gives us an indication of how close participants were able to mark in relation to each of the six targets. Measuring the angular difference and distance marked from the lab enables us to determine whether an inaccurately marked location is the result of a pointing error or a distance error. For example, a participant may have pointed in the right direction but underestimated the amount of tilt required to mark the location, giving us a distance error.

At the end of the study period the data logs were collected from each PDA and analysed to identify the marked points and the times taken to mark each target. The resulting marked positions were then compared with the exact target locations to identify the errors made by each participant when marking each of the targets.

Each participant was rewarded with a bookstore gift voucher at the end of the study.

5.3 Findings

<< Figure 12 about here >>

<< Figures 13 and 14 about here >>

All participants successfully used the system to mark at least one location during training, and again to mark each of the six specified points during the study.

These marked targets were analysed for each separate system to enable us to compare the marking accuracy and time taken to mark a location between each system, and between each of the targets.

Tables 13(a-d) show the mean and standard deviation of each measured value for each target on each system. Figure 12 shows a visual representation of the target areas marked by participants. There were no outlier participants.

When comparing the complete set of data from each system we can immediately see from the visual results that participants using the aerial view appear to have been more accurate in perceiving and selecting targets. However, the time they needed to mark these targets was longer. Participants using the arrow view and non-visual systems have taken less time to mark a point than those using the aerial view, but the absolute distance error and angle errors produced are higher. Further statistical analysis of each measurement shows several significant differences between the three systems. These findings are presented and discussed below.

5.3.1 Target absolute error distances

Fig. 14(a) illustrates the spread of absolute error distances recorded. Using GLM ANOVA it was found that the type of visual feedback used had a significant effect on the absolute distance error ($f = 25.66, p < 0.05$). There were, however, no significant differences between the six individual targets used.

Posthoc analysis using a Tukey test shows that there is a significant difference between the absolute distances of selected targets away from the actual targets when using the aerial view ($p < 0.042$), i.e: the aerial view allowed users to be more accurate in their marking. However, the distance errors were not significantly different between the arrow view and non-visual systems ($p > 0.3$): neither of these systems was more accurate than the other. The Tukey test also shows that when using the aerial view, participants were able to mark the closest targets (3 and 6) more accurately than the others, but there was no significant difference between any individual targets when using the arrow view or the non-visual system.

5.3.2 Angular errors

Fig. 14(b) shows the spread of angular errors for each target using each system. ANOVA between targets shows a significant effect on the angular error ($f = 15.45, p < 0.05$), indicating that some of the targets used were more difficult to point at than others. However, the direction of this error (i.e. to the left or the right of each target) was not significantly different between targets ($p > 0.3$).

The type of display used showed no significant effect on either the magnitude of the angular error or the direction of this error ($p > 0.2$): the angular errors recorded for each target were not related to the display type used. In addition, these angular errors were not significantly different between each display type ($p > 0.5$), indicating that errors in the participants' pointing accuracy were not caused by one particular display more than the others.

5.3.3 Distance marked from participant's position

Fig. 14(c) shows the range of marked distances from the participants' position in the lab. ANOVA indicates that the distance of the target from the participant's location has a significant effect on the absolute target error ($f = 19.04, p < 0.001$): targets further away from the participant were marked less accurately. In addition, the type of display used has a significant effect on the distance the participant marked from their position ($f = 7.82, p < 0.001$): participants using the arrow view and non-visual systems often marked locations further than was necessary.

A posthoc Tukey test shows that there is a significant difference between the distance of marked targets from the participant's position when using the aerial view system ($p < 0.02$): the distance marked when using the aerial view is generally lower than those of the arrow and non-visual displays.

5.3.4 Time taken to mark targets

Fig. 14(d) illustrates the times taken for each target using each view. ANOVA shows that the display type has a significant effect on the time taken to complete the task ($f = 27.62, p < 0.001$): the aerial view is slower than the arrow and non-visual displays. There were, however, no significant differences in times taken between the six individual targets ($f = 2.062, p < 0.07$), i.e: no one target was quicker to be marked than the others.

