Conference contribution :

Robinson, S., Eslambolchilar, P. & Jones, M. (2008). *Point-to-GeoBlog: Gestures and Sensors to Support User Generated Content Creation.* MobileHCI 2008, (pp. 197-206). ACM.

http://dx.doi.org/10.1145/1409240.1409262

Point-to-GeoBlog: Gestures and Sensors to Support User Generated Content Creation

Simon Robinson, Parisa Eslambolchilar, Matt Jones
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. No content is provided to the user in situ but, later, web materials including images, entries from other people's blogs and web pages are automatically placed on an interactive map for viewing on a larger screen device. We built two mobile prototypes to explore the approach – one combines gestures and visual map feedback; the other is more lightweight, allowing the user to simply point-and-tilt. We describe and motivate the approaches and present user studies that raise issues relevant to their design and to the wider class of device and service concerned with mobile spatial information access.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: input devices and strategies; interaction styles; prototyping

General Terms

Design, Experimentation, Human Factors

Keywords

User-generated content, mobile, geo-web, blogging, gestures, sensors

1. INTRODUCTION

We blog therefore we are. Increasingly, people are documenting their lives – reporting on their emotions, thoughts, plans and actions – to remember, make sense of or share their experiences [8].

Three trends in blogging motivate our work here. Firstly, there is the proliferation of short, frequent low-cost posts most obvious as 'status updates' in social networking sites such as *Facebook* or *Twitter*; what has been called nano-blogging. These sparse and frequent acts of authorship are particularly attractive to mobile users. Then there is the enriching of the blog medium. At first, blogs were text postings but now photos, videos and graffiti all play their part. Finally, while a person may use several web services – *Flickr* for photos, *Twitter* for status updates – there is a clear role for places where these materials can be drawn together. Part of the success of Facebook may be accounted by its integrating role; and, map based views, such as *Google MyMaps*, that enable users to add their own content, are likely to become extremely popular.

In this paper we explore lightweight approaches for casually gathering location-orientated material while mobile; sensor data is used both to collect and provide content; and, integrating map visualisation is used as the basis of the journey record.

Consider this example interaction, supported by our systems, to indicate 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 the paper. 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 [2, 19], 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.

We built two experimental sets of apparatus to explore options for supporting the interaction. The mobile element of the first combines a pointing gesture with visual feedback to indicate areas of interest. In the second, the mobile aspects require minimal attention from the user while *in situ*.

We begin in Section 2 by describing the first prototype. Section 3 then presents the study used to understand the value of the approach and the effectiveness of the interactions afforded. The lighter-weight prototype is described in Section 4 along with a study of its viability. Our work cuts-across developments in gesture interactions, search and the geo-web and we situate the approach in Section 5. Conclusions and pointers to future work complete the article.

2. PROTOTYPE 1: GESTURES AND VISUAL FEEDBACK

The system segregates the blogging 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 Figures 3, 4 and 5). In this section, we outline how these facilities are provided.

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

2.1 Mobile hardware

We use SHAKE (Sensing Hardware Accessory for Kinesthetic Expression, see [18]) 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 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.

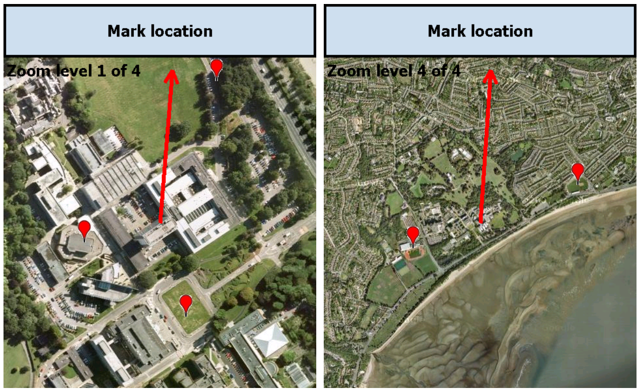


Figure 2: Sample screens from the device. Left: the default (minimum) zoom level. Right: maximum zoom level, with several locations marked.

