COLLABORATIVE IMAGE-BASED TAGGING AND INTERACTIVE MAPPING ON TABLETOPS

UNIVERSITY OF SYDNEY
SCHOOL OF INFORMATION TECHNOLOGIES

A thesis submitted in partial fulfilment of the requirements for the degree of Bachelor of Information Technology (Honours)

GLEN WHITAKER
SID: 200307788

SUPERVISOR: JUDY KAY

3rd NOVEMBER 2006
I certify that:

I have read and understood the University of Sydney Student Plagiarism: Coursework Policy and Procedure;

I understand that failure to comply with the Student Plagiarism: Coursework Policy and Procedure can lead to the University commencing proceedings against me for potential student misconduct under Chapter 8 of the University of Sydney By-Law 1999 (as amended);

this Work is substantially my own, and to the extent that any part of this Work is not my own I have indicated that it is not my own by Acknowledging the Source of that part or those parts of the Work.

Name: Glen Whitaker

Signature:          Date: 3/11/2006
Abstract

The move towards collocated collaborative systems and the novel user interface provided by tabletops has motivated research in collaborative applications with natural gesture-based interaction. The focus of this research was to explore methods for collocated collaborative image-based tagging of interactive maps exploiting the novelty of tabletop interfaces and enhancing their collaborative nature.

This research involved the development of a prototyped framework called MyPlanner. The MyPlanner framework builds upon the University of Sydney’s Cruiser Tabletop Sharepic photo sharing application. MyPlanner provides key elements for evaluating the image-based tagging and interactive mapping in the context of a planning framework. These necessary elements include: image-based tagging; alternative methods for tagging text to objects; identification and interaction with regions to add geographical depth to maps; map objects to provide geographical organisation of digital content. The elements are designed to be used by multiple users collaborating together to perform a range of design tasks.

MyPlanner integrated and explored various research issues, which were validated in qualitative user evaluations. These assessed five core elements of MyPlanner: usability of image-based tagging and interactive mapping in a collocated collaborative interface; performance of users completing design tasks using elements provided by MyPlanner; collocated collaboration between users performing design tasks on tabletops; effects of user short-term memory load when using tagging elements; comparing an alternative text-tagging method to a pre-existing tabletop text-based tagging system.

There are four main contributions of this research: interactive maps and tagging on tabletops, a potential new approach to file classification, and a new swipe gesture. Evaluations of the prototype showed the performance, usability and collaboration results for the potential supporting of tabletop interaction for organisation of content using tagging and mapping.
ACKNOWLEDGEMENTS

Firstly I would like to thank my supervisor Judy Kay for exceeding expectations of a supervisor and providing valuable input into the original design. I would also like to thank Trent Apted for his suggestions on approaching the research problem and Andrew Lum for his valuable comments on the presentation and discussion of this work.

I would like to extend thanks to all those who were willing and performed user evaluations giving up their personal time, providing vital data allowing this work to be validated.

Finally thanks must be given to the School of Information Technologies and the SITRG for making this work possible.
# TABLE OF CONTENTS

STUDENT PLAGIARISM: COMPLIANCE STATEMENT .................................................. I
ABSTRACT ............................................................................................................... I
ACKNOWLEDGEMENTS ....................................................................................... II
TABLE OF CONTENTS .......................................................................................... III
LIST OF FIGURES .................................................................................................. VI
LIST OF TABLES ..................................................................................................... IX

## CHAPTER 1 : INTRODUCTION ................................................................. 1
  1.1 MOTIVATION ............................................................................................... 3
  1.2 DEFINITION OF PROBLEM ...................................................................... 4
  1.3 RESEARCH CONTRIBUTION ...................................................................... 4
  1.4 THESIS OUTLINE ..................................................................................... 5

## CHAPTER 2 : BACKGROUND ................................................................. 7
  2.1 TABLETOPS ............................................................................................... 7
  2.1.1 Overview ............................................................................................. 7
  2.1.2 Hardware Features of Tabletop Interfaces ............................................. 8
  2.1.3 Collaboration ....................................................................................... 11
  2.1.4 Software Features of Tabletop Interfaces ............................................. 12
  2.2 VIRTUAL MUSEUMS/ENVIRONMENTS ............................................... 16
  2.3 PERSONALISATION .................................................................................. 19
  2.3.1 Overview ............................................................................................. 19
  2.3.2 Personalisation that influences Direction of MyPlanner .................... 19
  2.3.3 Tagging ............................................................................................... 20
  2.4 SUMMARY ............................................................................................... 22

## CHAPTER 3 : USER PERSPECTIVE .................................................... 24
  3.1 OVERVIEW ............................................................................................... 24
  3.2 HARDWARE ............................................................................................... 24
  3.2.1 Mimio Interactive ................................................................................ 25
  3.2.2 Top-down projection .......................................................................... 25
  3.3 GESTURES ................................................................................................. 26
  3.3.1 Overview ............................................................................................. 26
  3.3.2 Select .................................................................................................. 26
  3.3.3 Movement ........................................................................................... 26
  3.3.4 Dwell .................................................................................................. 27
  3.3.5 Flipping ................................................................................................ 27
  3.3.6 Rotation ............................................................................................... 28
  3.3.7 Resize .................................................................................................. 28
  3.3.8 Swipe ................................................................................................... 29
  3.4 ELEMENTS OF CRUISER ....................................................................... 30
  3.4.1 Image Objects ...................................................................................... 30
  3.4.2 Blackhole ............................................................................................. 30
  3.5 ELEMENTS OF MYPLANNER ............................................................ 31
  3.6 SUMMARY ............................................................................................... 31
# Table of Contents

**CHAPTER 4 : ARCHITECTURE OVERVIEW** ................................................................. 32

4.1 ARCHITECTURE ........................................................................................................ 33
4.2 IMAGES ...................................................................................................................... 34
4.3 BLACKHOLE ............................................................................................................ 34
4.4 NEW COMPONENTS ............................................................................................... 35
4.5 SUMMARY ............................................................................................................... 35

**CHAPTER 5 : INTERACTIVE MAPPING** ................................................................. 36

5.1 OVERVIEW .............................................................................................................. 36
5.2 ARCHITECTURE ...................................................................................................... 37
5.3 GESTURES ............................................................................................................... 37
5.4 USER INTERFACE .................................................................................................. 38
5.5 SUMMARY ............................................................................................................. 40

**CHAPTER 6 : REGION IDENTIFICATION AND INTERACTION** ......................... 42

6.1 OVERVIEW ............................................................................................................. 42
6.2 ARCHITECTURE ..................................................................................................... 42
6.3 GESTURES .............................................................................................................. 43
6.3.1 Region Creation Tool ......................................................................................... 43
6.3.2 Region Objects .................................................................................................. 44
6.4 USER INTERFACE .................................................................................................. 44
6.5 SUMMARY ............................................................................................................. 47

**CHAPTER 7 : IMAGE-BASED TAGGING** .............................................................. 48

7.1 OVERVIEW ............................................................................................................. 48
7.2 ARCHITECTURE ..................................................................................................... 48
7.3 GESTURES .............................................................................................................. 49
7.3.1 Creating Tags ..................................................................................................... 49
7.3.2 Interacting with Tags ......................................................................................... 49
7.4 USER INTERFACE .................................................................................................. 50
7.5 SUMMARY ............................................................................................................. 53

**CHAPTER 8 : TEXT-BASED TAGGING** ................................................................. 54

8.1 OVERVIEW ............................................................................................................. 54
8.2 ARCHITECTURE ..................................................................................................... 54
8.3 GESTURES .............................................................................................................. 55
8.3.1 Primary Text objects ......................................................................................... 55
8.3.2 Attached Text objects ...................................................................................... 56
8.4 USER INTERFACE .................................................................................................. 56
8.5 SUMMARY ............................................................................................................. 59

**CHAPTER 9 : METADATA STORAGE AND HANDLING** ...................................... 60

9.1 OVERVIEW ............................................................................................................. 60
9.2 ARCHITECTURE ..................................................................................................... 60
9.3 METADATA STORAGE ......................................................................................... 61
9.4 PROCESSING OF METADATA ............................................................................ 63
9.5 SUMMARY ............................................................................................................. 64

**CHAPTER 10 : EVALUATION DESIGN** ................................................................. 66

10.1 OBJECTIVES ....................................................................................................... 67
10.1.1 User Performance ............................................................................................ 67
10.1.2 Other Usability Aspects ................................................................................ 67
10.1.3 Collocated Collaboration ............................................................................... 68
10.1.4 Short-term memory load .............................................................................. 68
10.2 SCENARIO ............................................................................................................. 69
10.3 METHODOLOGY .................................................................................................. 69
10.3.1 Prototyping ..................................................................................................... 70
10.3.2 User Evaluations ........................................................................................... 71
10.3.3 Questionnaire ................................................................................................. 71
# Table of Contents

10.4 TEAMTAG-CONTROL .................................................................................................................. 72  
10.4.1 Overview ................................................................................................................................. 72  
10.4.2 Reason for Comparison ............................................................................................................. 73  
10.5 USER EVALUATIONS ................................................................................................................. 73  
10.5.1 Overview ................................................................................................................................. 73  
10.5.2 Objectives ............................................................................................................................... 74  
10.5.3 Procedure ............................................................................................................................... 74  
10.6 QUESTIONNAIRE ....................................................................................................................... 77  
10.6.1 Overview ................................................................................................................................. 77  
10.6.2 Objectives ............................................................................................................................... 78  
10.6.3 Questions ............................................................................................................................... 78  
10.7 SUMMARY ................................................................................................................................. 80  

CHAPTER 11 RESULTS ......................................................................................................................... 81  
11.1 RESULTS RELATING TO EVALUATION GOALS ......................................................................... 81  
11.1.1 User Performance .................................................................................................................... 81  
11.1.2 Other Usability ....................................................................................................................... 83  
11.1.3 Collocated Collaboration ....................................................................................................... 84  
11.1.4 Effects of Short-term memory load ....................................................................................... 85  
11.2 OBSERVATION RESULTS ......................................................................................................... 86  
11.2.1 Interactive Mapping ................................................................................................................ 86  
11.2.2 Image-Based Tagging ............................................................................................................. 87  
11.2.3 Region Creation ..................................................................................................................... 88  
11.2.4 Text Association ................................................................................................................... 89  
11.3 QUESTIONNAIRE RESULTS .................................................................................................... 90  
11.3.1 Interactive Mapping ................................................................................................................ 90  
11.3.2 Image-Based Tagging ............................................................................................................. 91  
11.3.3 Text Association ................................................................................................................... 91  
11.4 CONCLUSIONS ........................................................................................................................ 93  
11.5 SUMMARY ............................................................................................................................... 94  

CHAPTER 12 : FUTURE WORK ........................................................................................................... 96  

