Incidental interactions and large screen displays.

Garnham, Darius

Incidental Interactions and Large Screen Displays

Darius Garnham

Submitted to the University of Wales in
fulfilment of the requirements for the Degree of Master of Philosophy

Swansea University
Prifysgol Abertawe

Department of Computer Science
Swansea University

September 2009

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

Information and digital media surround us in many forms, but how do we discover this hidden gem of knowledge, and how can we present it in a form which is recognisable?

Every day we carry out various activities. Some are mundane, some are exciting. While we carry out these simple motions, we provide cues and information about such activities. The question which arises is how can we discover, and use these subtle actions which are carried out multiple times a day.

In this thesis we explore methods and techniques tailored to reveal the notion of incidental interaction, with the idea of incidental information which is used within such actions and how it relates to large screen public displays. We consider how these actions can be applied within the context of community communication. This has lead to the development of a visualisation system and a detection algorithm to record actions. We also consider information sourced from mobile devices and how it can be used with these displays.

Through the use of field studies we have developed prototype systems to demonstrate the feasibility of these ideas. We extend a digital video repository by incorporating computer vision techniques to allow us to explore detected people present. We also consider the novelty that is brought by repurposing mobile activities, within different locations.

Our experiment shows that incidental interaction and the incidental information it uses can be valuable as a communication tool and concept to support community information sharing. Parts of the results presented within this thesis have been published in [Garnham, 2008a, Garnham, 2008b].

Declaration

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed . ..... (candidate)

Date ....14/5/10.....

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .. .. (candidate)

Date ....14/5/10.....

Statement 2

I hereby give my consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ... . (candidate)

Date ....14/5/10.....

Acknowledgements

I wish to thank my parents and family who have always believed in me. They gave me strength and supported me in so many ways without which this project would not have come true. For all the family members that are no longer with us, they are not forgotten.

I would like to thank my supervisor Dr. Matt Jones for his close and continuous guidance, support, and friendship throughout this work. I am very grateful to him for advising me that continuing my education after my degree was the best choice. I also thank my second supervisor, Dr. George Buchannan for his kind support during this time.

I would also like to extend my appreciation to my colleagues within the *Fit Lab* group for their support and encouragement. Without them, the learning process would have been much tougher and more annoying.

I thank Swansea University's Department of Computer Science as a whole for providing me with the opportunity to pursue this work. I would also like to thank all the staff and postgraduates who have made this experience very enjoyable.

Dedicated to my father, Bryan Garnham

Table of Contents

List of Figures

List of Tables

Glossary

Notation	Description	Page List
2D	Two dimensional, having two dimensional geometry, which can be described by the (x,y) coordinate system.	20
3D	Three dimensional, having three dimensional geometry, which can be described by the (x,y,z) coordinate system.	19
A.I.	Artificial Intelligence	24
AdaBoost	Adaptive Boosting, is a machine learning algorithm, which can be used in conjunction with many other learning algorithms to help improve their performance.	94
API	Application programming interface	63
AR	Augmented Reality	105
BBC	British Broadcasting Corporation	62
Blob Detection	Is a technique for detecting points and/or regions in the image that are either brighter or darker than the surrounding parts.	98
Blogging	Is the recording of one's actions, in the form of a commentary, descriptions of events, or as composition of graphics or videos on a public website.	2
Bluetooth	A open wireless protocol for exchanging data over short distances from fixed and mobile devices	20
CCTV	Closed Circuit Television	24
Computer Vision	Is the study of methods which allow computers to obtain information from images.	24
CSCW	Computer supported cooperative work	1
DMA	Difference Motion Analysis	89
EPSRC	Engineering and Physical Sciences Research Council	11

Notation	Description	Page List
Ethnography	Is the study of groups or cultures over a period of time.	34
Face Detection	It is a computer technology that is used to identify the locations and sizes of human faces in digital images.	18
GB	Gigabyte	39
GHz	Gigahertz, One billion hertz	39
GIS	Geographic Information System	47
GPRS	General packet radio service	21
GPS	Global Positioning System	23
GUI	Graphical User Interface	19
HCI	Human Computer Interaction	9
HMM	Hidden Markov Models	82
HTTP	Hypertext transfer protocol	28
Instant Messenger Clients	Communication software which provides a form of real-time communication between two or more people based on typed text, but can also include video and sound.	16
IT	Information Technology	23
JDBC	Java Database Connector	63
LSPD	Large Screen Public Displays	6
Machine Learning	The process by which computer systems can be directed to improve their performance over time.	82
Machine Vision (MV)	Is the application of computer vision to industry and manufacturing domain.	24
MMS	Multi Media Messaging, is an extension of SMS which allows the transmission of pictures, sound or short low quality video.	15
Mobile Devices	Any device which is small form factor e.g. Mobile phones, PDA's, netbooks	1
MUI	Multi-user Interfaces	53

Notation	Description	Page List
MySQL	Is a open source relational database management system	63
OpenCV	Open Source Computer Vision, is a library of programming functions which can be used for real time computer vision	83
RAM	Random Access Memory	39
RFID	Radio Frequency Identification	22
SMS	Short Message Service	28
SP3	Service Pack 3	39
STM	Short Term Memory	41
SUS	System Usability Score	69
Telematics	Is the science of sending, receiving and storing information via telecommunication devices.	9
Texting	Text messaging, refers to the exchange of short written messages between mobile phones, over cellular networks.	2
Tweets	Tweets are text-based posts of text, used as communication between users of twitter	2
Ubiquity	Seeming to be everywhere at once	5
UI	User Interface	26
URL	Uniform Resource Locator, which is the address of a file or content on the internet	21
VR	Virtual Reality	50
WAP	Wireless Application Protocol	28
Wi-Fi	Refers to any wireless local area network products based on the IEEE 802.11 standards.	20
Widget	A widget is an element of a graphical user interface that displays an information arrangement changeable by the user.	19
WIMP	Windows, Icons, Menus, Pointing Device	18

Chapter 1

Introduction

Contents

Public information has always been of interest to the general public and government agencies [Johnstone, 2004]. The need to find out more information is just human curiosity. From that of the mundane local more personal level, for example this can be to find out what your neighbour is doing, what is happening in the street. This can continue onwards to the higher more public level, such as the country's or world news. This has been exacerbated by the internet and mobile devices, which facilitate and promote a quicker, mobile and more personalised form of information retrieval than was previously obtained.

Large screen displays have surrounded us since the late 20th century; from supermarkets as a display form of advertisement to large bulletin boards in Time Square. The difference now is that they are becoming more pervasive in public spaces around us due to better quality, size, resolution, and lower costs. This has resulted in them being an attractive multi-purpose commodity in public spaces, and allows for new forms of communication to become explored with the media.

Researchers have explored ways in which we can accommodate the need of the public for this information need, and have facilitated their ability to receive and obtain this through interactive displays. This has been from simple ambient objects to full interactive multi-touch screens allowing group activities to occur.

There has been a vast amount of research in the field of CSCW, which has considered the use of large screens displays in business locations such as offices [Greenberg and Rounding, 2001, McCarthy et al., 2001, Wichary et al., 2005, Fitzpatrick et al., 2002]. They consider how to facilitate workers with extra ways of sharing and collaborating with their colleagues and improve work productivity.

There has also been work which has taken this outside from laboratory and office settings, and has explored public spaces such as libraries, campuses and local villages [Koch, 2004, Mitchell, 2006, McDonald et al., 2008, Frohlich et al., 2009, Taylor et al., 2007]. They have explored the interaction techniques and impact that there work might have.

There has been research which has incorporated mobile devices with large screen displays, either as a form of control or as an extra screen [Ballagas et al., 2005b, Cheverst et al., 2005]. More interactive screens are starting to appear in public places such as bus stops and art galleries using wireless communication channels, allowing for a more personal and fun experience.

There is one primary concern which presents itself when placing media in public spaces; The information is normally singularly focused with a specific aim, be that of advertisement or general usage. The content which is displayed has either been controlled, such as being filtered for appropriateness by a controller, or has been left to the users of the system for self governance. When a moderator is used with such systems, this can result in restricting what content is represented, and can lead to possible limited interactions that could be achieved as a response.

Therefore we would like to explore what additional data can be obtained and how it can extend the current practices surrounding these displays, using this notion and application of this extra hidden information present.

The focus of the work is to look at the notion of incidental interaction. This is the idea of interactions which can take place or be initiated as a result of activities with electronic devices and displays. These activities can be associated or be around other unrelated activities than the current task at hand, with different aims and goals than this new perceived usage. These can be initiated either implicitly or explicitly when present at public large screen displays. This can range from being sensed by the display you pass, altering the information automatically oblivious to yourself, all the way to an interactive touch screen on the side of a building.

As an example, consider when we walk past a display; we gaze at it, for that fraction of time we assess its worth, considering do "I need this information?". This might be while we are carrying out another action such as texting on our phone or listening to music. The display of interest could sense your presence by technologies around you. For example, by using information such as your location, and the choice of music you were currently listening to, linked with formulated messages using such a system, like twitter. The tweets you upload to your local blogging site that has been combined with this extra information available to it, which it has detected, the display might then facilitate you in your future interactions that could occur.

Another example is that you are new to a location and wondering what to do on your visit, you pass a display, but unaware to yourself it had changed showing you new advertisement or recommendations that people have been searching for, related to your location on Google. Some of the words spark your interest, you have now changed your action as a result.

All these subtle interactions that are taking place in the foreground or periphery have the potential to provide us with extra information. Some of these subtle cues might activate a chain of events we are oblivious to. These can all be classified as some of incidental interaction that is some form of catalyst which brings to life other actions.

It is the communication between these actions which are of interest. What can be attained from the incidental information which is revealed from the interaction? What is the impact of revealing and sharing this information to a wider audience? What are the questions and guidelines that can be learnt for future exploration? One way of achieving this is through the visualisation of this new information to the user. By revealing this additional hidden information and exploring the impact which it brings to the situation. This allows for possible fostering of serendipity to occur revealed

from other activities applied and explored in new situations.

The use and application that incidental interaction offers to that of developing world contexts is that it can provide a helpful and accommodating facility to interaction. This will help the bridging of communication that can be obtained between display and mobile devices. We can help to close a divide of learning, from that of visual to literacy to visual representation of data. By exploring the use of existing technologies present at these locations we can consider extending the existing community traits.

Locations which are limited in electrical power provide an ideal location for asynchronous techniques to be explored. One example of this is the traditional story telling technique of spreading the word, and knowledge between communities in an area. Incidental interactions considers the general, higher picture of activities carried out within these communities and considers how, and in what ways we can provide and use the extra information provided.

This can be explored either through device oriented approaches, or by considering it in the wider context of device independent ideas incorporating the alteration of interaction practices.

In this thesis we explore how, what impact, and the possible value that can be obtained from repurposing mobile activities, placed in a public space. We are also interested in the possible influence that it has on that location in respect to community awareness and information intake. We extend this further by looking at how it is possible to sense presence at the display which could allow for dynamic alteration of the media being shown. This allows us to explore the notion of incidental interaction with respect to implicit and explicit interactions that could occur.

1.1 Research Contribution

We hypothesise that incidental interaction can alter, for the better and can provide additional benefits (e.g., such as useful and helpful information) to current activities. We will achieve this through explorative studies.

Below are the aims of the research:

- Explore and illustrate the notion of incidental interaction.
- Carry out usability studies to evaluate the usefulness of systems and to demonstrate the concept.

The contribution that this research brings is the findings, which includes a number of studies that have started to reveal the notion of incidental interaction. We explored this notion buy considering incidental information associated with the interaction between, mobile devices and large public screen displays and how it applies to public locations, fostering a sense of community.

Through the idea of technology probes we were able to deploy and evaluate systems which revealed the acceptance and possible usage of them. We explored the space of hidden information that lies within interactions, and ways in which we can reveal this to the general public through large public screen displays. To show the value and possible need that can be obtained from this hidden information, when applied it to that of information representation. More explicitly through the

capture of user data, we were able to develop and test an algorithm to acquire the possible hidden information which could exist.

We are demonstrating a form of information discovery and presentation to the user of a system. We show how one can interpret this notion, with in there daily activities, and the way it could enrich there current activities.

1.2 Thesis Structure

The rest of this thesis is structured as the following. Firstly we present the background in Chapter 2, detailing literature related to the project and any existing software. This will look at the forms of large screens, interaction techniques incorporated and the impact that they had on the locations which they reside.

Chapter 3 outlines the studies carried out and a description of the reason for the methodologies chosen. In Chapter 4 we describe one of the studies which considered reusing search queries on public displays. This involved creating and evaluating a visual collage system which was placed at three locations on a university campus. This study will consider the aspect of ``incidental information'', provided by mobile devices, conveyed through the use of ambient displays.

In Chapter 5, we describe a study which considered reusing a form of digital repository system on public displays, the ``Rhondda Lives''system, which was a cut down version based on ``StoryBank''[Jones et al., 2007b].. This involved creating and evaluating a system placed at three locations on a university campus with different content. This study will consider the aspect of ``incidental information'', provided by a community video repository, conveyed through the use of an interactive display. We will explore the view of users and items of interest which can be detected and captured, which in turn can be used for incidental information in public locations.

In Chapter 6, we consider visual processing as a technique to obtain detection at these displays. This will provide us with a way of sensing interactions taking place at these displays. The chapter will consider the development of an algorithm using information identified in chapter 5, and will result in the evaluation of the efficiency and possible applications of use.

Chapter 7 is the conclusion and future work that is associated to the project discussing an overall ending and any possible future routes that could be followed.

At the end in the appendix, there will be detailed any other additional information delivered/required in the project.

Chapter 2

Literature Review

Contents

In this chapter we will be covering the research which is related to my own work and which has inspired and guided this research. We will be interested with that of situated displays and the ways these can be utilised and combined to form ways of communication between user and display. We will be looking at the individual parts of interactivity, settings, interaction techniques and impact that can be associated with that of public large screen displays.

When reviewing the research which has been carried out on large public screen displays, we can see certain trends which have been explored. These have ranged from those of private related personal life, considering how we can use the displays which are present around them, to areas including the general public.

There has been a considerable amount of work which has propagated in the sub-field of computer supported cooperative work (CSCW), and the early work with video teleconferencing and other work which has explored and considered how displays can and have affected and effected work trends and productivity.

There has been other research that is more elaborate, with news grabbing ideas and locations using the latest technologies (such as multi-touch interfaces), which have considered more public, open environments such as museums, campuses and town centres.

With the ubiquity of mobile devices, people seem to be using these more as a multimedia communication device and tool. Research has been considering how these two separate technologies could be combined, and help promote in this voice between them.

This space and area is where our research can fit. We are considering the possible usage of technologies which surround these displays and which are with us everyday. By exploring the com-

munication which can be elicited from these technologies. This can hopefully contribute to the fostering and better understanding of the media in use, and explore the informational value that can be shared across these spaces.

We are trying to ascertain the value that this channel can provide. We want to see what the affective response of this media has on the social aspects of the local community. This would be explored by using technology probes which would help reveal these research questions and aspects.

We can group the work into subsections that consider the interactivity and settings which have resulted from such studies. In this literature review we will consider :

- The possible interactions that can take place.
- The settings of these installations.
- The techniques which have been used to elicit the research findings.
- The impact which can result from the media's content and the form of presentation to support community place.

In the next section we will discuss the interactivity related to types of displays that are currently available, and have explored possible usage scenarios.

2.1 Interactivity

In this section we will explore the interactivity of existing set-ups which have helped explore the ways in which we can create a dialogue between ourselves and LSPDs. This notion of interactivity, and that associated with current trends, considering the usage of LSPDs can be split into the following groups.

2.1.1 Research relating to ambient, non-interactive displays

In this section we will consider pervasive displays and set-ups which have explored the relaxed, and subtle ways in which information is revealed to the audience by means other than by lively and explicit ways. This could be imagined as more implicit requiring less conscious physical interaction. Along the lines of Mark Weiser's dream of ubiquitous computing
[Weiser, 1991, Weiser and Brown, 0924].

Displays can be of many shapes and forms, not just that of digital, in which we are interested in this body of research. They can range from items in the environment around a user which can represent and convey current information. This idea can be seen in some of the work that has been carried out in the sub-field of ambient and periphery displays.

There has been some work which has considered utilizing physical and often everyday objects. Weiser et al. [Weiser and Brown, 0924] use a tangible bit of dangling string exploring the idea of calm technology which engages both the centre and the periphery of our attention, and considers the fact it moves back and forth between the two.

The information perculator, "*was a new form of ambient display which was designed to give a rich medium of expression placed within an aesthetically pleasing decorative object* "[Heiner et al., 1999]. The display was formed by the controlled release of air bubbles rising up tubes of water. This resulted in a set of pixels which scrolled up the display allowing the visualisation of any small, black and white image to be displayed.

Within the idea of using objects around us in the location, Ishii et al. and Hiroshi et al. [Ishii et al., 1998, Hiroshi et al., 1998] considered initial explorations of peripheral awareness of digital information which they called bits in an augmented architectural space. Through the physical construction of various ambient displays they were able to explore information presented through subtle cues of sound, light and motion. In addition to this work there have been setups which have considered ambient fixtures, "Pinwheels"and the "Water Lamp"[Dahley et al., 1998].

They considered the potential of creating subtle and informative reactive environments using displays integrated within architectural spaces. Linked with this work, Ishii considered extending this notion of ambient media through the idea of tangible bits [Ishii and Ullmer, 1997]. "*The goal of tangible bits was to bridge the gaps both between cyberspace and the physical environment, as well as the foreground and background of human activities*". In their work they describe the idea of tangible user interfaces. These user interfaces "*would augment the real physical world by coupling digital information to everyday physical objects and environments*". They were ultimately concerned in "*seeking ways to turn each state of physical matter, not of only solid matter, but also liquids and gases within everyday architectural spaces into interfaces between people and digital information*".

"*Ambient displays are designed to work primarily in the periphery of a users awareness, moving to the centre of attention only when appropriate and desirable*"[Heiner et al., 1999]. Ambient displays are passive in nature and users glance when obtaining information from a display. The users do not use the displays as they would use computers, they perceive the displays and information displayed upon them. Such displays are inherently difficult to evaluate formally since they are designed not to distract the user.

Ambient displays can be a form of digital display, represented as objects which portray a sense of an information display. One such example is that of AuraOrb [Altosaar et al., 2006], which is a device that was created as an "*ambient notification appliance that deploys progressive turn taking techniques to minimise the possible notification disruptions*"when working at your computer. AuraOrb used "*social awareness cues, such as eye contact to detect the user's interest initially with an ambient light notification* ". Once touched the orb displays the waiting information. This type of display allows perceivance from the periphery of the user.

Holmquist et al. [Holmquist and Skog, 2003] have represented information through an electronic display considering how they can use them, in an art sense, to represent information. They have developed "*the concept of informative art as a way to integrate information visualisation in the everyday human environment*". Drawing ideas from work of the artists such as Dutch artist Piet Mondrian.

In InfoGallery [Gronbaek et al., 2006], they presented the information on a variety of surfaces in the library that were aesthetically attractive. They explored "*how to create attention about digital library resources in the physical library space and support serendipitous experiences of digital resources for visitors in the physical library*". This kind of serendipity was achieved through means

of informative art services which they created.

Mynatt et al. [Mynatt et al., 2001], describe the set up of two displays which allowed geographically separated family members to maintain awareness of each other. This was achieved through the use of digital family portraits on displays between them and their locations. The system allowed for the optional displaying of additional information such as weather to inform viewers of the display. The connecting of people which is in sense related to our work, that of foster community awareness. The initial presentation of the display, and the information present upon it are in a sense passive and ambient. This allowed for a connection to be formed between the relations and family members. It also allowed for the user to transition to an interactive state from which they could interact with the display by touch to access aditional features of the system.

In a system called Meeteetse [Brunette et al., 2005], they created a set of technologies which were *"designed to facilitate social well-being through place attachment"*. They considered how displays could benefit activities of the audience within current situations, more than was currently available. The developed components used were placed in the locations that of a house and a community centre. They found that usage of the multiple displays enhanced the social well being and lowered the barrier for participation in the community

InfoCanvas [Miller and Stasko, 2002] and Informative art, made use of dynamic pieces of electronic artwork to represent information in that appealed to the eyes. These displays were placed within a persons work environment or publicly displayed, which enabled at-a-glance information awareness to occur. They achieved this through the creation of an information collage displayed via a secondary monitor to users in a workplace.

Plaue et al. [Plaue et al., 2004], proposed the term *"peripheral display to describe peripheral awareness systems that may present multiple information items"*. They suggested that *"ambient displays describe peripheral awareness systems that generally convey only one piece of information"*. They evaluated the system of InfoCanvas by assessing how well people recalled information displayed upon it. They compared the performance with two other electronic information displays, *"a Web portal style and a text-based display. They found that participants noted and recalled significantly more information when presented by the InfoCanvas visualisation than by either of the other displays, despite having to learn the additional graphical representations employed by the InfoCanvas"*. This would suggest that picture representations would seem to connect more with the audience.

Plato's Cave [Marchese and Marchese, 2004], explores the situations with people that work in windowless offices, whereby they are deprived of the sight of their outside environment. They created an image stream installation within an office setting where a view of the surrounding neighbourhood is streamed from a camera, through a network, and projected as an ambient display on a blank wall in a hallway. This repurposing of captured footage of the outside allows the connection between places.

Barrington et al. [Barrington et al., 2006] consider how one could use music as a natural and flexible medium for ambient display, *"because it is considered as a desirable component of many human environments, public and private"*. They suggest that *"the relation of music to emotions suggests that it may be well-suited for communicating information about human affect"*. From their findings, it supports their hypothesis that musical effects are a promising method for ambient informational display. This work explored the use of audio effects as a means to construct a music based, ambient

display, exploring affective responses obtained from changing sampled music. They showed that their hypothesis, that musical effects, display systems can satisfy many of the important design criteria for ambient displays proposed in [Mankoff et al., 2003]. Barrington's work considers another avenue which could extend the possiblitlies of current ambient displays.

When evaluating the usage of ambient displays heuristics can present a quick and efficient way with few people to determine design problems. Jacob Nielsen's heuristics for interface design were not aimed at ambient displays and work has been carried out trying to resolve this. Mankoff et al. [Mankoff et al., 2003] take initial steps toward establishing ambient display design guidelines for evaluating ambient display usability. This extends those set out by Neilsen [Nielsen, 0924]. Neilsen's was focused on the users' task, where Mankoff suggested "*ambient displays are at the periphery of the users attention and main task, and have as primary goals relevancy of information, aesthetics, and ease of monitoring*"[Mankoff et al., 2003].

A more recent display that allows the user to perceive the information upon it, and which is more relevant to the types of ambient displays we are concerned with, is demonstrated in the reception area of Google headquarters in California. There resides a large public display which shows search terms which are entered by users across the world in real time, Live Query [LEE, 2002, Edwards, 0924]. It is with installations such as this which allow the user to become aware over time of them and the information they possess. They achieve a gradual awareness of the information presented. This can lead to an inferred understanding of the general thinking of the public around the world.

Commercially, large screen public displays are usually passive in nature. They are there presenting information to the public, this can be seen as an informative or directive device for some purpose associated with that establishment, or place of rest. Normally we place displays in a public or unused areas to brighten up the space, but how do we react in these locations to the new artefact which now inhabit the space, and how might we engage it to discover its splendour.

The idea of ambient displays can be applied to advertisement, as this offers potential revenue if the right sort of audience is obtained. One study which has considered the use of large public screens to do this is BluScreen [Rogers et al., 2007, Payne et al., 2006]. This was "*an experimental public advertisement system that was created to detect users through the presence of their Bluetooth enabled devices*". Through their bidding agent developed, they were able to build probabilistic models of both the behaviour of users who view the adverts and display relevant advertisements accordingly.

HCI research has been focusing primarily on foreground activity and neglecting the background [Buxton, 1995]. He describes a model which presents a unified framework for looking at human-computer interaction as well as technology mediated human-human interaction within telematics. However, subconsciously, people are constantly receiving in the periphery various information without attending to it explicitly. If anything unusual is noticed, it immediately comes to the centre of their attention.

In the next section we will consider work which has looked at a more explicit form of communication. This involves physical interaction on behalf of the audience.

2.1.2 Research relating to interactive large screen displays

In this section we will consider the ways in which large screen displays have been extended, and how we can turn these displays into a more interactive medium of communication. Moving on from that of ambient displays which can resemble an implicit sense of interactive display, we have interactive displays which incorporate and allow the audience to interact with the information present upon the screen in a more explicit style.

Most of the research suggests and proposes possible applications for interactive displays. This seems to be concerned and interested with that of advertisement and the entertainment domain. Interactive displays have originally been associated with use in museums and in entertainment parks, be it as a result of games produced by the games industry. Museums have been aimed more at kind of system, due to the need and want of providing a more engaging way of presenting information to the general public who explore the exhibits and collections they contain.

In a study carried out by Brignull et al. [Brignull and Rogers, 2003], they looked at the how people interacted with displays at social gatherings. From this they were able to identify three activity spaces from which they proposed interaction guidelines for. The three activity spaces were that of *"peripheral awareness activities, this was classified as typically eating, drinking and socialising elsewhere at the event. In general, people in these activity spaces are peripherally aware of the display's presence and do not know much about it"*. The next space was that of *"focal awareness activities. This was where people in these activity spaces are engaging in socialising activities associated with the display, for example talking about, gesturing to and watching the screen being used. In this stage they give the display more attention and learn more about it"*. The last space was *"direct interaction activities, in this activity space, an individual (or a group acting cooperatively) explicitly interact with the display"*. Some of these findings were also validated by our study [Garnham, 2008a].

One interesting aspect we found was that of the social affordance which Brignull called the 'honey-pot' effect. This is the *"progressive increase in the number of people in the immediate vicinity"* of the display, *"where people were creating a sociable 'buzz' in the area"*. Their work proposed some implications for design in respect to engaging people to interact with large public screen displays, such as *"that of encouraging people to cross the threshold from peripheral to that of focal awareness"*. One way they proposed of achieving this was *"to position the display near a traffic flow from which to draw a stream of people"*. This encouraged people to cross the thresholds to that of participation.

They described that *"participants need to be able to learn how to interact with the system vicariously, rather than be told or have to follow a set of instructions. They need to be able to simply walk up and use it, having watched others do the same. The interface needs to be clear to the person such that they are reassured that their interaction with it will be a low commitment activity, that will be quick to do and enjoyable"*.

Public displays can be used for the awareness of activities and co-ordination, providing information through the use of displays. This allows for us to engage an understanding of the situation in a group environment, and allows for the management of activities [Alan et al., 2005]. They carried out an exploration on the usage of placing screen displays upon office doors. This was carried out in a computing department and used displays as a form of notification to inform and update

occupants in the department of an occupants presence. They explored the possible avenue for supporting user interaction with situated public displays with the use of personal mobile phones using Bluetooth. This was under the EPSRC funded CASIDE project, which was concerned with investigating cooperative applications in situated display environments. Through the development of technology probes used to explore possible usage scenarios.

Moving on from the dream of ubiquitous computing, with relation to interactive large screens, Yvonne Rogers [Rogers, 2006], raises some interesting arguments against that of Weiser's dream, and talks about having the devices and systems be more proactive. This idea of pro-activity on behalf of the display device is quite an interesting proposition which is along the lines of our work. The use of incidental information can be considered as a catalyst for pro-activity.

2.1.3 Research relating to Incidental Interaction

There has been previous research which has considered low intentional and sensor based interactions. These have been classified under various names which umbrella, these less explicit actions such as calm interfaces and implicit interactions. These emphasise "*output that is non-intrusive and ecologically natural forms of input/control. In comparison, traditional interfaces have been concerned with controls and inputs of devices, low attention interfaces have been concerned with that of sensing and contextual interpretations*"[Dix, 2002].

There has been some interest within the wider field of HCI and mobility on that of incidental and implicit interaction [Dix, 2002, Jamie Lawrence, 2006, Newman et al., 1991, Gellersen et al., 1999]. Dix describes the idea of incidental interaction related to the focus on everyday activities whereby he describes how a move from traditional models of interaction would need to be considered.

Lawrence et al. [Jamie Lawrence, 2006] explored the class of HCI research not normally considered interesting. These were the interactions which are incidental rather than explicitly intentional. They discussed 2 projects which used the idea of co-prescence which involved sensing and logging of mobile phones within the loaction of large screen displays. They suggested "*a move from traditional modalities of the mobile phone and towards less attention-demanding, more autonomous, and more context-aware, interactions*".

Early work covered in the Pepys project [Newman et al., 1991] considerd the generation of autobiographies using automatic tracking. This can be considered as a form of incidental interaction. The staff wore a badge which detected their location within the building. At the end of the day the logs recorded were analysed and a personal diary for each person was produced. A person's purpose was to carry out everyday activities in the office, incidentally a diary entry was created and later revealed.

Another example of a project which has explored the facilitation of incidental interaction is that of the MediaCup [Gellersen et al., 1999, Beigl et al., 2001]. This was a base unit attached to a coffee mug, which detected and provided the movement, pressure, temperature and location of the cup. The use of the information gathered provideed some indication of the drinker's current activity and location which can be used for community awareness.

The interactions can be envisioned as the unintended uses or outcomes of encounters with interactive devices. They considered a more implicit way of interacting with the environment around

us and discussed this different way of interaction. They explored a more relaxed way that you wouldn't normally consider to be useful. They provide an interesting view and propose the notion of carrying things out with less effort, a kind of self-publicising of ideas that relies on co-presence. The inferring of actions, filtering of data and presentation of media catered to people present.

Incidental interaction is where an activity of a device is carried out as a result of some other activity without the user being aware of it. We extend this notion by also exploring the aspects of the interactions which could add additional value to the experience such as the media's context. We consider it within the scenario and considerations of developing world contexts.

2.1.4 Relationship to this thesis

In this section we will describe how our research fits into the existing literature and highlight the specific gap that it fills.

We have seen that ambient displays and the surrounding variations of that area have been concerned with helping extend our peripheral awareness. This considers bringing relevant information occasionally to the focus when needed. When we have this focus we can see there are various ways in which we can try and engage with this information. Previous research has either been focused on explicit needs of certain groups or perceived intended use.

Interactive displays have been concerned with personalised and small groups with localised activities. This has resulted in lots of special cases which reveal information about practise using technologies present and around these displays. Activities explored at these displays have been closely connected with the social setting and audience present and only give limited insights to those areas.

Our research tries to consider and builds upon previous work that has explored the idea of incidental interaction. We expand this notion to explore how it can not only be applied to the screen from which the user experiences the media, but also the media itself. We are interested in what types of media can be used and in what ways can we use them to effect the place at which we engage with the display? What are the possible ways in which the interaction can be altered and effected by this experience? We consider these questions from the perspective of mobile devices usage and the incorporation with large public screen displays.

This thesis and the contained body of research has considered the idea of incidental interaction. It has considered how we can combine technologies and interaction styles together. The combination of separated forms of communication helps explore the reuse of incidental information and how it fits in with the communication of community activities. Considering how the resulting effectiveness of using current situated and mobile technologies can be used together. This will consider the effect that reveals itself between mobile device activities and large screen usage in public spaces.

Our work considers both ambient and interactive displays. Using a technology probe field study approach we explore the ideas that combining mobile devices with a public display can present. The notion of incidental interaction is central to the thesis as we are interested in the possibilities it exhibits, and this will be detailed later how we achieve this exploratory idea.

In the next section we will explore the research which has considered the settings from which the

display studies have been located.

2.2 Settings of Studies

In this section we explore the effect that new situations exhibit when placed upon the chosen study locations. This has been split up further relating to certain situations and the questions they try to reveal.

When presenting information in a public location, we try to foster a sense of community awareness, and the acknowledgement of this information in the location needs to be appreciated and considered. This is especially needed when using peripheral displays as the information presented can influence and have repercussions.

The use of large screens in communities, and their use as shared user interfaces is not a new concept. Earlier work was revealed and discussed in the 1970s by Myron Kruger and his interactive environments [Krueger, 1991]. One such environment was that of VideoPlace [Krueger et al., 1985], where media was presented onto walls which allowed interaction through sensors on the floor and camera images recorded. Recent work has mainly focused on the support of collaboration between co-located or distributed users.

The work which has been carried out related to co-located support has considered interfaces which allow simultaneous use by more than one person, and to allow a large working area to fill the field of view of the user. Distributed systems have looked at the ways of connecting up two areas around the world, allowing a sense of community presence to be exchanged between them.

One such idea was that of Telemurals [Karahalios and Donath, 2004], this was a way to create public spaces for social gatherings. This was achieved by communicating abstract images with audio to act as a link between the disjointed places. Two dormitories were chosen for the creation of a shared communication link, where they explored the idea of social catalysts. The "*catalyst may be that of an experience, a common object like a sculpture or map, or a dramatic event*". They were influenced by William Whyte's work related to social triangulation [Whyte, 1988]. It was devised from his studies over sixteen years observing the workability and use of public spaces within New York City and other cities. Using time-lapse cameras, 35mm cameras, tele-photo lenses, and interviews, his group documented patterns of traffic and behaviour in selected public spaces.

Dourish et al. [Dourish and Bly, 1992], discuss the Portholes project, which considered a system for distributed awareness in a work group environment. Using the distribution and visualisation of camera video pictures allowing people to see what is going on surrounding them. They suggested that awareness across distance has meaning, it can lead positively toward communications and interactions and, most importantly, that it can contribute to a shared sense of community.

There has been a trend in investigating the use of these displays for publishing and matchmaking in communities, specifically addressing settings such as in conferences and exhibitions. There has been displays which have been created to address the issue of providing a common ground which allow a spark, leading to a possible start of a conversation. In the next section we will explore the work which has considered displays in public areas.

2.2.1 Research relating to public area displays

In this section we explore the work which has explicitly looked at the setting of public area displays. These have different aspects which have to be considered when looking at certain criteria when evaluating the work being done in contrast to that of limited access areas. We also have to consider the in between states which can occur with the notion of semi-public displays.

One of the problems with the settings of large screen displays is that of trying to engage the user to the displays itself. We can study the ways in which people can interact with a display. Research has considered this concern with which place can impose upon a display, and the engagement with which it receives. This was explored by McDonald et al. [McDonald et al., 2008]. They considered the setting of a conference, and explored the notion of "*supporting awareness in fluid social environments*". They highlighted three key aspects to consider these were "*control, context, and content, which are critical to how those technologies impact our social spaces and whether they are ultimately accepted or rejected*".

Congleton et al. [Congleton et al., 2008], building on from work such as Prospero and C4 (Context, Content, and Community Collage) [McCarthy et al., 2008], provided an interface to allow third-party developers to create new proactive display widgets, that allowed content to be visualised upon public large screens. They had concentrated on developing a framework which aimed to support the range of applications that are possible in the space of proactive displays. Previously each pro-active display application had to be constructed by hand for its purpose. "*A pro-active display can select content which is based on the set of users who have been detected nearby. The choices which are made in performing this mapping have a significant impact on the social interactions fostered by the display*"[Congleton et al., 2008].

There is a project which has been studying the possible usage and repurposing of large screens in an outdoor urban environment. They considered the notion of collective play which can be achieved through game involvement [O'Hara et al., 2008]. Games are normally played on a console or by individuals, but with the development of multi-player and the development of mobile gaming we are seeing a direction of some research considering the incorporation and use of large screens around us, to involve and make connections with people. Ballagas et al. [Ballagas et al., 2007], explored this idea through REXplorer, "*which was a mobile, pervasive spell-casting game designed for tourists*". "*The game used location sensing to create player encounters with spirits (historical figures) that were associated with historical buildings in an urban setting, using gestures created by waving around a mobile phone*"using camera-based motion estimation.

By exploring the idea which is presented to us using games, we can show how we can obtain a sense and feel for bringing people together using the excitement of the gaming experience. Using opportunities afforded by the mobility of mobile devices which allows for a range of settings to be explored. This is an interesting avenue to follow.

With the increase in physical forms of information such as notice boards becoming more digitised due to decreasing costs of displays and changeability, this has resulted in an influx of large screens being using around the world exploring possible usage scenarios, such as advertisement at bus stops. Projects are being created and funded, like that mentioned in the UrbanScreens concept developed by Mirjam Strupppek. This concept considers the use of screens in urban environments and how they are used [Struppek, 2007]. This ranges across many forms across rejuvenating towns

and cities, exploring possible games and activities which can take place.

One use of public screens for the community and public is that of CityWall [Peltonen et al., 2008]. It has been setup in a public building, allowing for the sharing and manipulation of images through a multi-touch screen, and the images transferred via e-mail or MMS texting. The items were selected and displayed based on their relevance to events in the city.

One project based in Lancaster, is the eCampus project [Storz et al., 2006a] which was a campus wide installation of networked displays, where several experiences have been made with displays at various locations in different contexts. Due to the location of some of the installations there was concern about the public nature, such as acceptance and privacy. One important observation that was made was that the quality of content is very important and that deployment and maintenance costs should not be underestimated.

Using the insights from eCampus, Muller et al. [Müller et al., 2007], considered "*the rate at which the information was updated, and used this as an important property to distinguish different kinds of information*". Their goal was to provide useful information for large user groups. They ran a 8 month study considering the idea of information chunks on a university campus providing useful information for the students and staff. Through the use of contextual enquiries with students and staff they "*gained models and categories for posters, noticeboards, stakeholders, and were able to deduce the respective information flow*".

From their results, an information flow model was formulated which was focused on how the information chunks were created, distributed and consumed in respect to the displays. They extended lessons learnt specific to situated public displays which covered, "*the rate of media update, identification of sources, filters and forwarders, the identification of actionables, changes and non-actionable information chunks and the location of and support planning for news and opportunism with reminder displays*".

We have considered research which has looked at the idea of an open environment, allowing more access to information and situations where the possible chance of interaction is greater. In the next section we will consider research that has focused on the settings with limited access.

2.2.2 Research relating to limited access, e.g., office-based displays

In this section we will discuss the work which has considered limited access with respect to the information displayed in more confined locations. The technologies which are present in the workplace or surrounding area have an effect on the ways in which people can interact with public displays.

In the setting of an office environment there are certain rules which have to be respected with regards to formality and presentation of oneself. This can be a little concealing and restrictive on the people located in this environment, and people seem to stick to themselves and don't end up communicating between social groups. There has been work which has considered this problem which we will now discuss.

The use of images as an ice breaker was considered by detecting and altering the information displayed at certain locations in an office environment. Their aim was to foster ways in the workplace

and worker for cohesion with in the group [McCarthy et al., 2008].

People like to have subtle communication channels available to them, be it via instant messenger clients, such as MSN messenger, ICQ, AOL to name a few. The exchange of information and communication between workers is crucial to try and foster work, and personal relationships.