2.2 Marking points of interest

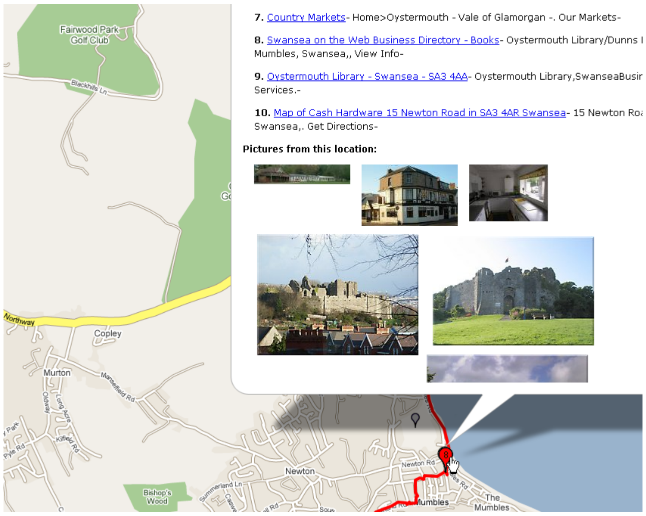
The system displays an aerial photo of the user's current location, overlaid with an arrow showing the direction they are currently facing, pinpointing the location that can be marked. To mark areas of interest, users point at the area of interest with the PDA in their hand. They then tilt the device 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 (a point we will return to in discussing the findings of the first study).

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.

Figure 3: As user hovers over a marker, a line is drawn to show where the gesture originated and content statistics are displayed. (Map data sourced via Google Maps' public API).


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 4: Clicking on a marker displays content results.

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. [12]). 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 is blank.


Swansea Museum, Victoria Road, Swansea, West Glamorgan, SA1 1SN

Swansea Heritage Net reveals the industrial, maritime and cultural histories of Wales' second city. The site comprises a wide variety of objects, artefacts, paintings and photographs, from the Swansea Museum Service collection; which includes pioneering photography, world-renowned porcelain, wartime memorabilia, an Egyptian mummy and a cursed ewer!

A visual journey through Wales' second city
The physical face of Swansea has undergone enormous change over the past century. Dylan Thomas may have called it an 'ugly, lovely town', but Swansea had witnessed many remarkable changes even before Dylan coined that famous phrase. With this website, we hope to show how both major and minor events have moulded Swansea into the ever-expanding city we see today.

Copyright © City & County of Swansea

1770 Objects Available

Figure 5: Sample webpage search result.

2.3 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, 100 metres squared in size, is defined.

2.3.1 *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. These location surrogates are then used as query terms in a series of searches. A web search engine (currently Google) supplies text and image results and further images are garnered from social networking sites.

2.3.2 *Map visualisation*

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. Figures 3, 4 and 5 illustrate the possible interactions and information retrieved.