CHAPTER 13 : CONCLUSION ............................................................................................................ 98  

REFERENCES ....................................................................................................................................... 101  

APPENDIX A : USER EVALUATION INFORMATION SHEET ............................................................. 105  
APPENDIX B : PARTICIPANT 1 CRIB SHEET ..................................................................................... 109  
APPENDIX C : PARTICIPANT 2 CRIB SHEET ..................................................................................... 112  
APPENDIX D : SAMPLE OF QUESTIONNAIRE RESULTS ................................................................. 115
LIST OF FIGURES


FIGURE 2: THE IMAGES ABOVE BOTH PRESENT GROUPS OF USERS COLLABORATIVELY USING A VARIETY OF APPLICATIONS ON DIRECT-TOUCH INTERFACES SUCH AS THE DIAMONDTOUCH TABLETOP [33] 9

FIGURE 3: REPRESENTATION OF AN ACTIVITY PLOT FROM STUDYING THE AREAS OF WHERE USERS INTERACT WITH OBJECTS ON THE TABLE, DEMONSTRATING TERRITORIALITY [35] ................. 11

FIGURE 4: BOTH IMAGES ABOVE REPRESENT A GESTURE FOR DEALING WITH PRECESSION ON DIRECT-TOUCH TABLETOP INTERFACES USING TWO FINGERS, DRAGGING FINGERS APART TO ZOOM [10] ... 15

FIGURE 5: PRESENTS THE USE OF A MID-POINT BETWEEN TWO FINGERS ON A DIRECT-TOUCH TABLETOP INTERFACE, WHERE THE MIDPOINT BECOMES A CURSOR FOR SELECTING OBJECTS THAT REQUIRE PRECISION [10] ......................................................................................................................... 15


FIGURE 7: SCREENSHOTS FROM A MIXED REALITY APPLICATION USED TO DISPLAY ARTEFACTS [40, 41] USED FOR 3D VIRTUAL MUSEUMS ........................................................................................................... 18

FIGURE 8: SCREENSHOT OF THE GOOGLE MAPS SYSTEM SHOWING A LOCATION HAVING A TAG AND INFORMATION ASSOCIATED WITH THAT TAG [4] ................................................................. 21

FIGURE 9: SCREENSHOT FROM WWMX BROWSER UI SHOWING RESULTS (RED CIRCULAR REFERENCE POINTS ON MAP AND THUMBNAIL IMAGES IN RIGHT PANEL) OF A QUERY POSED GEOGRAPHICALLY [38] ................................................................................................................................. 22


FIGURE 11: SEQUENCE OF SCREENSHOTS SHOWING THE SWIPE GESTURE TO ATTACH A PIECE OF TEXT TO AN IMAGE. LEFT SCREENSHOT SHOWS PRIMARY TEXT OBJECT POSITIONED ABOVE THE IMAGE TO HAVE TEXT TAGGED TO IT. MIDDLE SHOWS TEXT BEING TAGGED WHILE STILL MOVING THE PRIMARY TEXT OBJECT. RIGHT SCREENSHOT SHOWS THE FINAL STATE AFTER THE SINGLE GESTURE IS COMPLETE WITH THE PRIMARY TEXT OBJECT LOCATED BELOW THE IMAGE AND THE TEXT TAGGED ................................................................. 29

FIGURE 12: SCREEN CAPTURE OF IMAGE RESOURCES. IMAGE ON LEFT SHOWS ALL IMAGES FACE UP AND IMAGE ON RIGHT HAS THE BOTTOM LEFT IMAGE FLIPPED ......................................................................... 30

FIGURE 13: SCREEN CAPTURE OF BLACKHOLE GARBAGE COLLECTION OBJECT ................................................. 31

FIGURE 14: ARCHITECTURE DIAGRAM SHOWING THE NEW COMPONENTS RELATIONSHIP TO THE MAIN COMPONENTS OF THE CRUISER SHAREPIC PHOTO SHARING APPLICATION THAT MYPLANNER BUILDS OFF ................................................................. 33
List of Figures

FIGURE 15: ARCHITECTURE DIAGRAM FOR THE BLACKHOLE COMPONENT LINKING IN WITH THE ENHANCED VERSION OF IMAGE ................................................................. 35

FIGURE 16: ARCHITECTURE DIAGRAM SHOWING THE STRUCTURE OF THE MAP OBJECTS AND THE LAYOUT THAT THE MAP OBJECT DEPENDS ON FOR POSITIONING IMAGE-BASED TAGS AND REGIONS .......... 37

FIGURE 17: LEFT IMAGE SHOWS MAP OBJECT FLIPPED OVER (SEMI-TRANSPARENT REPRESENTATION) AND RIGHT IMAGE SHOWS THE INTERACTIVE (NON-TRANSPARENT) MAP OBJECT .......................... 38

FIGURE 18: SCREEN CAPTURE OF A MAP OBJECT WITH 5 IMAGE TAGS REPRESENTED BY THE SPHERICAL OBJECTS. TWO SEPARATE USER’S TAGS ARE REPRESENTED BY THE DIFFERENT COLOURS (RED FOR USER 1 AND BLUE FOR USER 2) .......................................................... 39

FIGURE 19: SCREEN CAPTURE OF A MAP RESOURCE WITH 2 OVERLAPPING REGIONS REPRESENTED BY THE SEMI-TRANSPARENT GREEN RECTANGLES ........................................ 39

FIGURE 20: IMAGE SHOWS DUPLICATE MAPS WITH THE MAP ON LEFT IN INTERACTIVE STATE AND THE MAP ON RIGHT IN NON-INTERACTIVE (FLIPPED) STATE .............................. 40

FIGURE 21: IMAGE SHOWS DUPLICATE MAPS WITH BOTH MAPS BEING IN INTERACTIVE STATE AND CONTAINING 4 IMAGE-BASED TAGS CREATED BY THE SAME USER .......................... 40

FIGURE 22: IMAGE SHOWS DUPLICATE MAPS WITH BOTH MAPS BEING IN INTERACTIVE STATE AND CONTAINING 4 IMAGE-BASED TAGS CREATED BY THE SAME USER AND A SINGLE REGION IDENTIFIED CONTAINING 3 OUT OF 4 IMAGE TAGS .................................. 40

FIGURE 23: ARCHITECTURE DIAGRAM SHOWING THE REGION CREATION TOOL AND REGION COMPONENTS AND THEIR RELATIONSHIPS WITH EACH OTHER AND THE PARENT IMAGE CLASS, WHICH THEY INHERIT FROM ........................................ 43

FIGURE 24: SCREEN CAPTURE OF THE REGION CREATION TOOL ........................................ 44

FIGURE 25: LEFT IMAGE SHOWS REGION CREATION TOOL, A MAP OBJECT WITH A SINGLE REGION AND 4 IMAGE TAGS. RIGHT IMAGE SHOWS THE REGION TOOL POSITIONED, SCALED AND ORIENTATED TO SUIT A NEW REGION TO BE CREATED .......................... 45

FIGURE 26: LEFT IMAGE SHOWS THE STATE OF THE APPLICATION AFTER THE REGION TOOL HAS BEEN DWELLED ON WITH THE NEW REGION CREATED UNDERNEATH. RIGHT SHOWS THE NEW REGION ON THE MAP WITH THE REGION CREATION TOOL MOVED TO THE SIDE TO BE ABLE TO VIEW THE NEW REGION ................................................................. 45

FIGURE 27: DUPLICATE MAPS SHOWING THE NEW REGION REPLICATED ON BOTH .......... 45

FIGURE 28: LEFT IS STATE OF APPLICATION WITH TWO OVERLAPPING REGIONS AND THE PRIMARY TEXT OBJECT THAT IS DESIRED TO TAG TO THE REGION JUST CREATED IN FIGURE 26 AND FIGURE 27. RIGHT IS STATE OF APPLICATION WITH REGION TO TAG TEXT WITH FLIPPED TO ENABLE TEXT TAGGING ......................................................... 46

FIGURE 29: LEFT IS STATE OF APPLICATION AFTER PRIMARY TEXT OBJECT MOVED ACROSS BACK OF REGION, TAGGING THE TEXT TO THE REGION. RIGHT IMAGE SHOWS THE HIDING OF TEXT TAGS WHEN REGION IS NOT FLIPPED ................................................................. 46

FIGURE 30: ARCHITECTURE DIAGRAM SHOWING THE PIN COMPONENT WHICH IS THE REPRESENTATION OBJECTS OF IMAGE-BASED TAGS .............................................. 49

FIGURE 31: SCREEN CAPTURE OF IMAGE RESOURCES. IMAGE ON LEFT SHOWS ALL IMAGES FACE UP AND IMAGE ON RIGHT HAS THE BOTTOM LEFT IMAGE FLIPPED ................................................................. 50

FIGURE 32: LEFT SHOWS AN IMAGE TO BE TAGGED TO THE MAP AND A COPY OF A MAP OBJECT. RIGHT SHOWS THE IMAGE POSITIONED OVER THE MAP, SEMI-TRANSPARENT TO PREVENT OBSCURING THE DESIRED TAG LOCATION ......................................................... 50

FIGURE 33: LEFT SHOWS THE STATE OF THE APPLICATION AFTER THE IMAGE HAS BEEN DWELLED ON TO CREATE THE IMAGE TAG RELATIVE TO THE DWELL POINT (IN THIS CASE NEAR CENTRE OF IMAGE). RIGHT SHOWS IMAGE MOVED OFF THE MAP, REVEALING THE NEW IMAGE-BASED TAG (PIN) REPRESENTED BY THE BLUE CIRCULAR REFERENCE POINT ......................................... 51
FIGURE 34: DUPLICATE MAPS SHOWING THE REPLICATION OF IMAGE-BASED TAGS (PINS) ON BOTH MAPS FOR 6 IMAGE TAGS PERFORMED BY 2 SEPARATE USERS. EACH USER’S TAGS REPRESENTED BY A DIFFERENCE COLOURED TAG OBJECT (RED FOR USER 1 AND BLUE FOR USER 2) ........................................52

FIGURE 35: DUPLICATE MAPS IN ALTERNATE INTERACTIVE STATES SHOWING THE REPLICATION OF IMAGE-BASED TAGS (PINS) ON BOTH MAPS FOR 6 IMAGE TAGS PERFORMED BY 2 SEPARATE USERS, REMAINING VISIBLE FOR BOTH STATES. EACH USER’S TAGS REPRESENTED BY A DIFFERENCE COLOURED TAG OBJECT (RED FOR USER 1 AND BLUE FOR USER 2) ........................................52

FIGURE 36: ARCHITECTURE DIAGRAM SHOWING THE TEXT TAG COMPONENT AND THE MODIFIED ENVIRONMENT COMPONENTS RELATIONSHIP ........................................................................55