People like to be notified of certain items of interest, related to what is happening around them. The Notification Collage [Greenberg and Rounding, 2001], was groupware created to explore the idea of notification and subtle ambient communication in a office environment. It allowed users to communicate over a network of desktop displays and large public situated displays. It allowed for research which considered the casual interactions that took place at these locations.

Kimura [Voida et al., 2002], offered additional features that provided an overview of the activities that a community of users undertook, as a form of shared awareness. Their research considered, "*to design an office that better supports knowledge workers and business professionals who interpret and transform information*". They "*augmented and integrated independent tools into a pervasive computing system that monitors the users interactions with the computer, an electronic whiteboard, and a variety of networked peripheral devices and data sources*".

Private and public information in a work environment is normally decided by the person, or set by the corporation from which the employees are governed. Personal media types are normally associated with some meaning or relevance to the user. There is a sense of ownership of the media provided. This can be provided and highlighted by the use of technological devices upon the person, such as the mobile phone. Muller et al. [Müller et al., 2007], were interested in maximising the utility of situated public displays. They devised an algorithm that took into consideration the presence of user and then customised a more personalised display of information to them.

When we consider corporate and work items, they are normally items which have been designated or provided by the company. There seems to be less sense of a personal ownership of the item, due to association with and the issue that the item has been provided by the place of employment.

Yamada et al. [Yamada et al., 2004], describe a system called "Yeti"which allowed for offices across locations and time zones to inform each other through digital content sharing. This was built upon the earlier work of Elizabeth Churchill on the Plasma Poster Network [Churchill et al., 2004a, Churchill et al.,]. The Plasma Posters used large screen digital displays acting as interactive poster boards. They were designed to allow "*informal content sharing within teams, groups, organizations and communities*". With "*use of the Plasma Posters, people would encounter digital content (e.g., text, Web pages, free form scribbles, images, movies) that had been posted by other community members, and could choose to engage with that content further*"[Nelson et al., 2004].

The concern of privacy can be related to any of the media which is presented upon a large public screen display. This brings in to contention the idea of governance, by where the persons themselves undertake a form of self censorship on the media and information presented. This could also be agreed upon by a collective group with the interest of the community at heart.

When one considers the place of a public display, which depending on location, time, reason and information provided can be sometimes classified as semi-public. Semi-public can be viewed in one sense as that of a small group of people rather than the larger audience. Huang et al. [Huang and Mynatt, 2003], explore this space by considering that of small, co-located groups in

a lab environment in which the members already possess some awareness of each others activities. ``*Compared to that of large public displays, loosely connected groups, designing semi-public displays mitigates typical problematic issues in sustaining relevant content for the display and minimizing privacy concerns that might occur*''.

2.2.3 Relationship to this thesis

In this thesis we consider the idea of creating a sense of community at these locations using the media present on mobile devices. We consider that of private and public situations which could highlight possible insights to acceptance and usage of incidental interactions that take place.

We have explored different types of settings which could be targeted for such studies. We have considered that of both ends of public displays research. That of public large displays which seem to becoming more prominent, and also that of large screens within limited access scenarios such as that of office based work. Our goal with this research is to expand upon these areas by exploring alternate locations from which these can take place. One such area was on the side of a building with the use of a projection display.

From the research described, the issue of setting has shown that some effect occurs on the possible interactions which take place at these displays. This can have a resulting effect on the uptake or appropriateness of media shown and the possible techniques where by the media is delivered. We offer a view of bridging between these separate extremes, that of mobile and situated settings. We propose by bringing in public activities into the public attention, we can build on and extend the idea of foster a sense of community.

Some research has considered the use of private media, others have looked to that of public. Some of the research has considered that of small groups and the locations, situations they inhabit. Not many have considered that of large groups of people present or surrounding the display itself. They have been concentrated upon a more personal or small groups. We are interested in exploring the interactions that are carried out in wider open spaces, by groups in the location of the screens.

We will consider aspects between that of large group participation and that of the smaller group activities and we will take under consideration how a wider, global, larger picture of interaction such as that employed through mobile activities shared across time and space, can be used in conjunction with that of a more close knit, limited space.

Exploring these two areas can offer the most potential for exploring the idea of incidental interactions. It will allow us to test the possibilities available in a developed world where we can test by exploring that of private/public areas more distinctively. We have to consider placement issues differently when considering undeveloped areas and developing contexts. Different responses may result as the display itself would change the style of interactions which could and already taking place. A more emotional attachment could be formed if it was placed for community gatherings and meetings.

In the next section we will explore interaction techniques which have been considered with public displays.

2.3 Interaction Techniques

One consideration when interacting with large screen displays is the location of the user relative to the display as this can affect the interaction that can take place. Research prototypes related to interactive large screen displays and some public art exhibits afford the use of direct interaction with the display surface itself. This is achieved through touch, such as touch screens using simple gestures and newer multi-touch systems, or have used pens as a device to interact with such as the technology associated with whiteboards.

This unfortunately leads to problems with scalability of users, as it normally requires a 1-to-1 mapping from input to screen which is not always appropriate for the situation. Early research that highlights some of these issues was considered with Tivoli [Pedersen et al., 1993]. The researchers considered the area of shared interactive surfaces, using a electronic whiteboard system which allowed for freehand input, which provided the sharing of drawings with others present at the meeting. Similar work of interacting with an interactive display system was, that of LiveBoard [Elrod et al., 1992]. This was developed to explore issues in "*supporting group meetings, presentations and remote collaboration*".

When we want to interact with objects that are around us, there are various modalities in which we can achieve this. The modalities such as touch, vision, and speech allow us to choose and select the appropriate way in which we engage with the object. By considering animals in the world and taking examples from evolution and human biology (physiology), we endeavour to replicate these examples and try to apply them to problems, adapting them to fit the needs present to us.

Multi-modal systems which incorporate the use of the various senses of the human have bean considered. A real time framework for multi-modal interaction was explored in the form of an interactive kiosk at a conference [Krahnstoever et al., 2002]. They considered the use of natural gestures and voice recognition when possible interaction could occur with a public display. This system was initialised by face detection of a user, which indicated presence before tracking hand and voice interpretation.

With the new ways presented to us by the pervasiveness of displays around us, we must try to incorporate these new aspects such as new technologies into our ecology of interaction techniques when one tries to engage with these displays. This can be explored by considering the individual components and looking at how they fit together.

Traditionally we have been accustomed to WIMP metaphor which was viewed in the earlier days as a more traditional way of interacting with computers credited to Douglas Engelbart and the human augmentation project in late 1960's. It became famous in the eighties by the Xerox Star system, from which all modern systems have evolved. The interaction has not much moved on as computers still require more traditional approaches to interaction with computers and these metaphors have been applied to multi-modal interactions across systems.

One such project which considered trying to use this metaphor as a way of exploring interaction with large screens was called Dynamo [Izadi et al., 2003]. They developed a system which allowed for collaboration on large interactive displays using a WIMP interface, where they considered the privacy and sharing issues within a classroom. This system looked at explicit control mechanisms to facilitate these interactions.

In a survey on interaction techniques for large screen displays [Badillo et al., 2006], they explored the interaction challenges that were present with large displays. They considered issues such as tracking the cursor, large screen target acquisition, task management, GUI widget for large screen displays and the modality that touch screens present. They then proposed some open challenges which are still pertinent such as "*the limitations of existing techniques, limited use of 3D interaction techniques due to immature implemented technologies at the time. Such problems of accurate spatial positioning, with respect to screen in everyday devices*".

Other questions open are, "*that associated with software and interface design guidelines, which were based on traditional interaction techniques of mouse and keyboard. The integration of proposed techniques with that of current working systems, can be linked to domain specific techniques also, depending on the nature of task. It can result in different goals and outcomes from using the technology*".

Other types of technologies have been considered in the past when trying to interact with a display, such as lasers and sensor devices. These allow a single user or multiple users depending on task required, but in the past have shown not to be that successful, but with new techniques further research is being developed [Cheng and Pulo, 2003, Myers et al., 2002, Bi et al., 2005].

2.3.1 Mobile devices

There is a variety of previous work that impacts this programme of research. Firstly the interaction between mobile devices and large displays has been the subject of a number of research projects. These include the "*use of vision and motion sensors to enable sophisticated, rich synchronous interaction between mobiles and displays*"[1]
[Dachselt and Buchholz, 2009, Jeon et al., 2006, Ballagas et al., 2005a, Ballagas et al., 2006].

Dachselt et al. [Dachselt and Buchholz, 2009], introduce the concept of throwing gestures to "*transfer media documents and even running interfaces to a large display*"using sensor-enabled, gesture-based mobile phones. They used gesture interactions on accelerometer enabled mobile phones. They used "*continuous and stepwise tilt gestures which they mapped to a variety of elementary interactions typical for media-centred applications*". "*They also introduced a throw gesture to connect a mobile and distant display by transferring data and even running interfaces back and forth, thus facilitating a natural flow of interaction*".

Jeon et al. [Jeon et al., 2006], used motion flow captured from camera on phone to help detect movement. The idea used "*software markers on the interaction object or on the cursor for the various interactive tasks*". They introduced "*three interaction techniques for selection, translation, scaling and rotation of objects using a cameraequipped hand-held device such as a mobile phone or a PDA for large shared environments*".

Ballagas et al. [Ballagas et al., 2005a, Ballagas et al., 2006] introduce the two techniques, the "Point & Shoot"and "Sweep". This first technique allowed "*the user to select objects using visual codes, this allowed the set up of an absolute coordinate system on the display surface which was used as reference, instead of tagging individual objects on the screen*". The second technique

[1] [Garnham, 2008a]

mentioned enables "*the user to use the phone like an optical mouse with multiple degrees of freedom which allows interaction without having to point the camera at the display*".

Some of the earlier work carried out with dual device interface designs, was with PDA and a interaction television [Robertson et al., 1996]. They suggested some interesting guidelines which could be taken into consideration for developing and interacting between multi-deviced systems in the future. One of the guidelines was the "*distribution of information across the appropriate devices*". This takes in the consideration of the strengths, weakness and the forms of information that can be presented upon them. This links into another guideline, the "*combination of devices, so that the ensemble provides more than each independent device*". They proposed "*that value is added if applications on two (or more) devices are cooperative and complementary to each other*". They emphasised that the "*information on different devices is coordinated, consistent, and up to date*"not to allow potential for confusion.

With the distribution of mobile devices in the world, research is considering the various ways in which we can use them for interaction with large display screens emanating in spaces around us. Some of these devices use the concept of markers as a point of reference. Using these markers they can apply 2D graphical transformations, before refreshing the image displayed upon the large screen [Ballagas et al., 2005b, Jeon et al., 2006, Vartiainen et al., 2006]. This allows for people to use their mobile devices to control and manipulate media present upon the screen which might be of interest.

Taking the idea and use of the camera which is becoming part of the mobile device, research has considered the use of colour, light and the display qualities of the device itself to interact [Miyaoku et al., 2004]. By having the display detect the camera itself, identified by a coloured image portrayed from the device itself, they were able to carry out a study on the viability of this as a way of interacting with a large screen display.

When we look closer at the mobile phone devices you can see that they are becoming more sophisticated with additional technical features being added. With the introduction and acceptance of the Bluetooth protocol in [Mettala, 1999], this allowes for the sharing of data across this wireless communication channel. The hermes photo display [Cheverst et al., 2005], considered the use of this communication channel as a way of sharing photo's between phone and interactive display in a public corridor.

The use of the Bluetooth protocol has resulted in a lot research considering the use of this technology. This is not the only wireless communication protocol that has been considered. With the increase in bandwidth and range of the wifi protocols (801.11a,b,g,n), other means of activation and sharing of information are possible. Aspects associated with the linking of web 2.0 services provided across this protocol offer many possibilities. The range of smart phones that are now circulating allow for the uptake and possible use of this channel, once the initial housekeeping, that of connection, activation, setup, and verification have been done.

It is interesting to consider the use that the public could one day have with these interaction possibilities that can be envisioned. In ContentCascade [Raj et al., 2004], they looked at a mechanism that allowed for the implicit download of meta content at various levels of detail onto the user's personal device. With the users acceptance to receive the meta-information the system provides automatically by the framework via Bluetooth in the periphery the information shown. This allowed the user to carry on their current activities. The data which got downloaded contains information

about the content visible at the display when present e.g. location, time, url, thumbnail of poster and booking information. With this information, the user can either engage with it synchronously there and then, or asynchronous by retrieval of information at a later date.

The idea of snap and grab, which is the exchange and sharing of contextualised multi-media packages via your phone when at a display of interest via the Bluetooth protocol using camera of phone [Maunder et al., 2007]. To make the interaction of obtaining information at a display more natural, they considered how people take photos, and applied this to the problem of obtaining information from a display they were interested in. You see a part of the screen which was of interest, you then take a picture of it using your picture phone and then send via Bluetooth, the system then analysis the image sent and in return sends you information pertaining to that image.

Mobile devices have also started to be included as a means of extending the user's experience when selecting and retrieving information from advertisement boards
[Boring et al., 2007, Scheible and Ojala, 2005, Peltonen et al., 2007, Narayanaswami et al., 2008]. The Shoot & Copy technique discussed by Boring et al. [Boring et al., 2007], "*does not require visual codes that interfere with content shown or reduce the screens real estate*", normally "*possible through the augmentation with visual markers and the recognition of their visual patterns*". Instead their prototype allowed "*users to capture an arbitrary region of a standard desktop screen, containing icons, which represent pieces of data*". "*The captured image was then analysed and a reference to the corresponding data was sent back to the mobile phone*".

In MobileLennin [Scheible and Ojala, 2005], they considered using SMS as a way of voting for favourite songs presented on a large screen, which allowed for mutual agreement on the next video played. In CityWall [Peltonen et al., 2007], they constructed a display in a city centre, "*which users could send their own media content using mobile devices*". They had the "*goal of showing information of events happening in the city*". They observed user groups who used "*mobile phones with upload capability during two large-scale events happening in the city*". Their findings revealed "*that this kind of combined use that of personal mobile devices and of a large public display as a publishing forum, used collaboratively with other users, created an unique setting that extends the groups feeling of participation in the events*".

Narayanaswami et al. [Narayanaswami et al., 2008], introduce the idea of "*pervasive symbiotic advertising, which applies device symbiosis to advertising*". They suggest that "*device symbiosis advocates collaboration between personal mobile devices and shared facility devices to produce a richer user experience*". By using the information stored on and around the user the display can provide advertisement customised to that person. They talk about an "*ecosystem for symbiotic targeted advertisement*", with the use of zones which can apply the use of profiles for dissemination of information.

Before Bluetooth became incorporated with mobile phones, the way of communicating was over GPRS, infra-red channels or using SMS protocol and then receiving and adapting the result from the response.

With the wii-mote from Nintendo has spawned a new form of interaction for games using gestures. This device incorporates the approach of using gyroscope readings interpreted into comparable gestures for actions. This type has been considered when interacting with large screens. The obvious resemblance which this can be seen with is in that of computer games and in [136, 2008], where they have considered this type of input and applied it when interacting with a large screen.

Gestural interaction has also led onto the exploration of how we can use the techniques and apply it to metaphors when sharing media, such as tossing information from person to screen for display [Scheible et al., 2008]. A study into the use of gestures for interaction with large screen walls was explored in the field of design [Kela et al., 2006]. In their study they considered the feasibility issues, and compared the different modalities for the different task required, using a range of devices.

In the next part we will discuss techniques related to presence detection at large screen displays.

2.3.2 Presence detection techniques

Previous large screen displays that have allowed for interaction have been of the form of a walk up and use concept. The activation of the interaction cycle is initiated by the users who were present at the display via a device worn or by presence detected from external sensors, such as pressure sensors on the floor. This is the form of explicit detection of users, by activation of them being there.

Some work has been done on rendering information based on a group of users located nearby a display. Groupcast [McCarthy et al., 2001], MusicFX [McCarthy and Anagnost, 1998]. These incorporated the use of user profiles which were used for reference in the system. WebPendle [Villar et al., 2003], allowed users to define a set of keywords, describing their interests. When one or more people are detected via the connection made wirelessly between the 'Pendle' and screen at the location of the display, a set of web pages based on these user keywords would be displayed. The system developed tried to show a set of content which is relevant to that group of people. These systems still involve some form of physical device to operate, and initiate the interaction itself.

Surfaces intended for community use, are the *walk up and use* interactive boards [Russell and Gossweiler, 2001, Wichary et al., 2005]. Blueboard [Russell and Gossweiler, 2001, Russell et al., 2004], tries to incorporate the idea of separate activities and uses, into an idea of "LISA (Large Information Scale Appliance)". This incorporated the uses of various activities into one appliance. This allowed for personal and group sharing with the use of a RFID badge system with information that gets detected as authentication. It then turns the large screen display into a device that delivers personal information, such as personal messages and calendar information, held on a network to the authenticated and identified user. This information is provided in a kiosk environment and shares information with others via other mediums such as e-mail. MERboard [Russell et al., 2004] extends this idea, to allow multiple people using a group tabbing system but was more work focused. Wichar et al. [Wichary et al., 2005] considered how a coffee corner located in work environment allowed ways to inform current employees and visitors of other people present at the company.

In UBWALL [Sekiguchi et al., 2005] with the use of RFID tags in combination with motion detection which allowed for the identification and positioning of persons near the display. The information changed in response due to an individual tag associated with a user and motion detected.

Previous frameworks have tried to use this notion of detection and are exploring how we can use devices to communicate between each other. Using such a framework allows and explores the notion of shared community media within a public space, one such exploration was ProD

[Congleton et al., 2008].

These previous studies have explored with the notion of an external entity present using such items as badge and other explicitly made objects. Another form of presence detection which people already have, which is a less obtrusive device and more personalised one is that of a mobile phone rather than a foreign object.

The work carried out by Nathan Eagle in his PhD thesis [Eagle, 2005], considered how we could use the detection of mobile phones using mobile phone masts and triangulating their location. This idea could be applied to the detection and inference of their position in relation to the display screen. Linked to this is that newer phones 2006 have the features provided by GPS modules. Using this information, the user while in view of satellites in an open outdoor public location would be able to locate themselves and a proactive display would be able to alter information dependant on the audience present.

When you pass large screen displays in public areas there are certain levels at which we engage with the display. These possible stages of interaction when engaging with displays can be considered as a form of presence detection at a display, that of implicit to explicit [Vogel and Balakrishnan, 2004]. In their study they described and identified "*a set of design principles and developed an interaction model for publicly located ambient displays that seamlessly move the users from implicit interaction with public information to explicit interaction with their personal information*".

There were four phases in their model. They were "*Ambient display, which allows the user the ability to get a sense of the overall information space with a quick glance*". "*Implicit interaction, produced by the peripheral notification when a user passes by inferred from body position and orientation*". "*Subtle interaction phase where an implicit cue such as pausing for a moment by the user which results in the system entering the subtle interaction phase, from where a more detailed description of the notifications and/or the current state of the available public information is displayed*". Finally there was the "*Personal Interaction, this was the last phase where by the user can touch the information items for more details, including personal information*". The ability of sharing a large display through explicit space partitioning, using a motion tracking system providing a high resolution location and orientation data for the users head, body, right hand, and selected fingers. It tracks the user by placement of small wireless passive markers on body parts they wished to track.

Most of the systems explored required some degree of pre-enrolment which provided individualised formation of public information on the display. This takes time and pre-cognitive effort. Visual sensing could take away pre-effort and facilitate in additional value which could be added to existing systems e.g. More people an indication of either novelty or interest of some form.

2.3.3 Computer Vision

Most large screen public display research has focussed on developed world contexts. When we consider the developing worlds, initiatives are slowly being initiated that are helping introduce IT systems into these locations. The research carried out in these places considers ways in which electronic equipment could be used to facilitate the needs of the local population.

When we consider people who are present in these locations we can see that they are visually

literate and can understand symbols. They also have ways of interacting and communicating which has been evolved over time. There is a high learning curve which imposes itself upon users of such systems and it is a concern when introducing new technologies. This is where with the use of computer vision techniques we can help with the uptake and usage of systems by the local communities. People take photos of objects that are of interest to them and they then store them in a form, either mentally or physically. This provides a reminder of that object where by later they can add meaning to it. This link formed with the idea of using imagery as a method to help interaction between systems and user. It offers a way to lessen the burden with that of interacting with the system explicitly.

Mobile phones are becoming more prominent in rural areas around the world, whether it be for a shared community usage e.g. looking for jobs or as off casts donated by developed countries. By considering how we can combine and think of usage scenarios to help local communities this could provide interesting research. We consider possible usage for such ideas in our research.

When we consider computer illiterate users, what other forms of interaction can be considered to allow communication between them and theses systems. In the past, pictures have been used to tell stories and communicate meaning for centuries, so using this idea why can't we use the physical landscape of the user's interaction to help with their communication.

The idea of combining different areas of research such as computer vision, A.I. and using cameras for detection has been considered before [Crowley et al., 2000]. They looked at various ways in which machine visionmachine vision could be used for user interfaces. They describe some projects which use basic detection to augment the users experience. There has also been the consideration of using vision based tracking systems for large screens [Nakanishi et al., 2002]. They considered using 2 cameras, using 2 near-infrared lights to catch the eye gaze of user, which they then triangulate the response and represent the interaction upon the screen.

Kameda et al. [Kameda et al., 2004], considered discovering hidden information using CCTV to determine the way people take notice of displays. One such interest relating to displays is that of how many people are looking at a display while walking pass [Smith et al., 2006, Smith et al., 2008]. Tracking of people seems to be of importance for security reasons and this constitutes a field of its own, but some work has been considered on how to detect and trace a group of people's movement [Haritaoglu et al., 1999]. The actions of a group are of interest as this can propose different social norms and activities, than that of the individual person. The filtering of a video stream in tele-conference situations has also been considered [Boyle et al., 2000].

2.3.4 Relationship to this thesis

In this thesis we consider some of the interaction techniques which have permeated within the work of large screen displays. They techniques have been used for direct, explicit interactions with a system. We can relate these types to our second study in chapter 5, where we try to extrapolate value from images of groups of users. We explore another avenue which considers that of a more relaxed and implicit view of interaction. We were focusing on mobile phone interactions which could take place with these displays. We also considered the incorporation of vision and presence to aid in these possible interactions. As we described most of the research has considered that of

developed worlds, we were trying to explore how this could be considered and implemented within a developing world context.

To explore the idea of developing contexts we have developed a prototype system using media gathered for use in an in-situ environment. This allows us to explore the possible idea in real world situations to gauge the uptake and use of such systems before possible deployment in the wild. We are interested in the incidental information which can be obtained from and through the usage of possible interaction techniques while at the system, and away from it.

We address the idea of presenting mobile usage at one location and present it to new locations. Doing so can help explore serendipitous events which could occur as a result of providing new information provided by different means such as mobile devices. Computer vision techniques have been concerned with that of surveillance or as a means of recording activities for posting on a display. We address the idea of exploring what additional information can be obtained from a scene and how this incidental information can be used to augment the situation and experience with the screen.

With respect to modern LSPD's, mobile devices, presence of oneself, and the use of vision building on from human senses we can find research which falls under these headings and our line of research seems to follow these headings.

The counting of people present at a display is of interest to us and could allow for extra information to be deduced. Using a number of cameras and the triangulation using a view of the cameras placed research has considered one way of achieving this [Kilambi et al., 2008]. We can take inspiration from this and apply to a real world situation.

We are focusing on mobile phone interactions using vision and presence detection. Considering how this can be applied to developing world contexts.

2.4 Impact of Interactive Systems

When considering previous research with respect to public large screen displays, what are the impacts that these studies have explored? how does placing these displays at locations, impact and change the atmosphere of these social settings? Does the displayed media reveal anything new at these locations?

Media can have an impact upon the audience and it can change the situation and space. Without these explorations we can't get a better understanding of technology uptake and practises at these locations. There are many levels in which the impact can be considered, each can reveal insightful information pertaining to community and public usage of media placed upon large screen displays.

With these studies which have taken place, one concern is that of the impact that they bring and place upon the location where they reside. We now explore some of these questions within the research.

2.4.1 Community

One common focus of large display research has been concerned with the degree to which placing displays in public settings improves a "sense of community". *"It is hypothesised that public displays can foster social encounters and provide an enhanced awareness within the community"*[2] [Brignull and Rogers, 2003, Churchill et al., , Churchill and Nelson, , Antonietta Grasso,]. One focus of the research has been, augmenting notice boards found in communal areas, halls and foyers of organisations. Plasma poster [Churchill et al., , Churchill and Nelson,] and Community wall [Antonietta Grasso,] were designed to enable people to post and annotate information onto a large public displays available to a community of users.

The use of interactive displays allow for new and engaging experiences. It allows a way of presenting media to an audience around it. When we consider people and around the world who are less fortunate. How do they communicate with each other in these locations? Some of them are textually illiterate but have a full vocabulary through images and other artefacts. Using this notion of vocabulary formed by pictures and using mobile devices which are shared throughout a community and by incorporating this with large screen displays in a local village to foster a focal point that allows a community to share stories and ideas [Jones et al., 2007b, Frohlich et al., 2009]. Jones et al. in a small Indian village created a digital repository for sharing stories captured on phones with the local residents. The visualisation was created using a collage technique to present the various stories to them stored upon an interactive display.

Similar work which has considered placing community provided information in a local village. The media was used to reveal what impact it had on that community, the displaying of sharable information to others in a community. Taylor et al. [Taylor et al., 2007], created a "*simple photo gallery application on a interactive display and deployed it in a central social point in a small village which displayed user generated photos and videos*". They were interested in investigating how public displays could complement existing non-digital displays to deliver information and help "*foster a sense of identity and history*". Their novel approach of enabling villagers to take responsibility and ownership of their own categories appeared to be successful in enabling user generated content and allowed the encouraging participation by distributing control.

Earlier work that considered the supporting and dynamical exchange of information and experiences between the community of people living in historical cities of arts and culture, their local cultural resources and foreign visitors was that of the Campiello project [Grasso et al., 1999]. They considered creating tools and technologies which better "*connected the members of the communities (local inhabitants; past, present and future tourists; organizers of cultural events = cultural managers)*".

The project was based on "*classical ideas of community networks, empowering a community by giving them the means for communication and making their communication persistent and visible for other members of the community*". They considered personal recommendation, using a paper UI which brought a focus on bringing the system to the places people are, which were then taken back to a server spot for input and feedback to a system. This feedback was represented back on a community wall for feedback.

The demographic group of pensioners are of consideration when we think about large screens as

[2] Garnham(2008).

they seem to be provided by loved ones to help with their decreasing eye sight. In a study considering how we can use these displays to facilitate in helping the support of the elder community [Brunette et al., 2005], letting them feel a part of the community. How do we design for human contentedness that can occur between people in a community? In
[Agamanolis, 2004, Churchill et al., 2004b] they considered these aspects and the placing of setups in locations and the connection afforded from them.

2.4.2 Presentation of Media

The way information is presented upon displays is important and can have a great impact on the acceptance and continual use of a system. The types of media which are used to construct a montage of information to the user have to be considered and representation needs to be considered. What are the types of media which can be accessed and created using mobile devices in situations associated with large screen public displays?, and what are the extra properties that we can extend to make these types of data more useful. In this section we will discuss some of the types which can and have been considered with respect to LSPD's.

2.4.2.1 Digital media usage upon and surrounding LSPDs

When we consider the content that has previously been displayed upon LSPD's, much of the previous work has been concerned with that of explicitly authored media for the display in use. It has also been aimed at particular uses related to the location and social audience present.

This has ranged from systems within organisations, looking at task centred, synchronous, collaborative work, to systems considering aspects of information sharing within groups and communities with asynchronous communication with the retrieving of images and media from the internet.

Some of the media used has been obtained from collection sites such as flickr [Inc., 0612b] and YouTube [Inc., 0612a]. Once the information is uploaded it can be used as reference and additional media for tour guides in the location. The time period of study can range from short periods using games as input to test input technique such as [Rohs et al.,], which uses a jigsaw to test out selection demonstrating novelty effects, to longer periods revealing information sharing practises [Carter et al., 2004, Churchill et al.,]. Most of the displays have used pictures, videos, text and other files common to most uses as test media for exchange and selection on the screens.

One thing of interest related to this media is that it has been personalised in one form or another albeit of individual user or business origin.

In a survey which covered the emergence of multimedia usage in the USA and Europe circa 2005-2006, it was revealed that camera and game applications were the mostly used on phones [Verkasalo, 2006]. With this insight in what ways could we extend and incorporate this in use with large screen displays. Using this information one way in which we could evaluate services provided through large screen displays, could be to incorporate the aspects of gaming experience and use of camera techniques mentioned in interaction techniques section.

Much of the content has been explicitly authored for the display in use and has particular uses related to that location and social audience. The media used has been a combination of pictures,

videos, text and other media types for exchange and selection upon the screen. This has ranged from systems within organisations considering task centred, synchronous, and collaborative work to systems considering aspects of information sharing within social groups and communities with asynchronous communication with the retrieval of images and media from the internet.

2.4.2.2 Public Services and forums

There is a current trend in the uptake of social network technologies such as facebook [fac, 0924], myspace [Media, 0924] and twitter [twi, 0924]. There seems to be interest in knowing about your next door neighbour, and friends. This is a social and community affordance. This kind of information has to be presented with caution due to privacy issues which exist. The use of information such as this can be enlightening and interesting and can affect people's opinion within locations.

A framework was discussed and developed which allows for multiuser communication, and interaction via public communication displays which is called WebWall [Ferscha and Vogl, 2002]. This system provided web access via HTTP, e-mail, SMS and WAP, most of which are available protocols used on mobile phones these days. They "*separate the data from access and display technology, allowing them to provide an open, flexible, extensible architecture that offers instant access to Internet information sources on an ad-hoc basis*".

The notion of collective content which is regarded as "*commonly owned or collective rather than purely personal. Content becomes collective when it is jointly created and used*"[Olsson et al., 2008]. Olsson et al. consider this in a study where they considered 4 user groups and their use of media in respect of sharing media. They describe a "*new axis of communality which describes the contents level of collectively*". These were that of "*Collective - Personal which supplements the conventional sharing level axis of Private - Public*". They propose that "*collective content is not primarily a matter of ownership, but instead, it is more a matter of perception of communality*".

By using the wrongly identified or altered meta-data associated with a media item, could lead to miss communication between the user and the actual use from which it was created. Meta-data is data about data and it allows us using technologies and techniques to facilitate in a more personal and automated representation of dynamic information that can be presented upon displays.

2.4.3 Relationship to this thesis

In this part we considered the impact with which these displays can have upon a community, be it for good or bad. We have also seen that it is important that the appropriate media is considered. We have discussed the ways in which new media related to social networks is starting to have impact on current practises. We have explored issues relating to media which might be used upon these displays.

All these issues have implications for our work. They highlight considerations and issues which need to be addressed when undertaking this type of research. The idea of community is quite close to one of our themes running through this research. It allows us to explore and look at the bigger picture for uses and implications of this research. It grounds us within that of previous work, where we are concerned with that of the use of usable situated displays.

2.5 Summary

In the previous sections we have explored the depth of work which has been considered related to LSPD's. This has included the usage and ways in which we can extend these practises. There seems to be a opening for the research related to that of incidental interaction as this is a topic which needs further consideration to reveal its contribution to the overall feel and place within the ecology of large screen public display interaction.

We have given a detailed description of the different types and forms of previous research that has contributed to the field of large screen public displays. They have considered from a range of scales, that from large to small displays and that of public to private forms of information presented. We have looked at the various techniques which exist for interacting and communicating between mobile devices and displays. We then considered the setting in which displays have been situated and the effect which has resulted from them.

This review has taken in a holistic view of the work which has been carried out. It's not exhausted as there is research always being undertaken. This has resulted from studies considering situated, public large screen technologies and work that has influenced its conception.

We have highlighted the area of incidental interaction which the closet form is that similar to work relating to implicit interaction and ambient intelligence. We will now explore this idea further within this thesis.

In the next chapter we will describe an overview and methodology of the studies which have been carried out to explore this space in the research literature.

Chapter 3

Research Methods

Contents

In the last chapter, we reviewed the previous research which has been carried out. We highlighted areas where there are avenues for our research to fit within. In this chapter we will discuss the methodology that was used to explore these areas of research. This will take the form of exploring the notions and methods available, and how we used them to elicit our findings. The overall technique will be a combination of descriptive and evaluative research.

3.1 Overview

The aim of this work is to consider and carry out research highlighting some of the aspects which pertain to the notion of incidental interaction. This will incorporate the discovery, and exploration of the notion associated with that of incidental interactions and the idea of incidental information held within. We will consider the ways in which it can be referred to in the context of public large screen displays.

Using these notions we will consider them from the perspectives of ambient and interactive screen displays. We will be using methods that have been discussed from the perspective of mobile interaction design [Jones and Marsden, 2006]. This provides an ideal progression, as a communication technique between displays and produced human media. Some of these techniques have backgrounds in work related situations, and that of human sciences, such as sociology and psychology.

As you can see in figure 3.1, we are exploring the combining of the two separate areas of research. The first area has been that constituted within situated display research. This has been concerned with how displays at locations can be used with considered information. The second area of research has been that concerned the use of mobile phone interactions and how they can be used with large screen displays. They have both been concerned with their own individual goals, this being

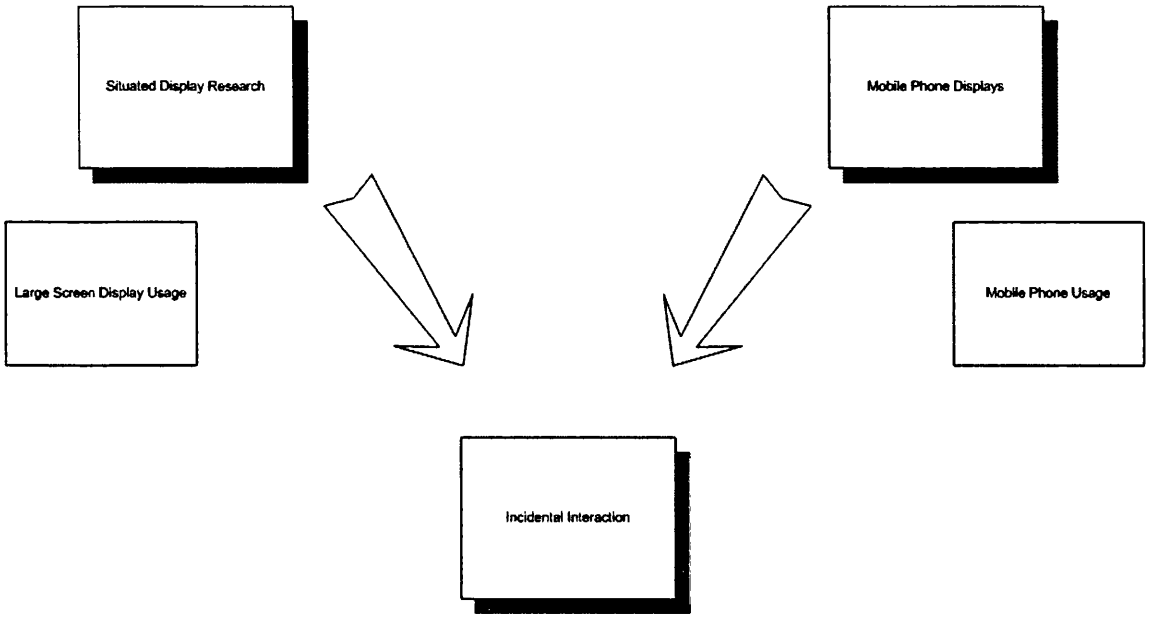


Figure 3.1: Formulation of research areas

interaction technique or specific information domains and topics. We look at the gap provided between them and consider combining the two individual aspects, exploring how this might benefit the interactions taking place.

We extend the idea shown in figure 3.1, by considering the components, which constitute to two separate technology probe field studies undertaken. This is shown in figure 3.2. As can be seen we split up the types of interactions and explore them separately under the notion of incidental information.

In technology probe field study 1, we explore the idea of non co-located media used. In this sense we use data that was obtained from a different location, time and place. In our case this is the mobile queries, which were then presented upon an established public display in different locations. The queries were not located at same place. The media used was asynchronous as the information was re-purposed from another time, the media's purpose was not associated with the location it is now presented in.

In technology probe field study 2, we explore the idea of co-located media used. In this sense we used data purposed for that location, and data recorded from the device was logged within that location. This was achieved synchronously through capture of logging data associated with media items selected and viewed. This could also be classified asynchronously as the data recorded could impact the future usage of new media shown.

In technology probe field study 2, we identify possible items of interest which could be used to add and be used as incidental information in a scene. This lead us to consider ways in which we could obtain this information. In chapter 6 we explore a way of obtaining the incidental information.

By considering these two studies and their individual attributes, we are better able to explore and show how incidental information resides and can be obtained.

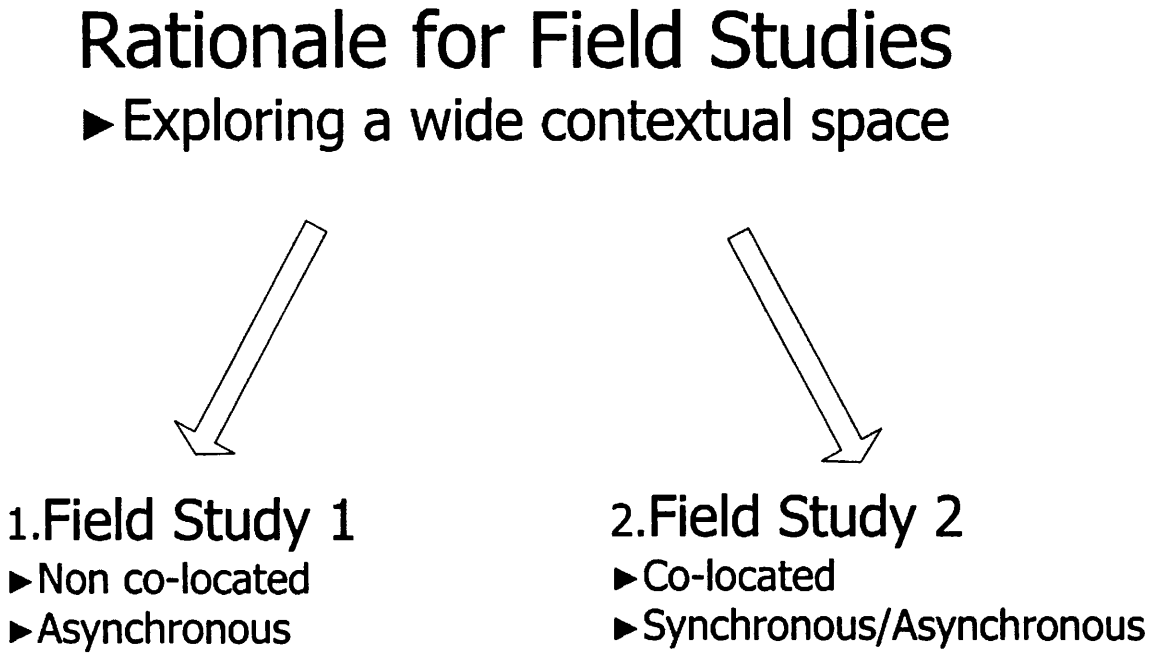


Figure 3.2: Rationale for studies

3.2 Methodology

For this research we carried out two technology probe field studies which will reveal and highlight informative details about interaction techniques and community acceptance within the idea of incidental information.