Posthoc analysis using a Tukey test shows that there is a significant difference between the times required to complete each marking task in the aerial and arrow view conditions ($p < 0.001$), with the participants using the aerial view taking significantly longer to mark each target. The test also shows that there is no significant difference between the times taken to mark each of the targets when no display is visible ($p > 0.35$). However, the Tukey test shows that there is a significant difference between the times required to complete the individual tasks using the aerial view for target 6 ($p < 0.02$). In addition, results from the arrow display showed a significant difference ($p < 0.016$) between the times taken to find targets 3, 4 and 6.

5.3.5 Additional feedback and observations

All participants using each system found pointing while looking at a location to be a natural way to request information. Participants using the non-visual system appreciated the low effort required to select a target, but several commented on the lack of accuracy in their marking. Two participants using this system stated that after using system two for their training they were confident selecting targets at various distance intervals, however three other participants noted that they were unable to recall the distances used in their training, so had to guess the tilt angles required during the study.

Participants using the arrow view were positive about the ease of marking locations, and several stated that it was easy to mark targets. However, three participants remarked on the difficulty in estimating distances accurately from their location, pointing out issues with height differences.

Participants using the aerial view were confident in their ability to identify a location from the aerial photo of their location, but many found that precise and accurate marking of each location was difficult due to the need to constantly compare the view in front of them with the aerial photo. Five participants spent a large amount of time (over 40 seconds) making their gesture very specific, whereas the maximum time taken for each of the other displays was 25 seconds, with the majority of the participants taking far less time than this.

5.4 Discussion

The use of an aerial photo for visual feedback to help the user more accurately indicate their areas of interest appears to have had both positive and negative impacts on the user experience. Visual location-specific feedback seemed to allow participants to position points of interest close to the level of precision required for the tasks; however, while the users of the aerial view have generally been much more accurate, this accuracy has come at the expense of time taken to mark a location. Trying to match up the aerial photo view with the physical surroundings is a potentially fiddly task. Using the arrow and no-display systems has allowed the users to mark places in less time, but at the expense of targeting accuracy.

The aerial view system is significantly slower than the other systems, but, especially with closer targets (below 50m), it is far more accurate. When we consider the arrow view and non-visual display results, there is a serious degradation in accuracy. Participants using the arrow view were, however, able to mark some targets accurately, and were able to do this faster than the participants using the aerial view. As some applications of this marking method may not need particularly accurate targeting, with perhaps the need to just specify a particular building rather then an accurate target, this method of interaction could be used for inexact or fuzzy targeting in some situations.

The results from the non-visual display are the least accurate of the three systems, and clearly allow only gross indications of distance, such as ‘near’, ‘middle’ and ‘far’. Despite this inaccuracy, it may still be useful in a mixed-mode system if users were able to switch between modes depending on their current status, from a non-visual marking method to a visualisation of their marked location if necessary.

Using the aerial view, targets closer to the user appear to have been easier to mark accurately. This is especially clear from Fig. 12, where targets 3 and 6 show a close spread of marked locations around the intended target. This is also present in the arrow view results for target 6, though not as clearly. A possible explanation for the inaccuracy of targets further away from the users is a lack of understanding of the tilt angles that were necessary to mark a location, possibly as a result of the small amount of training participants were given. Another possibility for this error could be a more general lack of distance perception or, rather, inability to convert visual distance perception into a discrete value in metres. This problem could be addressed with a system that gave feedback about object locations, or, as previously suggested, a multimodal device.

Errors in the pointing accuracy of our participants seem to suggest that they struggled to point in the correct direction, which is a surprising result given the simplicity of a pointing gesture. A likely explanation for this angle error is that an error that would be insignificant for nearer targets becomes far greater when projected further into the distance, making accurate pointing difficult beyond close range. Despite this, we would have expected much lower impacts on the angular error at the relatively short distances used for this study. Given that the angular errors that were recorded were present over a large number of participants, there are possible design implications for the design of future point-to-select systems, both non-model and model-based, if pointing can not be relied upon to allow accurate targeting of locations.