FIGURE 37: LEFT IMAGE SHOWS PRIMARY TEXT OBJECTS PILED IN THE ORIENTATION THEY APPEAR IN AT SYSTEM START UP. RIGHT IMAGE SHOWS PRIMARY TEXT OBJECTS USED FOR TAGGING TEXT TO IMAGES .......................................................................................................................... 56

FIGURE 38: ATTACHED TEXT OBJECTS ON THE BACK OF AN IMAGE REPRESENTING TEXT TAGS HAVE BEEN MADE ........................................................................................................................................ 57

FIGURE 39: SEQUENCE OF SCREENSHOTS SHOWING THE SWIPE GESTURE. LEFT SCREENSHOT SHOWS PRIMARY TEXT OBJECT POSITIONED ABOVE THE IMAGE TO HAVE TEXT TAGGED TO IT. MIDDLE SHOWS TEXT BEING TAGGED WHILE STILL MOVING THE PRIMARY TEXT OBJECT. RIGHT SCREENSHOT SHOWS THE FINAL STATE AFTER THE SINGLE GESTURE IS COMPLETE WITH THE PRIMARY TEXT OBJECT LOCATED BELOW THE IMAGE AND THE TEXT TAGGED ........................................ 58

FIGURE 40: SEQUENCE OF IMAGES FOR ATTACHING TEXT TO A REGION. LEFT SHOWS STATE OF REGION BEFORE TEXT TAGGED. MIDDLE SHOWS STATE OF REGION AFTER TEXT TAGGED. RIGHT SHOWS REGION FLIPPED, HIDING TAG OBJECT ............................................................................................................................... 58

FIGURE 41: ARCHITECTURE DIAGRAM SHOWING METADATA STORAGE COMPONENT (TAG COMPONENT) AND THE RELATIONSHIP TO THE BASIC IMAGE COMPONENT .......................................................... 61

FIGURE 42: SYNTAX DESCRIPTION FOR METADATA FILE STRUCTURE RELATING TO IMAGES ........................................ 61

FIGURE 43: EXAMPLE OF A METADATA FILE FOR AN IMAGE CALLED SITRGIMAGE1.JPG ....................................................... 62

FIGURE 44: SYNTAX DESCRIPTION FOR METADATA FILE STRUCTURE RELATING TO REGIONS ........................................... 63

FIGURE 45: DIAGRAMMATIC EXAMPLE OF HOW METADATA IS COMBINED WHEN IMAGE-TAGS WITHIN A REGION THAT CONTAINS TEXT WITH THE TEXT LISTING COMBINING BOTH IMAGE AND REGION TEXT ENTRIES, ASSOCIATING IDENTIFIERS FOR EACH ENTRY .......................................................................................................................... 64

FIGURE 46: VIDEO CAPTURE OF A PAIR OF PARTICIPANTS USING THE MyPlanner SYSTEM .................................................. 66

FIGURE 47: LEFT IMAGE SHOWS THE CENTRALISED CONTROL INTERFACE AND THE RIGHT IMAGE SHOWS THE REPLICATED CONTROLS INTERFACE FOR TEAMTAG [30] WHERE THE CONTROLS CONTAIN PIECES OF METADATA USED TO ASSOCIATE TEXT WITH IMAGES .......................................................................................................................... 72

FIGURE 48: EXAMPLE OF A LAYOUT OF 3 REGIONS ON A MAP OBJECT THAT PARTICIPANTS SHOULD TRY CREATE DURING THE USER EVALUATION PROCESS .......................................................................................................................... 76

FIGURE 49: VIDEO CAPTURE OF EVALUATION AFTER A USER HAD JUST PERFORMED AN IMAGE-BASED TAG (NEW IMAGE TAG/PIN LOCATED TOWARDS CENTRE OF SEMI-TRANSPARENT IMAGE) ................................................. 83

FIGURE 50: VIDEO CAPTURE OF USERS INTERACTING WITH DUPLICATE MAPS ................................................................. 87

FIGURE 51: VIDEO CAPTURE OF IMAGE TAGS (REFERENCE POINTS – DARK DOTS) ON THE MAP WITHIN 2 OF THE 3 REGIONS (REPRESENTED BY BLUE RECTANGLES) CIRCLED IN RED .......................................................................................................................... 88

FIGURE 52: VIDEO CAPTURE OF THE SWIPE GESTURE. LEFT SHOWS A COLLECTION OF TEXT OBJECTS THAT THE PARTICIPANT WANTED TO ASSOCIATE WITH THE FLIPPED IMAGE. RIGHT SHOWS THE TEXT OBJECTS IN A PILE THAT HAS BEEN FORMED BY MOVING EACH TEXT ACROSS THE BACK OF IMAGE TO DO THE TAG .......................................................................................................................... 91
LIST OF TABLES

TABLE 1: SUMMARY OF HARDWARE CAPABILITIES OF SOME OF THE TABLETOP IMPLEMENTATIONS [9, 16, 22, 27, 33, 35, 36] ................................................................................................................................. 10

TABLE 2: SUMMARY OF SOFTWARE CAPABILITIES OF SOME OF THE TABLETOP IMPLEMENTATIONS [9, 16, 22, 27, 33, 35, 36] ................................................................................................................................. 14

TABLE 3: USER EVALUATION TIMES TAKEN BY EVALUATORS TO COMPLETE THE TASK 1 USING EACH OF THE SYSTEMS. PAIRS 1-3 USED TEAMTAG-CONTROL FIRST AND 4-6 USED MYPLANNER FIRST. ....82

TABLE 4: AVERAGE TIMES TAKEN PER PAIR USING EACH OF THE SYSTEMS WHEN EACH SYSTEM WAS USED FIRST AND SECOND......................................................................................................................... 82

TABLE 5: PARTICIPANT (E.G. 1:A MEANS PAIR 1 USER A) PREFERENCES RELATING TO TEAMTAG-CONTROL AND MYPLANNER WITH PAIRS 1-3 USING THE TEAMTAG-CONTROL SYSTEM FIRST AND PAIRS 4-6 USING MYPLANNER FIRST......................................................................................................................... 84

TABLE 6: PARTICIPANT (E.G. 1:A MEANS PAIR 1 USER A) PREFERENCES RELATING TO THE TEAMTAG-CONTROL PILES AND MYPLANNER REGION METHODS OF ASSOCIATING TEXT TO MULTIPLE IMAGES (X SIGNIFIES USERS THAT LIKED THOSE METHODS).................................................................................. 92
Chapter 1: INTRODUCTION

This thesis explores the support of tabletop interaction for organisation of content using Tagging and Maps. Tabletop interfaces have the potential to provide a new form interface to pervasive computing environments. This thesis explores one particularly promising direction, which we call MyPlanner and is based upon maps with tagging. Tabletops make a table surface act as a user interface. A computer screen image is projected onto the table, providing the output to the users. The user can interact with the interface by using either a special pen-like device or, in the case of special table surfaces, the user’s fingers which the system detects. Tabletops present quite a different form of user interactions from conventional computers and require research to identify how to make natural interface elements and methods. This thesis explores new ways to interact with tabletop interfaces.

Tagging is a level of personalisation, where users can identify things such as what part of the work is theirs (e.g. users can attach their name to a part of the document indicating they did that), how things relate to another (e.g. the linking of a vase from Egypt with a Pharaoh by sharing a common word such as “Egypt”), and using them to better identify things for future references (e.g. the association of a word with an object such as a vase so that by searching that word in the future will return a reference to that particular vase). There has recently been a dramatic growth in systems that supported tagging as part of web based or conventional systems. However, there is little work on of these features in tabletop interfaces although there has been one project [28, 30]. To fill this gap, this thesis aims to explore this area of tagging in tabletop interfaces. There are currently systems that allow for tagging with text in tabletop interfaces through consecutively selecting metadata then object that is to be tagged.
1.1 Motivation

This thesis will explore these issues in the context of an environment to create a virtual museum. This makes a good context for collaborative tagging. We now describe that context in more detail before returning to our core goals of exploring tagging on tabletop interfaces. A virtual museum is one that takes digital form and the only constraints are with the technology that implements them. By constraints the digital collection of exhibits is not limited to physical space, only the storage space available. Digital museums can present all exhibits that are available through the backend database, instead of storing exhibits in store rooms due to lack of display space on the museum floor which prevents visitors from viewing them.

Virtual museums are available on personal computers over the Internet these represent attempts to make museums more accessible and present more exhibits than is possible in a physical museum, such as the “Virtual Museum of Canada” [6] Two of the more common usages of virtual museums are to provide a preview of the physical museum, and allowing remote visitors who do not have access to the physical museum to view the exhibits. However, the environment provided by a personal computer does not provide aspects of group collocated collaboration and a natural environment for users to interact with what is presented to them. MyPlanner, the system created in this thesis, aims to enable users to create their own personal museums using object tagging.

The core facility of MyPlanner is to provide a framework for users to create their own museums using image-based tagging and geographical organisation through interactive maps. A map is a 2-Dimensional layout as in a physical map or map-like display. Users will be able to assign images/objects to particular locations on a specialised map object. Images and regions will hold details relating to their location/region on the map where they were tagged, also storing their own text based tags for labelling and identification purposes. In the future users will be able to view their created museum by having their basic user model (in the context of MyPlanner this will be a list of words that indicate what are a user’s interests) entered and the system will present which regions/exhibits are of greater interest to the based on the tags.
The end goal of this research is a well-founded prototype system where a small collocated group of users (approximately 2-4 users) can collaboratively organise digital objects by using the tabletop interface. These tags will be images, and text, which will be associated with a location on a special digital image (e.g. the map of a museum). The interface of this prototype is evaluated by small groups of users collaboratively interacting with the interface.

1.1 Motivation

Considerable research has been conducted in areas of tabletops, personalisation and museums. Although research has integrated some concepts of personalisation and tabletops, there has been minimal research conducted in the area of providing image-based tagging and interactive mapping for organisation of digital content on tabletops.

Tabletop interfaces provide a novel computer interface which supports collocated collaboration amongst a group of users. Currently architects, engineers, fashion designers tend to use tables, spreading out images and planning things by dealing with physical pieces of paper. To take just one other motivating example, a group of users could geographically organise digital collections of photographs taken while on holiday using regions to identify sections of a trip.

There are numerous potential applications for using tagging on tabletops with users able to associate metadata with piles of digital information. There has been very little exploitation of interfaces to support tagging on tabletop interfaces [28, 30] and this involved just text linked to images. Maps provide an ideal context for implementing and evaluating the possibility of image-based tagging with respect to particular locations on an object. However, the concept of image-based tagging and adding regions to a map would provide geographical organisation of content instead of the hierarchical organisation seen in file systems. The appeal of exploring maps to organise and tag objects in a tabletop interface is that it seems to offer promise of easy gesture-based interaction.
1.2 Definition of Problem

The main focus of this thesis is enhancing collocated collaborative applications on tabletop interfaces through the evaluation of tagging to organise digital artefacts using interactive maps. The general question is how tabletops support organisation of content.