We chose to explore two types of interaction to explore this notion. It allowed us to prototype possible usage scenarios which could be undertaken in a developed world context, and allows us to consider how we could apply this to a developing world context.

The types of interaction considered were asynchronous, described in the first study (see chapter 4). Here we considered the use of recorded mobile activities, re-using the collected data with large screen displays. The second type was synchronous, described in the second study (see chapter 5), which was set up in a library setting. Were we recorded logs of the activities carried out on at a touch screen display. This incorporated images being obtained from a webcam and screen shots.

The two field studies which were undertaken considered the effective and affective impact on community awareness. These were undertaken in public locations and considered the use of user activities relating to a location. This incorporated the mobile search queries recorded on the devices present, at these locations. With this information we obtain a sense of local ambient community awareness.

The second field study furthers the interest with community awareness. This is achieved in one aspect by considering the display from which the information is presented upon. We consider how we can sense groups of people within the vicinity, implicitly with the use of computer vision techniques. With this captured data we will be able to gain and add value to the media being

presented. These two studies build upon each other, providing a base exploring the part and place incidental information has and can play upon public community awareness.

3.2.1 Technology Probes and Field Studies

For the proposed studies that will be used to explore the above mentioned notion, we will be undertaking a field study methodology incorporating the use of technology probes and altering where necessary.

Probes provide use with features that allow us to explore and capture data for a study. They allow us to capture artefacts, in our cases these are the responses of the users and logged data. They provide a way of recording account of the studied environment and participants, in our case the locations of the technology probes themselves.

One of the main reason for using probes is that they allow us to make the invisible visible. As described in Graham et al. [Graham et al., 2007] "*The act of participants engaging with Probes involves recording a point-of-view, while 'in-the-moment' and making visible, on one hand, particular actions, places, objects, people etc. and, on the other, wishes, desires, emotions and intentions.*", "*these momentary interactions may be logged and discussed later in the design process.*", "*This 'enforced visibility' is enabled by an artefact newly introduced into everyday life*", in our case technology probes themselves.

Probes also provide us with a way of disrupting "*the everyday practises of participants through enforcing an awareness and visibility of action previously absent*"described in Graham et al. [Graham et al., 2007], they believe "*that the 'disruption' that probes can bring, rather than being a concern, if responsibly addressed, may well be a strength: this 'disruption' is part of the 'working out' of how technology might affect them.*"

We consider the use of a probe method discussed in Hutchinson et al. [Hutchinson et al., 2003], where they define it as "*Technology probes are a particular type of probe that combine the social science goal of collecting information about the use and the users of the technology in a real-world setting, the engineering goal of field-testing the technology, and the design goal of inspiring users and designers to think of new kinds of technology to support their need and desires.*". This description of goals which this method tries to cover, is suited for the types of questions the field studies are trying to explore in relation to incidental information.

We use the idea of technology probes to explore the deployment in real use context's such as in the cafe, library and sports hall. We watch how it is used over the chosen period of time, and then reflect on the usage to gather information about the users and inspire ideas for new technologies.

We had developed and placed two technology probes, within the confines of the the two field studies. Both technology probes was placed at multiple locations in the field studies, described in detail in chapter 3 and 4 respectively. Each of the probes was designed to gather data about the surrounding audiences who viewed and interacted with the display and communication patterns while inspiring them to think about ways of communicating or using this new information presented.

For all the deployments, we wanted to be able to place the probes in high-traffic areas of possible users, where the users would hopefully look at them and use them. We were relativity successful

in doing this, but we had to respect the wishes of acceptance by them.

For our examples we are concerned with that of location relevance and the people who observe and what extra information can we deduced. The use of the technology probe field studies is to provoke a reasoned debate about this new area.

3.2.2 Supporting Methods

To compliment the methods chosen in the study, we incorporated evaluative methods such as questionnaires, which allowed for some quantitative data to be collected. We carried out contextual enquiries using in-situ interviews to gain some qualitative feedback related to their user experience of the system. We undertook some natural observations considering direct and indirect variations allowing us to get a sense of local environment. These methods are discussed in [Jones and Marsden, 2006].

One type of method which we will be considering is that of the ``Wizard-of-Oz''design approach, developed by Gould et al [Gould et al., 1983] and used by others such as Höysniemi et al. [Höysniemi et al., 2004], and White [White and Lutters, 2003] in their work. It is a form of interaction prototyping. This is where you take the place of the interaction system itself and act out the responses which would be presented to the user. In our studies we will use this on separate situations, one where the equipment has become unavailable, and when creating a scenario for interactions when developing the detection algorithm.

We undertook logging through the creation and recording of visual artefacts and user interface actions which could lead to later log analysis. We considered the use of focus groups of interviewees with the screens as we were interested in finding out how group dynamics played a part in the overall interactions which took place.

The use of ``creatively engaging methods''[Jones and Marsden, 2006], was considered, and we applied that of a technology probes to explore new experiences in a location. Here the prototypes were installed and the people's behaviour and surroundings were studied. Using an ethnography-lite discount methods such as that of quick-and-dirty. This involved carrying out a series of short and focused explorations within the field.

It has been considered that this type of methodology, especially with mobile device studies is suitable with the various situations, and real life scenarios that need to be explored. Therefore we have chosen to use this method for allowing the collection, and study of the situations where incidental interaction can be accommodated and explored.

There are other methodologies which could be considered for use when exploring these types of situations. These methods are that of the laboratory experiments, which we deemed are not suitable for understanding how users work in the real world. This style is normally constrained to a laboratory environments which can be stale, and are not out there in the environment of action.

``Often quick informal evaluations are needed, in which ideas are tested out with users and where the setting is not important. For these studies paper and pencil sketches are often adequate. Such studies take place early in design or at the time when decisions about screen design are being made. This will be the case for when we were designing the interfaces for the software used and

adapted. The important thing is that the ideas are at an early formative stage and the purpose of the evaluation is to aid rapid feedback which can be used to improve them"[Preece et al., 1994].

The attraction of using these methods, was that it allowed use to get out of the laboratory environment which can and has proven to effect the responses associated with studies. It also provided us with the ability to utilise and explore public locations, considering the responses of different groups.

One problem we admit with this methodology was that of obtaining numbers for participation. This is a reoccurring instance and factor, which can be found when undertaking research. The time at which one of the studies was undertaken coincided with that of the holidays, and exam preparations. This had a resulting factor on uptake.

We decided on small scale field studies, which looked at separate techniques. These were ambient and interactive large screen public displays. The population for the studies was a combination of university residents and visitors to the location at the time of the studies.

From this population, a selected amount of participants detailed in the studies were randomly picked from that of the population to not bias results from local people present. This allowed for dismissal of bias as they were randomly chosen. We wanted to engage the local inhabitants of that location to see what effects resulted from the introduction of these systems into their environment. We decided to locate the systems in locations, which already had, and were using these displays. Some of the displays were underutilised, and we wanted to see if a new interest, and appreciation for it could reassert itself.

3.2.3 Ethics

Dealing with the issues of ethics for the studies, we explained the experiment and asked them if they were willing to carry out such activities. By their initiated interest, acceptance and verbal agreement we then allowed them to carry out the experiment. Each participant was provided with a consent form which detailed what was required of them and provided them the option to stop the study at any time. The forms and additional documents are detailed in Appendix 8.4 - Ethics, section 8.4.1. Details were changed pertaining to the study it was used in.

For that obtaining to objectivity of the experiment, we carried out the same questionnaires and method in each new location for which the experiments took place. This allowed for consistency across spaces, with the information presentation and interaction techniques explored. We use the triangulation of various separate methods in the study to add a view of checking the results. Using the cross examination of results to validate our findings.

As discussed above there are a variety of methods which have been chosen and covered to facilitate in the discovery of information when considering incidental interaction. This selection allows for both quantitative and qualitative data to be revealed, which in turn helps to reinforce our notion. In the next section we will highlight some of the studies further.

3.3 Summary of the Studies

In this section we will give a short description of the studies, pointing out how they will reinforce our notion of incidental interaction.

Reusing Search Queries on Large Public Displays To explore the idea of incidental information, we proposed to use mobile activities captured on a mobile device recorded from another study. We then presented the data captured upon large screen displays, at public locations within a university campus to obtain the value with which they can uncover at these locations.

As we were exploring the notion of incidental information, we wanted to explore a way in which we could test this out. The way we achieved this was in the form of an ambient display which would allow people which had become accustomed to these positioned large screen public displays to experience this media.

We wanted to reveal what happened as a result from the introduction of this information. This involved observational activities combined with in-situ interviews. These combine as survey methods to illicit value and opinions from the people present at the display. This gave us insight into the current practices, values and view of the people surrounding and interacting with these displays.

Sensing Interactions & Visual Processing Building on from the insights revealed from the previous study, we started to explore what additional ways in which we could allow interaction to take place. We were also interested in the way the audience digested the information present on the screen. We were interested in detecting the number of people that were present at the display. We wanted to consider what extra information could be obtained to help add value to the dynamic media, present upon the screen.

To explore this, we undertook two separate parts of the study. For the first part we carried out image use and image capture of people present at an interactive display with the use of groups, which included interviews. This allowed for us to identify what was obtainable, in the images and provided some logged data about media interaction. For the second part we considered the development and testing of a detection algorithm which would be used to estimate the amount of people present at a display. For the evaluation of the algorithm we carried out simple experiments. Part one allowed for some early user feedback and got us the number of people needed to validate its effectiveness, and efficiency in part two.

We undertook a Wizard-of-Oz design approach when considering the initial testing of the detection algorithm which was being evaluated. The participants were given a scenario to envision, and were told to carry it out while interacting with the display while images were recorded. For the second part of this study we extended an existing collage system [Jones et al., 2007b]. Additional interface parts were added which incorporated our detection algorithm. We used a camera to obtain visualisations of the people, which was used to run against and test our algorithm against.

We also incorporated a logging system which recorded the media, and the various interactions which took place at the display. We logged the times of interactions, the media selected and played, metadata describing type, name and file based nuances.

We used a camera to acquire images of people present at the display. This was then stored as a scene image of people present at the display within the logged data. We also captured a screen shot of the interface itself, to see what was present at the time of interaction. This could be used as a cross reference to validate the interactions taken place for verification later on.

We are using this as an exemplar way of exhibiting what could possibly be achieved by considering the usage of computer vision techniques within that of detection of users of this kind of system. This considers the ways in which we can use incidental information for purposes to elicit new forms of action at those locations.

In the next chapter we will take a closer look at the study concerned with incidental information.

Chapter 4

Incidental Information with Public Spaces

Contents

In the last chapter we discussed the methodology with which we aim to carry out the research. In this chapter we will be exploring the notion of incidental information, through incidental interactions that occur, within the retrieval of information at a glance. We explore the effect of combining separate activities that are carried out on mobiles, with that of public, large screen displays, and the benefits they can bring to the viewers, when displaying information. Part of the work in the chapter has been published in [Garnham, 2008a].

4.1 Incidental Information

We are surrounded by little parts of information which are presented to us in textual, visual, or auditory form. We focus on the forms which can be visualised, upon and between mobile and large screen displays.

We define incidental information as information that arises from some other use or activity, which is not explicitly created for the task at hand, but can contribute to the overall activity being carried out at a particular time.

The information can reside across time and space. We are interested in the asynchronous possibilities, that present themselves when incorporating the use of mobile devices and large screen displays within this study. We can also consider that of synchronous, real time interactions which could take place, and the connection that it can bring to a place and it inhabitants. We will explore this in the next chapter.

Incidental information provides many opportunities which are present in everyday activities around us. For example, when we are shopping in a shopping centre, and a sign with an advertisement grabs our attention. Another example is that of information searching for information on the internet. While we are searching on the internet, as a result of information seen we are spurred on to another activity as a result. It is information you don't realise you can use, until it is presented to you.

4.2 Field Study - Incidental Mobile Information and Large Public Displays

``Mobiles are pervasive and large displays are becoming so. There is a clear and timely interest in extending the approaches and thinking regarding the interaction of these two platforms. Our focus, here, is to consider two elements, first, asynchronous combination of mobile and large-screen use, decoupling the two experiences (and potentially the user groups), and second, the form of the content displayed on the large screen, moving from explicitly authored material to the notion of 'incidental information', that is, information that arises from some other use or activity. Specifically, we take mobile search queries captured during previous studies and re-purpose them for large screen display [Arter et al., 2007, Jones et al., 2007a]''[Garnham, 2008a]. This work also considers the effect of the location of the large display on the utility and perception of the information.

This study contrasts with most studies of interaction with large screen public displays, by focusing on asynchronous, incidental interaction and how it can communicate a sense of place.

4.2.1 Aim

The aim of the study was to consider the value and types of reactions that may be provoked by unanticipated incidental information. Furthermore we wanted to consider how responses might differ in different locations. Finally we probed the possibilities of people interacting with the display via mobiles and other technologies.

The purpose of the probe was to explore whether other people's queries while mobile might be of use intrigue or indeed be a distraction to the large display viewers.

4.2.2 Methodology

In this section we will describe the methods used and the way in which we collected the data. We will also discuss the equipment used and the settings chosen for the study.

4.2.2.1 Procedure

The study involved the observation and questioning of 36 participants. The participants were chosen at random from people present in the location. This allowed us to gain feedback from a significant size of participants who viewed the display. We decided that the choice of study size we

considered was sufficient for the study. The number of participants was split up further, as shown in table 4.1.

Location	Sex		Total
	Male	Female	
Walkway (Outside Projection)	3	5	8
Theatre Cafe	12	10	22
Laboratory Communal Area	5	1	6

Table 4.1: The number of user, at the various locations

Included within this study was the interviewing in detail of 20 people from the cafe setting; 5 from outside and 2 inside the lab. There was a range of backgrounds, from student to business people to retired persons. This allowed for a broad diversity of the target audience and allowed the expression of different views and experiences.

With the size of our groups, we can only generalize for groups of similar locations, this being university settings. Different environmental settings could provide us with different insights.

Some of the participants refused to being interviewed in detail due to personal preference; therefore we had to obey their wishes. This resulted in the aforementioned numbers when compared with the total number of participants detailed in table 4.1.

The time allocation for the study is shown in table 4.2. With the length of the chosen locations, we carried the observations over a one week period. We first let the display run for at least 5 minutes, to allow the audience to get acquainted with it before we began to question them about it.

Location	Duration
Walkway (Outside Projection)	3 hours (6 - 9pm)
Theatre Cafe	4 hours (11am - 3pm)
Laboratory Communal Area	2 days

Table 4.2: Time Allocation for study at locations

We used various study methods to gain feedback from the areas of interest. These ranged from simple field observations in public places (naturalistic observations), to full interviews. Observations lasted for the duration of the studies, and included 30 minutes before and after the allotted times.The questionnaire had a total of eleven questions, we also provided an extra space for any additional feedback. See Appendix 8.1.1, for the questionnaire used. A discussion of the results is shown in 4.2.3.5.

We carried out the study in the following order of steps :-

- Observation of users responses to display
- Introduction of interviewer and Study Background
- Interview Questions

By following this order of steps each time, we were able to keep a consistency across locations. Initially we were observing peoples views of the screen, taking a bystander view point when ob-

serving them. This allowed us to gather a local awareness surrounding the social acceptance of the display. We also considered the use of anecdotal evidence, from which we could deduce whether people talked and eye witnessed engagement, visually or verbally of the system. Using subtle cues from body suggestions.

For the locations such as in the theatre's cafe, it involved us being present to observe the actions of people present. Outside in the walkway, we would walk by the viewers, and sometimes sit in the sun opposite the location before interviewing them about their observations. In the coffee break room, due to the area being small and the screen being located in the entrance corner we took observations of the visible actions as they entered and left the area in respect to the display, before informally questioning them later. We tried to blend in with the surroundings first.

We observed the reactions people associated with the changing words upon the display as they were fading in/out at the locations. We carried out a combination of informal interviews with people (asking about their opinions about the display, e.g. ``can any topics be seen?, do the words reveal anything, what interests you on the screen etc.'') and direct questioning in the form of a structured interview.

After we had observed for a period of time, we then introduced ourselves to the interviewees and introduced them to the study which we were carrying out. We then questioned them if they consented to taking part.

When we questioned the participants, the questions ranged from simple observational questions such as ``could they see the text?'', ``did they notice it?'', ``did they know what it was trying to show?''. This allowed us to gauge if they were actually interested and willing to engage with the display itself. We also considered the possible interaction selection and their preferences for future interaction. Questions such as ``how would they see themselves selecting, interacting with it in the future?'', ``would they think of using a mobile with it or any device if they could ?''. The results from these questions asked 8.1.1 are shown in the sections 4.2.3, 4.2.3.5.

4.2.2.2 Apparatus

In this section, the apparatus for the study will be discussed. This has been split up into hardware and software.

Hardware For the study we used various public large screens, from which we displayed the collage display upon. These consisted of a two 40 inch monitors. One was based in the laboratory common room, shown in figure 4.7 and one in the theatres cafe, shown in figure 4.6. There was also a projector system used for projection on the side of the building, this is shown in 4.8. The displays themselves were established fixtures at the chosen locations.

The screens used in the study were connected to an external laptop for the visualisation system. The laptop consisted of a Pentium 4 3GHz single processor, with 1.87 GB of RAM, running Windows XP Professional 32 bit, SP3. As we did not require the wiring up of additional input accessories, this specification fitted our needs for the running software outputted upon the screens.

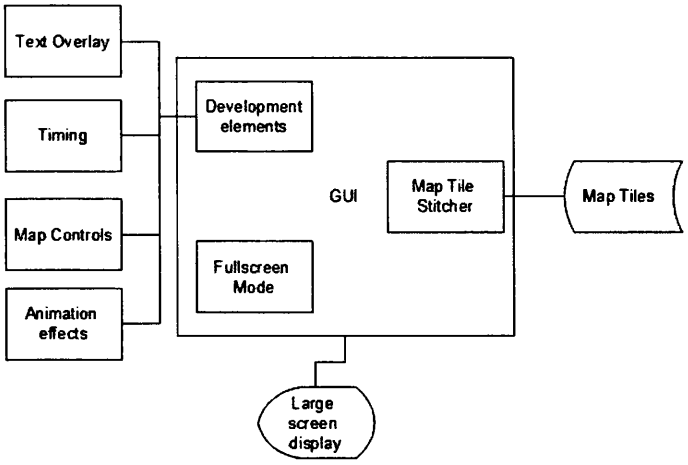


Figure 4.1: Query visualiser GUI components

Software The software we used for the system was created using the C# programming language and the Microsoft .NET framework. This allowed for quick development and prototyping to occur before finalising on the end layout. The overall layout of the interconnecting components which make up this software is shown in figure 4.1. The corresponding developmental elements are represented in figure 4.2

There are three main parts which communicate with each other to present the visualisation to the user. The map tile stitcher part constructs the overall image of the map using the downloaded map tiles. The development elements part, contains the customised items which are used for the running and visuals of the system. The full screen mode, deals with initialisation and presentation of the complete layout, set up by the options selected from the on development and screen elements.

The system was created to be displayed full screen, to allow the maximum screen space for the information to be displayed. It also provided a way that did not distract the user/audience from the display, by the association and resemblance of accustomed taskbar and other widgets associated with the desktop personal computer.

The initial screen provides the features to allow you to choose the time delay and aesthetics of the display. We used these features to find which colours and sizes were effective on the audience and locations. The initial setup screen is shown in figure 4.2.

The interface shown in figure 4.3 allowed for simple tests of display ideas to be prototyped and tested. Such ideas were that of zooming and clipping actions, and using media techniques for animation, such as stop motion. The display retrieves data stored on the local hard disk, which has already been collected such as the map tiles and search queries. The images are digital tile maps obtained from Microsoft's Live search server, and associates, which is then recombined to form the overlay image of the location displayed upon the screen, as illustrated in the top left image in figure 4.2.

The selection of the images, were chosen for their visual and locational relevance to the Swansea area and student population. These places are six locations which we thought would give a good range of emotional and associative responses from. This was also were popular places for students

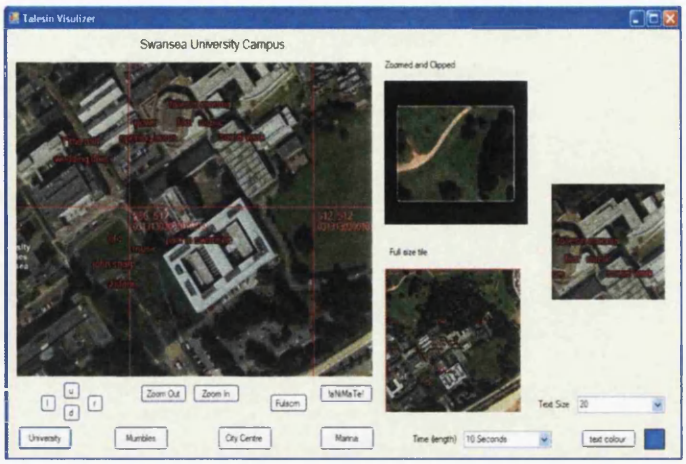


Figure 4.2: Query visualiser test GUI

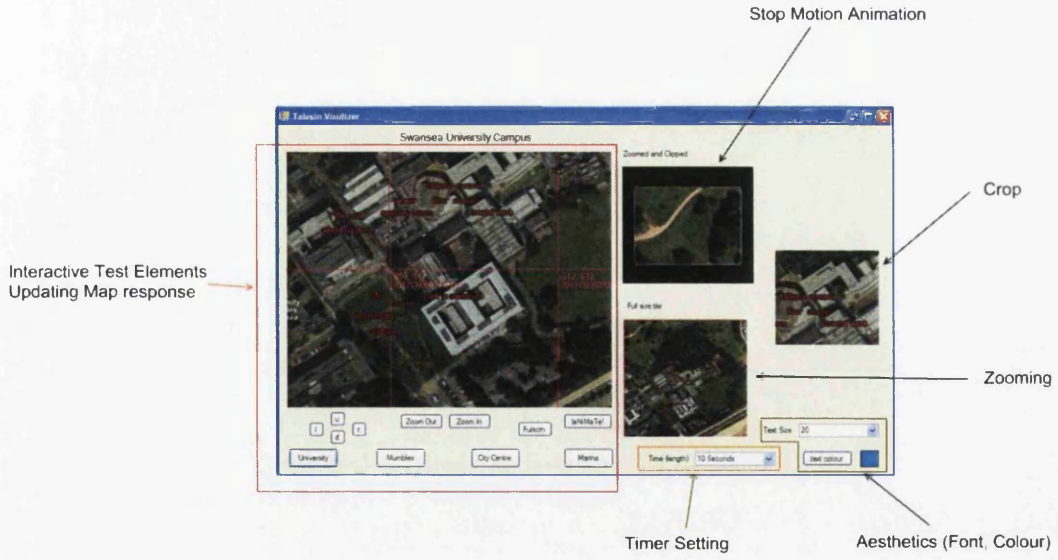


Figure 4.3: Query visualiser test GUI annotated

and visitors to locations, and would be able to give a feel for those locations.

These locations that were chosen, were the "University Campus", "Mumbles town centre", "Mumbles pier", "Uplands student housing", "Swansea city centre (Kingsway)"and the "Marina". These were limited to 10 seconds of viewing, with a 1.8 second fade in/out effect on either side.

We decided on 10 seconds for the duration of the location for an impact response that would be generated from the short period. We were trying to replicate the effect carried out similar to advertisements presented in broadcasting. The time period used, we felt, allowed enough time to interpret the words before disappearing. This was verified by initial prototyping carried out on the design of the system.

We also wanted to limit the number of places that were displayed, as not to overload the user with too many locations to remember. As this was already pushing the upper bounds of what could be recalled by the user. We refer to that of STM, which can hold 7±2 chunks of information at any one time. This became known as Miller's magic number, after his classic study to do with auditory sounds [Miller, 1956]. He was using acoustic memory, but STM also has two other parts, that of iconic memory (imagery) and working memory. The duration of STM can store information between 15 - 30 seconds depending on what type.

By taking this into consideration, we chose our times, so that they could provide a kind of snapshot of locations, from which we could allow the user to explore.

The queries were obtained from a previous study. It involved recording the GPS location of the place from which the query was expressed. The stored queries were then used and then mapped back on to the display in relation to the map tiles, linking the queries to their map location matching them up again.

With the tiles we had acquired from the Microsoft map server and the stored queries, which we had populated upon the map. We did not use the feature used in testing which allowed the user to move around the map using the mouse. We decided upon this because we were displaying them upon large passive public displays. For the display we were engaging how people would perceive and think about interacting with it, thus for this initial study we have not incorporated any input device for use with the large display version.

When considering the design of the information displayed, in this case the aesthetics of the displayed text, we considered whether the change, effected the viewers opinions of the display. We achieved this by including an update delay in the appearing queries upon the display. By fading in and out of text with the locations on the screen allowed us to give a feel of real time text update (changing), and the feel (resemblance) like a camera shutter fading in and out of focus of the area. e.g. The use of opening/fading in to reveal new location and new information, then back out again leaving the queries to the mind. We also added the feature to allow colour change of the text, to gauge different responses, and to select a colour which was viewable in the different locations. We finally decided upon a blue colour as this was visible with the surrounding colours.

The large screen display shows aerial overviews of locations around Swansea, UK. These included work, shopping, leisure and cultural areas. These are overlaid with mobile search queries gathered in those locations in an earlier study [Jones et al., 2007a]. The University Campus location is shown in 4.4.

Figure 4.4: Overlaid queries at university location

Every 14 seconds, the location being displayed was changed, revealing a new location and related queries. The queries themselves were updated every 7.5 seconds. This allowed for every query location, 2 separate collections of queries would be displayed. This continuous change and dynamic display was intended to maintain audience interest.

The cycle in which the screens were shown is shown in figure 4.5. When we first initialise the animation, it starts from the university location. The cycle of locations depicted follows. We added an advertisement screen to show where the research resigned from. This was added to provide a way of letting the audience find out more information, and to allow the viewers to have the screen feel like an advertisement screen once more. By adding in this screen, we allowed the bridging and closing of the conditional effect produced by interrupting their built up model of the use of the screens they had previously formed. So instead of them just thinking of the displayed images as a reel for advertisement, which was the way the current media displayed was being presented, the display of a new type of information could possibly change their current practises. This was revealed by some of the comments revealed in the findings.

4.2.2.3 Settings

For the choice of the settings, we had decided that we wanted to explore locations which have established large screen displays present and the types of reactions provoked by unanticipated incidental information could reveal. We achieved this through the deployment of the probe in three different locations on a university campus in order to elicit insights into how different areas might affect the responses received and observed. The locations were :

- Busy cafe inside a theatre (public, indoors).
- Walkway adjacent to theatre (public, outdoors).
- Common room of research Lab (semi-private, indoors).

In the two indoors locations we used existing on-site large screen plasma screens. These screens were already in use for public information and advertising purposes. The set-up in the cafe is shown

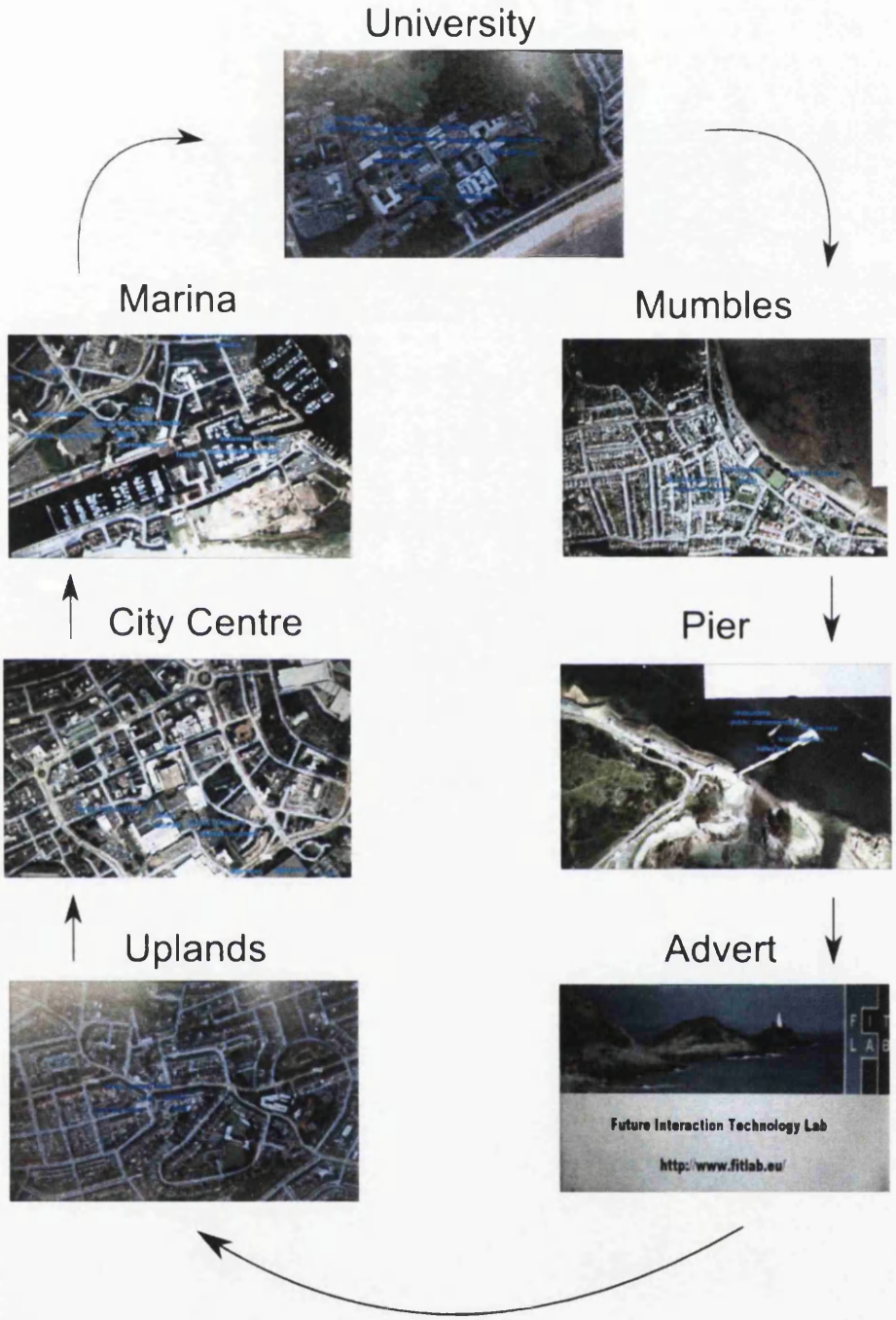


Figure 4.5: Cycle of locations

Figure 4.6: Cafe location

Figure 4.7: Common room location

Figure 4.8: Projected wall location

in figure 4.6 and the set-up in the common room in figure 4.7.

In the case of the outdoor venue, we projected the display onto the side of a building which was adjacent to a busy walkway on the campus. This resulted in obtaining the views of the people as they passed to see if they engaged or acknowledged the display. The projection displayed on the building was 20 x 10 feet, and is shown in figure 4.8.

Looking back at the groups and individuals who participated, we found that a more task oriented activity could of been more useful. The study only provided us with a survey response, but did provide some insightful data, such as the highlighted comments in the section 4.2.3.

4.2.3 Findings

In the following part we will discuss the results that were obtained from the study. This has been divided into the 3 locations mentioned earlier in this chapter.

4.2.3.1 Cafe

There were three categories of people, who were discovered at this location:

- General public visiting the venue, which also included lecturing staff on lunch break - These were people observed who initially ignored the display, as it was part of their everyday routine, who only watched it for quick updates of advertisement.
- Students - These were people who tried to spot their houses (a kind of game/entertainment), due to novelty of map upon screen.
- Catering/Art Centre - These were the workers, employees of the theatre, who were present for setup and observed for the complete study, and who have observed the questioning with audience and talked with each other about it.

From the categories described, you are showed the diversity that was present in the location. Each could reveal individual aspects common to themselves.

Initial Impressions In this part of the study we discovered that the context of the display was inappropriate to some degree. We found out that there were issues regarding timing of the user. We noticed that people were so busy, that most didn't have time to fully engage with the screen. They might have noticed the screen first in their peripheral view, as they first discovered it. This can be associated to location and line of sight they had with the display. We observed that when they were present for a longer period of relaxation, they were more ready for observing the display, without the time constraints that previously hindered them.

When we were considering the locations, we thought that this was a perfect location for engagement of social communication and collaboration. It was revealed that people did notice the display in some form, either visually or by indirect mention by others in that location, this is shown by the response in figure 4.9. This could lead as a result to the exploration of the notion of spread of word and topic about the display in the area. This in turn, could lead some people to tell their friends

about it, who then might observe the display next time they were in the theatres cafe, this can be seen as a kind of self advertisement. This idea was obtained from the post questioning of the study participants, one suggested this when discussing about "he'll tell his friends about it is their around today". This is an interesting idea which could be looked at.

One interesting idea which was expressed from the questioned participants, was that they suggested sometimes they seemed to get immersed into the displays representation, as they got use to the system. When they pondered for a while, they finally got a bearing on the map while exploring the space with text queries displayed upon it.

From the 22 participants, only 14 of the interviewed revealed that they initially engaged the display, without intervention (prompting). This was revealed during questioning, that it was due to them either checking for something new, or it caught their eye, where the others just stopped looking at it or were too pre-occupied to engage.

Responses When we questioned the participants about the display and what use they might see themselves using it for, most people concluded that it would be good as a kind of tourist board for displaying the information to new visitors to the area. The older participants that already knew the area suggested that they wouldn't be interested, while the younger ones would not mind. This response from the older people is quite significant, as it highlights the possible dismissal of adoption for newer interaction ideas, leaving it to the younger audience.

An interesting insight was expressed by the younger audience. They suggested that they would use such a system if it were present on a mobile device such as a phone. They said "especially if they were in a new location", this would allow them to get an overview of area, acting as kind of information device.

Discussions with the participants revealed the want for some more eye catching visuals. One interesting idea that was expressed, was that of making the representation 3D. This would add perspective and could be more engaging to more visual/object relating people, who would be able to make a better association/mapping to that area. This would also capture the eye of people, as the questioned said that they liked cool graphics (3D stuff), especially the younger people relating back to computer games once again.

Due to the diversity of people present in the study, we were able to obtain various views from the participants. One interesting view that was expressed was that from a business man, who suggested "if they were going to interact they would like some form of zooming facility which could allow them to pan in/out which allowed the grouping of text in that area, and would also give an overview when zoomed further out, (like abstract - most popular), and more detailed when closer in". This idea was influenced from stuff seen in business projects which incorporated GIS systems and the use of Google earth. The significance of this response was that it highlights the connections with passed experience of technologies, and how it can have an impact on the ideas of possible future interactions.

Value of Queries From the observations and questioning, we discovered that most people's first impressions, and final association would being some form of advertisement. They assumed it was somewhere in the local area, until they were explained too about its purpose, this would suggest

that words on there own without context, is not very useful, therefore refinements would suggest placing a naming overview of area or linking in with some context to it.

There was a 15/7 split, towards the idea of whether people thought it would be useful to see other people's ideas as mentioned earlier. The 15, thought it was interesting to see what other people were thinking. The other 7 just thought that they wouldn't consider using it.

From the participants, there was an expression for a need for clearer association to the text in that location (e.g. they would have liked a pin point location marked linked to a box with text box, which could be linked to more information and then finally expanded), or a kind of link, which gave more information to them, e.g. sent to their phone.

There were a couple of concerns related to the text displayed upon the screen which were related to mapping of words to locations and aesthetics of text displayed.

General Observations It was revealed that just using text on the background of the map produced a blurred effect. This happened as the text became merged into the image unless it was in a contrasting colour which was recognisable. This would suggest that visually, you would need more concrete/framed representation, which could be considered for easier recognition on the background image, by the audience.

4.2.3.2 Walkway

For this part of the study, the viewing angle and placement provided us with an opportunity to question passers by and onlookers, to and from the theatre. This allowed us to engage their feedback about the display.

Effecting Conditions From the observations taken, we found that a combination of time linked with the age of the participants had an influence on the audience.

One of the limiting factors, which influenced the possible audience was the lighting conditions. As the projector was more visible at night we decided to display it later at night, this resulted in an older audience viewing it. The average age of the group was 35 or older. This was in part to the performance, which was being performed at the time in the theatre, an Opera.

Initial Impressions There were a couple of associations people assumed from the display when they first saw it:

- As it was a map being shown, they assumed it's a location based system.
- They concluded that because it was being shown on a display used normally for Advertisement. They thought that it was just another one, until the started to fully engage with the display.

When we questioned the participants after they had viewed the visualisation for at least one complete cycle of the screens, they expressed that they probably wouldn't use it for long periods of time,

as they said "it would be a novelty item for them as it wouldn't fit into their normal activities". We can hypothesise that this was due to lack of interaction and relevance to there current activity. This was expressed by one of the participants on the walkway "it's not useful for what I was about to do", this could possible suggest that it was not part of their normal viewing habits, but this idea would need to be explored further to verify.

In relation to the size of the display, they expressed that they liked the dimensions of the display. It was visible from far away, compared to the large screens inside. As one commented "I didn't need my glasses", "It covered the side of the building".

Responses Related to the engagement of the display, five of the people that were questioned had noticed the display of the map on the side of the building. This was expressed by one of the participants, "it showed something new moving". From this we concluded that the movement of the words caught peoples eyes, which is the first part of the interaction step, capturing audience attention and engagement.

The other three were people who had previously stopped looking at it as we hypothesise they had previously built a conceptual model (Image) of what is displayed normally upon the display. When asked they revealed it was information about performances that was being shown, this resulted in them dismissing it.

From further questioning it was revealed that most people commented on the aesthetics more (the colour, shapes, font, and spelling), as this was the aspect that effected visualisation of the system and the questioned audience seemed interested in as a choice of preference, suggesting that this could be a deciding factor which influences their selection, and engagement process.

Value of Queries The group of participants we questioned in this location were less accepting of the incidental, apparently 'random', display of information. Some of the comments placed in the further comments were "it's just a mess on the screen", "I need something quick, why not a list?". From these types of responses, we made the conclusion that they would of rather liked more curated, structured information. This response may have been due to the need for quick, directed information while outside and on the move in contrast to the participants comfortably accommodated in the cafe.