3. EVALUATING THE SYSTEM

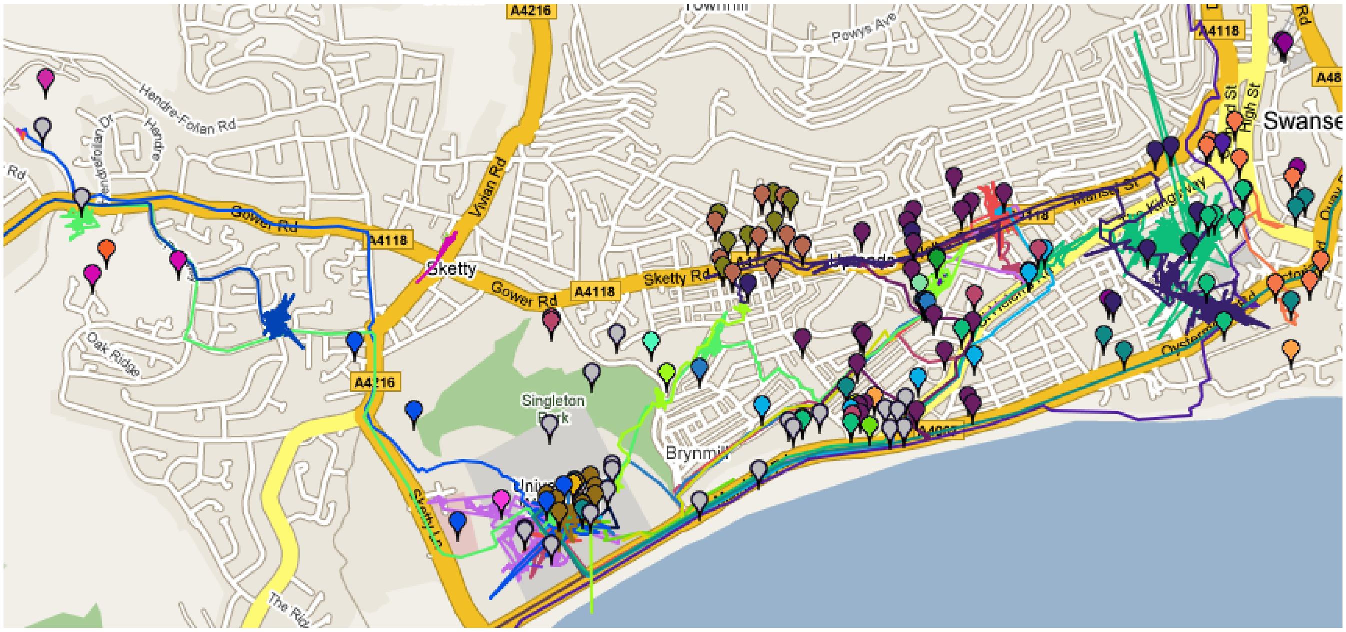


Figure 6: All journeys (shown as lines) and marked locations (as pins) within the main common journey area.

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.

3.1 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 previous first-hand knowledge of accelerometer-based interaction in a gaming context 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 [5]. 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.

3.2 Findings

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.

3.2.1 *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 6 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 7 shows the distribution of times taken to mark locations.

3.2.2 Mobile Task Load and Usability Ratings

Figure 8 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 9 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.

3.2.3 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.

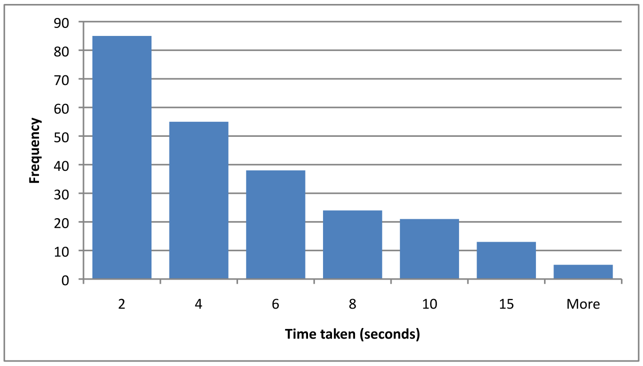


Figure 7: Time taken to mark locations.

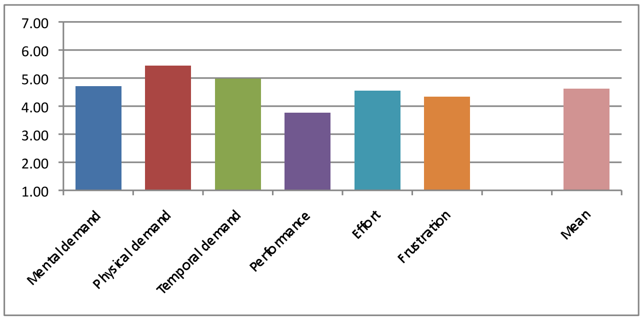


Figure 8: TLX ratings.