This thesis approaches the generic research question by focusing on how tabletops support organisation of content using tagging and interactive mapping. Therefore the main focus of this thesis is evaluating initial stages of geographical organisation of content on tabletop interfaces.

1.3 Research Contribution

The main areas relate to presentation of virtual museums, personalisation, tagging and interactive mapping in collocated collaborative environments. These contributions of this thesis are:

- Improving understanding in tabletop personalisation through the enhancement of tagging
- Explore the applications of interactive mapping in tabletop interfaces
- Replication of actions between objects on tabletop interfaces
- Association of metadata through tagging of objects on tabletop interfaces
- Exploring an alternative method of text-based tagging
- Evaluating collocated collaborative image-based tagging on tabletops
- Evaluating collocated geographical organisation of digital content on tabletops
- Explore a framework for designing and presenting museums on tabletops
- Exploring gestures for performing tagging of objects to locations and regions
- Discovery of a swipe gesture
1.4 Thesis Outline

This chapter has provided an explanation and a breakdown of the research question, in relation to the context used to evaluate the hypothesis.

Chapter 2 provides an insight into some of the most relevant aspects of research that forms the foundation of this thesis.

Chapter 3 outlines user interface elements of the prototype system used to evaluate the concept of collocated collaborative image-based tagging and interactive mapping on tabletops for organisation of digital artefacts.

Chapter 4 introduces the previously developed elements of the user interface and how the new elements provided by this thesis link with the original architecture.

Chapter 5 discusses the elements of interactive mapping provided by the prototype system as a context for image-based tagging and geographic organisation of digital content.

Chapter 6 discusses the elements of the system prototype that provide region identification on map objects and the interaction with these regions for geographic organisation.

Chapter 7 discusses the key element of image-based tagging in the context of the system prototype providing a framework for organisation of digital content using geographical organisation.

Chapter 8 describes an alternative method for text-based tagging implemented in the prototype system to add metadata to images and geographical structure.

Chapter 9 outlines the methods of storing and handling metadata to represent the variety of tag types used within the system prototype.
Chapter 10 outlines the plan followed for the evaluation process of the system prototype. This includes details of the design of the qualitative user evaluations and questionnaire used to validate the claim of collaborative use of image-based tagging and interactive mapping for organisation of digital artefacts.

Chapter 11 discusses the results and conclusions obtained from the qualitative user evaluations and questionnaires used to validate the elements of the prototype system.

Chapter 12 outlines potential future work and research that expands on the outcomes of this thesis combined with other tabletop research.
Chapter 2: BACKGROUND

The areas covered in this section are based on several of the key aspects that are drawn together to form the MyPlanner research. Tabletop interfaces are the main aspect of the research with the objectives focusing on certain pieces of functionality and use relating to tabletop interfaces. The functionality involves concepts relating to user personalisation with tagging being the main contribution of the research, which is a form of user personalisation. All this is being prototyped within the context of museums. Since the tabletop presents a different form of interface and user interaction to computers understanding of the current types of interfaces and frameworks for museums is required to know what is possible while maintaining the key. There is a contribution being made in combining tabletop interfaces with virtual museums since there is little evidence of presenting such information in this context on tabletop interfaces with personalisation and a framework for creating museums.

2.1 Tabletops

2.1.1 Overview

Tabletops are the use of a table that provides an interactive computing environment. Interaction with tabletops can take various forms such as physical touch-based interfaces and stylus devices. Research into tabletops has several areas including design, interaction, collaboration, distributed systems and ubiquitous computing. These areas of research make use of several methodologies to evaluate the systems and research contributions: observations, surveys, statistical/data analysis using existing materials, and ethnographic studies. The results of these studies are then used to provide support for the claims being made and form the foundation for further research to be performed.
2.1 Tabletops

2.1.2 Hardware Features of Tabletop Interfaces

There are many variations in current tabletop system implementations. Each system presents a variety of methods for performing tasks. An example of this is the rotation of objects can be performed differently under certain systems. These variations make it difficult to compare different systems since there is no standard to be compared against. Tabletops present several challenges for developing effective interfaces. For example the Lumisight Table [22, 27] and DiamondTouch [16, 33, 35, 36] are two of the possible designs for presenting information to the users.

The Lumisight Table makes use of internal projectors limiting the number of users to be the number of projectors installed in the table itself. The reason for the limitation is that each projector is used to present a different perspective of the contents of the table to each user. One of the difficulties with the Lumisight Table is when users are collaboratively using the table gestures (e.g. physically pointing) to a particular object on the table interface may relate to a different object to another user.

![Figure 1: The image on the left shows the internal elements of the Lumisight table, with multiple projectors giving each user a different perspective of the image being projected [22]. The image on the right shows the appearance of the Lumisight table [22]](image)

The DiamondTouch Table [16, 19, 32, 33, 35, 36] uses a different approach. This interface uses a single projector positioned above the table with the same image being presented to all users around the table. The number of users is not limited to the number of projectors present; instead the number of users depends on the number of specialised chairs/pads. These chairs/pads act as a circuit where each user is touching the table surface completing the circuit and registering a unique ID with the system to establish which user interacted with the system.
Research into the techniques for user interactions was extended by evaluating natural gestures [11, 16, 17, 19, 33, 42] and there implementation in tabletop systems. There has been consideration in use of speech as a method of interaction while interactively using a distributed tabletop [17]. However, in a collocated collaborative organisation application, speech input could be misinterpreted when users suggestively propose planning ideas. Wu, Shen, Ryall, Forlines and Balakrishnan et al. [42] evaluated gesture implementations in tabletop systems. Direct-Touch was the focus where the use of gestures within the system interaction was evaluated and some considerations outlined. Gesturing is an important part of tabletop interfaces since one of the key aspects of tables is to enable users to interact naturally with the table as if they were performing similar tasks using real objects. The introduction of relaxation in the gesturing and interaction is important since direct-touch interfaces would require the same posture to be maintained throughout the completion of the gesture. Relaxation is where after the initial registration of a gesture has been achieved, the user is allowed to complete a gesture through touch alone and is not required to maintain the exact posture. With the introduction of the image-based tagging and interactive mapping, there may be the requirement of new gestures to allow for users to naturally interact with the new features.

Hardware capabilities for some of the major classes of tabletop interfaces are outlined in Table 1 below. The summary below outlines three different styles of tabletops, the DiamondTouch, Lumisight, and the Cruiser/Mimio implementations. The Cruiser/Mimio implementation is analysed since that is the implementation that the MyPlanner project will be using. The capabilities of tabletops are quite varied. This
Table 1: Summary of Hardware capabilities of some of the Tabletop implementations [9, 16, 22, 27, 33, 35, 36]
2.1.3 Collaboration

Collaboration is one of the key aspects of tabletops that have attracted the attention of researchers. Through evaluation of how users collaborate while using tabletops, interfaces and interaction methods can be modified to enhance the users’ experience in performing tasks. Smeaton, Foley, Gurrin and Lee et al. [36] evaluate collaboration in relation to searching video using a tabletop interface. This paper identifies elements that should be considered when designing a collaborative environment on a tabletop: the division of tasks amongst group members, the degree of group awareness and the degree of user coordination. Group awareness refers to how aware a user is of other user activities in the system. This also extends to user coordination where user activities could conflict with another user if they do not coordinate their activities to allow both users to use the table accordingly.

![Activity Plots](image)

**Figure 3: Representation of an activity plot from studying the areas of where users interact with objects on the table, demonstrating territoriality [35]**

Morris, Paepcke and Winograd et al. [28] extended on research into collaboration by evaluating how users search digital media as a group. This also examined group awareness while the users interacted with the table to perform a set of defined queries. The data that is queried in this system has been tagged earlier which provides some interest with respect to this thesis. Morris, Paepcke and Winograd et al. [28] use pre-defined tags for users to collaborative search a collection. This is one aspect that MyPlanner will attempt to further by allowing users to add tags similar to the del.icio.us [1] Internet tagging service that allows user tagging.
The research performed by Morris, Paepcke and Winograd et al. [28] was later extended to evaluate another method of tagging using a list of predefined text tags and allowing users to drag and drop these onto images to tag them [29]. This concept of having tags listed in front of each of the users was the replication control. Users generally found the replication controls more comfortable to use than the original centralised controls that were introduced in Morris, Paepcke and Winograd et al. [28]. The idea of having these tags as objects to be dragged to an image and associated with that image is where MyPlanner will expand even further by associating another image/object to a specific location on a map (the map is a special form of image).

Understanding how users collaborate and interact with each other on tabletops is an area that is still being researched and understood. An aspect of collaboration include how users assign space on the interface [35] i.e. personal, shared and public space on the application’s interface during interaction. Further research has been carried out evaluations on collaborative coupling [37] to attempt in assisting the design of collaborative interfaces for tabletops. Tabletops are an interface that is more open to collaboration than regular desktop computers and understanding how users interact with the interface and each other is an important part of expanding the applications of tabletop systems.

### 2.1.4 Software Features of Tabletop Interfaces

Tabletop research has influenced development of applications that support user group collaboration and interaction. One application that has been explored on tabletops is the use of images and photo sharing software. The viewing of images is an integral part of MyPlanner since exhibits will be displayed as images and users interacting with these images in similar ways to those explored in various photoware. The Cruiser tabletop currently provides an application relating to the concept of photo sharing called Sharepic [9] which is the basis for developing MyPlanner.

One of the software applications that has been applied to tabletops is Photoware [14, 15, 18]. This is a system that is based on the collaborating and sharing of photos amongst groups of people. There has been research into sharing on several types of systems such as tabletops, desktop computers, and portable devices such as PDAs.
2.1 Tabletops

The concept of photoware has developed from research with other devices and is not just isolated to the use on tabletops. Evaluations of the way people use photoware on other devices, not just tabletops, have been conducted to further research and understanding. Therefore software developed on tabletops is not isolated to that one implementation; they can be implemented on other devices. In addition the opposite can also apply with pre-existing applications being applied to tabletops, with the possibility of modifications being required to allow for simultaneous interaction by multiple users and the restrictions of interacting only by touch or pen.