Positive feedback was also obtained from the questioned participants, who appreciated the large size of the display. They were also intrigued by the apparent dynamic nature of the display. This dynamic form contrasted with the canned advertisements that were usually projected onto the wall.

4.2.3.3 Laboratory Communal Room

We left the display shown in figure 4.7, running for 2 days due to the location and possible flow of people. In this time people from the laboratory could observe the screen. We observed peoples engagement with the screen and asked them to provide feedback related to what they had seen.

From the people who took part in the study, some initial impressions and feedback were expressed:

- "Was it related to Google Maps?"- Due to the increasing popularity of Google Earth / Maps, people straight away assumed that it must be that, as from past experience, this is probably the only geographical system, they have used or seen other than in earlier education.
- Stared for a while - This was in surprise as there was something different on the screen, thus they paid attention to it, as normally either in standby when people are not using it, due with concerns of power saving and screen burn associated with old plasma screens.
- "Would be useful to them but not in current place"- They thought that it was useful as a reminder/view of what people liked, but wasn't much use in its current location, as they were focused on their topics of interest, and mainly wouldn't serve any purpose in break room, with what was shown currently upon it.
- "What's that?"- Some people thought, 'what, uh' in amazement wondering, what and why some people had expressed these queries, as they thought that the meaning had nothing to do with that place, but obviously it did to that person, thus it engaged the person to think of strange connections?

While all of the people who viewed the screen in this location were residents of Swansea, some were able to see value in seeing familiar places. We observed some people trying to make sense of the changing queries - unlike specially authored location content, some queries appeared to have little to do with the place in question.

Responses When we questioned people in the room, the method we chose for collecting the information was to question them as they passed by. We would informally interview and question them when they were present in the room. This would coincide with their work breaks. This sometimes led to interesting responses. There was a range of emotions expressed, such as "leave me alone", to "I can see a point but ?", some were remarks but overall people liked the idea out of politeness, but could see potential and were interested in future developments.

Most common room users thought that it would be of most value in a more public place but indicated that if the queries were 'live' it could be an interesting background display in the common room even perhaps acting as a 'conversation piece' where groups might try to guess the need being the mobile search query or identify trends and patterns.

Possible Interactions When we questioned people about how they would interact with the screen, they said, "well it depends what for". Most still liked the idea of touch based interaction, based on ideas seen from the movies and VR set-ups and information kiosks in museums. Some suggested with current technological advances to the public of gestural interaction, with devices such as that incorporated with the new Nintendo console the 'Wii'. This uses a combination of accelerometers and infrared to control interaction through sensing users movement. The participants suggested "that you could use that, or similar technologies".

Some expressed that they would like some gadget to play with, as they said "in its current state, it wouldn't appeal as not enough interactivity", this would suggest that further development might could incorporate more interactivity.

From the study of these locations, earlier demos have shown that the system displayed did engage the audience and did make them want to interact with it. It also sparked off conversation topics related to and about the design and use of images, words and the context in which it's supposed to represent.

4.2.3.4 General issues of interaction and engagement

There were a varying number of patterns that revealed themselves from the study, these were:

- Link with past user experience of devices - These were people who had previously had experience with a computer and where already used to the WIMP philosophy for interacting with displays, and expressed ideas of touch screen linked with kiosks like they have used in theme parks and museums.
- Concerned with privacy and exposure - Some people suggested that they would become self-aware of themselves interacting in a public space, thus some would like the option to keep it to themselves (a personal selection and interaction device). Others relished on the idea of being centre of attention and would like the collaboration, interaction with people around them.
- Visual issues - This was to do with aesthetics of the displayed image (size, colour, position, representation of map and text upon it), this effected how they perceived the display, and ultimately overall experience from the system.
- Lack of interaction ideas - Most people didn't know how they wanted to interact with the displays, they were just happy to observe and if it had some form of interaction, take it for granted that was the way of interacting they would have to learn.

When we look back at the locations, there has ended up being some correlation between some of the areas that were used in the study, such as the theatre cafe and the lab common room. As one was based in a larger area, this resulted in a more social experience when in comparison with the lab.

In the laboratories communal room, the opportunity of questioning viewers was somewhat limited due to less frequent visits. An interesting comparison could be made with that of the theatre cafe, as both locations were using a screen the size of approximately 40 inch diameter. In the Theatre cafe and laboratories communal room, as the displays were the same size some of the comments reside for both, relating to the lighting and aesthetics of the screen and display. We also found out that the placement had a considerable effect of the viewing audience observed.

One interesting thing that was observed was that concerning the number of people present while observing. This presented a correlation to the direct response on the amount of interaction, engagement people were willing to give. The less amount of people that were around the more people were willing to do more, without fear of embarrassment, spectacle, attention, when compared to when lots of people around people were more conservative, except for the occasional extrovert.

Across all three venues in which we installed our display, some common underlying issues arose. Many people were concerned about approaching or engaging with the display in fear of losing the anonymity and privacy that is, paradoxically, a feature of being in public.

Several participants asked us about the possibility of using a private, handheld device (e.g. such as a mobile phone) to engage with the display without other people being aware of their interactions. We observed that the fewer people at a location, the more likely people were to approach the display and interact with it.

While the desire to blend in was most evident, some participants we approached relished the idea of being centre of attention. They suggested that they would like to collaborate and interact with people around them. We observed people who were intrigued by current users and the display, who were playing a kind of game (spot your house) on the maps which they recognised by features of buildings. As a result of people stopping and pointing at screen, more people within the location became more interested in the display and the current activities surrounding it.

The most enthusiastic responses to the probe came from those less than 30 years old. This age represents an audience who have been brought up with digital technologies, and who could imagine using such a system. The general response of older participants (aged 60 and over) is succinctly captured in this comment: ``oh we don't understand or have these new gadgets we'll leave it to our grand children''.

This reveals a gulf for acceptance and usage between the age categories, which needs to be addressed. Even though they would not or could not see them selves using it, they expressed strongly that it started conversation, by the display of different words, which pondered in their minds, or connected with past activities. As one lady expressed ``it brought back memories of a wedding I went to last week'', this was after word association with the text query ``wedding dress''she had previously saw on the display.

4.2.3.5 Summary of Interviewing

From the interviewing of the participants we were able to extrapolate the following information. The questionnaire was a 11 point questionnaire probing them about their experiences. The questions covered various aspects which covered areas such as qualitative feedback related to awareness and understanding (questions 1,2,11), usefulness and usability (questions 3,7), Aesthetics of actual output and functionality (questions 4,5,6,9,10).

The questionnaire can be found in appendix A section 8.1.1.

From the people who were interviewed and the questions asked we obtained feedback on the questions as shown in figure's 4.9, 4.10. This shows the responses to the study questionnaire, surveying peoples opinions about the public screen and the information displayed upon it. There was no responses for the laboratory due to the infrequent users and length of study period. This would be something to consider next time.

From the questionnaire certain trends appeared, people did notice the display projected on the side of the building, which was also visible from within the cafe. With regard to the usefulness and usability of the system, from our results we can see that the audience did find it useful, and if was available would use such a system if was available, though there were some concerns raised by age, this was expressed by one of the participants ``I would use the system, only if I could learn it quite well''. This comment was made by an elderly person who feared learning new technologies. Some of the questions the participants did not provide any answers or was not appropriate for the question.

These were put down as N/A values. We noticed this more towards the end of the questionnaire when people wanted to finish. The participants were more responsive to asked questions to obtain their ideas.

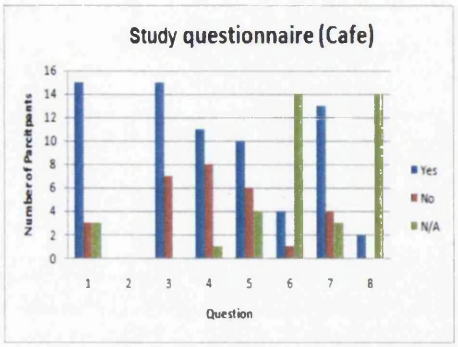


Figure 4.9: Survey questionnaire responses for cafe

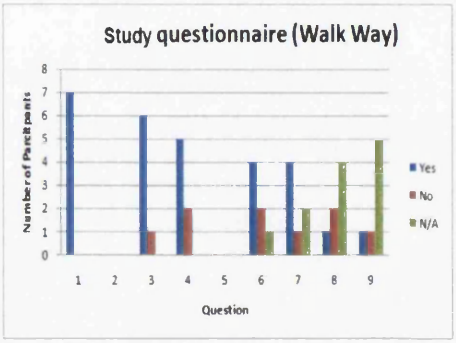


Figure 4.10: Survey questionnaire responses for walkway

4.2.4 Conclusion

The probe and study had three elements of novelty. These were the decoupling of the mobile and large screen interactions. The move to incidental rather than considered content. The exploration into the impact of location on the engagement.

The studies presented provided us with some themes to stimulate further discussion and work. Our findings suggested that respondents were intrigued by the notion of watching representations of the interactions of other, asynchronous and mobile users. While the mobile queries are clearly not valuable for those with specific, structured information, seeking needs their large screen presentation did give some viewer unexpected insights and lead to discussions with other people at the location.

We found that location clearly effected the engagement with the device in two ways. Firstly, there was the 'conditioning effect' - people are becoming used to large screens providing advertisement style information. This was expressed succinctly by one of the participants ``ah! something new''. In the future, more engaging displays will have to accommodate, or exploit this exposure. Secondly,

the form of ambient visualisation provided in our probe seems most suited to locations, like the cafe and common room, which allow people to become gradually aware of the display.

In progressing with this work we were looking at further ways of capturing other forms of incidental information via mobiles and other sensors and representing them on large screen displays. A further concern is a more direct interaction between a user's own mobile and the displays they encounter. So, for example we are considering how interactions captured by the display might be re-purposed on the mobile device.

Within the context of use at the 3 described locations, we were able to obtain the short term values that can be obtained from augmenting the normal media spaces at these places, from that of natural (original) media presented. This is expressed through some of the trends identified in section 4.2.3.4 We have also obtained the views, and discovered a possible need or not for displaying this kind of locational summary of public search queries to promote community awareness.

We have found that extending an idea, developed for small devices can discover additional benefits of connections in a (public) community setting. We did see some of the participants through the association of words recognised did form a personal connection. One of the responses was ``That's the name of the church I got married in''. On a more social connection, the image locations used did resonate with the public, as they could reflect and associate with an area that they knew.

4.3 Discussion

Interacting with large displays can be fun and engaging but only within the right context, and through the ease of interaction techniques. Within the context of queries being presented upon large displays, people found this kind of fun and intriguing but it was more of a novelty. The true question we need to answer is how can we facilitate in the development of these fuelled by this new idea of incidental information. To allow MUI's to be more contextually aware for users, in locations from where they're placed. Exploring how to incorporate these mobile devices to act as communicators/controllers to these displays/systems.

This suggests a keen area of research in the area of interaction through a more natural, less obtrusive way. This suggests the consideration of possible incorporation and study into gestural interaction, within contextual (context-aware) environments for multi-user environments, there by alleviating the burden upon learning complex techniques for natural affordance based interactions. This leads us to explore the notion of incidental interaction as one possible route to do so.

We have explored the notion of incidental information within this chapter, and we have explored a possible application and use for such information.

The findings from this early study reveal that by taking a different view on possible interpretations of usage of collected media, hidden ways of information dissemination can be explored on the influence and altered perceived understanding of a location, old or new.

We have decoupled the previous experiences that have been formed with these devices (mobile and situated), and considered the usage of incidental information their actions can provide.

This chapter explored the coupling of separate asynchronous activities, and considered how we

could combine these two forms to help promote a sense of a community to a location. We considered these technologies from the outside looking in, for possible exploration of use.

In the next chapter we will consider how we can obtain possible incidental information which may reside around the pubic screen itself considering the modalities, which can occur and lead to interactions. This will consider and explore possible synchronous/asynchronous actions which take place surrounding the interactions that occur with the system itself.

Chapter 5

Sensing Interactions

Contents

In the last chapter we discussed work related to that of incidental information in a search application and considered how and where it could be found. We explored one way of applying this kind of information to different situations and contexts that might exist.

There are many possibilities which present themselves in a way of linking a person and their actions with that of a large public screen. In the last chapter we concentrated upon information related to mobile devices and actions carried out as a means of incidental interaction. We considered how devices could be used and explored one type of incidental information that could be obtained. This is not the only way in which we can obtain incidental information present at these displays. Within this chapter we have decided to consider that of visual processing and the ways in which it can be used to facilitate the acquiring of incidental information.

In this chapter we will explore how we can extend the situations that occur with asynchronous data being placed on a public display. We focus specifically on that of detection of users present at the displays. We will explore the extra information that could possibly be obtained. We will be considering the aspects of the interactions, the physical location, form of the display, and the audience of an interactive responsive display. This can take the form of a passive display such as that explored in the last chapter, or of a proactive display system, by obtaining the attention and eliciting interactions which are taking place with the display.

The chapter will report on two studies, in sections 5.3, 6.1 respectively. The first considers a usage scenario for incidental information, using the images captured to explore what can be obtained from them. The second study will be concerned with the development of an algorithm to detect the identified items.

This algorithm will be used to test the concept and provide extra information from the exploration of incidental information presented and captured from visual recordings. We will be exploring questions about place and contextual relevance which occur in images captured. We will explore what is obtainable from the images and the value that can be achieved and added to that location's media, presented for present and future usage.

We will consider how we can use additional technologies to sense the presence of people at a display, to allow incidental interaction with a public display.

Using an adapted form of the StoryBank system described in section 5.2 allows us to explore in what ways, features can be obtained from a live environment. In the next section we shall describe an exploration of this notion of incidental information from incidental interactions which can take place from a glance.

5.1 Incidental Information Surrounding LSPD

The interactions which can be afforded by the positioning and use of large public displays offer many possible interpretations and alterations for the media displayed and collected from around them. With the media available, which is either presented on the screens or collected in some form, be it through tag based (e.g. barcode) reading or external devices (e.g. fob key ring) to the display, it can all be put to good use. The ways in which it could be useful are by the sensing of audience members or just making the display more amenable to the audience present in that location.

People who are present at the display are carrying out the action of information seeking. They are carrying out various subtle, passive actions, which could be explored. These actions could be interpreted and used as catalysts for additional interactions that might occur as a result.

There is so much information present in the environment, space around the display and user. This needs to be explored to find value that could be deduced for activities, and as a result bring information into the attention of the user.

We consider one aspect of this task by exploring the way in which we can make displays more proactive and engaging for the user, considering the incidental information that can be revealed as a result. A simple scenario to demonstrate this idea is:

> "While at a display people are interested in interacting with it on a visual and cognitive level. This can be a result of a possible playing scenario with media present on screen. Images observed and subtle cues possibly detected give an insight of people's attentions spans while viewing information, which can be dependent on time, and have to be able to conform to their organised schedule of life's tasks."

There are many possible opportunities which are presented by the use of visual processing in developed and developing world contexts.

5.2 Storybank

5.2.1 Introduction & background

The StoryBank system [Jones et al., 2007b] was conceptualised as a way of sharing digital stories created via mobiles phones with members of a local Indian village.

Instead of recreating from scratch, we have chosen to extend the features of the Rhondda Lives system, which was a cut down version of the original StoryBank. This allows us to have a base from which we can build components to augment the features of this system and explore its potential. By using a modular design methodology when creating the software we can apply the techniques explored here with that of new choices made in future representations then might present themselves.

By using this as a form of repository we have media, which can be used to obtain viewers and audiences which are required to help with the exploration of this notion of incidental interaction. This relates to the acquiring and augmentation of existing information and media to the user.

The StoryBank system has explored ways in which media can be used to explore the ideas and cohesion that video sharing can present in a developing world environment. We use this system to explore the usage scenarios of a central media repository in a community, and apply what was learnt to explore that of incidental interaction concerned with that of similar situations. These are situations where the audience is visually accommodating due to being functionally illiterate.

When we consider the usage of such a system in a developing world context, this has a considerable meaning and difference from that of a more modern, built up environment. The concerns and issues are changed from the purpose and aim that the system now represents. In a modern, developed country, the means by which people can obtain information are endless, due to the myriad of technologies which now surround us. This, in contrast, is not the same in developing countries, where every luxury must benefit a family or community to be of value.

5.2.2 The System

The system was created to act as a media repository for the media created by community members on their phones. This has taken the form of personal stories about a subject of their choice. The operation of the system revolves around separate components. These can be grouped into sections which compose of:

- Phone connection and communication (Uploading/Downloading)
- Menu systems and general selection screen
- Individual media playing on the interactive display
- Mobile phone software (Video Editing)

All of these parts form the overall running system present in the village.

We extended the Storybank software [Jones et al., 2007b] which provided a platform for displaying media which we used as an incentive to watch the display. We added a logging system to record media selected and incorporated a web camera to record people present at the display. This involved

additional screens for users to interact with (this is discussed later in section 5.3). The original system is shown in figures 5.1, 5.2, 5.3. These images are screen shots from the user's manual.

Figure 5.1: Original Storybank main screen[1]

Figure 5.2: Original Storybank selected media playing[1]

Figure 5.3: Number pad for selecting video explicitly[1]

5.2.2.1 Media (Videos and Images)

The types of media that were considered were those associated with the mobile phones available and present for the study. These were Nokia N95 handsets. They are camera smart phones with the processing power needed to allow for the capture and editing of media.

The videos were created by composing a collection of individual images taken with the phone.

[1] Pictures from Jones et. al [Jones et al., 2008]

These images were then coupled with sound layered over them. When the villager using the phone returned to the base station [2] the media was uploaded to the repository.

5.2.2.2 Collage Technique

The collage technique is a way of presenting information as a composition of media elements. This is formed by various videos of stories and images captured by users on phones in the community. They are then randomly positioned on screen, not overlapping to allow full screen utilisation for media presentation. Shown in figure 5.1, is the representation of this technique for the presentation of the media present on the interactive screen. This seems inspired by the work from Andruid Kerne's ``CollageMachine''[Kerne, 2001].

5.2.2.3 Rhondda Adaptation

The Storybank system was adapted and refined by Simon Robinson[3], for the Rhondda Lives project. These alterations were undertaken to make the system more passive. This resulted in the mobile phone component of the system being removed, as there were different goals that were required by that line of research.

The front end of the system is shown in figure 5.6. For this study we were trying to use the presence of audience members to explore possible new ways of augmenting the data being presented implicitly. Therefore having the mobile phone component would illicit different reactions and exhibit different observations than we were aiming for. Rather than interacting with the screen, they would use it as more of a reference and would be more engaged by their interactions and communications via their personal phones.

Using an adapted form of the Storybank system allows us to explore what features can be obtained from a live environment, and how we might obtain them.

In the next section we shall describe an exploration of this notion of incidental information from incidental interactions.

5.3 Exploring the Potential of Visual Incidental Information

This study explores the placement of an interactive display within a library environment. We consider the impact and findings of introducing such a system into that space. We also explore what types of information can be obtained and understood, relating to that of incidental information.

Due to the public location of the system, we decided to explore the feasibility that such a system could provide. This was evaluated by using questioning and answering of participants in the study carried out. This can be viewed as a type of technology probe [Hutchinson et al., 2003], and linked to cultural probes [Gaver and Dunne, 1999].

[2] The interactive display at the teaching centre in the village.

[3] A research assistant in the future interaction technologies research group at Swansea University

This part will report on a study taken to reveal value and opinions of exploring the validity of using face and object detection and image processing techniques as methods of interaction compared with traditional ways. This will be concerned alongside the issue of implicit versus that of explicit input modalities.

5.3.1 Aim

The aim of this research was to explore the possibilities of image based sensing when applied to adapting content playback on large screen displays.

The motivation for this research was to consider what could be revealed with images of viewers as they interact with displays. When considering the acceptability of such approaches, due to image capture and issues that might arise. Possible research questions that arise related to this type of interaction are:

- What are the possibilities that present themselves and issues with using viewer images recorded at displays?
- What is the acceptability of using computer vision techniques in public locations?
- Value of content adaptation and filtering, when exploring a sense of community, particularly in respect with that of ``viewer number''?
- What can be automatically deduced from scene images at a display?

5.3.2 Methodology

In this section we will explore how we undertook the research and how we aimed to explore this notion.

5.3.2.1 Procedure

We carried out the study in the following order of steps :-

Order of Steps

- Observation of users responses to display
- Introduction of interviewer and Study Background
- 5 minutes use with display
- Post usage questionnaire (SUS)
- Interview Questions

The display was left at floor height, similar to most kiosks in the surrounding area, which had the collage system run upon it. When the user first approaches the device they are left for a couple of seconds to make sure that they are actually interested in the information displayed. If they were, intrigued or interested and wanted to know more about the display and media shown, we obliged them with a discussion about the study that was taking place before they took part.

We asked them to engage with the device by spending time exploring the GUI. This allowed them to get accustomed to it. We then requested that they select and view one or more videos; depending on their time and interest with the topic being displayed. This allows the user to build a further understanding of the system, meanwhile learning something about the Rhondda story they have selected.

After they have watched at least one video, we then ask them to fill in a SUS questionnaire related to their perceived and actual experience of the kiosk. The questionnaire used and the explanation for calculating the results is presented in appendix A, section 8.1.2.1. This is a standardised self-reporting, post-session ratings questionnaire on the usability of a system [Tullis and Albert, 2008]. We then ask them to fill in a post questionnaire related to some of the ideas they experienced, and to detail any concerns they had about the system. This provided us with qualitative feedback relating to the system and various issues such as location, topic, and time. Both of these questionnaires are available in appendix A, section 8.1.2.

The post questionnaire included such questions as, "Which interaction do and would they prefer?(Explicit or Implicit)", "Would such a system that used vision techniques appeal to them?", "Did the image detract from their experience of the system?", "Did they mind their image being recorded?", "Did they think it would be useful being showed usage statics on the screen?", "Does the chance to change future interactions entice you to use the system?", "What would your concerns be with such a system?", "Could you see yourself using such a system in the future?". All the questions had a "why"element to them which provided us with qualitative responses.

The time taken for the complete session lasted on average between 10 - 20 minutes. This was dependant on the amount of media selected, and also the views and opinions expressed related to questions that arose.

5.3.2.2 Apparatus

Various apparatus were used for the studies which were set up. They are separated into the hardware and software used.

Hardware Due to the location that we had chosen, we considered conforming with the surrounding area. We decided on blending in with the surroundings. For the study we used a 15-inch touchscreen (Protouch T15), having a single web camera (Microsoft LifeCam VX-1000) placed above it. This setup is shown in figure 5.4. The screen was smaller than we wanted, but it was trade off between technologies available; a touch screen which was available and could blend in.

We used this touch screen as we were also considering that of the comparison with the modality of touch versus that of visually sensed.

The way we set this up allowed for the user to retrieve tactile feedback to their interactions. This also provided us with input for the logging system. This involved the recording of media used and captured images of scene and display screen itself.

Considering computer vision techniques provided us with a way of retrieving the responses of the participants present at the display. Using these techniques allows us to provide feedback to the system user by using the image recorded by a webcam placed on top of the screen.

The study was carried out using a Pentium 4 3GHz single processor with 1.87 GB of RAM, running Windows XP Professional 32 bit, SP3. We used external speakers for the library and the laptop's inbuilt speakers while testing out the system before deploying it.

Figure 5.4: Webcam positioned on a touch screen monitor, with *Rhondda Lives* running

The choice and use of this hardware made it similar to existing kiosk type systems that people who visit museums are used to. Thus we thought that this would relieve them of any burdens of usage. The camera was placed in the middle such as that on modern laptop screens for familiarity.

Software The design of the software and the interconnecting elements which make up the software are shown in figure 5.5. The software is made up of individual elements. The collage system, the logging system (which is incorporated into the collage system), an image taker which records the audience members, and the database viewer for reviewing the captured images and housekeeping of logged data.

We extended the Rhondda Lives version of this software based on the Storybank system explained in section 5.2.2.3. The front end of the system was adapted for the media provided by the ValleyKids [val, 2009] group, who had made digital stories in co-operation with BBC Wales and the National Screen and Sound Archive of Wales, with the project named "Rhondda Lives!" [Wales, 0924]. This was a simplified version. We were interested in only using it as a test-bed for acquiring images to test algorithms with, and to explore explicit vs. implicit interaction upon media adaptation. Thus we did not require the mobile phone part of the original Storybank system in this study. See figure 5.6, 5.7, 5.8, 5.9 for the various screens presented to the user of the system.

The image shown in figure 5.6 shows the main screen of the system that is presented to the user. The participant then selects either a filter icon on the left side or one of the playable media on the rest of the screen.

The screen shown in figure 5.7 asks for the user to select the option explicitly, indicating the number of people present at the display who were watching it. We designed the screen so that we could use

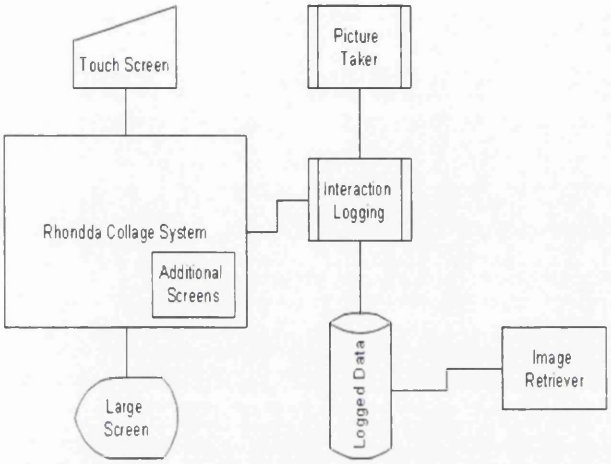


Figure 5.5: Elements of software

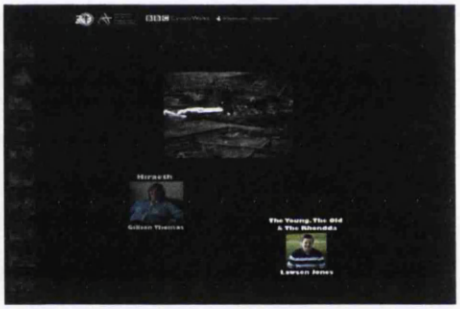


Figure 5.6: Main screen of *Rhondda Lives*

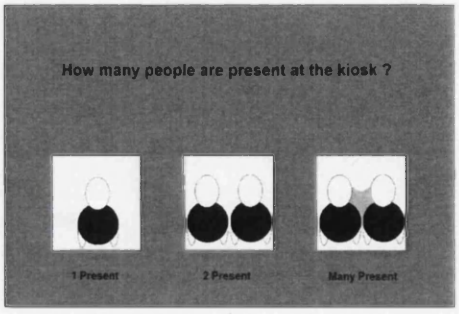


Figure 5.7: Explicit questioning of people present

the logged responses for later comparison and verification against our algorithm. From the results, this could lead to exploration of possible trends present in the data, related to time and the specific media involved.

Figure 5.8: Selected video playing

The screen shown in figure 5.8 shows the selected video playing, chosen by the participant.

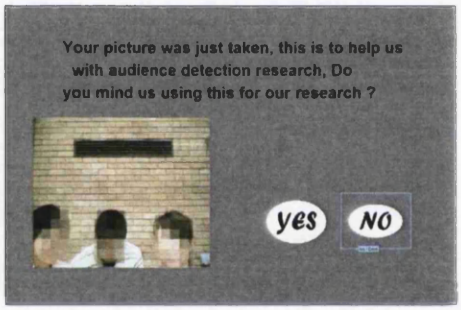


Figure 5.9: User image participation screen (pixelated faces for anonymity)

The screen shown in figure 5.9 shows the option that is provided to them after their chosen video has finished or exited. This screen provided the participant with the option to review the image that was recorded of them, while they were in the process of watching the video. We gave them the choice asking them whether we can use their image or not. This allows us to cover ethical issues that might arise. If accepted, the image is then added to the logged information in the database. Once the information is stored, it can then be retrieved and used as input to our algorithm.

Using the videos allowed us to capture images of the participants present at the screen, which were then used to test our algorithm and questions.

We added the additional features of recording user input via a web camera. We also developed and incorporated a logging system using a MySQL back end for real time communication and later perusal. The overall interaction cycle which was undertaken for every media (film) interaction is shown in figure 5.10.

The system was written in Java incorporating the JMF Media Framework for Multimedia features. The original Storybank system and adaptation was constructed using java, and as we were extending the legacy code from this system, we thought that it should be the same language. Using Java allows for cross platform compatibility, which is useful considering the various platforms which

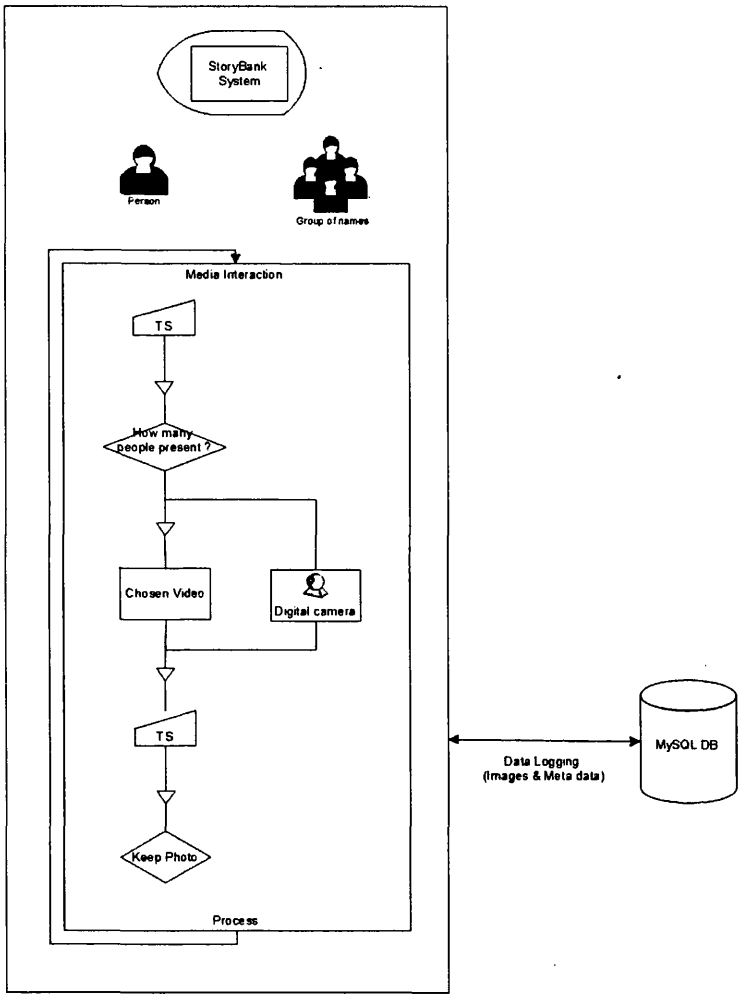


Figure 5.10: Interaction cycle (TS:Touch Screen)

might reside upon local hardware. We used the Java JDBC API to access the database via invoked methods for the logging of actions carried out within the study.

For verifying the data collected, especially the images for verification after the study, we created a small application to retrieve and view the recorded data stored in the database. This is shown in figure 5.11. This program allowed the selection of individual items, needed from the database.

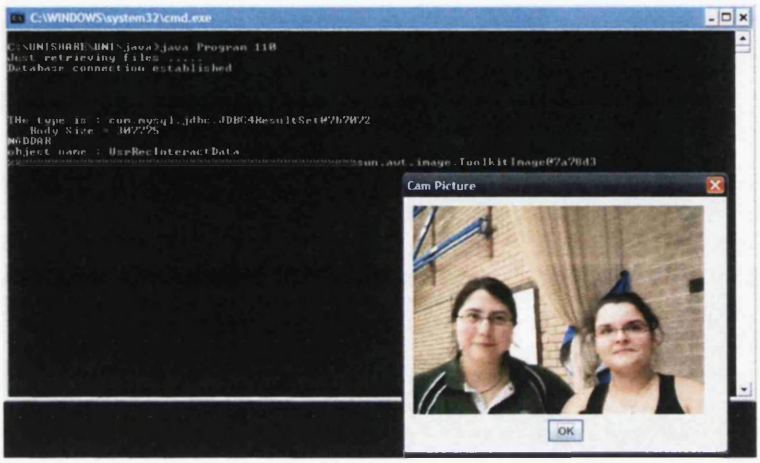


Figure 5.11: Data retrieval from database

5.3.2.3 Media

In this section we describe the media which was used, and what items of information were recorded. The media we used for the study consisted of vignettes. These are short videos which last around 3 minutes in total, which contain a personalised impression of the community member and setting of their chosen video, created for the Rhondda Lives project. The media provided from the ValleyKids project was displayed to the audience. The media was categorised into various topics pertaining to that location and community. These were chosen topics which were chosen by staff of the project and which they used to encode and reference the media by. The various media was categorised as follows:

- Childhood
- Education
- Work and Wages
- Friends and Family
- Cymraeg
- Porth
- Ferndale
- Tonypandy
- Treorchy

There are a total of 64 videos which are used within the system. The number was supplied by Rhondda Lives project. Individual videos are randomly selected from this collection, and are placed upon the screen ready to be selected for viewing.

The reason for logging the extra data, detailed below was to provide us with possible additional information from which we could use to alter future media with. This provided a starting place relating to the context of the video which could be used to reveal patterns and allow for interesting media (popular media), to be re-purposed for future interactions at a later date.

The interaction is initiated by the first press of the touch screen by the user. After this the system runs automatically until next interaction.

Logged Details Most of the users' interactions with the system were logged, which included

- Interaction start and end - This recorded the date and time of the interaction
- Scene image - Used for the analysis with the developed algorithm. To allow detection of people present at display. This was taken while video playing.
- Screen image - Used for additional meta-data. Used to see what was present at time of scene image recording.
- Filename of video - Used for additional meta-data. Used for reference by system
- Unique ID of video - Used for additional meta-data. Used for reference by system
- Video categories - Used for additional meta-data. Used of filtering and adaptation of future media
- Whether image was allowed to be saved or not - Used for additional meta-data. Used for ethics related to research & algorithm testing
- Person's number - Used for additional meta-data. Inputted by user and then compared with actual detected data as verification.

There was a lot of what we can consider to be rich information that was recorded. This could be used to reveal anything informative. Initially we will be considering the number of participants inputted by the user, the video they chose, and whether they approved us using their image for use. By using these simple pieces of information, we could over time start to see trends appear. From this we would be able to infer patterns that might reveal themselves about chosen topics and popularity. This then could be fed back in a cyclic fashion as possible input for a context awareness system. We could also use this information as an entry point and catalyst to other activities for bringing the community together.

From the collection of data recorded from the number of participants, we could use this as a verification tool, to compare actual numbers to that in the image, if allowed by participant in the option screen shown in figure 5.9. This can be used to verify actual selection and participant present. We would then run this against a detection algorithm to compare and validate the detection rate upon an unseen data sample.

We decided upon logging this additional information, to allow the adding of meta-data to the current videos. Adding extra information could add value and could be used by future systems for filtering the media.

5.3.2.4 Experiment Setup

There were three locations chosen for the study, each chosen to provide different social situations. This can be seen in work carried out by Churchill
[Churchill et al., 2004b] and in the eCampus project [Storz et al., 2006b]. These were the university library shown in figure 5.12, the university sports hall, and the postgraduate research lab shown in figure 5.13.

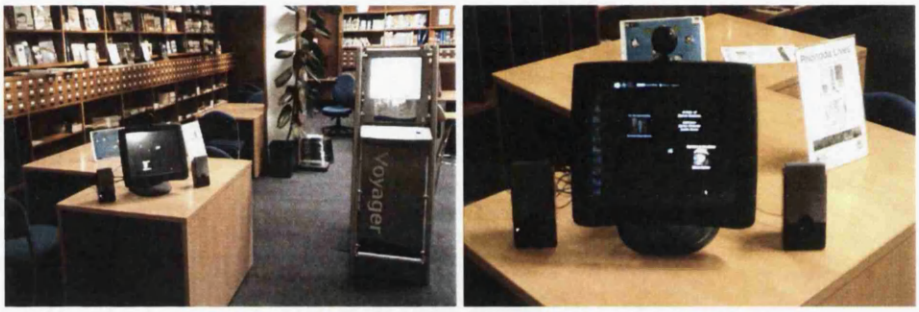


Figure 5.12: Library position of system

Figure 5.13: Postgraduate research laboratory position of system

The settings mentioned above incorporated different situations of engagement. The library was a high visibility area which had the potential for the most engagement. It was based in the busy foyer with high throughput, shown in figure 5.14. This allowed a variety of people ranging from students, lecturers and general public of all ages to participate.

The second location, due to maintenance in the library, required movement to a different part. This was in the PC area of the study hall shown in figure 5.14, right image. We carried out more logging in that area which all contributed to the library section of results.

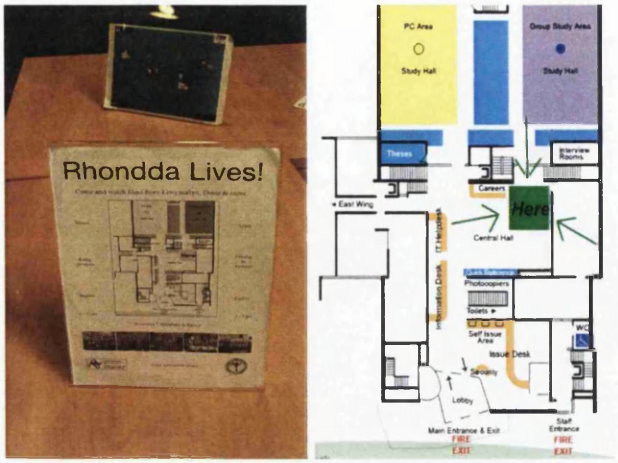


Figure 5.14: Advertisement of library location

The sports hall was also a high visibility area, but was catered for sports people (different demographic), and hopefully would illicit different responses. We also concentrated more on awareness of groups.

The last location was the PhD research lab, which provided the context of a test bed for the interface alterations before deployment. This also provided us with expert knowledge from users who have experienced touch screen technologies.

5.3.2.5 Participants

The types of people covered by this study would hopefully provide us with some insightful information. As mentioned previously, the library caters for a variety of public members which provides a wide variety of useful results. We didn't specifically go out to target specific people as this was exploratory probe, we just let the system run and engaged people who became interested on their own course.

This changed for the sports hall, as we were limited to certain times and depending on what time, various clubs and sports club members were present. The people that were used due to coincidental time were those in the badminton team. Members of this team were students studying a range of disciplines, the participants were mostly engineers and sports science students.

