iCapture: Facilitating Spontaneous User-Interaction with Pervasive Displays using Smart Devices

Keith Mitchell, Nicholas J.P. Race and Michael Suggitt*

February 10, 2006

Abstract. *The eCampus project at Lancaster University is an inter-disciplinary project aiming to deploy a wide range of situated displays across the University campus in order to create a large pervasive communications infrastructure. At present, we are conducting a series of parallel research activities in order to investigate how the pervasive communications infrastructure can support the daily needs of staff, students and visitors to the University. This paper introduces one of our current research investigations into how one is able to mediate spontaneous interaction with the pervasive display infrastructure through camera equipped mobile phones (i.e. smart devices).*

1 Introduction

The eCampus project at Lancaster University is currently in the process of creating a campus-wide pervasive communications infrastructure. More specifically, we are in the process of deploying a range of display technologies (plasma, LCD and projection systems), communications technologies (wi-fi and Bluetooth) and sensors (cameras, PIR, etc) within a number of public spaces (indoor and outdoor) across campus. Perhaps uniquely to eCampus, the infrastructure must satisfy some high level goals, namely to be available to all members of the University 24/7 365 days of the year and furthermore, to act as a research resource for all faculties of the University.

The duality related to supporting both a 'research' and a 'production' service poses many challenges both for design, deployment and on-going maintenance. To overcome some of these issues we have deployed a hardware platform in which each display can be controlled from multiple sources. This allows developers and content providers across campus to be able to use their own preferred tools and platforms(Mac, PC, Linux). This architecture not only offers redundancy but additionally allows us to toggle rapidly between production level (i.e. reliable) news and information services and a research infrastructure in order to facilitate researcher experiments and user trials.

At present, the eCampus infrastructure has been used for several new media and artistic performances [6] as well as offering local news and information services. This paper introduces one recent area of

*Network Research and Special Projects Unit, ISS, Computer Centre, Lancaster University, Lancaster, LA1 4YW, U.K. email: [k.mitchell, n.race, m.suggitt]@lancaster.ac.uk

research with respect to the use of camera equipped mobile phones for initiating user interaction with situated display environments. We describe an application which facilitates fluid and spontaneous user interaction with the infrastructure and highlight some of the challenges involved when small (displays) meets large (displays).

2 Facilitating User Interaction

The ultimate goal of ubicomp and pervasive computing systems is that they are invisible and become part of the fabric of everyday life [10]. Whilst at a high level this is an objective for the eCampus project, a dichotomy exists between this goal and the goal of attracting user attention at appropriate times and facilitating interaction with the systems in place. We are currently in the process of investigating a number of different interaction mechanisms and this paper describes one scenario based around the use of smart devices.

At present, there are conflicting trends within collaborative environments between delivering information and services to a large audience via situated displays and also being able to exploit mobility on smaller displays in order to exploit more personalised [1] content provision. It is clear that while large (potentially shared) displays provide greater choice, flexibility and presentation of content, mobile devices offer greater flexibility in terms of personalisation [7].

Large situated displays within public environments by their very nature afford themselves to the presentation of content which may be suited to attracting the attention of a wide audience, for example news, travel and information being displayed in train stations, airport lounges and motorway services. However, while useful for *pushing* information these may not be appropriate for offering *interactive* services directly. This may be due to the fact that it is often hard to attract user interaction in the first instance due to factors such as social embarrassment [3]. While the provision of *personalised* information may be an incentive to use interactive public displays the additional technical overhead required on the users's behalf (i.e. though configuring or authentication) may act as an additional deterrent [5]. For these situations, delivering content to personal devices may offer an attractive solution since these may already be personalised towards a user.

We chose to investigate what happens when small meets large, that is, could user interaction be facilitated by the use of smart devices such as camera phones. We therefore chose to leverage the best of both worlds by allowing multiple users to access and interact with a large, situated display using their own personal device. The goal was to provide a lightweight user interaction mechanism which would exploit the eCampus infrastructure in order to offer ubiquitous access to information and resources at any convenient time. Furthermore, we hoped to investigate the use of smart devices for initiating both public and private interactions and the trade-offs associated with each.

3 Context Of Use: Location , location, location!

Initial observations of the current infrastructure reveal that it is a non-trivial task to design captivating content for multiple public displays [9]. In fact, while peoples attention is automatically drawn to things which are novel in our environment [2, 4], we believe this novelty factor will soon disappear and something more compelling is required in order to stimulate user-interaction with pervasive display environments long term. Furthermore, we feel that spatial facets such as location and positioning are crucial when considering the style of interaction expected based on the display location. More specifically, the types of public spaces being targeted may offer significant and often contrasting challenges. For example, whether a display acts as a passive information screen or an interactive access point may point towards the preferred style of interaction (e.g. touch, gesture or mobile device). This interaction style may heavily influence the amount of attention required to maintain focus on the display itself.