Other software capabilities of tabletops are summarised in Table 2 for some of the many tabletop implementations. The summary shows just the core features of these major types of tabletop systems. These provide a comparison between some of the different tabletop interfaces. Since the Cruiser/Mimio is the basis of the MyPlanner project the features outlined below are related to general tabletop interfaces and more specifically the current application on the Cruiser.
Photoware and other tabletop applications that involve the interaction with images form the basis for MyPlanner to enhance the possible uses of these applications. These photoware applications have the potential to be used for planning documents, which is what MyPlanner aims to provide: a more effective framework to allow users to do more fluently.
2.1 Tabletops

Figure 4: Both images above represent a gesture for dealing with precession on direct-touch tabletop interfaces using two fingers, dragging fingers apart to zoom [10]

Figure 5: Presents the use of a mid-point between two fingers on a direct-touch tabletop interface, where the midpoint becomes a cursor for selecting objects that require precision [10]

The problem with precision is a problem that is both hardware and software. Software needs a method of dealing with the issue of user interaction precision. The human finger cannot easily select a single pixel because of the size which causes selection of many in the attempt of selecting that one distinct pixel. Benko, Wilson and Baudisch et al. [10] explored a variety of methods to address the problem of precision on tabletop interfaces (Figure 4 and Figure 5). The methods were aimed at direct-touch interfaces such as the DiamondTouch system where there were no stylus devices, like uPens, to help increase precision of interactions. Software applications on tabletops are required to consider implications of precision and the minimum size of objects that users can effectively interact with.
2.2 Virtual Museums/Environments

The primary function of a museum is to acquire, classify, preserve and research collections. The secondary function is the educational activity using collections such as exhibits [25]. The move towards digital museums is to enhance museums so that those functions can be improved. The area of virtual museums that apply to MyPlanner is the creation of a framework that allows users to design their own virtual museums while maintaining the important concepts behind what defines a virtual museum (the primary and secondary functions). With the scenario of creating virtual museums, some background is required to understand what constitutes a virtual museum, interaction methods, interfaces currently available and functionality currently available.

Koshizuka and Sakamura et al.[25] explored ways to build the Digital Museum, identifying some sub problems such as; museums with digital archives; openness of the museum material; presentation; personalisation; distribution. Koshizuka and Sakamura et al. [25] take two approaches of furthering the development of digital museums: virtual museums using Multi-Media Multi-User Dungeon (MMMUD) systems to develop a virtual environment; augmented-reality based on digital devices such as PDA’s and HMD to interact with physical exhibits.

Photorealistic virtual museums have become possible with the advances in 3D-modelling and virtual reality. Huang, Chen and Chung et al. [20] evaluated a virtual museum that was based on a desktop computer where users would interact with a photorealistic museum through a cube-like control device. This system was based on using augmented panorama to display the museum exhibits and although considered interactively successful it presented a low level of collaboration since only users situated around the screen could collaborate amongst each other. It did show how museum exhibits can be presented in a photorealistic way where the exhibits appear almost real to the users viewing them.
The advances in virtual reality have enabled virtual museums to present exhibits using 3D modelling in a way that users can view them in fine detail as shown in Figure 6 and Figure 7. However, with virtual reality, it is complex to set up museums which would usually require a skilled developer. Wojciechowski, Walczak, White and Cellary et al. [40, 41] evaluated the development of an architecture that museum staff can use to set up their own virtual museums. The architecture provides three key components to assist museum staff with the content production, content management and visualisation components. With presentation of the exhibits being through tabletop interfaces, the way that these exhibits are presented graphically is an important consideration in MyPlanner. 3D modelling is possible, except it presents difficulties with perspective, making a requirement for ways to present the exhibits to the users graphically.
Frameworks are being developed to assist in the production and design of virtual museums. Another framework that has been researched is for the rapid authoring of audio recordings using devices. Hull, Clayton and Melamed et al. [21] performed research on the development of a framework that allowed the design of audio recordings which are deployed on devices such as PDAs, and when the person using the PDA enters a designated area, a media recording plays what is designated for that physical region. This framework could be applied to virtual museums and various tours where visitors move around the museum with various audio recordings being associated to locations within the museum. This relates to MyPlanner by the use of regions to be used to assigning content as a PDA enters a designated region and the frameworks used to create these regions.

T. Kirner, C. Kirner, Kawamoto, Cantão, Pinto and Wazlawick et al. [24] researched the development of a collaborative environment that allowed users to create a virtual world and collaborate amongst other users on a distributed system. The system that was examined focuses on a client server model that creates a virtual world based on user requests. Collaboration provided by this system is in the form of messaging.
between users. Virtual environments such as this could be applied to a tabletop interface where groups of users can be distributed with multiple users at a single user interface. The Virtual Museum Project [24] involves the development of tools that make it possible for the creation of virtual worlds.

The main aspect of MyPlanner is a framework that allows users to create museums and plans through tagging. Research carried out in Hull, Clayton and Melamed et al. [21] and T. Kirner, C. Kirner, Kawamoto, Cantão, Pinto and Wazlawick et al. [24] provide some of the understanding and basics required to apply a framework to a tabletop interface that allows the creation of museums.

### 2.3 Personalisation

#### 2.3.1 Overview

Personalisation is an aspect that is required to be researched to understand the direction that MyPlanner is heading towards in addition to the concept of tagging which is a form of personalisation. The goal of this thesis is to evaluate a prototype system that provides tagging, which can be enhanced in future work to evaluate the inclusion of things such as user modelling and recommendations.

#### 2.3.2 Personalisation that influences Direction of MyPlanner

Systems that provide personalisation and recommendations should also allow users to scrutinise how those decisions were made. Bright, Kay, Ler, Ngo, Niu and Nuguid et al. [12] outlines a system that presents different detail levels to users based on user data collected. The recommendations are based on the amount of time spent by users viewing and/or interacting with an object. The system uses an ontological approach to link the artefacts together in order to evaluate a tour path for the user. The system makes use of Personis Lite which allows adaptive systems to manage evidence collected for the use in developing user models. The concept of an ontological

---

approach for linking museum artefacts is also mentioned in [23]. The more important issue of these papers is the use of an ontology and data gathered from usage to make recommendations to the users. The concept of recommendations made from the linking of key words is an important part of providing personalisation in MyPlanner with respect to using the produced museums/maps.

Personalisation research has been conducted on tabletop interfaces in addition to the regular systems that occur on PCs. Some research in tabletop interface personalisation includes the use of personal space and personalised audio channels. Personal space is something that was influenced by observing users interacting with the tabletops [35]. Observers noted that users interacting with the tables define three kinds of areas; personal space, shared space and common group space. Scott, Sheelagh, Carpendale and Inkpen et al. [35] introduce these three kinds of areas that users tend to establish during use. Through this research some tabletops have defined personal areas for users that are isolated and unable to be modified by other users. The Cruiser tabletop that is the focus of the thesis has personal space defined resembling the outcomes of the research by Scott, Sheelagh, Carpendale and Inkpen et al. [35]. The technique used for personal space in Cruiser is copying an image once it passes a certain point, which can be applied to the tagging functionality of a map with an image being copied once it is assigned to a particular region.

2.3.3 Tagging

Tagging is an area of personalisation and one of the main focuses of MyPlanner. Tagging systems exist on personal computer interfaces such as Flickr [3, 26, 34], Google Maps [4] (Figure 8) and del.ici.ous [1]. The reason for tagging on tabletops being different to these current applications is the idea of removing the keyboard and mouse to create a more natural interactive interface using gesturing instead of pressing buttons.
Tagging systems on personal computers have the main issue of users not tagging objects the same as others, resulting in difficulties in searching and managing of collections in social tagging systems. This vocabulary problem has had considerable research into methods of approaching it in a variety of ways, such as the use of complex ontology and formation of social network vocabularies [26, 39].

The concept of tagging is not limited to attaching text to objects. Tags can also involve attaching images or groups of objects to particular locations on another. Geographical location tags [38] are an aspect of tagging that involves attaching images and metadata to particular locations (Figure 9). For example users can attach images of a house with information about the house to a specific geographic location as part of a real estate system. Geographical tagging on personal computers is not the same as on tabletops due to the absence of a keyboard and mouse, since users can enter in addresses where the system can interpret freely to make a tag [4]. Even in systems that enable the identification of regions and associating text tags to the regions [7], the requirement of a keyboard to enter tags and search data means that similar functionality will require prototyping and evaluation on tabletops to make it suited to the more natural interactive interface.
Figure 9: Screenshot from WWMX browser UI showing results (red circular reference points on map and thumbnail images in right panel) of a query posed geographically [38]

The absence of keyboard and mouse in tabletop systems combined with the desire for collaboration effects the way tagging is performed compared to personal computer systems. MyPlanner explores tagging on interactive mapping, such as maps of buildings for designing museums or geographic maps for planning and reflecting on holidays.

2.4 Summary

Overall the this chapter provided insight into the key aspects of the thesis to be conducted in the area of tagging and interactive maps in the context of creating personalised virtual museums on tabletops. Some of the literature identified areas to be considered when designing and developing systems that focused on the use of certain technologies. Others examined the effectiveness of applications being implemented on new types of systems i.e. as tabletops.

The key main areas covered related to tabletop interfaces with a comparison between alternate systems (DiamondTouch and Lumisight). To understand the novelty of
tabletop interfaces the methods of interaction and presentation were outlined in a collocated collaborative environment. Tabletop interfaces present an alternative interface to the current virtual museum systems that use augmented reality. Therefore understanding of the requirements for a museum to be digitally presented to users was necessary since the scenario that the thesis will be evaluated with is the context of planning a virtual museum.

There are many systems that provide tagging and interactive maps, and few have focused on tabletop interfaces. This thesis focuses on tagging and maps in a collocated collaborative environment that doesn’t rely on the mouse and keyboard for interactions. Research conducted indicated that most systems rely on some keyboard input of text-based tags and backend processing to identify location tags.

Research conducted creates the foundation for understanding how image-based tagging can be applied in a collocated collaborative environment with text-based tagging providing contextual meaning to organisation of digital content using maps.
Chapter 3: USER PERSPECTIVE

3.1 Overview

This Chapter introduces the user interface of the prototype MyPlanner, focusing on the elements provided by the Cruiser tabletop [9]. The key components that create the user perspective that can be classified into two distinct categories, hardware and software. The hardware components consist mainly of technologies required to present the application and provide the methods of interaction. Hardware, such as Mimio whiteboard capturing sensor, Mimio pens and a digital data projector, are the visible aspects of the Cruiser Tabletop. These hardware elements were a given for this thesis.

The real motivation from a user perspective is the MyPlanner software, with its visible interaction components such as the map component, regions, text and images. Another set of elements of MyPlanner are the gestures which users perform to trigger events in the absence of the mouse and keyboard. The elements of software described in this chapter are parts of the Cruiser used by MyPlanner, with the new swipe gesture incorporated.

3.2 Hardware

Although this thesis did not make contributions to the tabletop hardware, this section presents an overview of that aspect to secure as a foundation for describing MyPlanner and the significant constraints that the tabletop hardware imposes on design of software like MyPlanner.
3.2 Hardware

3.2.1 Mimio Interactive

Interaction mechanisms for tabletop interfaces can involve direct touch or the use of stylus devices such as pens. The Cruiser tabletop that MyPlanner expands on makes use of Mimio Interactive whiteboard capturing technology (Figure 10) with the pens used as the stylus devices for interaction.