6 Conclusions and future work

The use of sensor data to mediate the combination of physical and digital experiences is a rich area for future research. In this work we illustrated how lightweight point-and-tilt gestures allied with location and orientation data can be used to generate interactive web maps, then studied in more depth the possibilities of these interactions. We described the approaches and presented three user studies to explore both visual feedback with point-and-tilt gestures as well as gesture only interactions.

In our initial studies, casual pointing and selecting was found to be an engaging task and people especially noted the attraction of putting down markers without looking having to look at the visual feedback: one of the users commented that '*Googling the real world*' was possible with nothing but a gesture. Our field prototypes allowed users to locate and mark locations with very little training, and find unexpected information even about familiar areas.

From our later study, it seems that visual feedback is too slow and cumbersome, and less-visual interaction is too inaccurate for the sort of scenarios we envisage. Tilting does provide some accuracy, but at the cost of time. In situations where a screen is available, tilting is a possible interaction method for marking locations; when no screen is available, or the display is limited, only broad categories of distance accuracy are possible. As noted previously, a combination of the three views presented here could offer the benefit of casual interaction when less accuracy is needed, and the opportunity to switch to a more precise interaction method when necessary.

Without a sophisticated model of the environment, it is clear that some form of feedback is still required in targeting areas of interest. The SHAKE devices we work with include haptic outputs and we are now considering how to employ these efficaciously. One possibility is to provide haptic feedback related to the density of geo-tagged content in an area. That is, as the user looks around an area, probing it with the device through pointing gestures, the SHAKE might vibrate depending on the amount of content that is known to be available. The user can then use this feedback along with their own view of the location to assess the current area being targeted by the system.

Another area of potential involves considering a wider vocabulary of gestures with the SHAKE. For instance when pointing at a location different gestures could be used to indicate the sorts of content the user is primarily interested in: an anticlockwise turning motion might denote the desire to find out about the history of a location; a series of up-and-down spoke like movement could show that the user wants content produced by people in their social network.

Previous work has shown the popularity of pointing to select targets, but it is evident from our results that correctly targeting distant locations may be problematic. Pointing accuracy degrades with distance, even when an aerial view is present to aid with precision. This is clearly an issue to address when considering the design of future point-to-select systems, both non-model-based designs and those with geo-spatial models. If a more casual interaction method is to be developed then targeting accuracy is essential to allow for the creation of usable devices.

7 Acknowledgements

This work is funded by EPSRC project EP/E042171/1, undertaken in collaboration with colleagues at Glasgow University: Roderick Murray-Smith, Stephen Brewster, John Williamson, Steven Strachan and Craig Stewart. Stephen Hughes provided the SHAKE devices used in our user studies.

References

[1] William Bamford, Paul Coulton, and Reuben Edwards. Space-time travel blogging using a mobile phone. In *ACE '07: Proceedings of the international conference on Advances in computer entertainment technology*, pages 1–8, New York, NY, USA, 2007. ACM.

[2] Russell Beale. Mobile blogging: experiences of technologically inspired design. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 225–230, New York, NY, USA, 2006. ACM.

[3] M. Egenhofer. Spatial information appliances: A next generation of geographic information systems. *First Brazilian Workshop on GeoInformatics*, October 1999.

[4] Fredrik Espinoza, Per Persson, Anna Sandin, Hanna NystrÃÂ¶m, Elenor Cacciatore, and Markus Bylund. GeoNotes: Social and navigational aspects of location-based information systems. In *Ubicomp 2001: Ubiquitous Computing*, volume 2201/2001 of *Lecture Notes in Computer Science*, pages 2–17. Springer Berlin / Heidelberg, 2001.

[5] Peter Fröhlich, Rainer Simon, Lynne Baillie, and Hermann Anegg. Comparing conceptual designs for mobile access to geo-spatial information. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 109–112, New York, NY, USA, 2006. ACM.

[6] Dominique Guinard, Sara Streng, and Hans Gellersen. Extending mobile devices with spatially arranged gateways to pervasive services. In *International Workshop on Pervasive Mobile Interaction Devices (PERMID 2007) at the International Conference on Pervasive Computing, PERVASIVE 2007*, May 2007.