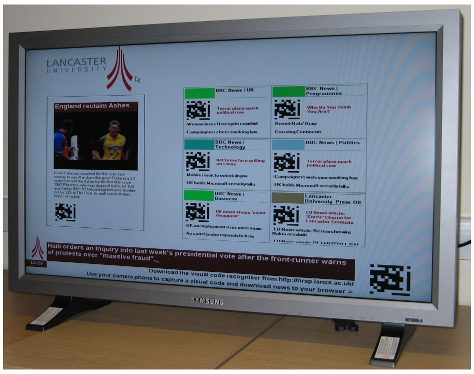
In order to establish the style of interaction likely to occur within the various campus surroundings, we investigated the notion of space in order to identify the following types of public spaces:

- Public - An open area of space in which any member of the public may enter without restriction, such as a square, etc.
- Private - An area in which restriction may apply, such as a staff member's office or lecture theatre.
- Transient - Spaces which people do not congregate easily or frequently, such as hallways, corridors or walkways.
- Social - These are generally public spaces but often have a purpose associated with them, for example, coffee shops or bars in which people arrange to meet. They may or may not have restrictions applied to them, such as the requirement to purchase a drink in order to sit in a cafe.
- Informative - These spaces are closely related to public spaces and transient spaces and include building foyers or areas close to notice boards, that is, spaces which are dedicated to or used for public announcements.

These categories may offer some hints towards helping designers understand and identify the requirements for the target deployment scenarios however there are likely to be many additional deployment specific constraints or factors to be considered. For example, in our University library - perhaps the busiest building on campus and therefore an ideal opportunity to engage many of the staff and students - we faced an interesting challenge. More specifically, we had to ensure that we did not develop an invasive technology or something which was likely cause noise, distraction or cause large numbers of people to congregate in a single location. This caused us to deploy a display within an *informative space* which offers little interaction and acts more as a passive information display. In the next section, we focus upon an application which we believe is suitable for *transient* spaces.

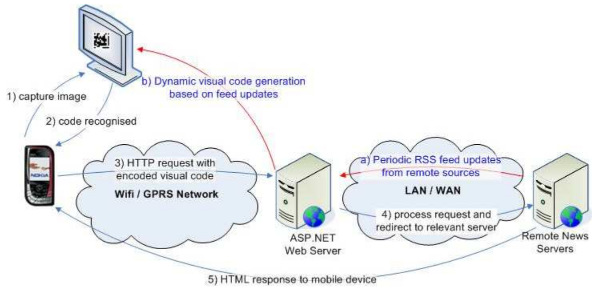
4 The Overall iCapture Application

The iCapture software architecture is presently based around an ASP.NET web application which resides on an eCampus server. This web application is primarily responsible for handling incoming web requests from mobile clients over HTTP. In order to satisfy these requests it must query our back-end application which is responsible both for processing RSS and aggregating RSS news feeds and



(a) eCampus Large Display

(b) iCapture Recogniser Application



(c) Overall iCapture Architecture

Figure 1: iCapture Situated Display System

also for controlling the large display user interface. Incoming requests from mobile clients contains a HTTP GET request with a string supplied by the mobile device which is used by the server to determine the specific RSS feed required. This URL sting parameter is used to redirect the mobile device to the relevant news server which responds with the appropriate content directly.

An eCampus large display is shown in figure 1 and comprises of two main regions. The lower part of the interface displays a news ticker which currently toggles between displaying content based on the BBC Front Page headlines [1] and the Lancaster University News headlines [2]. This region also offers information and a link in order to download the recogniser application. The content from the various feeds is retrieved and processed by the news aggregator before being displayed within the news ticker control. Additionally, the aggregator component retrieves other RSS feeds in order to populate content in the upper region of the situated display interface. At present, we have created a number of categories each showing a small number of news headlines, such as UK Headlines, politics, technology, sport and entertainment.

The aggregator component is also responsible for generating visual code objects which corespond to new content retrieved from RSS data feeds. A visual code object reflects an encoded representation of the URI, and therefore act as the anchor between the 'real' and 'virtual' world. By selecting a visual code using a camera phone, it enables the relevant news item to be shown on a users' mobile device.

The iCapture recogniser application resides on the mobile device and makes use of the visual code recognition algorithm described in [6] in order to capture a single image using the devices built-in

[1] http://newsrss.bbc.co.uk/rss/newsonline_uk_edition/front_page/rss.xml

[2] http://domino.lancs.ac.uk/info/lunews.nsf/rss2-0.xml

camera. The software produces a code information object for each detected code within a captured frame. This code is then passed as a parameter to our ASP.NET web server via HTTP which processes the request and responds with the appropriate news article. At present each web request results in a HTML response which is rendered locally on the mobile device handset.

5 Evaluation and Results

The original aim of this prototype was to facilitate spontaneous user-interaction with situated display environments. More specifically, to investigate if small personal devices could be used to interact with a large public displays without breaking the fixed temporal flow of the large display or interrupting the user's current activity.