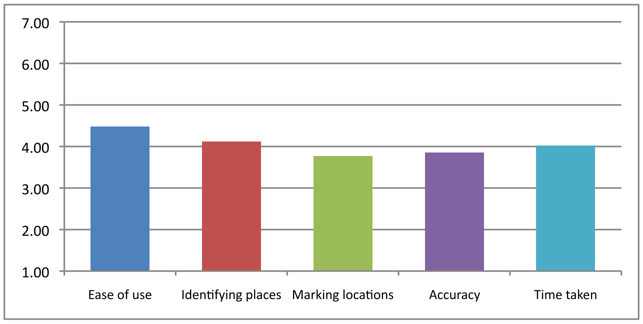


Figure 9: Feature ratings.

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.

3.2.4 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.

3.3 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 geo-tagged with a low degree of precision, it could be argued that approaches such ours are overly involved for these nearby queries. 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. 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 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. Before developing the gesture-visual system, we created a prototype that illustrates such a minimal attention user interface method and deployed it in a field-study.

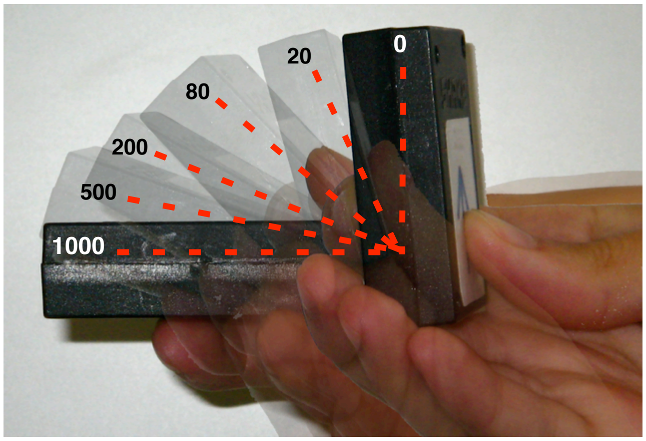


Figure 10: Indicating distance (metres) by degree of tilt.

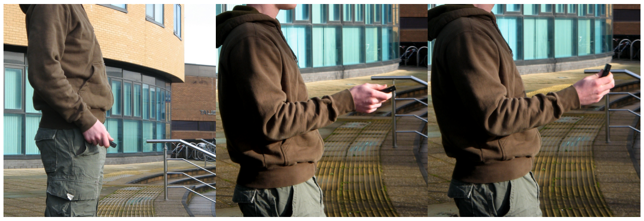


Figure 11: Left to right: the user takes the device from their side, points at a location and then tilts the SHAKE to indicate the target's distance.

The map creating and visualisation elements were essentially the same as those described in Section 2. 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 2.1 during the studies but no visual feedback was provided.

4.1 Marking points of interest

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 10). 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 11).

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.

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 staff members at a local university, 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 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.

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. RELATED WORK

Recent growths in blog content diversity have led to an increase in the variety of different uses for blogs, ranging from personal diaries to scrapbooks, and from news digests to discussion forums [11]. The latest revolutions in blogging have been driven by the availability of mobile devices, which can allow bloggers to update whilst on the move, at any time [2, 19]. Ames and Naaman [1], for instance, created *ZoneTag*, a mobile phone application that allows users to upload photos to Flickr. When uploading a photo, users are presented with a list of tags that the application believes are particularly relevant to the photo taken. These tag lists are sorted based on the user's physical location, and are also augmented with tags created by other people in the same location and tags from the user's social network. In analysis of users' tags, they found several user motivations for tagging, and noted that users often tagged their pictures to ease later retrieval, similar to the phased interaction method of this project.

Gesture-based control of mobile device and service functionality has received much attention – for example, in [6] voice recorder activation is achieved by bringing the device to the ear, obviating the need to press a button.

Other proposals combining gestures and other sensor data for smart spatial appliances have also been made. These, 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.

Wasinger *et al.* [17] 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 [3, 13, 12, 14, 15] discuss implementations of 'point-to-select' methods of interaction to retrieve information. This technique allows a user to simply point at an object to indicate an interest in it. Initially, Peter Fröhlich *et al.* [3] 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.* [13] 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.* [12] 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 [14] 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.