[7] Lars Hallnäs and Johan Redström. Slow technology – designing for reflection. *Personal Ubiquitous Comput.*, 5(3):201–212, 2001.

[8] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Human mental workload*, pages 139–183, 1988.

[9] Matt Jones, George Buchanan, Tzu-Chiang Cheng, and Preeti Jain. Changing the pace of search: Supporting "background" information seeking. *J. Am. Soc. Inf. Sci. Technol.*, 57(6):838–842, 2006.

[10] Bonnie A. Nardi, Diane J. Schiano, Michelle Gumbrecht, and Luke Swartz. Why we blog. *Commun. ACM*, 47(12):41–46, 2004.

[11] Trevor Pering, Yaw Anokwa, and Roy Want. Gesture connect: facilitating tangible interaction with a flick of the wrist. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 259–262, New York, NY, USA, 2007. ACM.

[12] Jun Rekimoto, , Takashi Miyaki, and Takaaki Ishizawa. LifeTag: WiFi-Based continuous location logging for life pattern analysis. In *Location- and Context-Awareness*, volume 4718/2007 of *Lecture Notes in Computer Science*, pages 35–49. Springer Berlin / Heidelberg, 2007.

[13] Simon Robinson, Parisa Eslambolchilar, and Matt Jones. Point-to-geoblog: gestures and sensors to support user generated content creation. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 197–206, New York, NY, USA, 2008. ACM.

[14] Enrico Rukzio, Karin Leichtenstern, Vic Callaghan, Paul Holleis, Albrecht Schmidt, and Jeannette Chin. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. In *UbiComp 2006: Ubiquitous Computing*, volume Volume 4206/2006, pages 87–104. Springer Berlin / Heidelberg, 2006.

[15] Rainer Simon and Peter Fröhlich. A mobile application framework for the geospatial web. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 381–390, New York, NY, USA, 2007. ACM.

[16] Rainer Simon, Peter Fröhlich, and Hermann Anegg. Beyond location based - the spatially aware mobile phone. In *Web and Wireless Geographical Information Systems*, volume Volume 4295/2006, pages 12–21. Springer Berlin / Heidelberg, 2006.

[17] Rainer Simon, Harald Kunczier, and Hermann Anegg. Towards orientation-aware location based mobile services. In *Location Based Services and Tele-Cartography*, pages 279–290. Springer Berlin Heidelberg, 2007.

[18] Steven Strachan, John Williamson, and Roderick Murray-Smith. Show me the way to monte carlo: density-based trajectory navigation. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1245–1248, New York, NY, USA, 2007. ACM.

[19] Rainer Wasinger, Christoph Stahl, and Antonio Krüger. M3I in a pedestrian navigation & exploration system. In *Human-Computer Interaction with Mobile Devices and Services*, volume Volume 2795/2003, pages 481–485. Springer Berlin / Heidelberg, 2003.

[20] John Williamson, Roderick Murray-Smith, and Stephen Hughes. Shoogle: excitatory multimodal interaction on mobile devices. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 121–124, New York, NY, USA, 2007. ACM.

Figure 1: The equipment in use. Inset: the SHAKE sensor pack.