In order to evaluate the application at this stage, we used the following scenario: An eCampus situated display is showing several news headlines as well as a live video stream of the BBC 1 o'clock news. As student Simon walks past the display, he notices a headline relating to the University's weekend sports results. He wishes to read more about this article but is currently on his way to a lecture and so does not have time to stop and look at the display. He therefore uses his camera enabled phone to take a quick photo of the visual code which is displayed next to the news headline he is interested in and continues on his way. Since the iCapture application automatically downloads the article he is able to read this as he walks or view the content at his leisure anytime later.

At present, we have tested the iCapture application with several members of the networking team within the department. We have been initially interested to see if this style of application is primarily easy to use and also able to offer useful functionality. Thus far, the results have been positive with all users agreeing that the interaction method is simple and intuitive. In general most users were successfully able to capture an image on their first attempt. This is due to the feedback offered by the recogniser on the mobile handset, shown in figure 1b. Some negative comments were made relating to the overall time taken between capturing an image and receiving the content on the device. At present this is approximately 15 seconds for the initial request and subsequent requests are satisfied within 8 - 10 seconds. The extra time taken for the initial request is the time taken for the handset to establish a GPRS connection with the service provider, a process beyond our control. We also received some feedback in terms of potential improvements, which are described in the following section.

6 Conclusion and Future Work

Pervasive computing and communications technologies are being rapidly deployed and becoming more familiar within contemporary society. However, these presently focus on marketing/advertising deployments or scenarios such as train stations in which they are used to represent context related information such as timetables, local news. However, their ubiquity could be used to stimulate more compelling and useful user interaction on a large scale. This paper discusses an initial application scenario in which users are able to rapidly gain access to information through the use of camera enabled mobile phones which make use of the visual code recognition software [8].

We are currently investigating several areas of future work based on our initial experiment and trial with users from the department. With respect to usability, we are modifying the recogniser such that the application asks the user if they wish to retrieve the content now or save as a bookmark (for later retrieval). This will allow the user to control when content is retrieved since at present, recognising a visual code automatically launches the mobile device web browser. This will also allow the visual code markers to act like virtual bookmarks embedded within the University environment and this is being implemented as a result of user feedback.

We are currently porting the Symbian based visual code recognition software to .NET in order for the services to be made available from Windows Mobile devices, such as the i-mate SP5 and SP5m[3] which have built in wi-fi and 802.1x authentication. We then aim to exploit this for more personalised service delivery via the campus wireless LAN and to explore the issues relating to the small and large display user interaction with respect to public versus private content. For example, upon recognising a code successfully, should the server display the content on the large display (public) or return the result to the requesting client (private) or some combination of these (such as content on the public display and navigation on the private display). It would be interesting to investigate the trade-offs more closely when considering public and private information retrieval within situated display environments. These modifications followed by a wider deployment outside of the lab will enable us to carry out some long term studies and improvements.

References

[1] Agamanolis, S. (2002) Designing displays for Human Connectedness. In: Workshop on Public, Community and Situated Displays at CSCW02, New Orleans. 4.

[2] Berlyne, D.E., Conflict, Arousal and Curiosity, McGraw-Hill, 1960.

[3] Brignall, H. Rogers, Y., ”Enticing People to Interact with Large Public Displays in Public Spaces”, Proceedings of INTERACT’ 03, Zurich, September 2003. pp. 17-2

[4] Fleck, R. (2003) How the move to physical user interfacees can make human computer interaction a more enjoyable experience. In: Workshop on Real World User Interfaces, Mobile HCI 2003, Udine, Italy

[5] Izadi, S., Brignull, H., Rodden, T., Rogers, Y., and Underwood, M. 2003. Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In Proceedings of the 16th Annual ACM Symposium on User interface Software and Technology (Vancouver, Canada, November 02 - 05, 2003). UIST ’03.

[6] Metamorphosis, Interactive Art
http://domino.lancs.ac.uk/info/lunews.nsf/I/97F2B7D7B37636BA8025703C00559183

[7] Paek, T., Agrawala, M., Basu, S., Drucker, S., Kristjansson, T., Logan, R., Toyama, K., and Wilson, A. 2004. Toward universal mobile interaction for shared displays. In Proceedings of the 2004 ACM Conference on CSCW (Chicago, Illinois, USA, November 06 - 10, 2004).

[3] http://www.i-mate.com

[8] Rohs, Michael., Real-World Interaction with Camera-Phones, 2nd International Symposium on Ubiquitous Computing Systems (UCS 2004), Tokyo, Japan, November 2004.

[9] Ben Salem, Matthias Rauterberg (Technische Universiteit Eindhoven, Aesthetics as a key dimension for ubiquitous entertainment, in The 2nd International Workshop on Ubiquitous Home, Kyoto, Japan.

[10] Mark Weiser, ”The Computer for the 21st Century”, Scientific American, 265(3):94-104, September 1991