![Components of the Mimio whiteboard capturing system](image)

Figure 10: Components of the Mimio whiteboard capturing system; the sensor, four pen stylus devices and an eraser [5]

Each person using the system has a pen which enables them to perform gestures to trigger events. The original Mimio drivers have been previously modified to enable the interactions to be registered with applications running on the Cruiser tabletop (University of Sydney). Interactions are sent using infra-red and ultrasound to track the positions of the stylus pens.

3.2.2 Top-down projection

There are two main alternatives for projecting the computer image onto a table surface. The Cruiser uses top-down projection, which is a digital data projector positioned above the table, projecting the image down onto the table through reflection off a mirror.

The use of top-down projection allows a regular table to act as the interface, making the users feel more comfortable. Comfort is the result of making the interface feel as
natural as an actual table and when the system is not active, the table can be used as usual.

3.3 Gestures

3.3.1 Overview

Tabletop interfaces focus on making user interaction natural, removing the presence of keyboards and mouse. This results in requiring alternative methods for user interaction. The Cruiser tabletop makes use of pen stylus devices to enable user interaction.

Gestures are actions performed by users using the interaction method to trigger events. These gestures include selection, movement, dwelling, flipping, rotation, resizing of objects and the new swipe gesture. These are the base gestures provided by the Cruiser Tabletop and have been incorporated into the new elements provided by MyPlanner, allowing for the additional functionality required. This section describes the foundation gestures in Cruiser as these were the starting point for designing MyPlanner.

3.3.2 Select

The select gesture is the highest level of interaction. This gesture identifies an object by making contact with the table in the area representing an object. To perform any of the other interactive gestures provided by MyPlanner, selection of an object is required before the extended attributes of the other gestures can be performed.

3.3.3 Movement

The movement of objects represented in applications on tabletop interfaces is a critical part of providing a cognitively natural device. Tabletop interface applications attempt to make the movement of objects similar to the movement of physical objects such as sliding a piece of paper across a regular table surface.
This gesture is performed by applying pressure with the interaction device (direct touch or stylus device) and dragging it across the table while maintaining contact with the desired object. While this interaction is being performed the object that is selected is moves relative to the interaction device’s position, remaining with it.

### 3.3.4 Dwell

The dwell gesture is an extension on the selection gesture. This gesture is the application of consistent pressure on the interaction device on the same location for a pre-determined short amount of time.

The absence of keyboard and mouse dwell provides an added dimension to the number of gestures available. While select is defined by pressure on an object, dwell provides time as a dimension of interaction. The time dimension relates to how long pressure is applied to the device while remaining stationary.

Audio events are an element of the original photo sharing application implemented on the Cruiser Tabletop. To record audio, dwelling on an image (that is not located in the area of an unflipped map) will initiate the recording of audio. Once the necessary audio is recorded, dwelling on the image again while under the same conditions will stop the recording process, attaching a yellow numbered circle representing the audio track. Audio can be replayed by selecting the appropriate audio circle and dwelling on it.

### 3.3.5 Flipping

The environment presented on the Cruiser Tabletop is 3-Dimensional space. Therefore the back of objects can provide an additional dimension/view for the object. This alternative view can provide added functionality or modality, such as dwelling on one side can trigger a different event from the other side.

The flipping gesture enables the changing of views, increasing the number of events that can be triggered. In MyPlanner, the flip gesture is used to toggle the map between
interactive states, and images that have been flipped over can have text tagged to them.

Flipping is performed by selecting one of the triangular regions that are positioned towards the centre of each edge of an object. Once the triangular region is selected, dragging the pen to the opposite side will flip the image/object.

### 3.3.6 Rotation

Perspective is an important issue for objects on tabletop interfaces. Allowing objects to be rotated provides a method of addressing the orientation problem. Rotation of objects on the Cruiser Tabletop allows for 360 degree rotation.

Rotation allows participants the ability to view images/objects and read text more fluently. If an object containing text is oriented such that it appears upside down to a participant, rotation can be performed until the orientation enables. Rotation is performed on Cruiser by selecting the object in one of the corner triangular regions and dragging the corner in either clockwise or anti-clockwise direction from the original position to rotate the object. The object then rotates, seemingly because that corner stick to the users’ finger (pen stylus), rotating the image.

### 3.3.7 Resize

Precision of interactions and readability of small text is an area of tabletop interfaces where a variety of methods exist across the different tabletop implementations. The Cruiser allows users to resize objects to assist in the precision when viewing or selecting small parts of objects. The resizing also increases the size any text contained within that object, improving the readability if the size is the only attribute making it unclear.

Resizing of objects is performed by selecting a corner triangular region of an object and dragging the corner in a specific direction. Dragging the corner away from the centre of the object increases the size, and dragging towards the centre decreases the
size of the object. These corner triangular regions are also used for rotation of the object.

### 3.3.8 Swipe

The MyPlanner has introduced a new gesture to the collection previously mentioned. This gesture is referred to as the swipe gesture. The gesture is an enhancement of the movement gesture. Swipe gesture is the movement of an object across another with no particular final target area. During the movement process once the moving object passes over the area of another an event is triggered. Figure 11 shows an example of the swipe gesture to tag text to an image.

![Figure 11: Sequence of screenshots showing the swipe gesture to attach a piece of text to an image. Left screenshot shows primary text object positioned above the image to have text tagged to it. Middle shows text being tagged while still moving the primary text object. Right screenshot shows the final state after the single gesture is complete with the primary text object located below the image and the text tagged.](image)

Relating swipe to Fitts’ law [2], one can predict that the amount of time a user takes to perform the gesture is relatively low, due to the final target area of the moving object being very large.

There has been no references located that outline a gesture that is similar to the swipe in functionality. Therefore, the swipe is a new gesture that evolved from natural user interaction with the MyPlanner prototype.
3.4 Elements of Cruiser

3.4.1 Image Objects

MyPlanner depends on images to represent exhibits and sections of documents to be used in creating virtual museums, zoos and other planning tasks. Image representation is an element previously provided by the photo sharing application implemented on the Cruiser Tabletop. Figure 12 shows an example of images and representation of an image when flipped.

![Figure 12: Screen capture of image resources. Image on left shows all images face up and image on right has the bottom left image flipped](image)

Performing the dwell gesture can trigger different events. These depend on the location of the image relative to other objects and whether the image is flipped. The events include the recording of audio, playback of audio and tagging the image to a location on a map.

The recording of audio is only enabled when the image is not positioned directly above a region, pin or map object. If the image is over either of these the dwell gesture performs an image tag.

3.4.2 Blackhole

The original Sharepic [9] application provides a special object, referred to as the black hole, to replace the rubbish bin seen in the Windows Operating system. The blackhole (Figure 13) provides a method of clutter control and undo. Clutter is an issue on tabletop applications whenever the user has the ability to create new objects.
Any objects moved into the black hole are removed from view. When regions, piles and attached text are put into the black hole they are deleted and the data stored in the related images are removed and metadata files updated. Deletion of pins and regions is replicated between maps with any deletion of a pin or region on one map will affect any copies of the map deleting the matching pin or region.

### 3.5 Elements of MyPlanner

MyPlanner introduces new elements to the user interface. The new elements are discussed in more detail in later chapters:

- Chapter 5: Interactive Mapping
- Chapter 6: Region Identification and Interaction
- Chapter 7: Image-based Tagging
- Chapter 8: Text-based Tagging
- Chapter 9: Metadata Storage and Handling

### 3.6 Summary

The user interface at the tabletop is an important aspect of this thesis. The elements available and presentation of events to the users are an integral part of the collaborative nature of tabletops. This chapter has introduced a high level description of the main elements of the Cruiser such as gestures. The discussion of the swipe gesture introduced as part of MyPlanner, is discussed as part of the available interactions.
Chapter 4 : Architecture Overview

The MyPlanner prototype branches from the original Sharepic photo sharing application [9] running on the Cruiser Tabletop. Some of the original components had to be modified to provide the additional functionality required to evaluate the thesis objectives using a prototyping methodology. Other aspects such as live images from remote computers have been excluded since they are not necessary for evaluations.

The modified components include the basic images, generalised resources, the environment and layouts for associated objects. New elements link into these and expand the current functionality to enable evaluation of image-based tagging and interactive mapping in the context of a framework for designing documents, maps, holidays, museums and zoos.

This chapter introduces the general architecture of MyPlanner, explaining it in relation to the Sharepic and what has been modified of the existing elements. The new elements are discussed in later chapters in detail.
4.1 Architecture

The basic architecture builds from an application called Sharepic [9] which is a photo sharing application developed for the Cruiser Tabletop. The diagram above represents a general outline of how each of the new elements relating to a modified version of the basic image class. Other components such as the basic resource and environment were modified to provide the basic functionality for the MyPlanner prototype. The elements, presented in yellow, represent the new elements required for the MyPlanner.

Modifications to the Environment (upper right of Figure 14) component involved adding functionality to handle text, image tags, regions and maps. When the application starts up, the environment is initialised and loads all special objects such as the Maps and Region Creation Tool.
4.2 Images

The Resource class required some modifications with aspects from Image being moved to it, allowing for the Environment to load the initial text objects. This also enabled the generic declaration of Maps and Regions to be stored in the Environment, minimising the amount of active resources required to be searched when making tags and replications.

4.2 Images

The original image structure provides the loading of basic image files into the system, representing them on the tabletop surface. The triangular widgets for rotation and resizing are inherited from the Resource class which is the parent.

Images in the MyPlanner framework are modified to include gestures for associating tags with images and storing metadata inside the image instance. Inside the image class modifications were made to the dwell gesture to allow for two alternative events to occur depending on the location of the image: if the image was over a map or region an image tag is performed, instead of the audio recording.

The image class is the base class for the text, region, region tool, map and pin objects. Each of these components uses the basic image processing to load the pictorial representations of these objects. The underlying functionality for gestures and events are provided by the child classes.

4.3 Blackhole

The blackhole (Figure 13) is a special garbage collection object. Figure 14 shows how the new elements relate to the original system components and so the figure does not include the blackhole. The blackhole was a separate structure that was not modified for use in MyPlanner. The inclusion of the blackhole object is to provide clutter control.
Text, regions or pins placed in the blackhole are removed from the system, clearing any metadata linked to that object instance. Any errors made during interaction can be placed in the blackhole for deletion as a form of undoing the last interaction. This is illustrated in Figure 15.