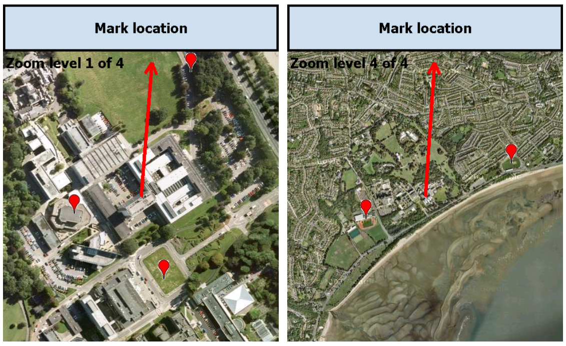
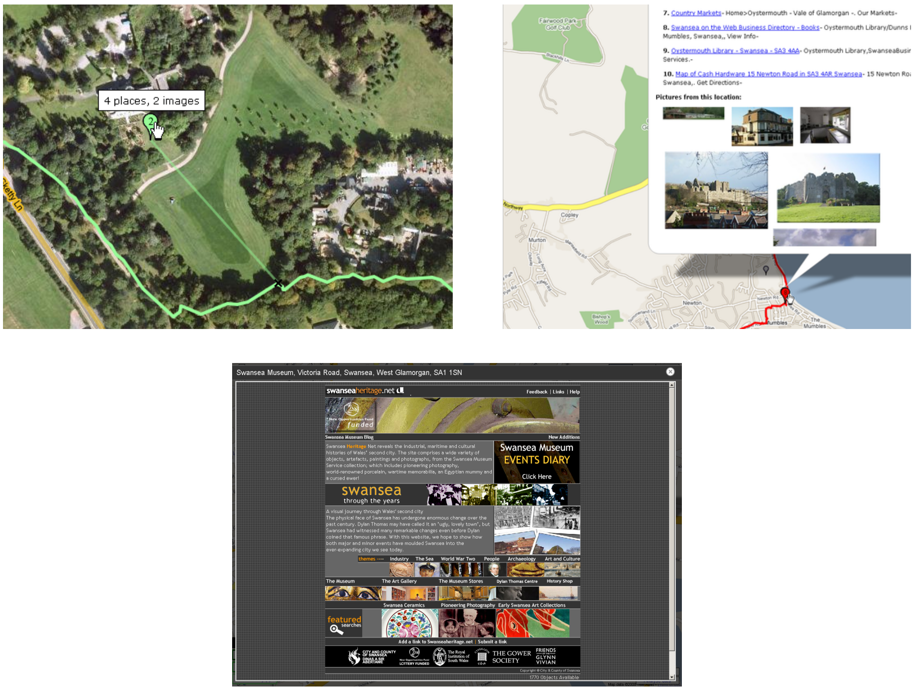


Figure 2: Sample screens from the device. Left: the default (minimum) zoom level. Right: maximum zoom level, with several locations marked.



7. Country Markets- Home>Oystermouth - Vale of Glamorgan -. Our Markets-

8. Swansea on the Web Business Directory - Books- Oystermouth Library/Dunns
Mumbles, Swansea,, View Info-

9. Oystermouth Library - Swansea - SA3 4AA- Oystermouth Library,SwanseaBusi
Services.-

10. Map of Cash Hardware 15 Newton Road in SA3 4AR Swansea- 15 Newton Ro
Swansea,. Get Directions-

Pictures from this location:

Figure 3: Upper left: Clicking on a marker displays content results. Right: As user hovers over a marker, a line is drawn to show where the gesture originated and content statistics are displayed. Lower: Sample webpage search result.

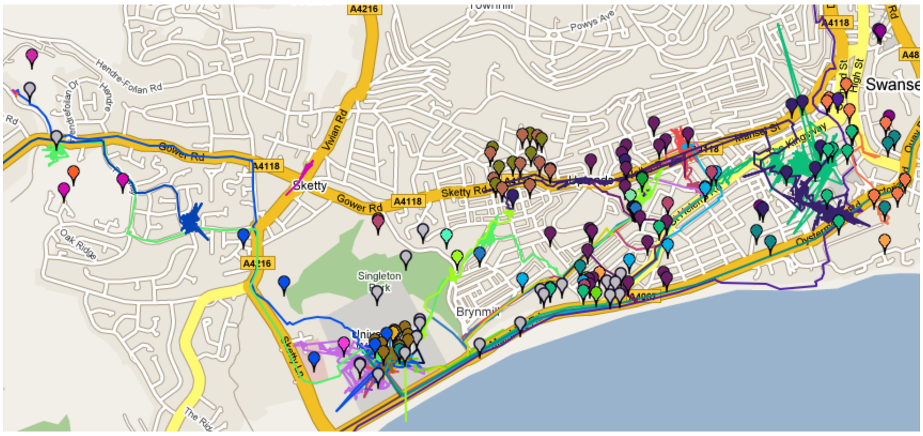


Figure 4: All journeys (shown as lines) and marked locations (as pins) within the main common journey area.