The use of sports club members allows us to obtain the views of a different demographic than that obtained from a library situation, and provide different experiences. Their feedback could provide different insights for future implementations. The features provided by the display, and the ways of interacting with such devices may result in different behaviour due to current activity and relevance.

Finally the post-graduate laboratory provided us with computer science students who where computer literate, and were more intrigued to explore the features of the system.

We have discussed the settings and the way in which we carried out the study. In the next section, we will discuss the findings.

5.3.3 Findings

We have consolidated and analysed the data we have collected. This has resulted in data from questionnaires (post and SUS) and logged usage data related to media used. Finally there are the stored images recorded. These are screen images the user sees and camera recordings looking out from the screen.

5.3.3.1 Acceptance of image capture

From initial review of the data, the findings suggest the likelihood under certain circumstances that people did not mind having an automated detection system change the media, as long as privacy issues were covered. This is shown in our findings by some responses of participants in section 5.3.3.4, and interaction preference shown in figure 5.28. The next step would be to develop and run the system utilizing the collected data to get feedback on the improvements that this has brought about.

From the questioning pertaining to the use of auto tracking recorded by a web camera, our results suggest that the majority of the younger participants did not mind being tracked. This is as long as they understood the applications goal and the usage that their image might receive. There was one participant who had initial concerns, which was expressed by negative sentiments, "I don't like my photo to be taken without my knowing!"and "That's a bit creepy".

5.3.3.2 Analysis of images captured from study locations

We explored three locations which are described in section 5.3.2.4. These locations tested the placement, and possible impact that could be received upon the interaction taking place with the system. This has allowed us to explore various audiences at different locations. This explores the idea and usage of vision techniques to automate the logging of user habits, and what impact it has on the user's experience with the system.

For the study we wanted to explore what extra information could be obtained from the display itself, to help in the adaptation of media present upon it.

There was a total of 42 images used to visualise and classify the observations. In table 5.1, we show the breakdown of the amount of images used for our classifications shown in table 5.2. We felt that this amount of unique images allowed for a sufficient indication to various classifications that could be made.

Location	Number of Pictures
Library	10
Laboratory	19
Sports Hall	13
Total	42 Images

Table 5.1: Breakdown of the number of images used in classification

After exploring the images which were captured, we classify items of interest from our observations. We can classify them as follows:

Classification		Example
User:		
	Direct Observation (Stare)	see Appendix C, figure 8.16
	Subtle Observation (Glance)	see Appendix C, figure 8.17
	Engagement (Enjoyment)	see Appendix C, figure 8.18
	Engagement (Informative)	see Appendix C, figure 8.19
	Looking Away (Partial)	see Appendix C, figure 8.20
	Looking Away (Full)	see Appendix C, figure 8.21
	Distraction (Human)	see Appendix C, figure 8.22
	Distraction (Other)	see Appendix C, figure 8.23
	Grouping Orientation (Back & Front)	see Appendix C, figure 8.24
	Grouping Orientation (Side by Side)	see Appendix C, figure 8.25
Environment:		
	Objects (Inanimate)	see Appendix C, figure 8.26
	Objects (Animate)	see Appendix C, figure 8.27
	Illumination	see Appendix C, figure 8.28
	Text	see Appendix C, figure 8.29
	Activity (Human)	see Appendix C, figure 8.30
	Activity (Other)	see Appendix C, figure 8.31

Table 5.2: Classification of image specific areas of interest

As shown in the list in table 5.2, we can split the data up into the audience at the screen or items in the background. All of these provide salient information which can be used and combined to add context to existing activities and media. It can also be used as provisional incidental information once used. As we are interested in the viewers of the display to see what information could be obtained, we are interested in identifying people present at the display which results in considering the user and items surrounding them as identifiable objects that can be classified.

In figure 5.15, we show a breakdown of the number of images identified with the classified items of interest detailed in table 5.2. Some of the images contained more than one classification, this is why there are various levels of image numbers.

Figure 5.15: Breakdown of individual image classifications

Focus and attention of the users was not always centred upon the display due to outside distractions. The ability to focus upon who, the number of people attending, and watching the screen is of interest when we place it within the context of incidental information. With the numbers of viewers, over time we can provide a gradual rating system for that time and place with respect to popularity and mood within that location.

Using collected information we can then highlight and provide feedback to other locations, which in turn can provide a semi-collaborative and networked demographic of mood surrounding a location.

We will now present some of the images which were obtained from the study.

Library

Part 1 We recorded pictures obtained from the webcam placed upon the screen looking out from it. In figure 5.16 is the typical setup with which people interacted with the display.

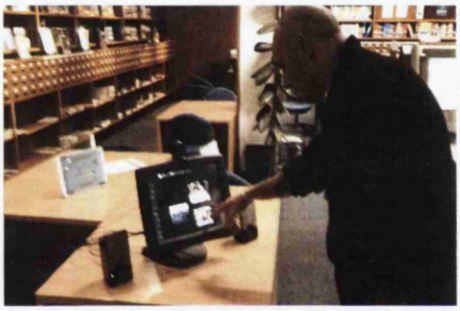


Figure 5.16: Interaction by user at the interactive display

While the user was interacting with the system the aforementioned interaction cycle shown in figure 5.10 is taking place. This records images varying in audience numbers. The base image (when people are not interacting) is shown in figure 5.17

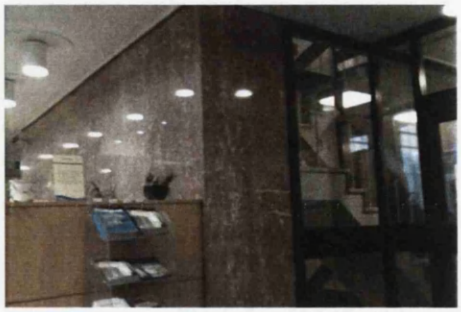


Figure 5.17: View from interactive display - destination 1

Part 2 Due to renovation in the library, and we wanted to explore another location. To see what impact this location had, we took another view. The new location is shown in figure 5.18

Figure 5.18: View from interactive display - destination 2

The images captured from this area are shown in figure 5.19

In the location of the library we explored the aspect of a public walkway. Using this information has allowed us to gauge whether there is any additional information that could be obtained by considering surrounding background information.

We have the user/s of the system in the foreground, but what about the background activity? It was observed that people pass the display with a periphery glance, but unless there is a group present or a friendly face they recognise, they were unwilling to participate within the interaction of the system. Presented here is an initial novelty of something new, but conformance fits it and that engagement is lost.

When we consider the images from the library, we can see there are various items which can be used for, and can be considered as inputs (incidental information fragments) for incidental interactions.

In the library we can see that the salience of the passer by, in the background with a glance, provides one opportunity to acquire additional information.

The images showed the gathering of people once someone has initiated an interest with the display. Information present in the background, such as names or text upon posters can be used to provide additional information about the surrounding area. This is useful for eliciting the surrounding context and location of where the interaction is taking place.

Smaller numbers of people seemed to be attracted to the screen. This is either due to its form factor or purpose of which they seemed to stop. The issue with the topic of media chose which only seemed interesting to a few people.

Figure 5.19: Images captured at destination 2

Sports Hall The recorded images from the library present additional problems and exhibit extra information which could be potentially useful and worthwhile exploring.

Figure 5.20: Images with extra background information

In figure 5.20 you can see there are various backgrounds which can be envisioned. Not only can you see individual people interacting with a display, but additional people present surrounding the main interactor with the system. Current systems developed are interested in interacting and consider the interaction with one person or a group in the vicinity. We try to achieve this by exploring what can be done visually.

If the display itself could detect the audience present, how might this help change and customise

the information presented upon it? In the images which are recorded what extra information can be revealed from them which could be beneficial and helpful for future interactions? The images also present the environment in which they are placed, and by inferring information sensed around them, this can be used to change information presented.

With this extra information presented to us which could be explored we have a research question presented, How can we obtain this extra information? and how can this information be used? We have shown that there is extra information that needs consideration with respect to interactions with large public screen displays. We explore this by developing an algorithm to explore the validity of the aforementioned ideas.

Research Laboratory Images obtained from within the research lab, demonstrate some of the graphical issues and possible incidental information, which can be obtained from the scenes that surround the user of the screen. We can see areas where the effects of light are visible within a scene and issues related to occlusion. Incidental information can be obtained by additional items in the background such as new people to the scene or contextual cues which can be obtained from items present.

Figure 5.21: Images with extra background information from lab position 1

We raised the camera so that it was at eye level for the second location. This is shown in figure 5.22.

By analysing the the images obtained at these locations we are able to answer one of the aims set out in this study. This was "What can be automatically deduced from scene images at a display". we have highlighted items of possible interest which can be used to provide information from a scene and be used for later purposes, in our case, future media that might be presented.

Figure 5.22: Images with extra background information from lab position 2

5.3.3.3 Logged Information

In table 5.3 the number of logged interactions that were undertaken while doing the study is shown. The logged interactions cover the details described earlier about the individually recorded items. We classified a log action to be one complete cycle of the video viewing process, as depicted in figure 5.10.

Location	Groups	Total Participants	Logged interactions	Logged interaction allowing Image usage	Time	
Library	18	25	84	53	14 hrs	
Sports Hall	8	21	15	12	4 hrs	
Postgraduate Laboratory	7	11	17	11	1 hr 30 minutes	
			116	76	19 1/2 hrs	Total

Table 5.3: Number of Interactions

We classify a group as in this context as 1 or more, people watching the display. There was a mix in the number of group sizes, ranging from 1 to 4. From the logged interaction, we can see the around 65 percent, allowed the image to be stored at the display. This would suggest that they weren't afraid from using such a system, in its current context. The time allocated took into

consideration the sessions which were undertaken to ask the questionnaires and other gathered information required.

As you can see from the breakdown of results from the logged interactions, the overall time for the study was around 19 1/2 hours in total. This was built up from 2 days in the library. 7 hours each day consisting of two 2 hour stints and one three hour stint, 9-11am, 12 - 3pm, 4-6pm respectively. Two days were carried out at the sports hall. 2 hours each day, between 5-7pm. 1 day was carried out in the post graduate research laboratory, consisting of 1 1/2 hours, at 2pm.

With the variety of times, we were able to gage what effect the location had upon social acceptance and to consider the types of people with whom might use such a system.

With the logged data we retained from the study (116 interactions logged from the running system) we can explore the timings between the media used and the length a person spent viewing the display. This is shown in figures 5.23, 5.24 respectively.

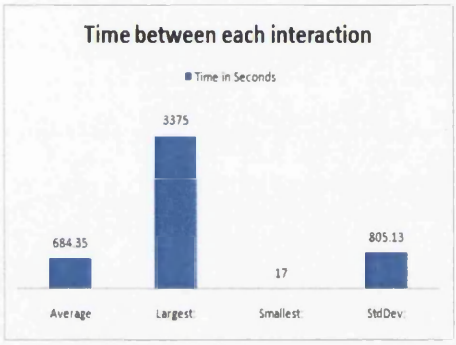


Figure 5.23: The time between each interaction

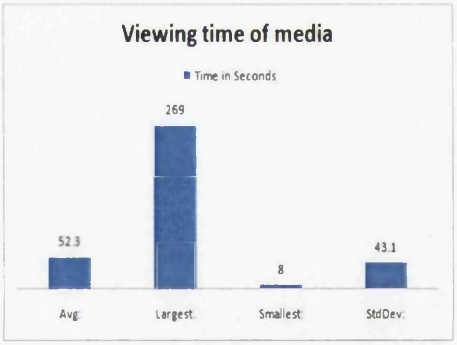


Figure 5.24: The viewing time of the media

When we consider the videos displayed our results reveal the appropriateness of the selected media across the locations. It also shows the types of media that people chose to view at the locations. In figure 5.25, we show the overall frequencies of the selected categories. Provided with the results is the amount of interactions from the users that allowed images to be stored. The statistics highlight which videos were more popular with regards, to topic. With an extended period of time this might show and help provide informative details about community preferences relating to preferred media, at different times and locations.

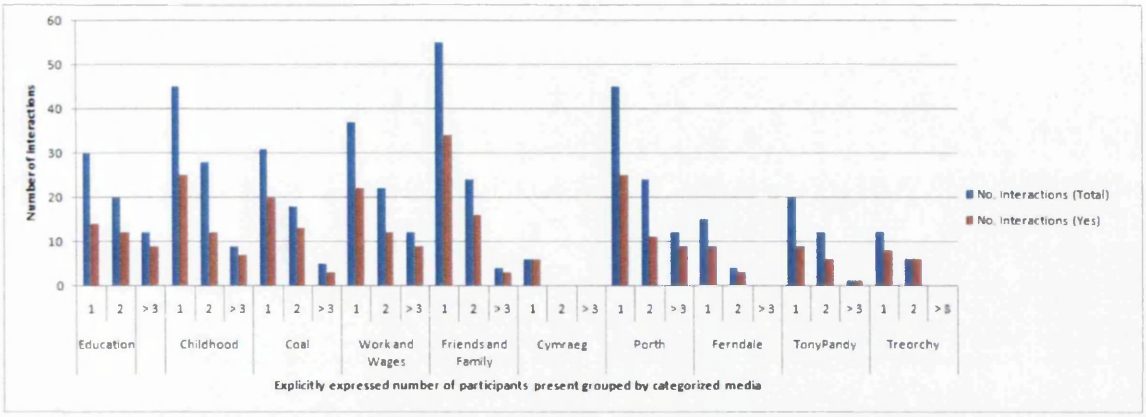


Figure 5.25: Category selection

There is a lot going on in figure 5.25. For each interaction that takes place in a certain category we show, how many interactions contained a certain amount of viewers when displaying the videos. This is categorised by 1, 2 or more people in the horizontal axis for each category. In the columns of these split up categories we show the total amount of interaction which relate to, and correspond to each category and the amount of people present. We compare the total amount of interactions taken place, against the amount of interactions from which the users consented to allowing usage of their images. This correlation could reveal something about whether group dynamics altered the acceptance of image acquisition than solo viewing.

5.3.3.4 Qualitative Feedback - Questionnaires, Observational & Usability Issues

When the participants were questioned about the possibility of filtering the content by the number of people who had previously selected and watched the video. We found that people endorsed the idea. It was revealed that they have been accustomed to this form of filtering, which is exhibited by other media experiences on the web. A similar idea is used by ``YouTube'', when using view numbers. This idea leads to some interesting issues brought up by the participants; that most of them did think that a higher number would have an impact on their decision of future video selection. Some outliers did say that ``it depended on situation and goal''.

From the studies we found that the subject matter of the videos was not interesting to most of the participants. From the observations of the two demographics, it seemed to engage an audience with more of a technical background when compared with the rest. When we enquired them about their subjects or backgrounds, it was revealed that indeed that they were engineers or a technically oriented subject. One hypothesis for this observation was due to the presence of a touch screen and camera. We also found that people with a humanities related background, were interested in the study due to different reasons. They were made interested in the topic and the purpose suggested by the system and the advertisement.

Building on from the original Rhondda collage system, we added an additional screen asking the participant for input with regards to asking them to explicitly indicate ``how many people were present at the display with them''. This is shown in figure 5.7. Initial feedback about this feature

was that it did not suggest that the additional screen effected their overall interaction with the system. One response that was found was that of shock when an image of them was presented with the option to not keep after the video. The audience members appreciated the reason and possible usage, once they were told about it. With the added benefit of knowing that they were able to delete the image if necessary.

One interesting observation from interviews with the sports people, was that they seemed to exhibit greater acceptance of the concept, and were more enthusiastic when trying the system. There was also a higher rating and more positive feedback when using the system. This could possibly indicate earlier adopters of newer technologies and practises.

Due to the placement of the display and the issue of being in a public space where the sound might not be loud enough. This can be associated to noise restrictions at the locations, public noise level issues, some of the users did suggest using headphones instead of speakers. We originally considered this in the design of the study, but we decided and had the assumption that having them would affect usage of the system. We didn't think that the people using the system would tolerate wearing headphones, due to being public. It turns out, that some didn't mind about possible embarrassment. This would be something to consider in the future. This has proved a worthy survey resulting in a fruitful collection of data, which can help reveal some interesting views and perceptions that can be formed with user studies with situated media experiences. It can also help explore the way in which we can engage the user to interact with such a system.

One additional, item we asked the users of the systems after they had finished was to fill in a SUS questionnaire. The results revealed an acceptance, and usage of such a system, in these locations studied, as shown in figure 5.26. These results are broken down further in appendix A, section 8.1.2.2. A surprise was that the sport hall was the highest acceptance. This we hypothesise, is due to the novelty factor, but also due to the number of participants involved. It has given us some indications as to which locations could be used for future use.

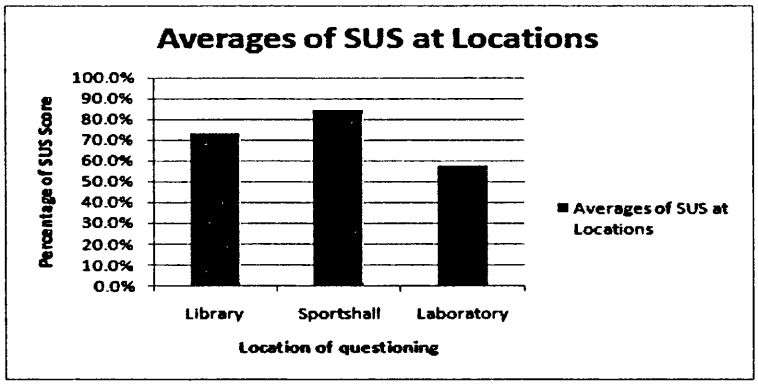


Figure 5.26: SUS averages

From our findings we see a similar trend across participants. They expressed interest in the idea of using image capture to help future interactions with a display. They also could see the benefits which it could bring to people located in the vicinity. This was expressed by one of the researchers from sociology, who said "Yes it's useful to know how many people have been watching it, then if I heard people talking about it then I would want to go and watch", they also expressed concerns

about acceptance due to chosen technology for the elder audience, "It seems like a good system because I belong to the computer age. But what about the more elderly users who might be unfamiliar with touch screens?". This an interesting point which would need to be explored further.

Some of the users linked this idea to a possible application of this research, the idea of advertisement. They could see it as a sort of information display board for people new to a location. Some expressed that they have seen similar work, such as in bus stops. This idea though is based upon Bluetooth device detection.

One problem that was expressed by the users was that without their interactions being actually, physically visualised before them, it was hard for them to fully envision our system. This could be additional work which could be looked at, so that they could fully appreciate the experience.

We compared the choice of interaction techniques from across the studies. Our results revealed that people were willing to use the system presented, and they considered that camera usage was acceptable as a way of exploring this modality of interaction. Overall they preferred the touch screen due to past experience with different systems. One possible explanation could be associated with that of losing self control. By being implicitly detected instead of the user explicitly initiating the interaction by themselves.

As shown in figure 5.27 and broken down further in figure 5.28, you can see the overall preference chosen for interaction by participants of the system. The results suggested that a touch screen was the more preferable interaction style for proposed changing of media.

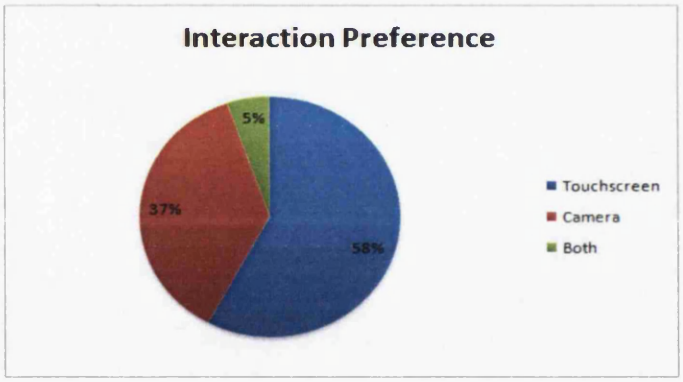


Figure 5.27: Revealed interaction preference

5.3.3.5 Uses of visual content

From the images that were taken, the additional information that surrounds the user could possibly add value and could be looked at. For our study we were concerned on obtaining a number representation of a group present engaging with the media on the display. Future work could consider the incidental information present in the background. This can be used for inferential deduction and to add additional value to the current image being taken or used.

The notion of automated tagging of media to facilitate in the display of community awareness media has highlighted some interesting issues. Some of these being accustomed to hardware and

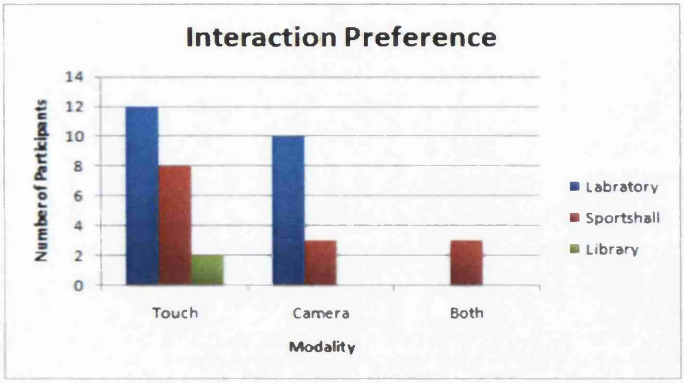


Figure 5.28: Revealed interaction preference breakdown

the premises associated with them, which have looked at the use of Web cameras and surveillance.

One avenue where such a system could be envisioned and incorporated into a real world situation would be the domain of advertisement, where the media would be more fluid.

5.3.3.6 Impact on algorithms

After reviewing the data collected, we found that using face and crowd images to detect items in a scene, which can be used to affect future content displayed, there were many limitations with the current setup which would have to be addressed in the future. Most of them are still large problems associated with in the vision community, issues regarding light and depth to name a few. This would require further research, but as our research was just a study of the concept to gain insight of a possible application from this domain, we think that this was sufficient to show the concept and initial prototyping to explore the notion.

When considering the feasibility of the algorithm, we found that for basic images that are not too noisy, the system seemed to provide us with results on the test bed. We have discovered from early tests, which further development would need to be required. One way in which we could extend the reliability of the system is by bringing in additional sensory data like, wireless options and possible additional cameras to help provide more information.

5.3.4 Discussion

In the study we found that due to the use of one camera, our viewing angle was limited to the scope of the cameras view. Future exploration could consider multiple cameras. Our concern and idea was to consider developing world situations where limitations of equipment and technologies exist. The factors such as limited electricity outlets would need to be considered and that of security.

While we have explored the possibilities of image based sensing to adapt content playback on large screen display, considering the acceptability and response to potential use of this data, further study would be required. We have revealed that there is possibility and potential in using such a system.

There would need to be more development and exploration into ethical and technical issues when developing such a system.

One of the main concerns was with the use of the image. Participants had feelings resulting from the fear of miss-use. We explained how and why we were going to use the captured images. Once a choice was presented to the participant on whether to keep the image, their initial apprehension on first sight of their image after watching the video, seemed to disappear. This leads to a question, would this same attitude / pattern be translated to the mobile scenario? Or has the uptake of such social trends (technologies), such as twitter and Facebook changed the attitude towards this? A deeper question is, even though people seem to be apprehensive at first, once given a choice, normally end up accepting and allow continuation? These are all interesting ideas which could be explored further.

The research questions we aimed to answer from this study, we have found out that there are ethical issues that have risen, which would need to be categorised for, which we did cover briefly by allowing explicit choice. We have revealed that people do value media that might be adaptable by viewing numbers, as revealed by participants in our study. It was suggested that they were influenced by such suggestions of popularity of an item, when viewing existing media repositories such as YouTube. This was succinctly put by one of the participants comments "I normally choose the higher numbered one with best picture". This highlights two possible factors which impact upon choice made for selection. By extending this notion and exploring it could reveal some interesting matters relating to community and public practices.

From the mentioned studies we have seen that topic, location and the social cohesion of an area have an impact on the uptake of and usage of such a system. We found in our study that people, especially the younger audience were more willing to interact when a camera was present.

It was hard to obtain participants for the study. There was an expression of future usage if it was in a more appropriate location with more aesthetically pleasing layout than the current setup. This was partially due to hardware and location issues which would need to be considered in future work.

The notion of losing ones self control to an information display system had split reviews. This is backed up by the results of interaction technique preferred, shown in figure 5.27 In some ways the findings revealed that some didn't mind having the system take control, as the users commented that it would be less of a hindrance when they were busy. Some had reservations, and expressed that they didn't want to leave current practice, which involves the physical touch of a system. It was revealed that they liked to have feedback of it as instant visual or tactile feedback.

One interesting idea which was revealed from this research, was that the system exhibited and fostered the identity and history of a village through the media. It allowed a way of people to gain insight to different contextual information, that they would not normally see, unless actively seeking it out. By placing this communication medium at different locations, allowed us to explore how the uptake of such information can be received within different public spaces.

We have seen the impact of introducing an interactive digital repository prototype with a web camera can have, and how automated capture can invoke such interest, but the question is how we can extend this interaction?. We can explore how to use current activities using mobile devices and how we can incorporate this by extending the message (voice) of the community.

The use of collected images obtained can be used with or against a detection algorithm to detect

presence of people in the scene. The identifications made from the images could facilitate in the automation for rating the media. We have considered this idea as this was one of the purposes of the study.

One way we can extend the interface is by placing values with the media and by altering the filtering menu giving the option to view by most interesting. This can be gauged by number of people present at display when video is playing. This idea could be extended further by having similar media items transferable to user devices for later usage. A visual representation of similar clustered media could also be explored.

There are some issues with adding viewing figures related to the media viewed by people present. If people join because other people join, we take this to be an indication of popularity as a side effect brought on by group activity this could be an interesting vein to explore. We could view all the people and see if there is a cut off after a certain period, which could indicate a possible novelty factor. Or take a certain distinct number as the target compared to all present.

One possible future plan is to illicit further information regarding the automatic collection and augmentation of community media by using the feedback from the study. We will consider porting/extending this system to mobile devices and consider what additional features can be tapped. One such route is to extend this using a game scenario as exhibited in Story Mashup [Scheible et al., 2007].

One route is to extend this work by using the collected data obtained from the study and to consider the reuse of the information to help promote the media which was popular at the time of study, and to investigate selection changes. Linking it in with social applications such as Twitter and Facebook, seeing what can be done to promote such community repositories are other areas of potential future work.

In a different direction we could change the topic of media used relating it to the student feel, as was mentioned by some participants for relevancy. This idea could link and track this as a sort of game where a treasure hunt for a community message could be waiting. In addition, a time element could be introduced to stop the study, this could be used them to explore the limit and possible impact this time interval had upon the study. Lives taking from the gaming metaphor could be used with this. If a location (GPS tracking) is not reached or shared media is not passed between users by time set a life is then lost. When a certain level of life is depleted, the information is stored at that location to be picked up, by somebody (a digital marker/tag). This incorporates game theory, social habits and location awareness domains coupled with mobile devices into that of supporting and exploring community communication (transitional media).

We have identified and highlighted some of the possible items, and cues which can be used for identification and can be classified as a form of incidental information. We are now interested in how we can identify and acquire these forms. The way we propose is by incorporating computer vision techniques, and in the next chapter we will develop and explore possible ways of achieving this acquisition.

5.4 Summary of Sensing Study

When we look back at the aims which were set forth at the beginning of these field study, we can see that we were able to explore most of these issues. In section 5.3, we explored the acceptability of the use computer vision techniques in a public location. This was carried out in the form of a web camera attached upon a touch screen, using questionnaires to illicit feedback of the audiences experiences. The results suggested an acceptance to the idea and use of these techniques.

From the images obtained in section 5.3, we were able to analyse and highlight potential context related items of interest, which could be used for automatic detection in a scene. We will go on to develop algorithms to achieve this in chapter 6, section 6.1. In the process we explored some of the potential techniques, and existing algorithms which could be used. We started to explore the idea of content adaptation and filtering of the media, by logging the data from the actions of the user. This was discussed in section 5.3.2.3.

By using the data captured as input, possible further development could be taken to extend the current system, to see what effect this new information has on the participants over a extended period of time. Overall we have explored the possibilities that present themselves when using data in the form of incidental information presented by people's actions (in logged data), and the issues that arise when using images of users at the display.

From this exploration into the idea and development of a technique for detecting audience members at a display, and the possible usage which can be revealed from the data for facilitating in possible dynamic changes to the media presented, we have shown that the acquisition of extra information can be obtained but further development would need to be carried out to make attainable.

With respect to the notion of Incidental Interaction, the findings show ways in which with the use of computer vision techniques we are able to obtain the possible actions of users present at displays. With the detection of users, we could provide ways to infer possible interactions that could, and might be used to highlight and classify actions taking place. This can lead to obtaining extra information allowing us to augment existing contexts of use, which reside in a scene of interest. The application of such detection techniques can provide for a more unobtrusive interaction style, while providing the benefit of collected information for extending current information needs of the user, by the provision of this incidental information.

When we explore the concerns resulting to ethics, it was revealed from our studies and that of others, that people are conscious and present concerns of privacy. This was related to the usage of the recorded media that was to do with themselves. Though as it was revealed by our research, people were willing to participate, once they were informed and explained to about the purpose of the study. It was revealed that, as long as the study was in the interest of research and we promised to delete recorded media after use, they allowed continuation.

When we look back at the scenario described at the beginning of this chapter, we gave a simple account of a busy, organised person. The interactions that are afforded by the notion of incidental interaction can be that of a more relaxed pace which transcribe over time. It can be used for short term occurrences, but to obtain and explore full benefits, an extended study would need to be undertaken. We have offered a form of interaction which can work with the user of a system, by providing community gathered information over a time period in the form of a information kiosk.

From our observation, we could see that people were interested in current activities that were going on around them. By providing them also with a history of that location can provide benefits, such an association with a location they had not previously experienced. There is also an issue of additional information, which can be provided to the user of the system. Just by their presence they are helping provide information to the system, alleviating them of effort on their behalf.

5.5 Conclusion

In this study we have discovered the impact of introducing an interactive display containing a media repository of media associated with another place (the Rhondda lives videos) see section 5.3.2.3, coupled with an image capture device has on an interaction. We have highlighted salient features that can be obtained and identified within a scene, see section 5.3.3.2 and can be incorporated as a form of incidental information.

We were able to log and record participants interactions with a digital repository device, which was used to highlight interactions with media selected, discussed in section 5.3.3.3. By using the media as a catalyst between user and system we were able to record images which could be used for detection of audience members.

We found that their was a preference for touch as the form of interaction rather than being sensed by the system, this was shown in figures 5.3.3.4, 5.28. The digital video repository was liked and usable by the participants of the system, this was shown by the qualitative remarks made in section 5.3.3.4 and the results of the SUS questionnaires shown in figures 5.26, 8.1, 8.2, 8.3, 8.4, 8.5.

We discovered that there was a privacy concern with regards, to capturing images of people interacting at the display, as shown by the qualitative remarks made in 5.3.3.4. This was a re-occurring pattern observed across the three locations.

In the next chapter we discuss the development of an algorithm to acquire the identified items of interest from in a scene.

Chapter 6

Visual Processing

Contents

6.1 Algorithm Development - Detecting Presence at an Interactive Display

In the last chapter we discussed some of the issues and concerns that arose from capturing images, for use with computer vision techniques. In this chapter we will discuss the development of algorithms which can be used to obtain the identification of the audience numbers in front of the display.

6.1.1 Work related to choice and usage of technique

The detection of people in a scene is a fundamental problem within the computer vision field. Although our system does not consider this specific issue in depth, there is still a need for basic techniques from these domains which can be used to obtain a basic visual detection system. Research related to these systems has considered the correctness of identifying objects for detection ultimately leading to tracking.

There has been a considerable amount of work that has been undertaken with regards to object detection, Dixon et al. [Dixon et al., 2007], object tracking which is detailed review in [Yilmaz et al., 2006], and the detection of surveillance systems [Haritaoglu et al., 1999, Masoud and Papanikolopoulos, 2001, Elgammal et al., 2002]. These incorporated various techniques considering, Bayesian Networks, machine learning and HMMs. We wanted to explore a less computational intensive strategy.

In some cases ours will be similar, as they try to use vision techniques to identify number of people. We are interested and our strategy is more in gaining a number of people actually viewing the scene.

This will explore the application and use of this idea within the context of large screen displays. We are not just trying to identify every occurrence of a person. The search space is reduced, as we are interested mainly in people in the foreground, but also to explore additional salient features which might be present in the background which could prove informative.

We were considering the ``what if a system, which is less accurate but has a higher detection rate'', of people. Would and could this be useful in revealing community awareness compared to having a specific face identifier, but not an overall awareness of people in a scene?

The algorithm which we used from the OpenCV library was developed in the work of Viola et al. [Viola and Jones, 2001, Viola and Jones, 2004] and extended in Phung et al.
[Phung and Bouzerdoum, 2007], where they were considering image feature edge density in real time processing. We will use some of the features in our system, via the OpenCV library which implements ideas from these [Lienhart and Maydt, 2002].

Work carried out by Cheng et al. [Cheng and Takatsuka, 2005], considered the use of distance with the use of a camera as a form of direct interaction using a monocular camera. They used the finger of a user as a point of reference, then using the camera did worked out the angle from the eye's along the finger to position on screen. this allowed for a point and select feature to be carried out.

The use of cameras have been considered with addition to public displays, but these have been concerned more today in activities of people for additional properties to that of current information [Antonietta Grasso,]. Flickner et al. [Haritaoglu and Flickner, 2002], propose a real-time computer vision system for detection of viewing habits for advertising company data mining. This runs similar to ours, but they considered the use of eye tracking as an alternative way of obtaining human presence.

The way in which our work differs from the previous work is that it applies the idea of combining the two techniques to be used into a public community awareness system, and we consider the notion of incidental interaction.

There has been other development which has considered other forms of implicit interactions using sensors [Dietz et al., 2004], for the dynamic change of information display. In CWall
[Antonietta Grasso,], they use camera technology similar to ours for the detection of users present at the display. They consider this as another way of sensing people present at the screen looking at the display. They used face detection to freeze information for the people reading near bye. Ours is concerns with counting these people and seeing the effect it has on the groups/community members when augmenting information present on the display, not just as a sensory device. We use the extra information which is provided for free.

A notion of a swarm of people described in [Nguyen et al., 2006], which considers the use of overall audience and movement in the scene rather than subtle points for interaction. Social aspects such as awareness can play a part in the interaction with systems [Erickson and Kellogg, 2000], and trying to discover this helps when looking at accountability, awareness in a social-technical system [Nguyen and Mynatt,]. Invisibility of ubiquitous systems is often cited [Weiser and Brown, 0924], and newer views with change in regards to direction have been proposed [Rogers, 2006].

What constitutes a glance? A glance[1]. This is a notion that can be associated with people when moving and present in a public space. This allows them to formulate whether or not to follow up their curiosity by another look or even to gaze for a longer period. We extend this notion by allowing the audience to provide input, which can facilitate in the notion of a group conscious for decision and to explore how using this notion we can promote community awareness through public display interactions

6.1.2 Aim

The aim of this part of the research was to develop and test algorithms which could be used to estimate the amount of people present at the display. By using the number of people present at the display we could explore the value that computer vision techniques can bring to the notion of incidental information capture. Using this information to augment the media with additional context cues to the existing data related to and from the user or environmental cues.

The study aimed to consider the value and the potential of using computer vision as an additional input modality, and to test the efficiency of three algorithms (techniques) for use in a multimedia vision system. The best and most feasible algorithm would then be used and incorporated with a display in a community setting. This will be used to explore the potential of information dissemination within a community context.

6.1.3 Methodology

To explore the feasibility and potential that the algorithms can present we used a iterative design method to develop the algorithm.

6.1.3.1 Procedure

To record and gain feedback from the prototyped algorithms, we carried out the study in the following order of steps :-

- Introduction of interviewer and Study Background
- Participants were given scenario to play out (see appendix c, section 8.3.2)
- Scenario was played out while system recorded images

The study involved the observation, recording and questioning of 12 people. They consisted of researchers within the department, who had different research backgrounds. This allowed for the expression of different views. The research areas encompassed by the individuals ranged from theory to graphics. With the size of our group, we can only generalize for small groups. Visual feedback was obtained from the output from the algorithms, we also conducted field notes of observations taken. We also carried out questioning at the end of the study, which helped probe the idea of whether using a system would be of value to them.

[1] is described as "a quick look", by Collins

In total there were 102 images recorded, which constituted as the audience viewing the interactive display and sample group for the feasibility tests of the algorithms. Before images were taken, the audience members were informed about the study, the purpose and were given a scenario from which they should enact their response to. See appendix 8.3.2 for more detail.

For capture of the images we initially carried this out using a paper prototype to represent the interactive display the user would see and interact with. This incorporated a Wizard-of-Oz approach to acquire the images which we used to test and develop the algorithms upon. After the participants had been informed, we began the study, which involved taking photo's as they played out their scenario. The whole study was carried out over a 2 hour period. In this time, we observed a variety of participants, as the audience numbers alternated. In table 6.1, we show the breakdown of the different sized groups for test on algorithms. Using these changing sized groups allowed us to simulate and capture the possible types of interactions that could be observed in a more public location.

Number of people	Occurrence
1	11
2	24
3	32
4	15
5	4
6	16

Table 6.1: Number of occurrences of group sizes in collected images

We first collected images for use on the algorithms, we then ran tests producing statistics detailing the results and comparison of them performance wise detailed in section 6.1.4.

These images contained differing numbers of people to see how effective they were using in-situ testing. Once we had evaluated the accuracy of the algorithms, we then extended and applied the usage of this detection system with a public display. It will implicitly detect the number of people present at the display. The result would be visualised and represented in the changing of the rating of displayed objects, and provide meta-data to the underlying data recorded. This can be combined with a traditional touch system, which would allow users to interact akin to web rating systems.

For the results and the experiment there were a few assumptions that we made.

- People do not stand completely still when viewing a display, which results in subtle movement.
- A face needs to recognisable with some body feature which can be used to be classified (shoulder, body)
- When a user is viewing a display, and engaged they are facing full frontal towards the screen.
- Peoples head position might be at angles in relation to viewing of the screen

Concerns *The approach of this study was to consider how computer vision systems could bring value to group community awareness within world situations, where technology is limited or is not*

feasible due to security and safety. Therefore we only considered a monocular view rather than more. This trade off we feel is acceptable due to the price and power that could be occurred. However for future evaluation once the detection algorithm is sufficient then more cameras could be considered.

For the process of evaluating the algorithms we went through the images and visually verified the actual number of audience members. We classified the detected people using TP - True Positive, FP - False Positive, F - False Negative, classification of results.

6.1.3.2 Apparatus

Hardware We kept the same equipment used for the capture stage. This allowed for consistency of the hardware and readings across studies. We used a single web camera (Microsoft LifeCam VX-1000) for the capture of the images again. The study was carried out upon a P4 3GHz Single Processor, with 1.87 GB of RAM, running Windows XP Professional SP3, running visual studio 2005 for software development. This was the mobile platform which allowed us to set up where needed.