4.4 New Components

Each of the new components are linked into the original Cruiser Sharepic photo sharing application through the Resource, Image and Environment classes. These main classes provide the interface for the MyPlanner framework as a plugin. The new components are discussed in more detail in their related chapters:

- Chapter 5: Interactive Mapping
- Chapter 6: Region Identification and Interaction
- Chapter 7: Image-based Tagging
- Chapter 8: Text-based Tagging
- Chapter 9: Metadata Storage and Handling

4.5 Summary

This chapter introduced the relationship between new elements required to provide image-based tagging and interactive mapping functionality, and the Cruiser Sharepic photo sharing application. The key elements such as image representation and clutter control were described on an architectural. The combination of original and new components provides the architectural structure for the MyPlanner framework supporting image-based tagging and interactive mapping.
Chapter 5: INTERACTIVE MAPPING

5.1 Overview

This chapter discusses the mapping elements forming the MyMaps subsystem of MyPlanner. An interactive map is the representation of a map as an image on a tabletop interface. The architecture and interactions are some of the aspects covered, expanding on the generic outline introduced in Chapter 4. Maps represent a geographic form of organisation instead of the hierarchical method that is common in things such as file systems.

Maps are familiar aspects relating to planning tasks. MyPlanner is evaluated in the context of tagging images to maps. Image tags support the association of images to locations and regions. This provides the functionality for users to design documents, museums, building layouts and holidays using maps, by marking regions that have one or more associated tags and then an image, when associated with that region acquires these tags.

The research question is whether maps can serve as a ‘geographic’ method of organisation. In comparison, file systems use a hierarchical tree structure. Geographic organisation in MyPlanner uses regions to identify sub folders and the map to provide the root.
5.2 Architecture

MyMaps is the subsystem of MyPlanner with respect to interactive maps, represented by Figure 16. The map component is an extension from image. Maps are special environment objects loaded at runtime, creating the same number of map objects as users defined in the application configuration file. Image representations of the maps are loaded from a designated folder.

The association of images and regions to the map require a layout that maintains the relative location and rotation of that object. The May Layout component provides the functionality to maintain an image tag and region in their relative map position by scaling, rotating and translating the position of the child object (image tag or region). The layout components provided by the original Cruiser photo sharing application had limited support of layouts that maintain relative position, therefore requiring the Map Layout class extension.

![Figure 16: Architecture diagram showing the structure of the Map objects and the layout that the map object depends on for positioning image-based tags and regions](image)

5.3 Gestures

Interaction with the map is done by using the gestures provided by the Cruiser tabletop. Flipping a map toggles the interactive state of the object, hiding any regions and pushing the image tags behind so that they can only be viewed but not interacted with.
When a map is rotated, resized or moved any image tags and regions are rotated, scaled and translated to maintain proportional size and relative position to the centre of the map.

5.4 User Interface

Interactive maps are represented by an image loaded from file, specified in the application configuration file. The map is loaded at start up, creating a copy for each user. The maps can be modified by any of the users, allowing for sharing of a single map.

Maps have an interactive and non-interactive side. Figure 17 shows the different representations of a map in non-interactive and interactive state. Toggling interactive state is done by flipping the map. Interactive state is represented by the transparency level of the map image. Non-interactive maps are semi-transparent, hiding any regions (Figure 20). Interactive state is a non-transparent version of the image loaded from file, showing image-based tags and regions above the surface of the map.

![Figure 17: Left image shows map object flipped over (semi-transparent representation) and right image shows the interactive (non-transparent) map object](image)

Image-based tags made on a map are copied to all duplicate maps. If one user creates an image tag on a map the image tag is replicated to all duplicate maps using that user’s image tag representation. Figure 18 shows an example of five image tags on a map created by two different users, with each tag being represented by the user’s tag image i.e. red spherical image for user one and blue spherical image for user 2. Additional examples with two duplicate maps are shown in Figure 20, Figure 21, and
Figure 22, where the interactive states change the visibility of regions and interactivity of image tags.

Figure 18: Screen capture of a map object with 5 image tags represented by the spherical objects. Two separate user's tags are represented by the different colours (red for user 1 and blue for user 2)

Regions can be assigned to a map, as shown in Figure 19, with two overlapping regions being tagged to the map. These are created by dwelling on the region creation tool (Chapter 6) while it is positioned over the map. When a region is created it is replicated to the duplicate maps. Regions created maintain a relative position to the centre of the map it is connected.

Figure 19: Screen capture of a map resource with 2 overlapping regions represented by the semi-transparent green rectangles

The position of regions and image tags are scaled, rotated and transformed to reflect any changes made to the map. The size of the image tags and regions is scaled and rotated to reflect changes to the map.
5.5 Summary

This chapter discussed the details of interactive maps. The map component is a special object that is used to evaluate image-based tagging and collocated collaborative planning applications on tabletop interfaces. Maps represent a metaphor for geographic organisation of content.
In comparison to a hierarchical organisation, maps represent the root and regions would represent the substructure. Geographic organisation is more appropriate for planning tasks such as organising the layout of a building, than using a hierarchical structure.


Chapter 6 : Region Identification and Interaction

6.1 Overview

This chapter introduces the creation and interaction with regions on the maps. Regions in geographical organisation add depth similar to folders in a hierarchical file system. Regions can be overlapped creating sub regions the similar to sub folders can be created in a file system.

The main elements of MyPlanner discussed in this chapter are the region creation tool and region objects. The regions architecture and user interactive capabilities of the elements are described, such as individual gestures, the structure of the components with relation to the original Cruiser photo sharing application, and events that take place in the application. There are some objects that have restricted operations available and when attached to a map their location can only be modified by placing them in the blackhole to delete.

6.2 Architecture

The region creation tool (Region tool) is inspired by the Frame and Smart Frame (image capturing tools that take a capture of the screen, creating a new image) of the original Cruiser Sharepic photo sharing application. However, the implementation of the region tool inherits only from the image class instead of the Frame related class structure. The Frame class is a capturing mechanism that had sections of code that are dedicated to the creation of images using the underlying pixels to form an image. Therefore the region creation tool uses the frame as inspiration in structure and interaction, without inheriting from it.
6.3 Gestures

The Tag Component represented in the architecture diagram (Figure 23) is the metadata class, storing data such as map location, orientation, scale, and text. When changes are made to the data, the new data is persisted to file. This metadata storage class is discussed in detail in Chapter 9.

The region tool creates new instances of regions attaching them to the underlying map, copying the region to duplicate maps. Region objects are a simplified version of images that provides functionality for identifying a section of the map that has specific metadata associated that defines it from other sections of the map.

![Architecture diagram showing the region creation tool and region components and their relationships with each other and the parent Image class, which they inherit from](image)

6.3 Gestures

6.3.1 Region Creation Tool

The region creation tool inherits the characteristics of the Frame and Smart Frame. This includes the use of specific gestures that have been evaluated previously in Chapter 3. Gestures such as rotate, resize and movement are used to position the region tool to reflect the desired region that will be created. There is no event triggered by performing the flip gesture on the region tool.

When the region tool is in the desired orientation, scale and relative map position, the dwell gesture triggers the creation event. Dwelling on the centre of the tool will create a region relative to that dwell point.
6.3.2 Region Objects

Regions are objects that are attached to a map. These objects have limited gestures available. The gestures that trigger events are the flip and movement gestures. Flipping a region allows tagging, which adds the necessary text metadata to the region and any images tagged to the map within that area of the region. When a region is not flipped, text is hidden from view as shown in Figure 29, this prevents obscuring parts of the map.

Deletion of regions requires the use of the movement gesture. If a region is moved off the map and into the blackhole garbage collection object, the region and all associative metadata is deleted.

6.4 User Interface

The regions and region creation tool compose the elements required for region identification and interactions. Regions and region creation tool are represented by images loaded from a file specified in the application configuration file. The images used can vary depending on user preference, which affects the aspect ratio of the tool and regions to reflect the image representation.

The region creation tool (Figure 24) unlike the Frame and Smart Frame does not always appear on top of images and text objects in the application. The design decision is to allow for the entire interface to be used for placing objects without being obstructed by the tool.

Figure 24: Screen capture of the region creation tool
To create a region object, the region creation tool is positioned, resized and rotated to represent the desired area. The approximate centre region creation tool is dwelled on to create the object while the tool is above a map. The creation is possible when the dwell position is over the map and there is no image (excluding image-based tags and other regions) between the tool and the map. When a region is created and attached to a map, a copy is made and attached to each of the duplicate map objects. The process of region creation is graphically demonstrated in the sequence of images shown in Figure 25 and Figure 26, the duplicate maps containing the new region shown Figure 27.

Figure 25: Left image shows region creation tool, a map object with a single region and 4 image tags. Right image shows the region tool positioned, scaled and orientated to suit a new region to be created

Figure 26: Left image shows the state of the application after the region tool has been dwelled on with the new region created underneath. Right shows the new region on the map with the region creation tool moved to the side to be able to view the new region

Figure 27: Duplicate maps showing the new region replicated on both
Created regions can have metadata tagged to them by flipping them over and dragging text objects onto the back, as shown by Figure 28 and Figure 29. Tagged text will appear as coloured copies of the text object. When the region is not flipped over the text is hidden to minimise clutter and obscuring of the map, as shown in Figure 29.

![Figure 28: Left is state of application with two overlapping regions and the primary text object that is desired to tag to the region just created in Figure 26 and Figure 27. Right is state of application with region to tag text with flipped to enable text tagging](image)

![Figure 29: Left is state of application after primary text object moved across back of region, tagging the text to the region. Right image shows the hiding of text tags when region is not flipped](image)

Region objects cannot be resized, rotated or relocated by moving. If users desire a region location changed it requires recreation. This is a current limitation of the system to minimise errors when accidentally selecting and moving a region could cause adverse effects. These include moving the region unintendedly and allocating the region to a position that is not on the map.
6.5 Summary

This chapter provided explanation of the elements for identifying regions, and adding dimensional depth to the geographic organisation provided by the interactive maps. The region creation tool and region objects were the focus, outlining the architecture, user interface and the relationship between the two elements of the system.
Chapter 7: IMAGE-BASED TAGGING

7.1 Overview

This chapter outlines the key component of image-based tagging. An image tag is the reference point on a map. The image can be referred to at anytime by performing a gesture on the reference point or pin image. The reference points are called pin images, symbolic of the attachment of a pin on a notice board.

Images can be tagged to numerous locations with all pins referring to that single object. Pins are used to minimise clutter and confusion on the map, and the image representing a pin varies between users adding indicates which user placed it.

7.2 Architecture

The Pin component provides a representation of image-based tags on the maps, indicating the location of a tag. Pin components inherit the image processing from the image component, overwriting the gestures (Figure 30).

The gestures that result in an event when performed on a pin object are movement and dwell. The rotate and resize gestures are disabled for this component since the object is small in size, causing difficulties with interaction if these gestures were enabled. This is due to the widgets consuming the majority of the image once it reaches a particular size.