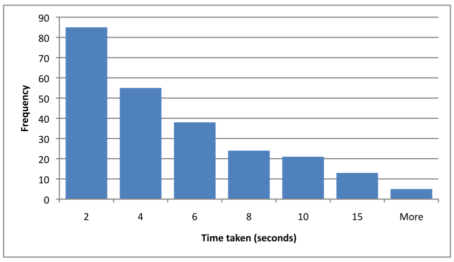


Figure 5: Time taken to mark locations.

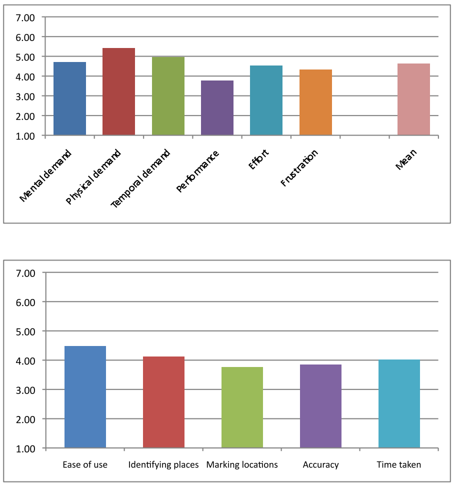


Figure 6: Top: TLX ratings. Bottom: Feature ratings.

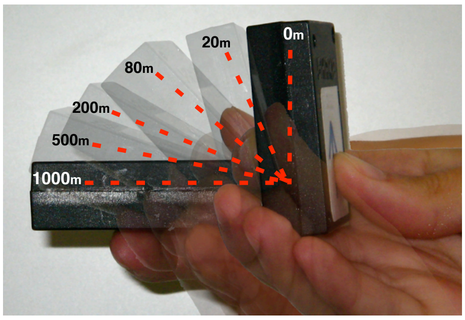


Figure 7: Indicating distance (metres) by degree of tilt.

Figure 8: Left to right: the user takes the device from their side, points at a location and then tilts the SHAKE to indicate the target's distance.

Figure 9: PDA screen displays from the three test systems. Left: system one: aerial view with moving arrow and two locations marked. Centre: system two: arrow view showing a user marking a location almost 125m from their present position. Right: system three: non-visual display – only the number of locations marked so far is displayed.

Target number and description	Distance (metres)	Viewing angle (degrees)
1: Banner on building balcony	140	-22
2: Flag on side of building	81	-58
3: Box on building roof	23	+27
4: Archway leading out of campus	174	-19
5: Tree on campus lawn	108	+31
6: Tree in building courtyard	46	-42

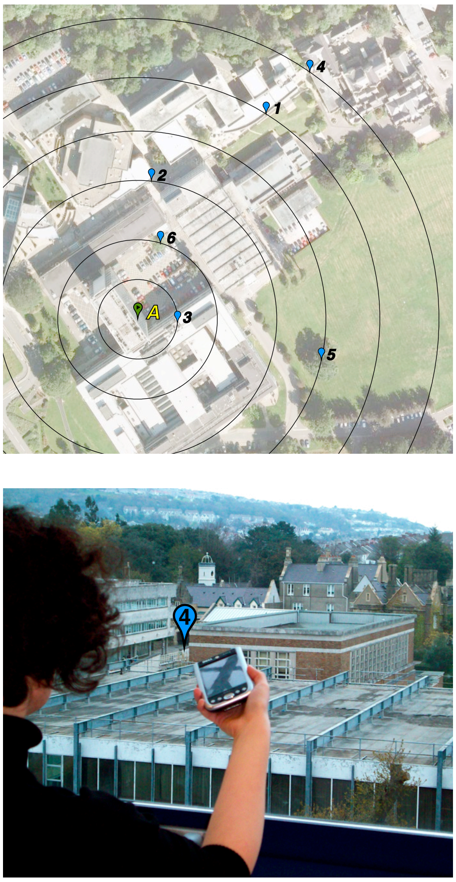


Figure 10: Top: Distance and viewing angle (from participants' viewpoint, perpendicular to lab) of each of the target points. Middle: the six targets used for our study. Point *A* shows the position of the participant in the lab. Bottom: Marking a target from the lab.