Software

User Interface For the development and evaluation of the system, we had to create some GUI's to allow administration activities to be carried out, such as logging, and feedback, with regards to the study. These are shown and discussed within this section.

For the evaluation of the system, we incorporated Wizard-of-Oz effect, whereby we presented a white piece of paper to represent the screen that they would then interact with.

When the GUI is first loaded, the camera gets initiated ready for use, and the screen image holders are created ready to store intermediate steps. They allowed us to see at certain steps of the algorithms. This provided us with a way to validate the steps of the algorithm. See figure 6.2.

We extended the initial prototype; we considered explicit input from the observer of the system, which could be used to validate the presence of the visually detected representations of audience members watching the display. In the background the system automatically assigned values (meta-data) to the recorded images of the people detected, which could then be used later for recall, comparison and validation of the algorithm against the images.

We considered various applications and designs from which we could create to engage and record information of the user's activities to test and evaluate the effectiveness of the algorithm. This resulted in the GUIs chosen.

A breakdown of the software is shown in figure 6.1. It shows hows the GUI's are related to each other. GUI 1 encompasses the other two, which are selectable from upon it. We will now discuss the 3 GUI's in detail.

Both GUI's 2 and 3 were constructed to explore possible applications which could be explored, as a result of the information collected.

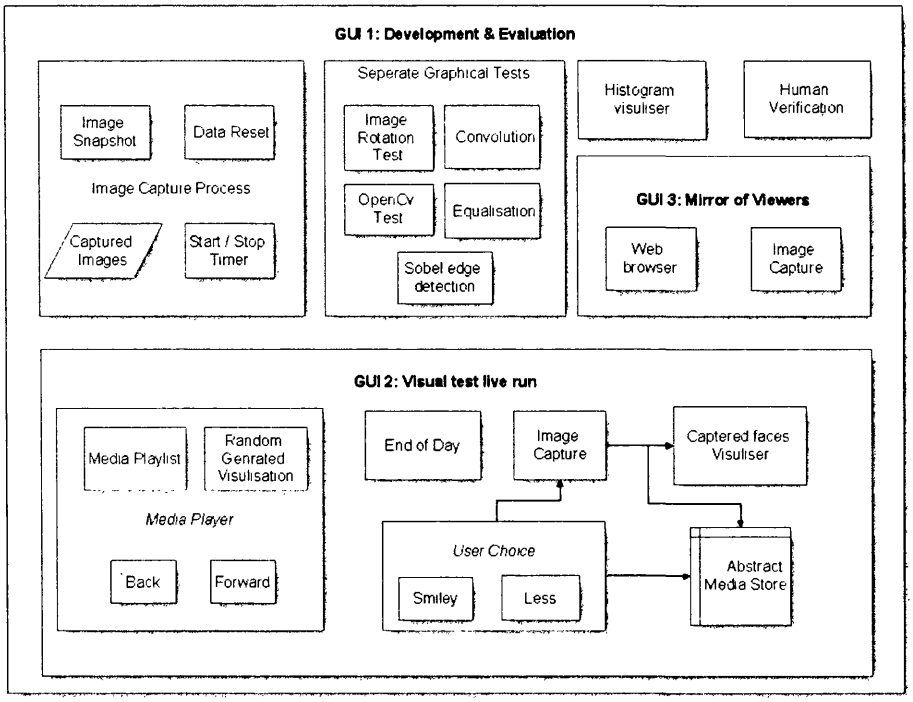


Figure 6.1: Development GUI breakdown of software elements

GUI 1: Development & Evaluation This is the main GUI, which will be used by the observer running the system. This is shown in figure 6.2.

The elements on the interface are presented, the two histograms of the current 2 latest images for reference are displayed in bottom right of figure 6.2. On the top right there is also a way provided to allow the observer to log detail about observations. This could be used for validation of the system at a later time. The 2 top left images represent the base image (left) and the current image (right) which are used by the algorithms. The other screens are some of the steps carried out by the BT(D/S)08 algorithms. This was achievable by splitting up and showing intermediate results as the algorithm was run. These areas are also used by the other, separate graphical tests as a place for the resultant images to be shown. With respect to the OpenCV wrapper using the library, the output only returned the detected patterns, so no intermediate steps could be envisaged.

GUI 2: Visual Test Live Running This GUI was created to visually inspect and compare the algorithms running, see figure 6.3. It demonstrates an interactive system which plays through a collection of media stored locally. It records the users present at the display detected by the recorded image, and visualises the detected people upon the screen. The images start to get recorded upon the activation initiated by the button presses of either the smile (more) or sad (less). This would emulate that of explicit activation with the incidental capture of faces assumed to be people present, in the form of a rating system. The system records their choice (more/less), which then gets added to a collection of meta-data stored with the image for latter reference and usage in a design abstract media type. We considered various ways in which we could store the data recorded. The various ways in which we considered storing the data recorded was, as C.S.V files, and excel document,

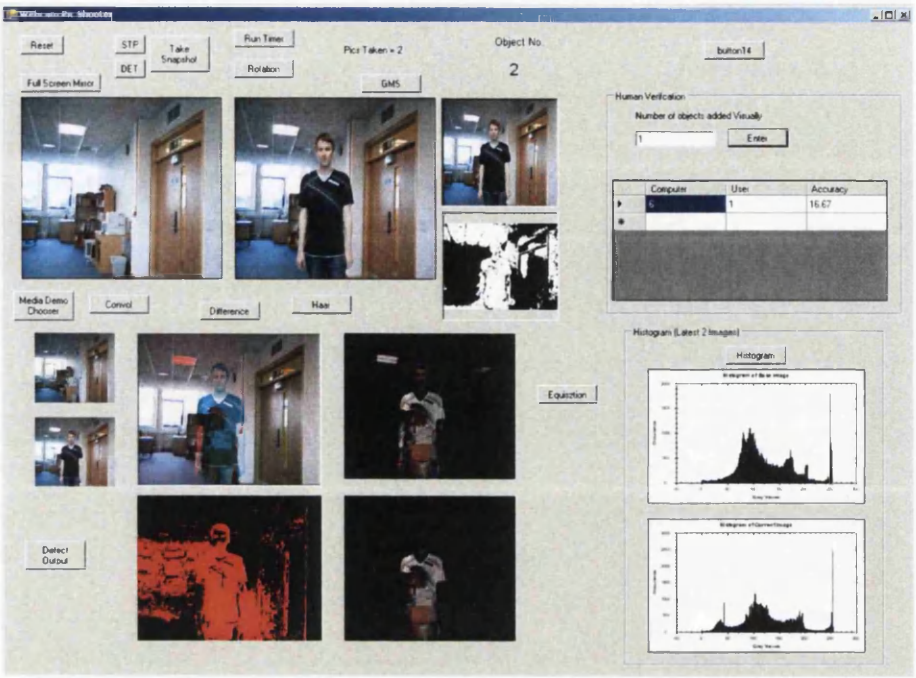


Figure 6.2: Development GUI

but due to the possible amount and complexity of the logged data, we felt that this was best served using the facilities provided by a DB. We decided upon a database form, which would allow for quick access and retrieval. Further development, would be needed to be considered and evaluated.

GUI 3: Mirror of Viewers In this GUI, we explore what reactions can be explored, and the input it has on the perceived media by taking the notion of mirroring the user when interaction is taking place. Similar work was also considered in [Roussel et al., 2004], were they considered security issues and the representations of presence that was detected at the display. When the user of the system got closer the information present upon the screen became more visible in the reflection, and when they did the opposite, further away, it resulted in a more blurred image. They explored how these interactions might change as a result of self reflection.

We started to explore a similar idea, and we created our own representation of this idea, see figure 6.4. The proposed idea would be to have, incidental information of one's activities possibly displayed, or that of the general community. The lack of person present was a result of a delayed reaction, due to user leaving the scene, the system will slowly fade back out ready for next user.

In the top right of figure 6.4, the representation shows a possible avenue by which an extended form can be constructed by using the number of people (inferred) in the presence of the display. More related topics of interest, related to videos, or the video itself. This would involve and allow the ability to have more control of the system to engage more with the idea of use of incidental information with the audience.

We used the BBC website displaying a general page upon it providing a peripheral awareness of the current topics around the world. Depending on presence of people at the display, then the media

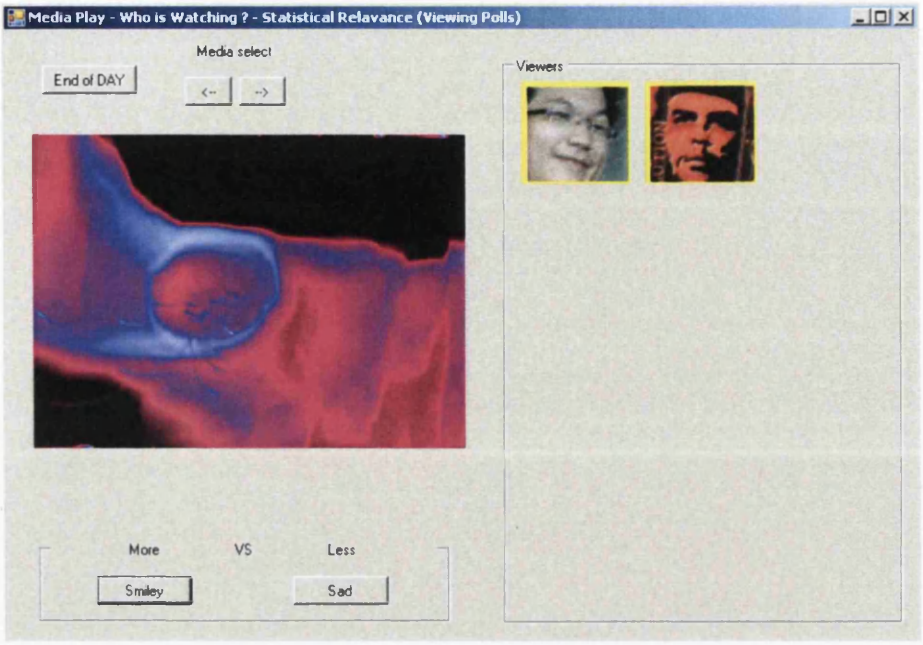


Figure 6.3: Parallel algorithms GUI

Figure 6.4: Self reflection

can change, related to what people chose. This could be considered through sensor detection, or extending the idea of mining viewing habits, the history of audience who were and had previously been present at the display.

We have explored the prototypes of the interfaces which have been developed. The main GUI (GUI 1), would be used by the observer to initiate the image capture, and the other two (GUI 2, 3) being possible extensions. In the next section we will describe the algorithms developed to detect, and which will be used by these GUI's for their input.

We will now go through the algorithms explaining their technicalities.

Algorithms

Technical We incorporated the techniques of various machine and computer vision techniques to isolate and count the members in a crowd. For the basic detection of objects we decided upon using Blob detection which was facilitated by the AForge.NET Graphic Image Library [afo, 0924]. By using this coupled with other techniques, allowed for the detection of movement. It could be used for the activation and obtaining of silhouettes representing objects in the scene, which we would take to represent as audience members.

We have also incorporated the use of an open-source wrapper library (OpenCVDotNet) [Ben-Israel, 2008] which works with OpenCV to provide functionality to .net languages, so that we could access the Haar face detection algorithm. This allowed us to develop in C# under the .net hosted environment. From the Microsoft OpenCV Computer Vision Library, we considered using the Haar like features detection algorithm for the detection of faces in a scene, as this was an efficient machine learning technique using boosting. Using this in combination with the initial object detection algorithm developed, would allow us to create a basic vision system. These two processes were used for evaluation of the detection of audience members within a scene.

With the use of OpenCV (OpenCV.Net) library, we considered full frontal face detection of people for real time detection. We used this feature as a validation for detection of people present in the scene, and with comparison with that of our own developed algorithm. Once the blob detection had detected initial objects, face detection could be used to validate, and provide clarification to the detection problem. The incorporation of the two techniques would help to provide object detection, leading to the counting / tallying of audience numbers.

We considered the technique of DMA, which looks at using the difference between consecutive frames as input. We found this technique not being particularly useful and applicable for static images. When applied with the assumption that many people don't stand completely still, it could be considered. Therefore it would allow for basic subtle movement / detection once initially identified. This would work for continuous shooting, but we are just concerned with time shots. For this algorithm we considered a static version (simple case) and a dynamic version over consecutive frames, which could lead onto motion tracking (possible future work), but we just considered the base test of algorithms initially.

We will now explain the 3 algorithms BTS08, BTD08 and Face detection individually.

BTS08 This algorithm which was created and then extended to evaluate and test against the OpenCV algorithm.

This basic algorithm requires the use of a base image, which will be compared to and used for the difference to be calculated between the latest frame and base image. The base image will be kept constant, allowing for the detection / presence of new objects (assumption people in our case) in a scene,

The base algorithm is discussed below:

```
Algorithm Basic Layout

-> Difference
-> To GreyScale
-> Threshold at 15
-> Opening Morphology
-> Blob Filtering (size 50 x 50 pixels) {Assumption larger objects People}
-> Gaussian Blur (1.4 {Gaussian sigma}),(5 {kernel})
-> Blob Counter
-> Get Rectangles from blob counter
-> Add bounding Rectangles to picture ( > 20 x 20 pixels)
```

The algorithm is a combination of various graphic techniques, which when combined allow for motion detection. One of the limitations resulting from using the provided library is that, it required the images of the inputs to be converted to grey scale, which removes possible information present with colour. This is a resulting limitation from using the AForge.NET library.

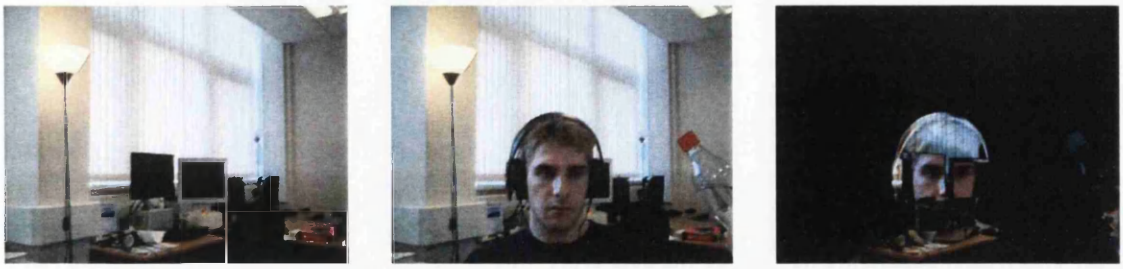
One of the advantages of converting the image to grey scale allows for easier detection of certain simpler features (edges), which can be built up to represent objects once grouped.

Going through the algorithm [2] :-

Difference - *This allows for the detection of different pixels. This is areas which are slightly different between the two images. An example is shown in figure 6.5. We take this to constitute movement, and by leading to the detection of a possible object new in the scene.*

To GreyScale - *The version implementation which we used of the AForge.NET library (version 1.6.2) required that the image was converted into a grey scale for the processing to continue. We also decided to switch to grey, to reduce problems of detecting different colours. The resulting image is shown in figure 6.6. By using the grey scale image, we were also able, as a bonus to visualise the images pixel quality by the generation of a histogram. This is shown in figure 6.7. This can be used to select the threshold intensity value, and also the visualisation of possible equalisation of the values.*

[2] Definitions acquired from [Shih, 2009], [Cyganek, 2007] & Wikipedia



(a) Base Image (b) Current Image (c) Difference Image

Figure 6.5: Difference in images: Before & after

(a) Before (b) After

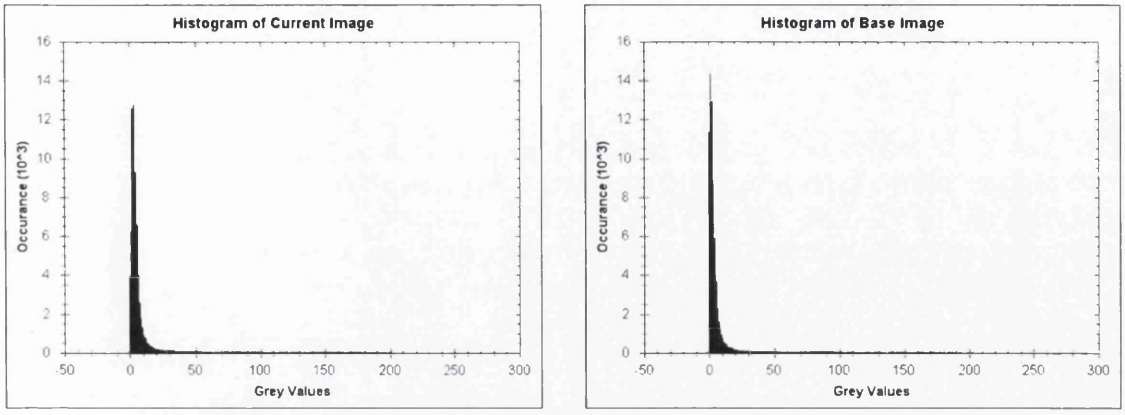
Figure 6.6: Difference in images after grey scale conversation

Threshold at 15 - *This is a changeable variable which allows us to indicate the intensity of the pixel that we would consider to allow through. This in turn could be used to represent an object, if we were looking for features, such as edges. By using the threshold we can reduce some of the noise in the image. The brighter the intensity, this can be used to reveal the change in the image. We considered an intensity of 15, for initial testing. The result of this is shown in figure 6.8. For further development see appendix 8.2.3, where values altered, which produced an image with less noise at a result of items lost.*

Opening Morphology - *This incorporates operations on the pixels which reduces possible noise that may be presented in the image. The use of opening results in the removal of small objects from the foreground (usually taken as the dark pixels) of an image, and placing them in the background. This is formulised in mathematical morphology shown in equation 6.1 , where $\ominus$ and $\oplus$ denote erosion and dilation, respectively. The sequence is dilation followed by erosion.*

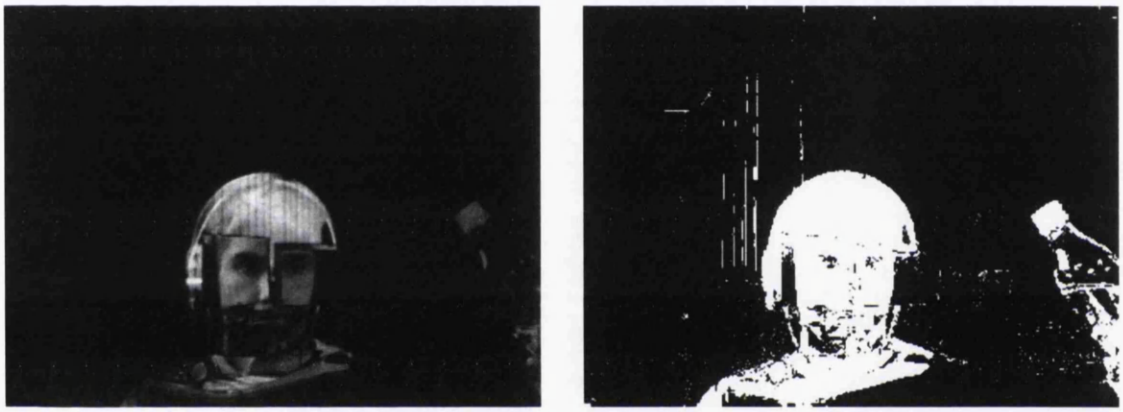
$$A \circ B = (A \ominus B) \oplus B \tag{6.1}$$

The results of applying this to the evolving image is shown in figure 6.9.



(a) Colour (b) Grey

Figure 6.7: Histograms of pixel intensity on images

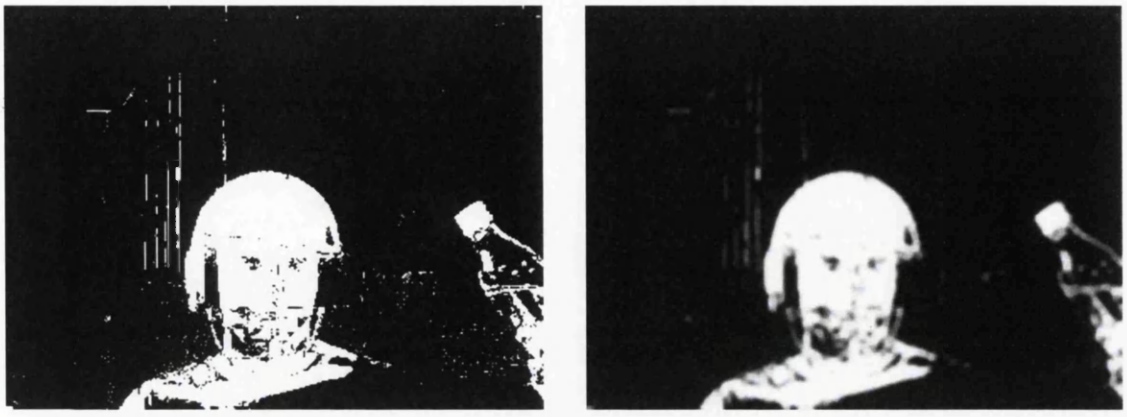


(a) Before (b) After

Figure 6.8: Difference in images after applying a threshold of 15

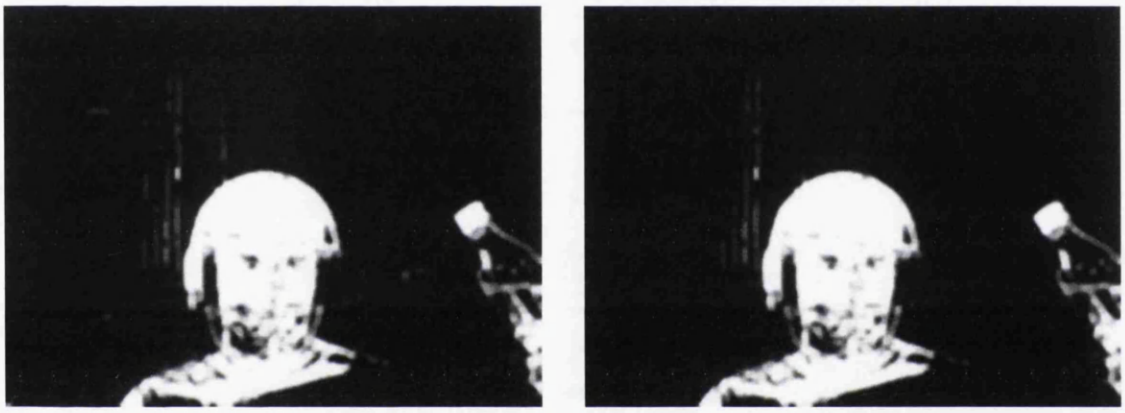
Blob Filtering (size 50 x 50 pixels) - *Filters the blob using a image segmentation technique which are to small (restricting the distance from camera for detection) to be classified as person. This process detects blobs which are bigger than* 50^2 *pixels in size. We decided that people present in the display would consume more of the real estate of the screen. Also by not picking up noise in the scene by very small changes which might exist. We are interested in people present, the silhouette. The results of the applied filter is shown in figure 6.10.*

Gaussian Blur - *We use another image noise reduction technique. The Gaussian blur is a type of image-blurring filter that uses a Gaussian function, for calculating the transformation which is applied to each pixel in the image. The equation of a Gaussian function in two dimensions, is the product of two such gaussians, one per direction (x,y). This is represented in the equation 6.2, where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and s is the standard deviation of the Gaussian distribution. Each pixel's new value*



(a) Before (b) After

Figure 6.9: Difference in images after applying opening morphology



(a) Before Image (b) After

Figure 6.10: Difference in images after applying Blob filter

is set to a weighted average of that pixel's neighbourhood.

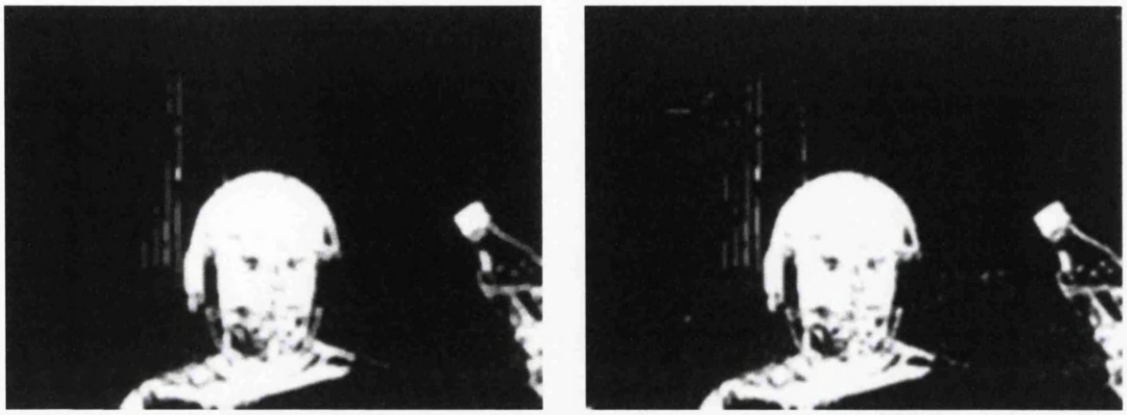
$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{6.2}$$

The results from applying this filter is shown in figure 6.11. This blurs in any little artefacts which might still remain in the image.

Blob counter - *This feature counts and stores the blobs that passed the filtering stage earlier.*

This incorporated the techniques of blob detection. *One of the first and also most common blob detectors is based on the Laplacian of the Gaussian (LoG). Given an input image f(x,y), this image is convolved by a Gaussian kernel shown in equation 6.3.*

$$g(x,y,t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/(2t)} \tag{6.3}$$



(a) Before (b) After

Figure 6.11: Difference in images after applying a Gaussian blur filter

At a certain scale t to give a scale-space representation $L(x, y, t) = g(x, y, t) * f(x, y)$. Then, the Laplacian operator shown in equation 6.4

$$\nabla^2 L = L_{xx} + L_{yy} \tag{6.4}$$

is computed, which usually results in strong positive responses for dark blobs of extent $\sqrt{t}$ and strong negative responses for bright blobs of similar size. *A main problem when applying this operator at a single scale, however, is that the operator response is strongly dependent on the relationship between the size of the blob structures in the image domain and the size of the Gaussian kernel used for pre-smoothing.*

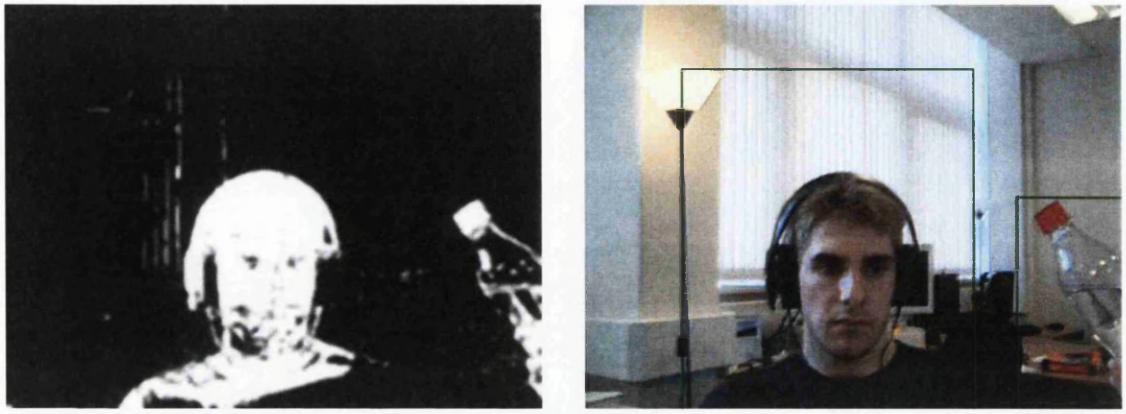
However we wanted to lower the possible input which could be read as noise. Detections over a certain size 50^2 pixels was sufficient, because the detection of people present in front of the screen would result in larger objects we hypothesised.

Get Rectangles - *This part of the algorithm uses the stored blobs found in stage Blob Counter. Using the information gathered, such as location and size, forms the bounding rectangles which contained the blobs area.*

Add Bounding Rectangles to Picture - *Using the bounding boxes of the blobs that were stored in stage get rectangles, we then drew then upon the original image taken by the camera. See figure 6.5 (b), resulting in a new image with the highlighted, newly identified areas visible upon the image. This is shown in figure 6.12 (b). Any object left which was greater than 20^2 pixels is highlighted.*

Above was a breakdown of the algorithm which we used as algorithm 1, in the next part we will describe the second algorithm which applied this technique differently. We used a different input for the base reference.

BTD08 For this algorithm we altered the referenced image for the difference step.



(a) Before (b) After

Figure 6.12: Difference in images: Before & after

The base algorithm is used on every consecutive frame, which allowed for continuous update of the background image rather than using a static single image captured at beginning. The use of an updated reference image when used to compare the images will provide short term answer to the problem with long term static objects in the scene when used for comparison against the difference. Another possible way of alleviating this problem would be to aggregate the average over time, before comparing the final image.

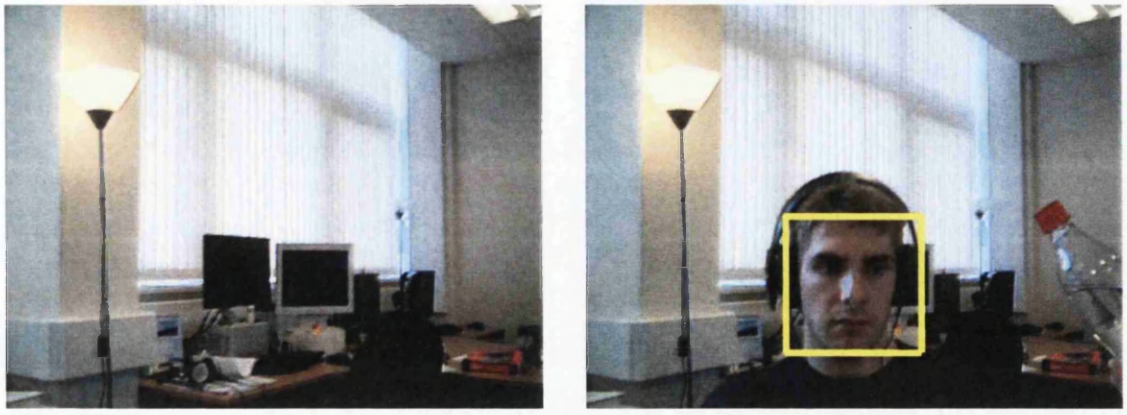
In the next part we will describe the algorithm used from the OpenCV graphics library.

Face Detection - OpenCV The speed that was seen from the detection of this technique which used Haar-like features using a boosted cascade classifier, allowed for quick detection of frontal faces. The image which resulted after applying this algorithm is shown in figure 6.13. The training set which was used, accompanied the OpenCV library, which was in XML format haarcascade_frontalface_alt_tree.xml. It detailed the results of a stump-based 20x20 gentle adaboost frontal face detector, designed by Rainer Lienhart.

The algorithm considers the use of simple intensities variation along the horizontal, then vertical of the image, which is used for detecting features of certain parts of the face. When combined, form and identifies whether a face is present, when compared to the training set. The algorithm then accumulates the rectangles of face detected per run on the image. We then retrieve and draw the identified rectangles, to draw highlights around the faces present in the image.

We found that, it wasn't useful for rotated or occluded parts in the image as it needed to build up parts to get the resulted face.

We tried to explore the problem with rotation, we considered rotating the image around 360 degrees. We would then test the image and record the various face detected which would need to be validated, when the detection of one or more was made. We would then need to check to see if images were the same, eliminating duplication of the same face leaving only unique faces. By rotating the image, we are able to identify any faces which, might be at angles to the normal. Some of the audience members might not be viewing the image in a true vertical stance, some might be



(a) Before (b) After

Figure 6.13: Difference in images after applying the OpenCV face detection algorithm: before & after

viewing from the side or obtuse angle, and this might not be detected. Initial trials are shown in figure 6.14. Part a shows various angles to demonstrate the idea. In part b, we show the temporary detected faces from the rotated images, before duplicate comparison and removal.

Setting The location that was chosen to be used in the development and testing of the algorithm was in the research lab. This allowed for fellow researchers to engage with the display, and provide feedback on the effectiveness of the algorithms. This allowed us to conclude which one was more appropriate for the task in hand.

As a controlled experiment for testing the effectiveness of the algorithms, we set up an area of a perimeter of approximately $2m^2$, a width of 278cm by a height of 208.5cm which produced the view of the camera. The camera was positioned at a distance of 356cm from a wall, at a height of 155cm off the ground. This created an area for the controlled experiment, from which we could then use. This allowed for use to gauge the effectiveness of the algorithms and set a ground for future development.

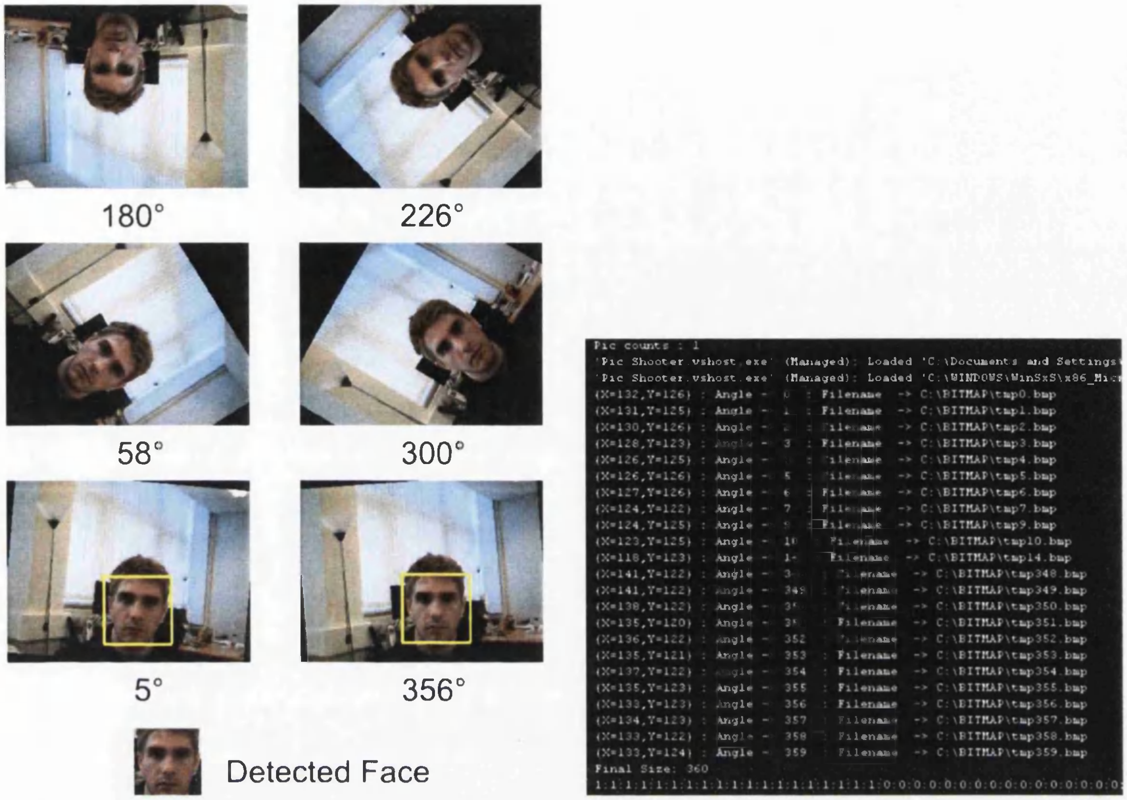
In the next section we will discuss the findings.

6.1.4 Findings

In this section we present the initial findings from this study, with the intention to give a sense of the robustness of these algorithms and the responses which can be obtained from audience members present at the display.

6.1.4.1 Comparison of Techniques

In the results, we were only considered with images which had one or more audience members looking (glancing) at the display, as this was one of the aims of the study.



(a) Image Rotation & face detected (b) Command line ouput of detected faces

Figure 6.14: Face rotation through 360 degrees using OpenCV face algorithm

Algorithm	Percentage of identified people
BTS08	100
BTD08	98
Face Detection	88.2

Table 6.2: People identified: one or more glancing/facing display detection rates

Detection Rates When we compare the results, we can see that the blob detection techniques (BTS08, BTD08) have a higher detection rate, than that of the face detection. One possible reason for this is due to the fact that their detection space is larger than that of the Haar-like features used for face detection. See tables 6.2, 6.3.

For the study we can break down the 102 images into the individual numbers of audience members lying within them (the occurrence of group size) shown in table 6.1. This helped us to explore the images more and we were able to count 331 instances of people lying within the collection of images. We then were able to identify and classify them to the classifications mentioned in section 6.1.3.1, these findings are shown in figure 6.17.

We can see from the results that though the blob detections identified more noise, when verified by the observer, identified more in the scene compared with a specific feature such as a face. By

Algorithm	Percentage of verified people
BTS08	59
BTD08	66
Face Detection	46

Table 6.3: Viewing audience members identified: detection rates

identifying the average, a higher number of viewers were identified. This is an interesting dilemma which presents itself.

From the initial usage of the blob technique developed from machine vision, there were limitations. The silhouettes of people when combined with links form a singular object, therefore false detections (less accurate result occurred, due to occlusion issues). This is shown in figures 6.15, 6.16. The result of the BTS08 is shown in figure 6.15. The result of the BTD08 algorithm is shown in figure 6.16. These images were acquired at runtime of the algorithm. This is where the incorporation and possible usage of face detection techniques for the detection of human audience could help. By using it as a second technique to verify false positives within a scene.

Figure 6.15: (a)Occlusion, shadows, (b)Grouping, (c)False positives

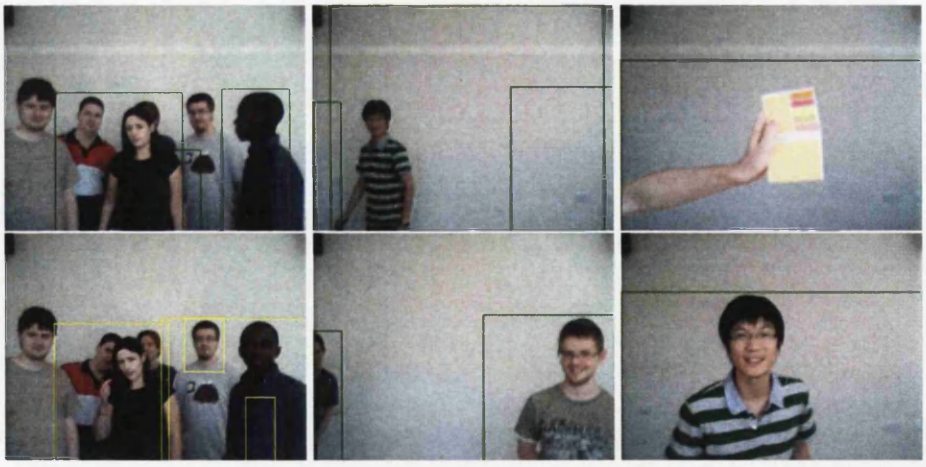


Figure 6.16: (a)Occlusion, grouping & false positives, (b)Shadows & identification (c)False positives & groupings

When we consider the results in table 6.2, which the results were collected by using the algorithms, it is clear that the detection rate for any occurrence of people in the scene is great due to noise. This resulted in false positives being identified which resulted in a high percentage of identified people.