Similarly, Strachan *et al.* [16] 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.

These approaches demonstrate 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 [4] 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.* [9] describe a similar system in which users can connect to and control physical objects using electronic tags and simple gestures.

Rukzio *et al.* [10] 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. They all require virtual location models in order to be able to ascertain the user's target points. Whilst these authors provide valuable insights into possible methods and uses of location-based interaction, the aim of this project is to provide users with similar data but without the need for complex location models and visibility calculations.

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. In [7], another form of delayed-search is presented. 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.

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 paper we illustrated how lightweight point-and-tilt gestures allied with location and orientation data can be used to generate interactive web maps. We described the approach and presented user studies to explore both visual feedback with point-and-tilt gestures as well as gesture only interactions.

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 users commented that '*Googling the real world*' was possible with nothing but a gesture. Point-to-GeoBlog allowed users to locate and mark locations with very little training, and find unexpected information even about familiar areas.

It seems that visual feedback is too cumbersome for the sort of scenarios we envisage. However, 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 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.

While we have focused on a two-phased scenario, the low cost gestures used here could be also put to use in real-time scenarios. Consider, for instance, the nano-blogging status updates of Facebook or Twitter. Instead of entering a text update while mobile, a user could gesture at a location and be provided with a list of interesting recommendations that they could simply select or customise. For example, when our Sam, in the paper's introduction, points at the houses in Singapore the list might include: "Sam is near Arab Street and the beautiful mosque"; and, "Sam is close to the famous Arab street coffee lounge". The updates could include hyperlinks to allow the users friends to find out more from the web resources used to create the recommendations. Such recommendations are already used for photo annotation in *ZoneTag* [1].

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.

8. REFERENCES

[1] M. Ames and M. Naaman. Why we tag: motivations for annotation in mobile and online media. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 971–980, New York, NY, USA, 2007. ACM.

[2] R. 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] P. Fröhlich, R. Simon, L. Baillie, and H. 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.

[4] D. Guinard, S. Streng, and H. 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.

[5] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Human mental workload*, pages 139–183, 1988.

[6] K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing techniques for mobile interaction. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 91–100, New York, NY, USA, 2000. ACM.

[7] M. Jones, G. Buchanan, T.-C. Cheng, and P. Jain. Changing the pace of search: Supporting "background" information seeking. *J. Am. Soc. Inf. Sci. Technol.*, 57(6):838–842, 2006.

[8] B. A. Nardi, D. J. Schiano, M. Gumbrecht, and L. Swartz. Why we blog. *Commun. ACM*, 47(12):41–46, 2004.

[9] T. Pering, Y. Anokwa, and R. 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.

[10] E. Rukzio, K. Leichtenstern, V. Callaghan, P. Holleis, A. Schmidt, and J. 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.

[11] D. J. Schiano, B. A. Nardi, M. Gumbrecht, and L. Swartz. Blogging by the rest of us. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1143–1146, New York, NY, USA, 2004. ACM.

[12] R. Simon and P. 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.

[13] R. Simon, P. Fröhlich, and H. 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.

[14] R. Simon, P. Fröhlich, G. Obernberger, and E. Wittowetz. The point to discover geowand. In *9th International Conference on Ubiquitous Computing (UbiComp 07)*, Innsbruck, Austria, 2007. Springer LNCS.

[15] R. Simon, H. Kunczier, and H. Anegg. Towards orientation-aware location based mobile services. In *Location Based Services and TeleCartography*, pages 279–290. Springer Berlin Heidelberg, 2007.

[16] S. Strachan, J. Williamson, and R. 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.

[17] R. Wasinger, C. Stahl, and A. 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.

[18] J. Williamson, R. Murray-Smith, and S. 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.

[19] C.-I. Wu, C.-M. J. Teng, Y.-C. Chen, T.-Y. Lin, H.-H. Chu, and J. Y.-J. Hsu. Point-of-capture archiving and editing of personal experiences from a mobile device. *Personal Ubiquitous Comput.*, 11(4):235–249, 2007.