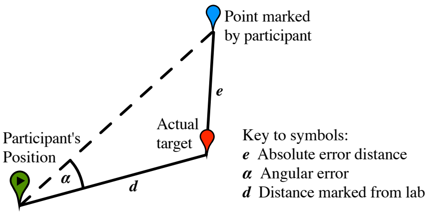


Figure 11: Accuracy measurements recorded for each point marked by participants.

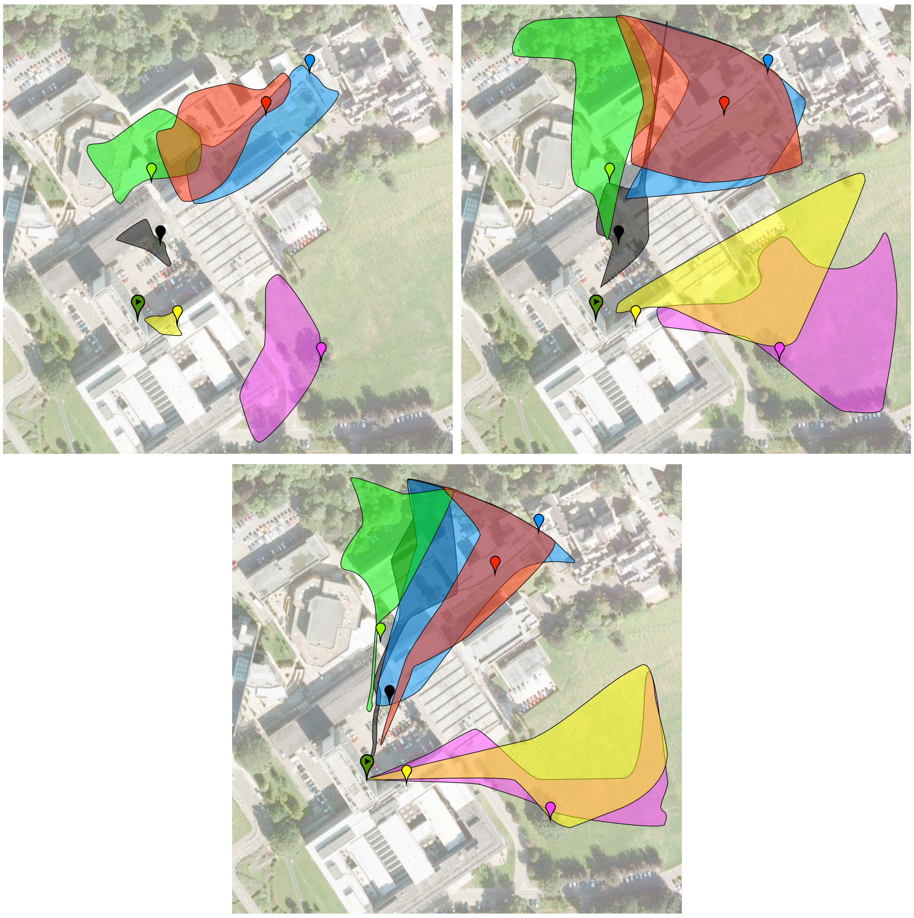


Figure 12: Targeted area ranges for each system, showing the intended targets and the outer edge of the actual target areas selected when data from all participants is combined. Upper left: aerial view, Right: arrow view, Lower: non-visual.

System	Aerial view		Arrow view		Non-visual	
Target	Mean	St. d.	Mean	St. d.	Mean	St. d.
1	38.79	22.04	40.96	20.60	55.63	26.65
2	23.33	12.92	55.78	28.59	50.04	29.64
3	11.94	4.31	41.24	54.54	96.85	52.21
4	42.46	36.92	44.94	29.52	48.63	43.92
5	26.30	18.24	62.84	18.00	67.32	29.95
6	8.58	6.51	33.47	38.98	68.90	48.10
All	25.23	23.11	46.54	34.33	64.56	41.68

(a) Absolute error distance from each target (metres).