This is shown in figure 6.17. The problem was a result of considering the use of a static background as a base image. Subtle changes accumulate into one overall change. An answer to this would be to split up the image further, by identifying features.

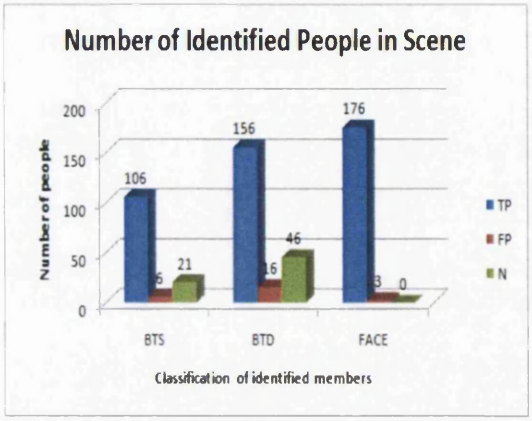


Figure 6.17: Classification of individual audience members

However one interesting observation that was revealed was the performance of the face detection in these scenes. It was lower than one would have expected. As mentioned in previous work, people have built upon this for detecting faces in pedestrians, however I was considering the base functionality and comparing it with that of proposed motion detection as a catalyst for future interactions, and to see which would be appropriate to continue with.

When we consider the results in table 6.3, which shows the results of the 3 algorithms when considering people actually looking at the display. The images were verified subjectively by the study observer using the classifications discussed earlier. We can see that there is a more realistic representation on the detection of viewers present at the display as detection of movement had been identified. This verified one of the assumptions that people, when standing at displays their movement isn't constant. However there was a lot of noise still detected within the image, which would have to be considered in the future.

We used Pearson's correlation coefficient [Field, 2005] to see if there was any correlation between the actual number of people and our algorithms, see table 6.4 for results.

Pearson's correlation is a method, which provides you with the ability of measuring the correlation between two variables of interest. It is given by the dividing the covariance of the two variables by the product of their standard deviations, and is shown in equation 6.5.

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \tag{6.5}$$

The value should lie between the values of -1 and +1. Nought being no correlation between the values. Values of $\pm.1$ represent a small effect, $\pm.3$ is a medium effect and $\pm.5$ being a large effect.

Our results show quite a low percentage on our algorithms compared with that of the OpenCV algorithm. This suggests that there was a considerable amount of noise in relation to the detection of people within the scene. With further development could consider and resolve these issues. For

our purpose, comparing against OpenCV, and our BTS algorithm allowed us to gage our algorithm with the base one created. This allowed us to see how our basic detection algorithm faced up against face detection, for the identification of average viewing numbers.

One interesting result we observed from this study was that the OpenCV implementation did not result in such a high detection rate as initially thought. This is shown in the correlation between the identified objects/faces and the actual verified objects/faces in the scene.

Algorithm	Correlation
BTS08	0.015
BTD08	0.215
Face Detection	0.426

Table 6.4: Correlations between detected and actual audience numbers

We will now consider the runtime performances of the algorithms.

Runtime Performance We wanted to consider the performance issues of running the 3 algorithms. In table 6.5, the average runtime performances over the 102 images are presented.

Algorithm	Seconds
BTS08	9.31
BTD08	9.14
Face Detection (OpenCV)	10.69

Table 6.5: Algorithm runtime: rates

We see from the results that the BT(S/D)08 algorithms were slightly faster than the OpenCV case. This is suspected due to the OpenCV implementation and lock of image operations, which are needed to be carried out.

We will now discuss the observations and themes found in the results.

Observations/themes within results Throughout the analysis phase of the results, there were themes that kept on appearing on the obtained images and from the algorithms applied.

These ranged from blob grouping issues due to light conditions (shadows), occlusion of full bodies, to missed data again with issues with intensities eg, such as skin colour. This is shown in figure 6.15. There were cases where subtle movement which was larger than then designated threshold used to remove noise, created some false positives. These did not constitute as an identifiable person as identified when subjectively going through results.

As we have demonstrated the factors that have influenced the final results. We considered certain settings with the threshold and blob sizes, which resulted in identifications but not totally conclusive for some situations. This would suggest that further time and the use of other techniques need to be considered to overcome these limitations in visual information systems. We see from the results

that the BT(S/D)08 algorithms were slightly faster than the OpenCV case. This is suspected due to the OpenCV implementation and lock of image operations, which are needed to be carried out.

From the collected images within the study, from our results we found that due to the limitations of the control experiments area bounds, only a maximum of 6 participants were present at any one time within the view of the camera. Any more would of been occluded due from view of the camera, due to placement and view angle.

6.1.4.2 Audience

From the observations of the users, we firstly detected there was a sense of novelty from the proposed system as with most things that are new. There was interest shown for the system from which they expressed "wanted to give it a go". One limitation with using the Wizard-of-Oz approach was that unless there was complete isolation with the system (which is very hard), the audience became easily distracted by either observer or fellow audience members present. Some of these issues were also identified and verified by some of the observations made in the study carried out in chapter 4. This limitation could be associated to the fact that they had to imagine rather than physically interacting with an tangible object. Future implementations would cover this be having a physical/visual counterpart incorporated, and we explored this issue with the proposed design of GUI 2, discussed in section 6.1.3.2.

Self-reflectivity From the individuals who participated, we found that using the Wizard-of-Oz was limited, but it served out purpose for collecting the images. The lack of real working system had an affect on the responses made, as there was no real life tangible media to interact with as highlighted by one of the participants response "but there's nothing there, its not real", which resulted in a lifeless response. It was with help of the wizard behind(studier) to keep the system running. This kind of feedback is notable but is one of the limitations that this method presents.

In the next section we conclude our findings from this study.

6.1.5 Conclusion

We have seen from the analysis of these three computer vision techniques that the problem of crowd number detection techniques requires further consideration. From our results we can show that neither algorithm was totally accurate to be used in a live situation, the results are shown in 6.3, only in a situation where a very base approximation needs to be obtained. Therefore in the current individual states they are limited in their effectiveness. A proposed combination of the various techniques, considering a hybrid of face and blob detection, could help overcome limitations presented for certain goals, like the one we tried to explore.

From the results obtained from our study, shown in figure 6.17, we can that face detection identified more individual people from the scenes, whereas the BTS and BTD algorithms suffered from noise symptoms, the identification of false positives and non-identified people.

Public displays in their nature are placed in locations which allows for public awareness. These locations normally have some form of security surveillance watching over them, or in a place

nearby. By extending the current uses from that of security, they could be used to provide a helpful means for advertisement and community awareness. By using their incidental information and applying it in a different context, adding semantics within that location.

The way computer vision can be used to facilitate this idea, is by the acquisition of audience members and the use of this identified quantity. Therefore help can be used and re-purposed back in circulation to the display and reused in other activities, the incidental information which is resulted as a bi-product of the current actions of users present at these displays.

The results presented from initial tests highlight some issues, but require further development to fully evaluate the current trade off's occurred from the differing algorithms, and possibly the incorporation of other sensory devices might provide additional benefits.

We have explored the use of background subtraction, coupled with blob detection to explore the identification of groups of people. We have proposed the incorporation of face detection to help verify presence of people. The ideas are acquired from across different areas. This has helped in the exploration and feasibility of the proposed algorithm. This is early work, but provides a possible route to explore. We have considered ways in which complement and consider another view at the existing literature that exists.

For the detection of audience members present at the public display when engaging, we have identified that the use of face detection can facilitate in counting members for use on dynamic displays. Blob detection techniques are useful as an catalyst for basic detection systems, but on their own can lead to issues highlighted.

In the next section we will discuss future avenues that could possibly be considered from light of what was revealed from this field study.

6.1.6 Discussion

The feasibility of using image detection algorithms as a way of obtaining extra information is feasible as demonstrated by our results. We need to find more accurate and efficient algorithms for the exploration of the image details to be more effective to be incorporated within such media systems. This is why the exploration of multi-modal interfaces could be an alternative at the moment.

The algorithms that are present look at specific features, such as that of face, segmentation, and basic shapes. It is with the combination of these separate techniques, we hypothesise will help in the definition and identification when exploring the presence of the people present at a public large screen display for future media augmentation.

We have seen that a variety of ambient conditions can greatly affect the reliability of detection methods for gauging the detection of people at displays. We have seen that due to noise in acquired images that a proposed combination (hybrid) of different techniques would need to be considered to alleviate the limitations of current techniques. One of the considerations with this work was to see if there could be benefits from the use of implicit interaction such as considered with detection of crowd numbers (a many to many mapping) compared with that of a more explicit (one too many mapping) interaction, and to some degree the initial building blocks and issues have been highlighted.

This study raised a lot of questions which need and could be followed up. We have demonstrated even though the detection of faces is quick, it is still not 100% reliable enough to be used in a real situation due to the limiting results that could be produced. Although if we want to know if a crowd is solo, the size of the blob might work. There is a vast literature on face detection, but we were interested in, how might we interact with this space, rather than on the technology.

Being interested in what additional information could be added to the media, raises possibilities for questions on how might the media be better used and catered for in the future. How could it be used within intelligent environments when interacting with mobile devices? Therefore a proposed phase would be to explore group to group interactions incorporating mobile devices within public areas such as cafe's, bars, shopping malls, and leisure centres. The possible interactions would allow audience members to communicate on the periphery, and access information by bringing in real life social metaphors to the digital awareness.

By incorporating face detection and similar detection techniques, we can facilitate in the production and acquiring of incidental information, such as the logged data exampled in chapter 5, encompassed within the idea of the incidental interactions that occur surrounding public displays. This type of information can be obtained from either the initiator of interaction or as a bi-product of background presence.

In the next chapter we will discuss and bring a close to this thesis.

Chapter 7

Conclusion and Reflections

Contents

In the previous chapter we considered the types of incidental information, that result from incidental interactions in the form of an interactive public display, and how we can use computer vision techniques to facilitate in its gathering.

In this chapter we will reflect on the information revealed from the studies, and draw a close to our research demonstrating how our research contributes to the existing literature. We will also express future avenues for the research. This research has been successful in applying and considering the idea of incidental interactions, from that of mobile phone with large screen usage. It has opened up a view of incidental information which can be a considerable feature and catalyst to the concept of incidental interaction. We have highlighted it from the aspect of mobile devices and public screen displays.

We have started to explore what is achievable, using this view and style of interaction to that of previous ones. It is an area which may prove insightful in future areas, such as AR scenarios (overlays), and intelligent environments.

We have shown the possibilities and possible avenues for a more relaxed, passive way of information dissemination and data reuse. This allows the user to focus more on their current actions, but still being presented with information, which facilitates in a sense of community.

With this use of incidental information providing input and functioning as helpers to a incidental interaction paradigm (style of interaction).

When we look back to the aims set out at the beginning, we can say that we have achieved what we set out to explore. We will now describe what was achieved.

7.1 Summary of Achievements

The main aim of this research was to explore, and expand the knowledge surrounding incidental interactions and to develop systems which take into consideration the notion of incidental interactions, with large screen displays. With such an idea, we have considered the issues and possible applications that could be envisioned. These studies explored aspects and considerations with that of asynchronous and synchronous communications which can take place with and surrounding the users of these systems.

We explored and illustrated the notion of incidental interaction. This has resulted in us carrying out usability studies to evaluate the usefulness of systems and to demonstrate the concept. We were able to demonstrate that incidental interaction can alter and supply additional benefits to users and their current activities. Possibly changing them after new associations made from new information revealed.

Through the use of computer vision techniques and mobile activities we were able to highlight what kinds and in what ways incidental information can be revealed from such activities. We explored the impact which resulted from showing this kind of information to the wider audience across different locations and what needs to be considered upon future studies.

Such a notion and technology not only offers users a novel and intuitive experience of interacting with multimedia data, such as text, imagery, audio, but it also provides convenient techniques which need to be explored further. It offers new opportunities for communicating between existing and newer devices.

Two publications have resulted from this research, which were presented at the *Computer Human Interaction conference 2008 Workshop on Designing and Evaluating Mobile Phone-Based Interaction with Public Displays* [Garnham, 2008a] and *MobileHCI 2008 Doctorial Consortium* [Garnham, 2008b] conferences.

Through this research we have shown that we can re-purpose information obtained on small devices and by displaying it through public displays can foster group activity different to existing social norms. The introduction of mobile activities incorporated with large screen displays offers a new user experiences, that might of been hidden before.

We have identified potential areas of interest, and explored ways of obtaining this new information and the ways in which it can be understood as incidental information. We have started to explore the social activity which is promoted through incidental interactions with information surrounding them.

Through the various studies undertaken we have shown the feasibility and use that incidental interaction and the incidental information within can be brought it to the forefront of interaction with large screen displays.

7.1.1 Technical Achievements

We have developed a range of prototypes to represent and demonstrate the idea of incidental interaction and the incidental information it can consider. We achieved this using the idea of tag-

ging (recording the information and adding reference), re-purposing (changing the visual presentation) and collating (bringing the information together), which resulted in bringing mobile and large screens together. It also provided us the chance of using visual sensing to detect users at the described locations, which could be used for change of community media.

We have developed two detection algorithms which we have compared with that of the OpenCV face detection. This has helped in the exploration of the feasibility of the two types of techniques used (object detection and Face detection). We were able to explore the potential use, and whether they were comparable, in the detection of users present in a scene.

We have extended the features of an existing collage system, which we used as one of the prototype systems to illicit actions of users. The logged details provided from these prototypes provide us with input for future development. We have created ways of presenting, collecting and re-purposing incidental information to facilitate in the dispersion of the idea that incidental interactions can foster.

Software issues were thoroughly researched, with the Java and C# programming languages ending up being selected to create the prototypes for the studies. This was part due to legacy extensions that needed to be considered and to consider possible systems for future use.

The technical feasibility of the ideas has been demonstrated, in that they provide a stepping stone from which can extend in the future.We have demonstrated the adaptability of the system allowing the updating of elements in the system. The house keeping functionalities including the ability to log interactions of the users for later perusal, stored within a database.

7.2 Review of Conclusions from Studies

The exploration of incidental interaction has been successful in that we have shown that it has potential for further exploration to see what can be achieved. We have shown possible scenarios where this can be achieved, and the potential the idea has, within certain contexts and situations.

In study 1 in chapter 4, we revealed the acceptance and novel aspects that reveal themselves as a by product of re-purposing mobile device activities upon large screen displays. We revealed some interesting social nuances that are associated with new media introduced into a public place. Such activities such as observational and inquisitive tendencies to explore a new item of interest were exhibited. One of the main ideas from this study was that of decoupling the experiences formed in their normal activities, that of mobile and large screen interaction and explore them within a new context. We achieved this by combining the two separate interactions and exploring the results produced.

In study 2 in chapter 5, we explored the acceptance and interactions afforded by an interactive display screen using a video repository as the catalyst to obtain incidental information at the location. Through logged data, images recorded and questionnaires we demonstrated what and how we could obtain incidental information within another activity.

In study 3 in chapter 6, we came to the conclusion that by a combination of methods would be currently required, to be able to capture the number of people present at a displays location. We demonstrated the development of a basic algorithm for counting people in a scene. This obtained

information could be provided as a by product obtained at an interactive display. We explored this notion within the study. Were we showed the develop of the idea.

Overall the three studies combine to frame the idea of incidental interaction and incidental information within the context of mobile and large screen display interactions. Hopefully we have highlighted some interesting ideas and work which can be built upon in the future.

Incidental Interactions successfully demonstrates a novel concept that helps facilitate in a more relaxed way of information discovery and dissemination. By considering this extra information and the ways in which we can use it, future interactions can be discovered, which were once hidden from sight. Unveiling the hidden.

We live with information in the periphery and by embracing the notion of incidental interaction we can hopefully bring the periphery into the foreground to provide that needed help, once in a while. There is so much potential which can be explored, we have only explored the surface of what can be achieved.

7.3 Further Work

While the developed algorithm was used to construct and test the concept, better technologies and techniques now exist which use neural networks and various other A.I. techniques. All though there is some debate still over the effectiveness of these algorithms, none are completely perfect, we can strive to make considerations on the features required until then.

We used the "Storybank"system as the platform for the media, this time as it provided a perfect bed for the story telling metaphor associated with communities. The media can be changed in the future, and could incorporate more social networking technologies for the developed world's to explore change in context and style of communication, but we have to consider what is important in developing countries also. Were mobile phones are more prevalent than broadband.

It is in situations like these where we have to consider the incidental interactions that occur and what can be achieved. We demonstrated a simple consideration of this by using a single camera web camera, which could easily be changed with that of a mobile phones. The text of a SMS between community members, could easily take the place of search queries on a map interface.

We presented incidental related large screen displays and interactive displays. We can extend this beyond these technologies and extend incidental interaction considering the engagement which is presented through it and how it affects situations. We briefly considered using the incidental information as a counting mechanism for people, In what other ways could we extend this? These are just some of the possible ideas that need exploring.

User interaction may be improved through more collaborative means. This may be achieved through considering groupware, and considering using modern social networks to inform and communicate. By considering the current community activities and subtle cues of users, we can spur on future developments to augment and extend them.

Chapter 8

Appendices

Contents

In this chapter we detail the additional material used, relating to the studies and other work carried out.

8.1 Appendix A - Interview Questions, and Questionnaires

8.1.1 Chapter 4 - Study Questions

This is the post questionnaire used, which allowed us to acquire qualitative feedback from the users.

Study - Public awareness of search queries

1. Did you notice the display on the projector outside?

2. What do you think the display is showing?

3. Do you think it is useful?

4. Is the text visible with the background image?

5. Would you like to see these being updated slower?

6. Less or More information associated with the text queries?

7. Could you see yourself using such a system, if it was available to you?

8. What could you imagine yourself doing with such a system?

9. In what ways could you imagine interacting with the display?

□Touch □Voice □Device (Attached to display) □Device (Mobile, e.g. Phone / PDA)
□Other _____

10. What feature do you think could be added to the system?

11. Any other comments?

8.1.2 Chapter 5 - Study Questions

8.1.2.1 SUS Questionnaire

Figure 8.1 shows the adapted SUS form, which we asked the users of the system to fill in, once they had finished one interaction.

For the calculation of the SUS scores, we used the provided explanation on pages 138 - 139 and 149 in [Tullis and Albert, 2008]. ``The questionnaire consisted of ten statements, from which half were positively worded and the other half were negatively worded. It uses a 5 point scale, which provides us with an overall usability and user satisfaction index. The intent is not look at the ratings individually, but only consider the combined rating. It is convenient to think of the SUS scores as percentages on a scale 0 - 100''.

``To calculate a SUS score, we first sum the score contributions from each item. Each item's score contribution will range from 0 - 4. For item numbers 1, 3, 5, 7, 9, the score contribution is the scale position minus 1. For item numbers 2, 4, 6, 8, 10, the contribution is 5 minus the scale position. Once the values have been worked out, we then multiply the sum of the scores 2.5 to obtain the overall SUS score''.

The scores calculated in section 8.1.2.2, show the results for the people who chose to fill in the questionnaire. Some decided not to fill in the form due to other reasons, such as time and willingness. We had to respect their wishes.

8.1.2.2 Further breakdown of SUS questionnaire charts

Figure 8.2, shows the number of people who were willing and filled in the questionnaires.

In Figure 8.3, we present a breakdown of the individual participants showing their SUS score in the library.

In Figure 8.4, we present a breakdown of the individual participants showing their SUS score in the sportshall.

In Figure 8.5, we present a breakdown of the individual participants showing their SUS score in the laboratory.

8.1.2.3 User Post-Questionnaire

This is the post questionnaire used, which allowed us to acquire qualitative feedback from the users. These questions were asked by the interviewer, and the responses recorded.

Choice of interaction type questionnaire

1. Which interaction technique would you prefer and why? (Explicit (Touch) or Implicit (Camera))

2. Do you think that visually showing you the viewing figures associated with previous interactions, could influence your current selection?

3. Does the current filter options provide enough option?

4. Would you prefer to filter by viewing numbers?

5. Giving you the ability and knowing, that you could influence the future media that could be displayed, does this entice you to interact more with the system?

6. Would such a system appeal to you and why?

7. What are your concerns with such a system? and why?

8. Did you mind the additional screen asking how many people present? and why?

9. Did the image taken at the screen detract from your experience of the system? and why?

10. Did the image taken at the screen enhance your experience of the system? and why?

11. Would you mind using your face or image to help change the content that is being displayed before you? (using inference and association with contextual use) and why?

12. Could you see yourself using such a system again?

13. How would you like us to improve the system?

14. Is there value, or usefulness of knowing how many people are present at the screen? and why?

8.2 Appendix B - Detection Images

For brevity of space, below is a sample of results that were obtained, highlighting some of the aforementioned issues.

8.2.1 Face Detection (OpenCV)

In this section are the images which were acquired at runtime of the algorithm.

8.2.2 Initial Training Set Images Recognition Testing

In this section are the images which were acquired when initial development of the algorithms took place.

8.2.3 BTS & BTD '08 Further Development

We decided to refine the algorithm, by exploring further what we could alter. We changed the order to see if we could remove some noise traces in the background which were giving false ratings. The altered algorithm is shown below.

```
Algorithm Basic Layout

-> To GreyScale
-> Difference
-> Threshold at 50
-> Gausian Blur (1.4 {gausian sigma}),(5 {kernal})
-> Opening Morphology
-> Blob Filtering (size 50 x 50 pixels) {Assumption larger objects People}
-> Blob Counter
-> Get Rectangles from blob counter
-> Add bounding Rectangles to picture ( > 20 x 20 pixels)
```

After exploring we found that we could reduce the noise in the scene by increasing the threshold of the algorithm. The new results are shown in figures 8.8, 8.10, 8.11, 8.12, 8.13, 8.14, 8.15. We can see that the noise, which was due to lighting conditions was reduced further, providing a more accurate blob on larger objects. One problem was that the loss of smaller non-connected blobs, this resulted in loss of possible identifications.

To GreyScale The resulting image is shown in figure 8.8. By using the grey scale image, we were also able, as a bonus to visualise the images pixel quality by the generation of a histogram. This is shown in figure 8.9.

Difference This is shown in figure 8.10.

Threshold at 50 The result of this is shown in figure 8.11.

Gaussian Blur *The results from applying this filter is shown in figure 8.12.*

Opening Morphology The results of applying this to the evolving image is shown in figure 8.13.

Blob Filtering (size 50 x 50 pixels) The results of the applied filter is shown in figure 8.14.

Blob counter Same as before

Get Rectangles Same as before

Add Bounding Rectangles to Picture This is shown in figure 8.15.

Participant ID: _____ Site: _____ Date: ___/___/___

System Usability Scale

Instructions: For each of the following statements, mark one box that best describes your reactions to the kiosk *today*.

		Strongly Disagree				Strongly Agree
1.	I think that I would like to use this kiosk frequently.	☐	☐	☐	☐	☐
2.	I found this kiosk unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought this kiosk was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need assistance to be able to use this kiosk.	☐	☐	☐	☐	☐
5.	I found the various functions in this kiosk were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this kiosk.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this kiosk very quickly.	☐	☐	☐	☐	☐
8.	I found this kiosk very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using this kiosk.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this kiosk.	☐	☐	☐	☐	☐

Please provide any comments about this kiosk:

Figure 8.1: SUS Questionnaire

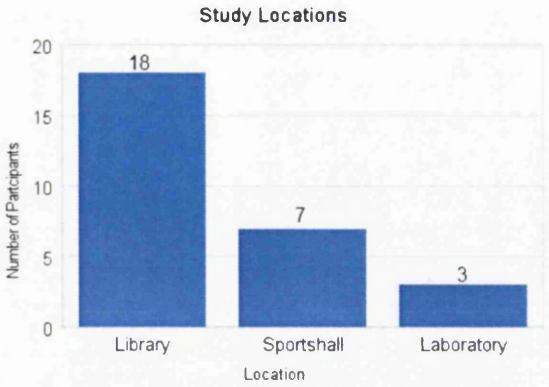


Figure 8.2: SUS Participants numbers at locations

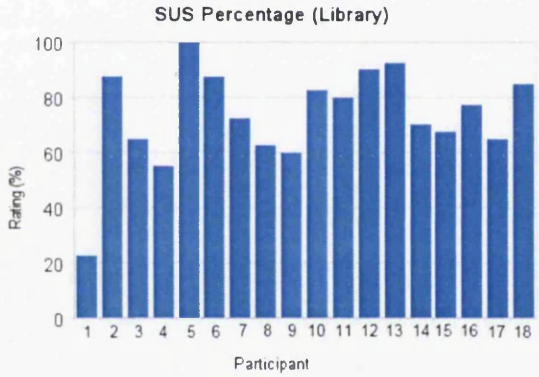


Figure 8.3: SUS Participants averages within Library

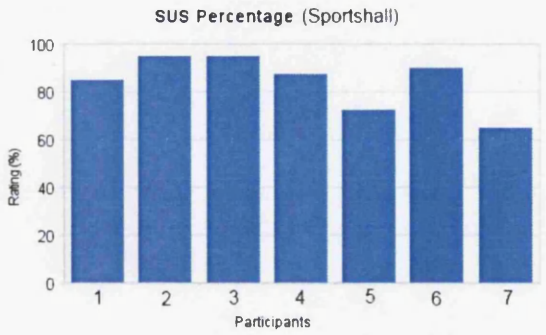


Figure 8.4: SUS Participants averages within Sportshall

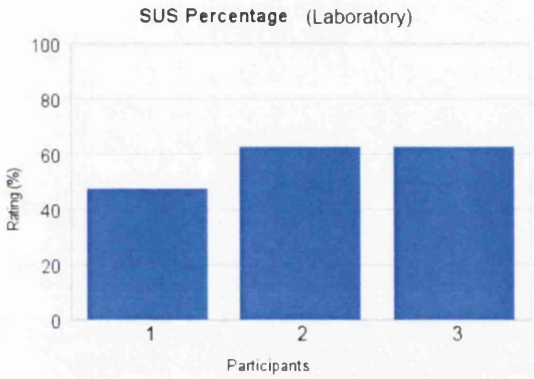
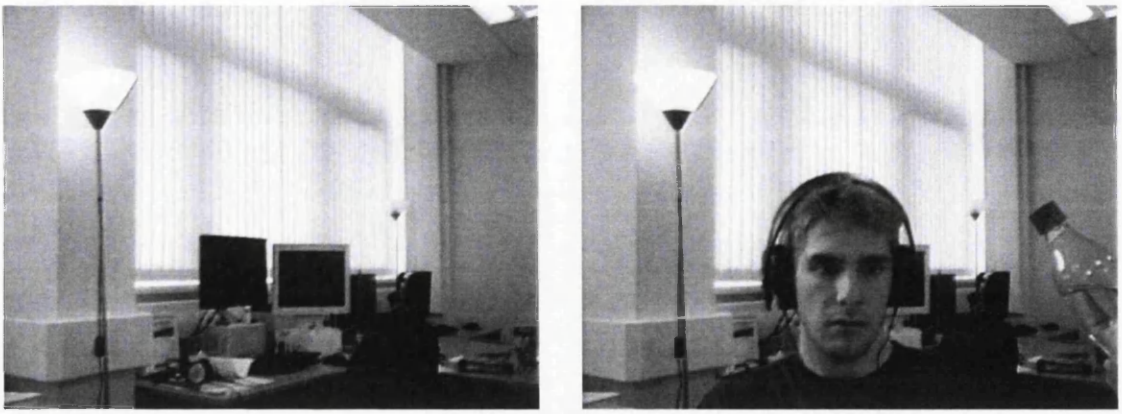


Figure 8.5: SUS Participants averages within Laboratory

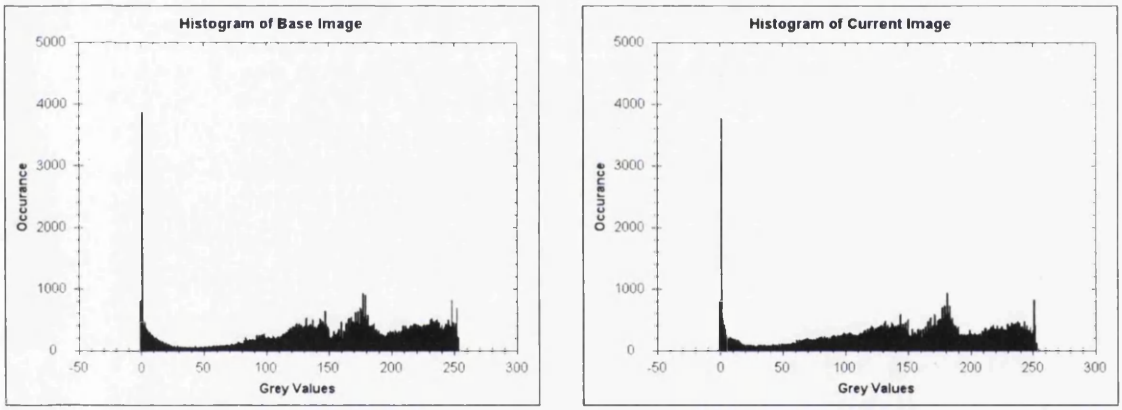
Figure 8.6: (a)Identification (b)Duplication & missed faces (c)Occlusion & false positive

Figure 8.7: (a)Base image, (b)Region, (c)Highlighted



(a) Base (b) Current

Figure 8.8: Difference in images after grey scale conversation



(a) Base Image (b) Current Image

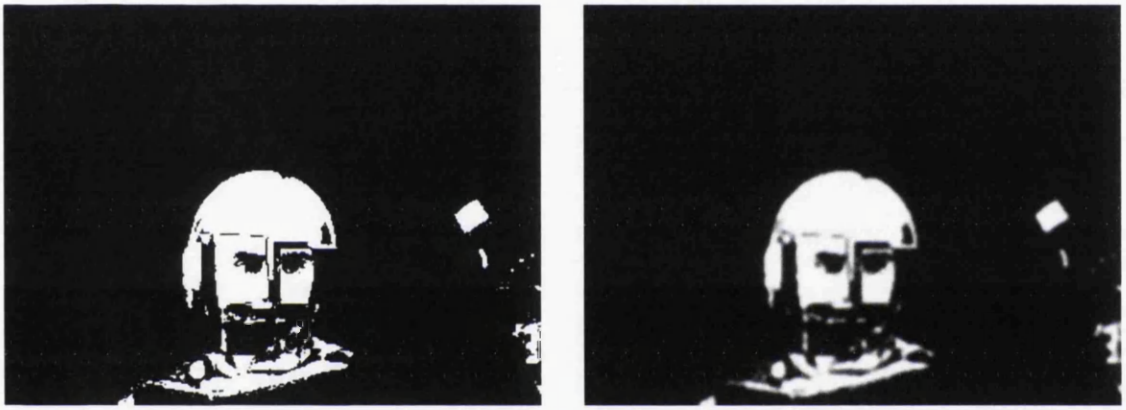
Figure 8.9: Histograms of pixel intensity on grey scale images

(a) Base Image Grey (b) Current Image Current Grey (c) Differnce of Grey Image

Figure 8.10: Difference in images: Before & after

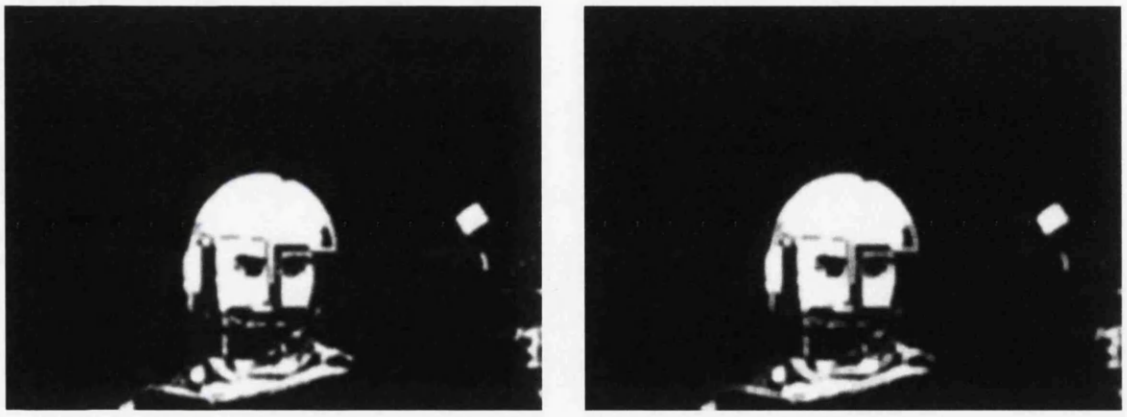
(a) Before (b) After

Figure 8.11: Difference in images after applying a threshold of 50



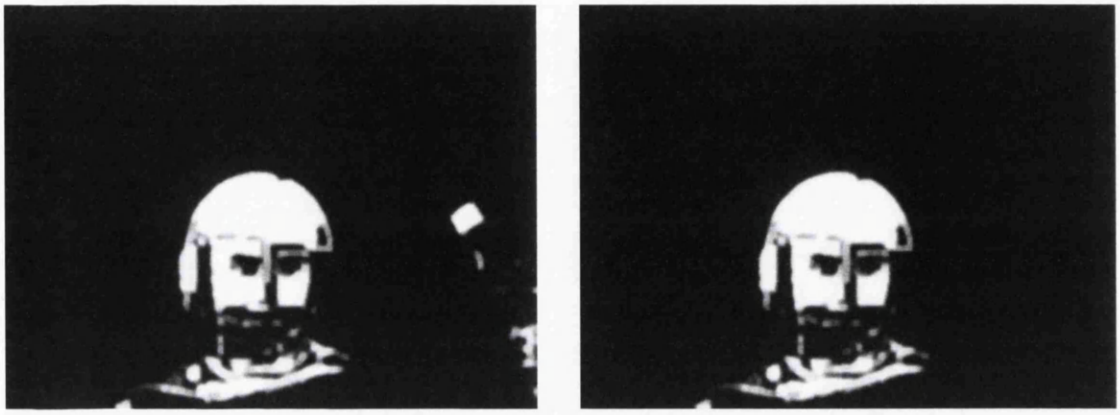
(a) Before (b) After

Figure 8.12: Difference in images after applying a Gaussian blur filter



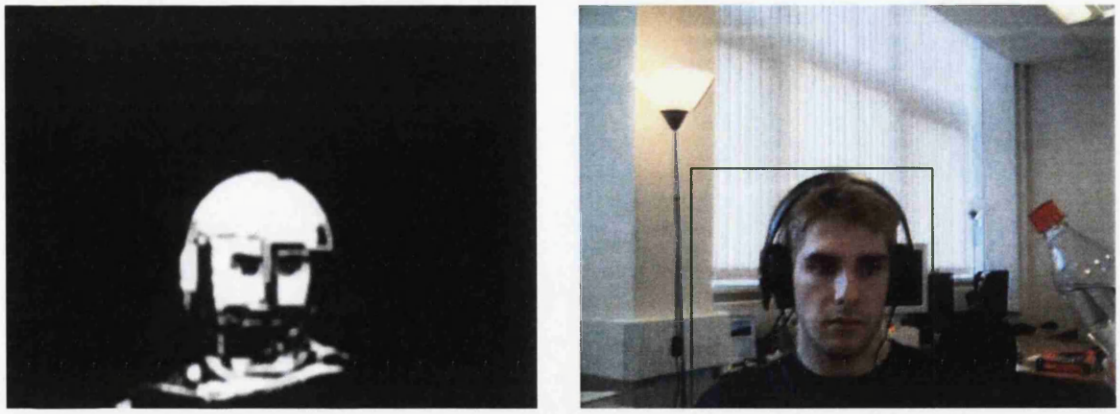
(a) Before (b) After

Figure 8.13: Difference in images after applying Opening morphology



(a) Before (b) After

Figure 8.14: Difference in images after applying Blob filter



(a) Before (b) After

Figure 8.15: Difference in images: Before & after

8.3 Appendix C - Extra

8.3.1 Classification Images for Chapter 5

In this section we show examples of the parts that we used to classify the images.

Figure 8.16: Examples of direct observation (stare)

Figure 8.17: Examples of subtle observation (glance)

The engagement (enjoyment) of the system is gauged by the detection of a smile in the users face. This would suggest some external factor is influencing there emotional state. Shown in figure 8.18

Figure 8.18: Examples of engagement (enjoyment)

We classify looking away (partial) as blinking, body and face towards, display but eyes diverted slightly. Shown in figure 8.20.

We classify looking away (partial) as having the body or face turned 90 degrees away from display. Shown in figure 8.21.

The distraction of human / observer can be classified as looking at another human presence beyond the camera. Shown in figure 8.22

Figure 8.19: Examples of engagement (informative)

Figure 8.20: Examples of looking away (partial)

The distraction of other can be classified as looking at web camera. Shown in figure 8.23

The type of users who are grouped back and front, shown in figure 8.24.

The type of users who are grouped side by side, shown in figure 8.25

The types of inanimate objects are items such as windows, ventilation shafts, book shelves, office equipment. All these types of objects can provide valuable information pertaining to the setting. Shown in figure 8.26.

The types of animate objects are items such as background people not engaged with the display itself. Shown in figure 8.27

The types of illumination are classified as light sources which might reside in the scene. E.g. lights and windows. Shown in figure 8.28.

Figure 8.21: Examples of looking away (full)

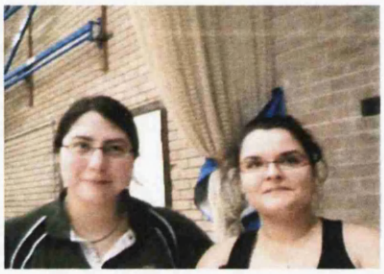


Figure 8.22: Examples of distraction (human/observer)

Figure 8.23: Examples of distraction (other)

Figure 8.24: Examples of grouping orientation (back and front)

Figure 8.25: Examples of grouping orientation (side by side)

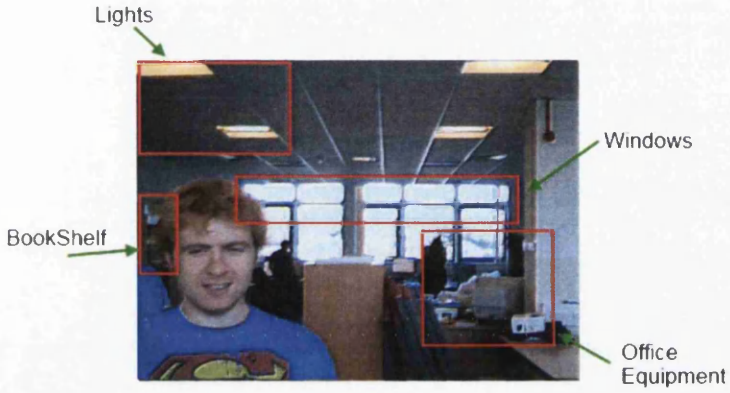
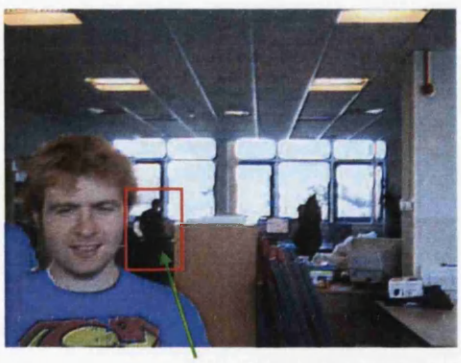


Figure 8.26: Examples of objects (inanimate)



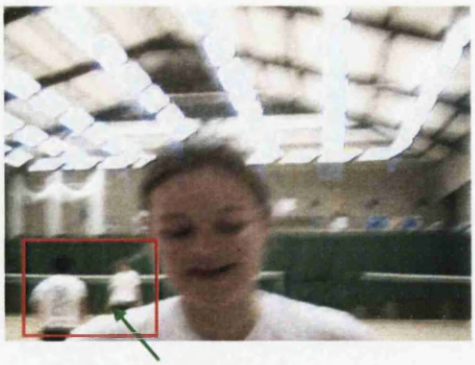


Figure 8.27: Examples of objects (animate)

Figure 8.28: Examples of illumination

Text in the image can be used for possible extra meaning. Shown in figure 8.29.

Figure 8.29: Examples of text

Activity (human), can be identified by surroundings and background knowledge. Shown in figure 8.30.

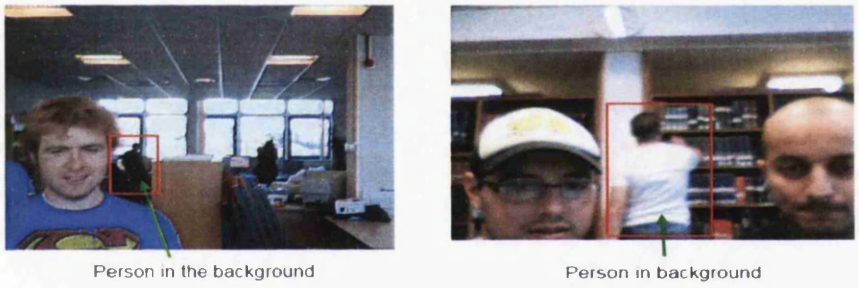


Figure 8.30: Examples of activity (human)

Figure 8.31: Examples of activity (other)

8.3.2 Image capture scenario

This is the scenario which we asked the user to play out when being captured.

- Designated person *You are new to a location, and you see an information kiosk that is displaying videos to you at the display. You become interested in one, and you want to watch it. You then do so.*
- Others *You see a person interested, you also want to view it.*

This will allow us to gage and emulate a response, which allows us to obtain people to test our algorithms against.

8.4 Appendix D - Ethics

8.4.1 Consent Forms

For the departmental supplied research consent form, which has to be provided to participant see figure 8.32, 8.33.

Swansea University – Computer Science Department

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title
Incidental Interactions and Large Screen Displays.

Researcher
Mr. D Garnham

Experiment Purpose
The purpose of this experiment is to observe and record the ways that users interaction with the interactive display.

Participant Recruitment and Selection
Undergraduate and graduate students and general public from Swansea University are being recruited for this experiment.

Procedure
This session will require about an 10-20 minutes of your time. You will be asked to fill in a short questionnaire about your experience with the display and then asked you answer some questions.

This is not a test – our objective is to find out how you approach the interaction.

Data Collection
A questionnaire will be used to gauge your experience at the end of the session. As you interact with the display task your screen actions will be recorded by the computer. The interviewer will take notes. The system will take a picture of you and your surroundings for use by the system.

Data Archiving/Destruction
Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager. You can request to have data removed at the start of the study.

Figure 8.32: Ethics rights form - page 1

Confidentiality

Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort

There is no likelihood of discomfort or risk associated with participation.

Researcher

Mr. D Garnham is working on his doctorate in the Computer Science Department at the Swansea University. This study will contribute to his research on programming by demonstration. His supervisor is Dr Matt Jones.

D. Garnham can be contacted in room 401 Faraday Tower, Swansea University. His phone number is (+44) (0)1792 205678 (or internal extension 4566) and his email address is csdg@swansea.ac.uk

Finding out about Results

The Participants can find out the results of the study by contacting the researcher after October 1, 2009.

Agreement

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____ _____

Participant Date

_____ _____

Investigator/Witness Date

A copy of this consent form has been given to you to keep for your records and reference.

Figure 8.33: Ethics rights form - page 2

Bibliography

[136, 2008] (2008). Using a mobile phone as a "wii-like" controller for playing games on a large public display. *Int. J. Comput. Games Technol.*, 2008(2):1–6.

[afo, 0924] (Last accessed on 2009.09.24). Aforge.net library. `http://www.aforgenet.com`.

[fac, 0924] (Last accessed on 2009.09.24). Facebook. `http://www.facebook.com` , `http://en.wikipedia.org/wiki/Facebook`. http://www.facebook.com , http://en.wikipedia.org/wiki/Facebook.

[twi, 0924] (Last accessed on 2009.09.24). Twitter. `http://twitter.com/` , `http://en.wikipedia.org/wiki/Twitter`.

[val, 2009] (Last accessed on February 23rd, 12:34 2009). Valleyskids home page. `http://valleyskids.org/`.

[Agamanolis, 2004] Agamanolis, S. (2004). Designing displays for human connectedness.

[Alan et al., 2005] Alan, K. C., Dix, A., Fitton, D., Kray, C., Rouncefield, M., Saslis-lagoudakis, G., and Sheridan, J. G. (2005). Exploring mobile phone interaction with situated displays.

[Altosaar et al., 2006] Altosaar, M., Vertegaal, R., Sohn, C., and Cheng, D. (2006). Auraorb: using social awareness cues in the design of progressive notification appliances. In *OZCHI '06: Proceedings of the 18th Australia conference on Computer-Human Interaction*, pages 159–166, New York, NY, USA. ACM.

[Antonietta Grasso,] Antonietta Grasso, Martin Muehlenbrock, F. R. D. S. Supporting communities of practice with large screen displays, chapter 11. `http://www.muehlenbrock.de/publications/Muehlenbrock-CSCW%20Chapter%2011.pdf`.

[Arter et al., 2007] Arter, D., Buchanan, G., and Jones, M. (2007). Incidental information and mobile search. In *Proceedings 9th International Conference on Mobile Devices and Services (Mobile HCI 07)*.

[Badillo et al., 2006] Badillo, B., Bowman, D. A., McConnel, W., Ni, T., and Silva, M. G. (03 October 2006). Literature survey on interaction techniques for large displays. Technical report, Virginia Tech.

[Ballagas et al., 2006] Ballagas, R., Borchers, J., Rohs, M., and Sheridan, J. G. (2006). The smart phone: A ubiquitous input device. *IEEE Pervasive Computing*, 5(1):70.

[Ballagas et al., 2005a] Ballagas, R., Rohs, M., and Sheridan, J. G. (2005a). Mobile phones as pointing devices. In *Pervasive Mobile Interaction Devices (PERMID 2005) Workshop at Pervasive 2005*.

[Ballagas et al., 2005b] Ballagas, R., Rohs, M., and Sheridan, J. G. (2005b). Sweep and point and shoot: phonecam-based interactions for large public displays. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1200–1203, New York, NY, USA. ACM Press.

[Ballagas et al., 2007] Ballagas, R. A., Kratz, S. G., Borchers, J., Yu, E., Walz, S. P., Fuhr, C. O., Hovestadt, L., and Tann, M. (2007). Rexplorer: a mobile, pervasive spell-casting game for tourists. In *CHI '07: CHI '07 extended abstracts on Human factors in computing systems*, pages 1929–1934, New York, NY, USA. ACM.

[Barrington et al., 2006] Barrington, L., Lyons, M. J., Diegmann, D., and Abe, S. (2006). Ambient display using musical effects. In *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces*, pages 372–374, New York, NY, USA. ACM.

[Beigl et al., 2001] Beigl, M., Gellersen, H.-W., and Schmidt, A. (2001). Mediacups: experience with design and use of computer-augmented everyday artefacts. *Comput. Netw.*, 35(4):401–409.

[Ben-Israel, 2008] Ben-Israel, E. (Last accessed on 16th July 2008). Opencvdotnet - .net framework wrapper for intel's opencv package. Google code Project - `http://code.google.com/p/opencvdotnet/`.

[Bi et al., 2005] Bi, X., Shi, Y., Chen, X., and Xiang, P. (2005). Facilitating interaction with large displays in smart spaces. In *sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligence*, pages 105–110, New York, NY, USA. ACM Press.

[Boring et al., 2007] Boring, S., Altendorfer, M., Broll, G., Hilliges, O., and Butz, A. (2007). Shoot & copy: phonecam-based information transfer from public displays onto mobile phones. In *Mobility '07: Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*, pages 24–31, New York, NY, USA. ACM.

[Boyle et al., 2000] Boyle, M., Edwards, C., and Greenberg, S. (2000). The effects of filtered video on awareness and privacy. In *CSCW '00: Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 1–10, New York, NY, USA. ACM.

[Brignull and Rogers, 2003] Brignull, H. and Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In *INTERACT*.

[Brunette et al., 2005] Brunette, K., Eisenstadt, M., Pukinskis, E., and Ryan, W. (2005). Meeteetse: social well-being through place attachment. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 2065–2069, New York, NY, USA. ACM.

[Buxton, 1995] Buxton, W. (1995). Integrating the periphery and context: A new model of telematics. volume Proceedings of Graphics Interface, pages 239–246.

[Carter et al., 2004] Carter, S., Churchill, E., Denoue, L., Helfman, J., and Nelson, L. (2004). Digital graffiti: public annotation of multimedia content. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1207–1210, New York, NY, USA. ACM Press.

[Cheng and Pulo, 2003] Cheng, K. and Pulo, K. (2003). Direct interaction with large-scale display systems using infrared laser tracking devices. In *APVis '03: Proceedings of the Asia-Pacific symposium on Information visualisation*, pages 67–74, Darlinghurst, Australia, Australia. Australian Computer Society, Inc.

[Cheng and Takatsuka, 2005] Cheng, K. and Takatsuka, M. (2005). Real-time monocular tracking of view frustum for large screen human-computer interaction. In *ACSC '05: Proceedings of the Twenty-eighth Australasian conference on Computer Science*, pages 125–133, Darlinghurst, Australia, Australia. Australian Computer Society, Inc.

[Cheverst et al., 2005] Cheverst, K., Dix, A., Fitton, D., Kray, C., Rouncefield, M., Sas, C., Saslis-Lagoudakis, G., and Sheridan, J. G. (2005). Exploring bluetooth based mobile phone interaction with the hermes photo display. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 47–54, New York, NY, USA. ACM Press.

[Churchill and Nelson,] Churchill, E. and Nelson, L. From media spaces to emplaced media: Digital poster boards and community connectedness.

[Churchill et al.,] Churchill, E., Nelson, L., Denoue, L., and Girgensohn, A. The plasma poster network: Posting multimedia content in public places.

[Churchill et al., 2004a] Churchill, E. F., Nelson, L., Denoue, L., Helfman, J., and Murphy, P. (2004a). Sharing multimedia content with interactive public displays: a case study. In *DIS '04: Proceedings of the 2004 conference on Designing interactive systems*, pages 7–16, New York, NY, USA. ACM Press.

[Churchill et al., 2004b] Churchill, E. F., Nelson, L., Denoue, L., Helfman, J., and Murphy, P. (2004b). Sharing multimedia content with interactive public displays: a casestudy. In *DIS '04: Proceedings of the 2004 conference on Designing interactivesystems*, pages 7–16, New York, NY, USA. ACM Press.

[Congleton et al., 2008] Congleton, B., Ackerman, M. S., and Newman, M. W. (2008). The prod framework for proactive displays. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 221–230, New York, NY, USA. ACM.

[Crowley et al., 2000] Crowley, J. L., Coutaz, J., and Bérard, F. (2000). Perceptual user interfaces: things that see. *Commun. ACM*, 43(3):54–ff.

[Cyganek, 2007] Cyganek, B. (2007). *An Introduction to 3D Computer Vision Techniques and Algorithms*. John Wiley & Sons.

[Dachselt and Buchholz, 2009] Dachselt, R. and Buchholz, R. (2009). Natural throw and tilt interaction between mobile phones and distant displays. In *CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 3253–3258, New York, NY, USA. ACM.

[Dahley et al., 1998] Dahley, A., Wisneski, C., and Ishii, H. (1998). Water lamp and pinwheels: ambient projection of digital information into architectural space. In *CHI '98: CHI 98 conference summary on Human factors in computing systems*, pages 269–270, New York, NY, USA. ACM.

[Dietz et al., 2004] Dietz, P., Raskar, R., Booth, S., van Baar, J., Wittenburg, K., and Knep, B. (2004). Multi-projectors and implicit interaction in persuasive public displays. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 209–217, New York, NY, USA. ACM.

[Dix, 2002] Dix, A. (2002). Beyond intention - pushing boundaries with incidental interaction. In *In Proceedings of Building Bridges: Interdisciplinary Context-Sensitive Computing.*

[Dixon et al., 2007] Dixon, M., Heckel, F., Pless, R., and Smart, W. (2007). Faster and more accurate face detection on mobile robots using geometric constraints. *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 1041–1046.

[Dourish and Bly, 1992] Dourish, P. and Bly, S. (1992). Portholes: supporting awareness in a distributed work group. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 541–547, New York, NY, USA. ACM.

[Eagle, 2005] Eagle, N. (2005). *Machine Perception and Learning of Complex Social Systems.* PhD thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology, Massachusetts, US.

[Edwards, 0924] Edwards, D. (Last accessed on 2009.09.24). Hey! i just saw my search go by! `http://xooglers.blogspot.com/2006/03/hey-i-just-saw-my-search-go-by.html`. blog.

[Elgammal et al., 2002] Elgammal, A., Duraiswami, R., Harwood, D., and Davis, L. (2002). Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of the IEEE*, 90(7):1151–1163.

[Elrod et al., 1992] Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J., and Welch, B. (1992). Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 599–607, New York, NY, USA. ACM Press.

[Erickson and Kellogg, 2000] Erickson, T. and Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.*, 7(1):59–83.

[Ferscha and Vogl, 2002] Ferscha, A. and Vogl, S. (2002). Pervasive web access via public communication walls. In *Pervasive '02: Proceedings of the First International Conference on Pervasive Computing*, pages 84–97, London, UK. Springer-Verlag.

[Field, 2005] Field, A. (2005). *Discovering Statistics Using SPSS.* SAGE Publications.

[Fitzpatrick et al., 2002] Fitzpatrick, G., Kaplan, S., Mansfield, T., David, A., and Segall, B. (2002). Supporting public availability and accessibility with elvin: Experiences and reflections. *Comput. Supported Coop. Work*, 11(3):447–474.

[Frohlich et al., 2009] Frohlich, D. M., Rachovides, D., Riga, K., Bhat, R., Frank, M., Edirisinghe, E., Wickramanayaka, D., Jones, M., and Harwood, W. (2009). Storybank: mobile digital storytelling in a development context. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1761–1770, New York, NY, USA. ACM.

[Garnham, 2008a] Garnham, D., B. G. . J. M. ((2008).a). Incidental mobile information and large public displays. *ACM Computer Human Interaction conference 2008 Workshop on Designing and Evaluating Mobile Phone-Based Interaction with Public Displays.*

[Garnham, 2008b] Garnham, D. (2008b). Implicit participation from a glance. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 553–554, New York, NY, USA. ACM.

[Gaver and Dunne, 1999] Gaver, W. and Dunne, A. (1999). Projected realities: conceptual design for cultural effect. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 600–607, New York, NY, USA. ACM.

[Gellersen et al., 1999] Gellersen, H.-W., Beigl, M., and Krull, H. (1999). The mediacup: Awareness technology embedded in a everyday object. In *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 308–310, London, UK. Springer-Verlag.

[Gould et al., 1983] Gould, J. D., Conti, J., and Hovanyecz, T. (1983). Composing letters with a simulated listening typewriter. *Commun. ACM*, 26(4):295–308.

[Graham et al., 2007] Graham, C., Rouncefield, M., Gibbs, M., Vetere, F., and Cheverst, K. (2007). How probes work. In *OZCHI '07: Proceedings of the 19th Australasian conference on Computer-Human Interaction*, pages 29–37, New York, NY, USA. ACM.

[Grasso et al., 1999] Grasso, A., Koch, M., and Snowdon, D. (1999). Campiello— new user interface approaches for community networks. *SIGGROUP Bull.*, 20(2):15–17.

[Greenberg and Rounding, 2001] Greenberg, S. and Rounding, M. (2001). The notification collage: posting information to public and personal displays. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 514–521, New York, NY, USA. ACM Press.

[Gronbaek et al., 2006] Gronbaek, K., Rohde, A., Sundararajah, B., and Bech-Petersen, S. (2006). Infogallery: informative art services for physical library spaces. In *JCDL '06: Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries*, pages 21–30, New York, NY, USA. ACM.

[Haritaoglu and Flickner, 2002] Haritaoglu, I. and Flickner, M. (2002). Attentive billboards: Towards to video based customer behavior. *wacv*, 0:127.

[Haritaoglu et al., 1999] Haritaoglu, I., Harwood, D., and Davis, L. (1999). Hydra: multiple people detection and tracking using silhouettes. *Visual Surveillance, 1999. Second IEEE Workshop on, (VS'99)*, pages 6–13.

[Heiner et al., 1999] Heiner, J. M., Hudson, S. E., and Tanaka, K. (1999). The information percolator: ambient information display in a decorative object. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 141–148, New York, NY, USA. ACM.

[Hiroshi et al., 1998] Hiroshi, C. W., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., and Yarin, P. (1998). Ambient displays: Turning architectural space into an interface between people and digital information. pages 22–32. Springer.

[Holmquist and Skog, 2003] Holmquist, L. E. and Skog, T. (2003). Informative art: information visualization in everyday environments. In *GRAPHITE '03: Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 229–235, New York, NY, USA. ACM.

[Höysniemi et al., 2004] Höysniemi, J., Hämäläinen, P., and Turkki, L. (2004). Wizard of oz prototyping of computer vision based action games for children. In *IDC '04: Proceedings of the 2004 conference on Interaction design and children*, pages 27–34, New York, NY, USA. ACM.

[Huang and Mynatt, 2003] Huang, E. M. and Mynatt, E. D. (2003). Semi-public displays for small, co-located groups. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 49–56, New York, NY, USA. ACM.

[Hutchinson et al., 2003] Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., and Eiderbäck, B. (2003). Technology probes: inspiring design for and with families. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, New York, NY, USA. ACM.

[Inc., 0612a] Inc., G. (Last accessed on 2009.06.12a). Youtube. `http://www.youtube.com/`, `http://en.wikipedia.org/wiki/YouTube`.

[Inc., 0612b] Inc., Y. (Last accessed on 2009.06.12b). Flickr. `http://www.flickr.com/` , `http://en.wikipedia.org/wiki/Flickr`.

[Ishii and Ullmer, 1997] Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA. ACM.

[Ishii et al., 1998] Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., and Yarin, P. (1998). ambientroom: integrating ambient media with architectural space. In *CHI '98: CHI 98 conference summary on Human factors in computing systems*, pages 173–174, New York, NY, USA. ACM.

[Izadi et al., 2003] 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 *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 159–168, New York, NY, USA. ACM Press.

[Jamie Lawrence, 2006] Jamie Lawrence, Terry R. Payne, R. V. K. (MIRW 2006). Exploiting incidental interactions between mobile devices.

[Jeon et al., 2006] Jeon, S., Hwang, J., Kim, G. J., and Billinghurst, M. (2006). Interaction techniques in large display environments using hand-held devices. In *VRST '06: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 100–103, New York, NY, USA. ACM Press.

[Johnstone, 2004] Johnstone, D., B. M. . T. M. (2004). Bringing human information behavior into information systems research: an application of systems modelling. *Information Research*, 9(4) paper 191.

[Jones et al., 2007a] Jones, M., Buchanan, G., Harper, R., and Xech, P.-L. (2007a). Questions not answers: a novel mobile search technique. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 155–158, New York, NY, USA. ACM Press.

[Jones et al., 2008] Jones, M., Harwood, W., Bainbridge, D., Buchanan, G., Frohlich, D., Rachovides, D., Frank, M., and Lalmas, M. (2008). "narrowcast yourself": designing for community storytelling in a rural indian context. In *DIS '08: Proceedings of the 7th ACM conference on Designing interactive systems*, pages 369–378, New York, NY, USA. ACM.

[Jones et al., 2007b] Jones, M., Harwood, W., Buchanan, G., and Lalmas, M. (2007b). Storybank: an indian village community digital library. In *JCDL '07: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, pages 257–258, New York, NY, USA. ACM.

[Jones and Marsden, 2006] Jones, M. and Marsden, G. (2006). *Mobile Interaction Design*. Wiley.

[Kameda et al., 2004] Kameda, Y., Takemasa, T., and Ohta, Y. (2004). Outdoor see-through vision utilizing surveillance cameras. In *ISMAR '04: Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 151–160, Washington, DC, USA. IEEE Computer Society.

[Karahalios and Donath, 2004] Karahalios, K. and Donath, J. (2004). Telemurals: linking remote spaces with social catalysts. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 615–622, New York, NY, USA. ACM.

[Kela et al., 2006] Kela, J., Korpipää, P., Mäntyjärvi, J., Kallio, S., Savino, G., Jozzo, L., and Marca, D. (2006). Accelerometer-based gesture control for a design environment. *Personal Ubiquitous Comput.*, 10(5):285–299.

[Kerne, 2001] Kerne, A. (2001). *Collagemachine: a model of "interface ecology"*. PhD thesis. Adviser-Ken Perlin.

[Kilambi et al., 2008] Kilambi, P., Ribnick, E., Joshi, A. J., Masoud, O., and Papanikolopoulos, N. (2008). Estimating pedestrian counts in groups. *Comput. Vis. Image Underst.*, 110(1):43–59.

[Koch, 2004] Koch, M. (2004). Building community mirrors with public shared displays. In *Proc. eChallenges e-2004 Conference*, Wien, sterreich.

[Krahnstoever et al., 2002] Krahnstoever, N., Kettebekov, S., Yeasin, M., and Sharma, R. (2002). A real-time framework for natural multimodal interaction with large screen displays. In *Proc. of Fourth Intl. Conference on Multimodal Interfaces (ICMI 2002), Pittsburgh, PA, USA.*

[Krueger, 1991] Krueger, M. W. (1991). *Artificial Reality 2 (2nd Edition)*. Addison-Wesley Professional.

[Krueger et al., 1985] Krueger, M. W., Gionfriddo, T., and Hinrichsen, K. (1985). Videoplace—an artificial reality. *SIGCHI Bull.*, 16(4):35–40.

[LEE, 2002] LEE, J. . (2002). Postcards from planet google. The New York Times, technology article.

[Lienhart and Maydt, 2002] Lienhart, R. and Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *IEEE ICIP 2002*, pages 900–903.

[Mankoff et al., 2003] Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., and Ames, M. (2003). Heuristic evaluation of ambient displays. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 169–176, New York, NY, USA. ACM.

[Marchese and Marchese, 2004] Marchese, S. and Marchese, F. (2004). Plato's cave: an image stream installation within an office setting. pages 954–958.

[Masoud and Papanikolopoulos, 2001] Masoud, O. and Papanikolopoulos, N. (2001). A novel method for tracking and counting pedestrians in real-time using a single camera.

[Maunder et al., 2007] Maunder, A., Marsden, G., and Harper, R. (2007). Creating and sharing multi-media packages using large situated public displays and mobile phones. In *MobileHCI '07: Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*, pages 222–225, New York, NY, USA. ACM.

[McCarthy and Anagnost, 1998] McCarthy, J. F. and Anagnost, T. D. (1998). Musicfx: an arbiter of group preferences for computer supported collaborative workouts. In *CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work*, pages 363–372, New York, NY, USA. ACM.

[McCarthy et al., 2008] McCarthy, J. F., Congleton, B., and Harper, F. M. (2008). The context, content & community collage: sharing personal digital media in the physical workplace. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 97–106, New York, NY, USA. ACM.

[McCarthy et al., 2001] McCarthy, J. F., Costa, T. J., and Liongosari, E. S. (2001). Unicast, outcast & groupcast: Three steps toward ubiquitous, peripheral displays. In *UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing*, pages 332–345, London, UK. Springer-Verlag.

[McDonald et al., 2008] McDonald, D. W., McCarthy, J. F., Soroczak, S., Nguyen, D. H., and Rashid, A. M. (2008). Proactive displays: Supporting awareness in fluid social environments. *ACM Trans. Comput.-Hum. Interact.*, 14(4):1–31.

[Media, 0924] Media, F. I. (Last accessed on 2009.09.24). Myspace. `http://www.myspace.com/`, `http://en.wikipedia.org/wiki/MySpace`.

[Mettala, 1999] Mettala, R. (1999). Bluetooth protocol architecture version 1.0. White Paper. 1.C.120/1.0.

[Miller, 1956] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. One of the 100 most influential papers in cognitive science: http://cogsci.umn.edu/millennium/final.html.

[Miller and Stasko, 2002] Miller, T. and Stasko, J. (2002). Artistically conveying peripheral information with the infocanvas. In *In: Proceedings of the Working Conference on Advanced Visual Interfaces(AVI2002*, pages 43–50.

[Mitchell, 2006] Mitchell, K. & Race, N. J. (2006). Oi! capturing user attention within pervasive display environments. In *Proceedings of Workshop on pervasive display infrastructures, interfaces and applications*, Dublin, Ireland.

[Miyaoku et al., 2004] Miyaoku, K., Higashino, S., and Tonomura, Y. (2004). C-blink: a hue-difference-based light signal marker for large screen interaction via any mobile terminal. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 147–156, New York, NY, USA. ACM Press.

[Müller et al., 2007] Müller, J., Krüger, A., and Kuflik, T. (2007). Maximizing the utility of situated public displays. In *UM '07: Proceedings of the 11th international conference on User Modeling*, pages 395–399, Berlin, Heidelberg. Springer-Verlag.

[Müller et al., 2007] Müller, J., Paczkowski, O., and Krüger, A. (2007). Situated public news and reminder displays. pages 248–265.

[Myers et al., 2002] Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., and Long, A. C. (2002). Interacting at a distance: measuring the performance of laser pointers and other devices. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 33–40, New York, NY, USA. ACM Press.

[Mynatt et al., 2001] Mynatt, E. D., Rowan, J., Craighill, S., and Jacobs, A. (2001). Digital family portraits: supporting peace of mind for extended family members. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 333–340, New York, NY, USA. ACM.

[Nakanishi et al., 2002] Nakanishi, Y., Fujii, T., Kiatjima, K., Sato, Y., and Koike, H. (2002). Vision-based face tracking system for large displays. In *UbiComp '02: Proceedings of the 4th international conference on Ubiquitous Computing*, pages 152–159, London, UK. Springer-Verlag.

[Narayanaswami et al., 2008] Narayanaswami, C., Coffman, D., Lee, M. C., Moon, Y. S., Han, J. H., Jang, H. K., McFaddin, S., Paik, Y. S., Kim, J. H., Lee, J. K., Park, J. W., and Soroker, D. (2008). Pervasive symbiotic advertising. In *HotMobile '08: Proceedings of the 9th workshop on Mobile computing systems and applications*, pages 80–85, New York, NY, USA. ACM.

[Nelson et al., 2004] Nelson, L., Churchill, E. F., Denoue, L., Helfman, J., and Murphy, P. (2004). Gooey interfaces: an approach for rapidly repurposing digital content. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1293–1296, New York, NY, USA. ACM.

[Newman et al., 1991] Newman, W. M., Eldridge, M. A., and Lamming, M. G. (1991). Pepys: generating autobiographies by automatic tracking. In *ECSCW'91: Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work*, pages 175–188, Norwell, MA, USA. Kluwer Academic Publishers.

[Nguyen and Mynatt,] Nguyen, D. H. and Mynatt, E. D. Privacy mirrors: Understanding and shaping socio-technical ubiquitous computing systems.

[Nguyen et al., 2006] Nguyen, Q., Novakowski, S., Boyd, J. E., Jacob, C., and Hushlak, G. (2006). Motion swarms: video interaction for art in complex environments. In *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 461–469, New York, NY, USA. ACM.

[Nielsen, 0924] Nielsen, J. (Last accessed on 2009.09.24). Heuristic evaluation. `http://www.useit.com/papers/heuristic`.

[O'Hara et al., 2008] O'Hara, K., Glancy, M., and Robertshaw, S. (2008). Understanding collective play in an urban screen game. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 67–76, New York, NY, USA. ACM.

[Olsson et al., 2008] Olsson, T., Toivola, H., and Väänänen-Vainio-Mattila, K. (2008). Exploring characteristics of collective content: a field study with four user communities. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 2967–2972, New York, NY, USA. ACM.

[Payne et al., 2006] Payne, T., David, E., Jennings, N. R., and Sharifi, M. (2006). Auction mechanisms for efficient advertisement selection on public displays. In *Proceedings of European Conference on Artificial Intelligence*, pages 285–289.

[Pedersen et al., 1993] Pedersen, E. R., McCall, K., Moran, T. P., and Halasz, F. G. (1993). Tivoli: an electronic whiteboard for informal workgroup meetings. In *CHI '93: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 391–398, New York, NY, USA. ACM Press.

[Peltonen et al., 2008] Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., and Saarikko, P. (2008). It's mine, don't touch!: interactions at a large multitouch display in a city centre. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1285–1294, New York, NY, USA. ACM.

[Peltonen et al., 2007] Peltonen, P., Salovaara, A., Jacucci, G., Ilmonen, T., Ardito, C., Saarikko, P., and Batra, V. (2007). Extending large-scale event participation with user-created mobile media on a public display. In *MUM '07: Proceedings of the 6th international conference on Mobile and ubiquitous multimedia*, pages 131–138, New York, NY, USA. ACM.

[Phung and Bouzerdoum, 2007] Phung, S. and Bouzerdoum, A. (2007). Detecting people in images: An edge density approach. *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 1:I–1229–I–1232.

[Plaue et al., 2004] Plaue, C., Miller, T., and Stasko, J. (2004). Is a picture worth a thousand words?: an evaluation of information awareness displays. In *GI '04: Proceedings of Graphics Interface 2004*, pages 117–126, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada. Canadian Human-Computer Communications Society.

[Preece et al., 1994] Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., and Carey, T. (1994). *Human Computer Interaction*. Addison-Wesley.

[Raj et al., 2004] Raj, H., Gossweiler, R., and Milojicic, D. S. (2004). Contentcascade incremental content exchange between public displays and personal devices. In *MobiQuitous*, pages 374–381. IEEE Computer Society.

[Robertson et al., 1996] Robertson, S., Wharton, C., Ashworth, C., and Franzke, M. (1996). Dual device user interface design: Pdas and interactive television. In *CHI '96: Proceedings of the*

SIGCHI conference on Human factors in computing systems, pages 79–86, New York, NY, USA. ACM Press.

[Rogers et al., 2007] Rogers, A., David, E., Payne, T. R., and Jennings, N. R. (2007). An advanced bidding agent for advertisement selection on public displays. In *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–8, New York, NY, USA. ACM.

[Rogers, 2006] Rogers, Y. (2006). Moving on from weiser's vision of calm computing: Engaging ubicomp experiences. In *Ubicomp*, pages 404–421.

[Rohs et al.,] Rohs, M., Sheridan, J. G., and et al. Direct manipulation techniques for large displays using camera phones.

[Roussel et al., 2004] Roussel, N., Evans, H., and Hansen, H. (2004). Mirrorspace: using proximity as an interface to video-mediated communication. In *Proc. of Pervasive 2004 Conference, volume 3001 of Lecture Notes in Computer Science*, pages 345–350. Springer.

[Russell and Gossweiler, 2001] Russell, D. M. and Gossweiler, R. (2001). On the design of personal & communal large information scale appliances. In *UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing*, pages 354–361, London, UK. Springer-Verlag.

[Russell et al., 2004] Russell, D. M., Trimble, J. P., and Dieberger, A. (2004). The use patterns of large, interactive display surfaces: Case studies of media design and use for blueboard and merboard. *hicss*, 04:40098b.

[Scheible and Ojala, 2005] Scheible, J. and Ojala, T. (2005). Mobilenin combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 199–208, New York, NY, USA. ACM.

[Scheible et al., 2008] Scheible, J., Ojala, T., and Coulton, P. (2008). Mobitoss: a novel gesture based interface for creating and sharing mobile multimedia art on large public displays. In *MM '08: Proceeding of the 16th ACM international conference on Multimedia*, pages 957–960, New York, NY, USA. ACM.

[Scheible et al., 2007] Scheible, J., Tuulos, V. H., and Ojala, T. (2007). Story mashup: design and evaluation of novel interactive storytelling game for mobile and web users. In *MUM '07: Proceedings of the 6th international conference on Mobile and ubiquitous multimedia*, pages 139–148, New York, NY, USA. ACM.

[Sekiguchi et al., 2005] Sekiguchi, M., Naito, H., Ueda, A., Ozaki, T., and Yamasawa, M. (2005). "ubwall", ubiquitous wall changes an ordinary wall into the smart ambience. In *sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligence*, pages 47–50, New York, NY, USA. ACM Press.

[Shih, 2009] Shih, F. Y. (2009). *Image Processing and Mathematical Morphology Fundamentals and Applications*. CRC Press.

[Smith et al., 2006] Smith, K., Ba, S. O., Gatica-Perez, D., and Odobez, J.-M. (2006). Tracking the multi person wandering visual focus of attention. In *ICMI '06: Proceedings of the 8th international conference on Multimodal interfaces*, pages 265–272, New York, NY, USA. ACM.

[Smith et al., 2008] Smith, K., Ba, S. O., Odobez, J.-M., and Gatica-Perez, D. (2008). Tracking the visual focus of attention for a varying number of wandering people. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(7):1212–1229.

[Storz et al., 2006a] Storz, O., Friday, A., Davies, N., Finney, J., Sas, C., and Sheridan, J. (2006a). Public ubiquitous computing systems: Lessons from the e-campus display deployments. *IEEE Pervasive Computing*, 5(3):40–47.

[Storz et al., 2006b] Storz, O., Friday, A., Davies, N., Finney, J., Sas, C., and Sheridan, J. (2006b). Public ubiquitous computing systems: Lessons from the e-campus display deployments. *IEEE Pervasive Computing*, 5(3):40–47.

[Struppek, 2007] Struppek, M. (Last accessed on 10th June 2007). Urban screens. `http://www.urbanscreens.org/`.

[Taylor et al., 2007] Taylor, N., Cheverst, K., Fitton, D., Race, N. J. P., Rouncefield, M., and Graham, C. (2007). Probing communities: study of a village photo display. In *OZCHI '07: Proceedings of the 19th Australasian conference on Computer-Human Interaction*, pages 17–24, New York, NY, USA. ACM.

[Tullis and Albert, 2008] Tullis, T. and Albert, B. (2008). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics (Interactive Technologies): Collecting, Analyzing, and Presenting ... Kaufmann Series in Interactive Technologies)*. Morgan Kaufmann.

[Vartiainen et al., 2006] Vartiainen, P., Chande, S., and Rämö, K. (2006). Mobile visual interaction: enhancing local communication and collaboration with visual interactions. In *MUM '06: Proceedings of the 5th international conference on Mobile and ubiquitous multimedia*, page 4, New York, NY, USA. ACM Press.

[Verkasalo, 2006] Verkasalo, H. (2006). Empirical observations on the emergence of mobile multimedia services and applications in the u.s. and europe. In *MUM '06: Proceedings of the 5th international conference on Mobile and ubiquitous multimedia*, page 3, New York, NY, USA. ACM.

[Villar et al., 2003] Villar, N., Schmidt, A., Kortuem, G., and Gellersen, H.-W. (2003). Interacting with proactive public displays. *Computers & Graphics*, 27(6):849 – 857.

[Viola and Jones, 2001] Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features.

[Viola and Jones, 2004] Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154.

[Vogel and Balakrishnan, 2004] Vogel, D. and Balakrishnan, R. (2004). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 137–146, New York, NY, USA. ACM Press.

[Voida et al., 2002] Voida, S., Mynatt, E., MacIntyre, B., and Corso, G. (2002). Integrating virtual and physical context to support knowledge workers.

[Wales, 0924] Wales, B. (Last Accessed on 2009.09.24). Rhondda lives! `http://www.bbc.co.uk/wales/audiovideo/sites/galleries/rhonddalives/`.

[Weiser, 1991] Weiser, M. (1991). The computer for the twenty-first century. *Scientific American*, 265(3):94–104.

[Weiser and Brown, 0924] Weiser, M. and Brown, J. S. (Last accessed on 2009.09.24). The coming age of calm technology. `http://www.ubiq.com/hypertext/weiser/acmfuture2endnote.htm`.

[White and Lutters, 2003] White, K. F. and Lutters, W. G. (2003). Behind the curtain: lessons learned from a wizard of oz field experiment. *SIGGROUP Bull.*, 24(3):129–135.

[Whyte, 1988] Whyte, W. H. (1988). *City: Rediscovering the Center. NewYork: Doubleday.*

[Wichary et al., 2005] Wichary, M., Gunawan, L., den Ende, N. V., Hjortzberg-Nordlund, Q., Matysiak, A., Janssen, R., and Sun, X. (2005). Vista: interactive coffee-corner display. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1062–1077, New York, NY, USA. ACM Press.

[Yamada et al., 2004] Yamada, T., Shingu, J., Churchill, E., Nelson, L., Helfman, J., and Murphy, P. (2004). Who cares?: reflecting who is reading what on distributed community bulletin boards. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 109–118, New York, NY, USA. ACM.

[Yilmaz et al., 2006] Yilmaz, A., Javed, O., and Shah, M. (2006). Object tracking: A survey. *ACM Comput. Surv.*, 38(4):13.