System	Aerial view		Arrow view		Non-visual	
Target	Mean	St. d.	Mean	St. d.	Mean	St. d.
1	14.17	13.72	14.61	10.60	11.41	5.85
2	17.15	10.54	14.66	10.18	8.81	6.21
3	8.10	4.81	16.15	6.85	31.11	33.52
4	7.66	6.74	14.80	10.05	14.30	13.22
5	6.72	8.00	13.27	7.90	12.20	9.16
6	13.84	11.85	8.51	7.31	9.76	5.82
All	11.27	10.24	13.67	8.99	14.60	17.00

(b) Angular error from target viewing angle (degrees).

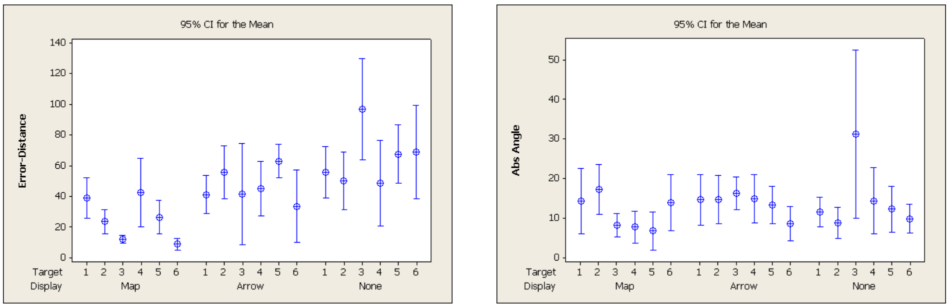
System	Aerial view		Arrow view		Non-visual	
Target	Mean	St. d.	Mean	St. d.	Mean	St. d.
1	123.28	26.78	145.26	28.73	122.09	55.79
2	90.18	16.38	125.13	39.49	107.13	51.72
3	13.82	7.78	53.00	59.35	107.14	69.64
4	137.09	37.45	156.00	30.60	148.07	47.04
5	91.99	10.58	113.31	47.37	111.17	67.76
6	45.56	4.91	72.68	43.45	100.60	64.23
All	83.65	47.30	110.90	55.69	116.03	59.93

(c) Distance marked from lab (metres)

System	Aerial view		Arrow view		Non-visual	
Target	Mean	St. d.	Mean	St. d.	Mean	St. d.
1	25.89	11.53	11.29	4.09	9.86	6.56
2	27.01	16.40	9.51	3.64	6.90	5.51
3	21.80	13.20	7.71	3.23	4.64	3.46
4	23.85	14.15	6.59	2.46	5.94	1.93
5	20.83	13.13	10.17	5.03	7.06	5.66
6	10.24	5.90	7.77	3.15	7.72	8.95
All	21.60	13.57	8.84	3.92	7.02	5.81

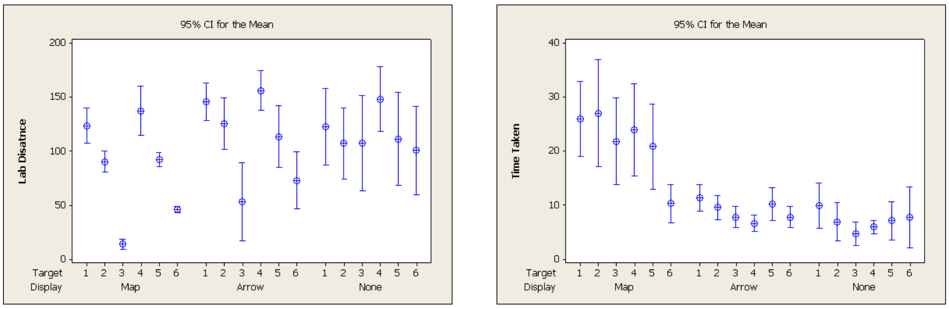
(d) Time taken to mark each target (seconds).

Figure 13: Summary of results from each measurement over each of our test systems.



(a) Absolute error distances.

(b) Deviations from the target viewing angles.



(c) Distance marked from participant's location.

(d) Time taken to mark each target

Figure 14: Interval plots including each measurement for each target on each system.