Rotation and resize are not needed since the pins have a generic representation. The representation by default is a coloured spherical object for each user. Images used for the pins are defined in the application configuration file.
Each pin that is created stores the data in an instance of the Tag Component that each image uses to store metadata. When a pin is made the location and map data is stored in the image and persisted to file.

The original image structure provided by the Cruiser photo sharing application has been modified to provide alternative events for the dwell gesture. The dwell gesture has two options of either making an image tag or providing audio recording. Replication of image tags across duplicate maps has also been added to the image functionality when a tag is made.

Figure 30: Architecture diagram showing the pin component which is the representation objects of image-based tags

### 7.3 Gestures

#### 7.3.1 Creating Tags

Creating a tag requires the dwell gesture performed while the image is situated above a map without any images or the region creation tool between the dwell point and the map. The dwell gesture for images is not only used for image-based tagging, it is used for recording audio and playing it back when the image does not satisfy one of the conditions required for completing the tag.

#### 7.3.2 Interacting with Tags

Pins are similar to regions with rotate, resize and movement. The pins cannot be resized and rotated. However, they do scale and rotate relative to the map they are tagged to.
Pins cannot be flipped over like regions and images. To view a tag, the pin is dwelled on to bring the image to the location of the tag. By removing rotation, resize and flipping the widgets that are used for these gestures are disabled enabling users to select any part of the object to move or dwell.

Pins can be moved to the blackhole for deletion. If the pin is moved and not placed in the blackhole it will be automatically moved back to the position of the tag.

### 7.4 User Interface

The process of image-based tagging is when an image object is dwelled on while situated above an instance of a map. Image objects (Figure 31) are representations of image files as objects on the application interface.

![Figure 31: Screen capture of image resources. Image on left shows all images face up and image on right has the bottom left image flipped](image)

There are a number of conditions required to be able to perform an image tag. This is due to the limited interactive gestures available for triggering events. The dwell gesture that was originally only used for audio recording in the Cruiser Sharepic application is used for the image tagging process as well. The conditions required for an image tag are:

- Centre of image within area of map;
- Dwell point within area of map;
- No images positioned between the dwell point and the map;
- The region creation tool is not positioned between the dwell point and the map.
The process for creating an image tag is to position an image above the map and dwell on a point where the tag is to be allocated relative to the map. To assist in the tagging process, images appear semi-transparent while over the map to avoid confusion and allow users to view where they want a tag. Figure 32 and Figure 33 show a sequence for creating an image tag on a map.

![Figure 32: Left shows an image to be tagged to the map and a copy of a map object. Right shows the image positioned over the map, semi-transparent to prevent obscuring the desired tag location.](image1)

![Figure 33: Left shows the state of the application after the image has been dwelled on to create the image tag relative to the dwell point (in this case near centre of image). Right shows image moved off the map, revealing the new image-based tag (pin) represented by the blue circular reference point.](image2)

The pin image representation can vary between users depending on the declarations made in the application configuration file. For example, one user might have a blue spherical image, as shown above, as their image tag representation and another user might have a red spherical image. The use of different tag representations between users providing identification of what each user has tagged when using replicated maps, as shown in Figure 34.
7.4 User Interface

Figure 34: Duplicate maps showing the replication of image-based tags (pins) on both maps for 6 image tags performed by 2 separate users. Each user’s tags represented by a difference coloured tag object (red for user 1 and blue for user 2)

Image tags are visible on the map when it is in both interactive and non-interactive state, unlike regions, as shown in Figure 35. The difference is when a map is in non-interactive state (flipped over and semi-transparent) the image tags cannot be selected or deleted. This minimises the load on user short-term memory for the locations of image tags, while minimising clutter and confusion when interacting with the maps.

Figure 35: Duplicate maps in alternate interactive states showing the replication of image-based tags (pins) on both maps for 6 image tags performed by 2 separate users, remaining visible for both states. Each user’s tags represented by a difference coloured tag object (red for user 1 and blue for user 2)

Similar to regions, image tags cannot be relocated using the move gesture. To change the location of an image tag the user must delete and the recreate in the new position. This is to eliminate possible accidental relocation of image tags. Deletion of a pin on one map is replicated, removing the copies from duplicate maps. To delete an image tag the pin is dragged into the blackhole, removing all metadata stored for that tag.
7.5 Summary

In this chapter the elements required to provide image-based tagging on the tabletop interface were introduced. Architecture, representation and interaction were outlined to provide a background in how the image tags are structured.

Image tags provide the allocation of visual documents to locations in a geographic organisational structure. Creating an image tag stores the location, orientation and regions metadata with the image. This is similar to the association of location, timestamp, name and path to a file in a hierarchical file system. Image tagging is a method of organising content combined with maps to provide a key element for a framework that allows collocated collaborative planning tasks.
Chapter 8: TEXT-BASED TAGGING

8.1 Overview

This chapter outlines the key component of the prototype of text-based tagging as a subsystem called MyTags. A text tag is the association of a piece of text to an object. Text tagging allows labelling of images and regions lowering the user memory load. Without text labels, users would have to remember what each region represented and any labels or notes they desire to associate with it.

Text objects provide pieces of metadata that can be attached to objects, giving meaning to organisation of the digital collection. The association of text to images and regions is similar to associating a name and notes with a file or folder in a file system. MyTags provides two alternative text objects for tagging. Each of these provide a different aspect of the tagging process; primary text objects are used for tagging text to an object and attached text objects symbolise a tag.

8.2 Architecture

Text tagging requires the implementation of a special object. This object represents a single line of text on a rectangular image. Text Component is an inherent version of Image Component that is permanently flipped. Figure 36 shows the architecture for the Text Component, which represents both primary and attached text objects.

There are two different representations of text objects; primary text objects and attached text objects. Primary text objects are those that are used for tagging the pieces of text to images and regions and represented by a semi-transparent white background image with the text on it. The primary text can be moved anywhere on the
interface and is not dependent on the location of another object. These primary images are loaded from a predefined file that contains a list of available text tags. The environment component is responsible for creating copies of primary text objects for each user.

Attached text objects represent a text tag. These have less functionality compared to the primary text objects and are created for each text tag made. Alternate representations of the two types of text objects are to prevent confusion when using the interface.

The Tag Component stores the text metadata when a text tag is performed. Each image and region contains an instance of this component. Changes made to the metadata contained within the instance of an object’s Tag Component are persisted to file. The Tag Component (Metadata Component) is discussed in detail in Chapter 9.

8.3 Gestures

8.3.1 Primary Text objects

The primary text objects are used to associate pieces of text with images. Gestures available for use on these objects include rotation, resize and movement. Movement of the primary text objects is used to trigger a text tag when moved over a flipped image or region.
Performing the rotation and resize gesture affects the size and orientation of text represented within the object. This is used to allow readability of text to suit users positioned around the interface. The gestures are performed using the triangular widgets.

**8.3.2 Attached Text objects**

Attached text objects maintain the ability to resize and rotate for readability. Rotation and movement are temporary gestures, that change the position and state of the object until an interaction is performed on a different object.

Moving an attached text object over the back of an image or region that it is not attached to will not tag text. The movement gesture is used for deletion, by moving the tag object into the blackhole removing any associated metadata.

**8.4 User Interface**

Primary text objects are used to assign pieces of metadata to images and regions. When the application starts up the primary text objects are loaded into piles which are duplicated for each user. Users can move, rotate and resize these text objects, organising them to suits their needs.

![Figure 37: Left image shows primary text objects piled in the orientation they appear in at system start up. Right image shows primary text objects used for tagging text to images](image)

Attached text objects are copies of the primary objects that are linked to an image or region symbolising a tag. These appear as orange backed copies that show when an
image or region is flipped. Numerous pieces of text can be tagged to a single image or region, each being represented by an attached text object.

![Diagram of attached text objects on the back of an image](image)

**Figure 38: Attached text objects on the back of an image representing text tags have been made**

Text tagging is performed by dragging the primary text object over the back of an image or region. The movement and swipe gestures can be used to perform a text tag. The movement gesture consists of moving the primary text object onto the area representing the back of the image or region. This will create the attached text object and link it to the object, storing any data about the tag in the appropriate metadata files.

Throwing a primary text object across the interface will cause any images or regions flipped over within the path of the object to be tagged with the text. This method can allow users to perform multiple text associations in a single gesture. The swipe gesture is a similar concept to the throwing, except it involves moving the text across the back of an image without releasing it to prevent it from using system momentum functionality continuing the objects movement. **Figure 39 shows the swipe gesture in a three stage sequence of moving a primary text object (white text object) across the back of an image to tag text.**
8.4 User Interface

Figure 39: Sequence of screenshots showing the swipe gesture. Left screenshot shows primary text object positioned above the image to have text tagged to it. Middle shows text being tagged while still moving the primary text object. Right screenshot shows the final state after the single gesture is complete with the primary text object located below the image and the text tagged.

The primary text objects (white text objects) can be placed in the blackhole, without deleting any text tags made with the same piece of text. This allows clutter control since primary text objects are duplicated for a copy of each piece of text per user creating a large number of objects on the interface. When the attached text objects are moved into the blackhole, any metadata associated with that tag is removed.

Text tags can also be performed on regions, as shown in Figure 40. Flipping a region allows primary text objects to be moved onto the back of the region to create a tag. When a region is not flipped over the attached text objects are hidden since regions are semi-transparent and would obscure the view of the underlying map.

Figure 40: Sequence of images for attaching text to a region. Left shows state of region before text tagged. Middle shows state of region after text tagged. Right shows region flipped, hiding tag object.

When a region is tagged with text any images that have been tagged to the map within the region will inherit the tag. Image-based tags within a region inherit metadata about
that region; this includes any text tags. Copying of text can be used for multiple text
tags, allowing for users to organise images by categorised.

8.5 Summary

This chapter introduced the elements of text-based tagging as a subsystem called
MyTags. Text tagging relates to collocated collaborative organisation of content by
attaching labels to better identify objects and their relationship to others. The two
different text object representations were outlined with the architecture, user interface
and the gestures that are incorporated to provide text-tagging functionality.
Chapter 9: Metadata Storage and Handling

9.1 Overview

This chapter outlines the storage and processing of metadata. Metadata is maintained in instances of a tag component holding the contents which are persisted to file. These contents are persisted for every update that occurs to ensure persistence of any data. The purpose of metadata is to maintain and store details of tags and the links between certain objects.

Metadata includes regions, maps, relative map position, rotation, scale and text. Storage of this data allows file persistence of the system state, which provides the ability to reload the state of the last tag update. When the system is loaded tags can be reloaded.

9.2 Architecture

The Tag Component (Figure 41) represents the instance that stores the metadata with each child object of the Image class. Tag Component provides the necessary methods for reading and writing the metadata stored within a file instance. For each piece of metadata there is a variety of search methods that allow checks for pieces of data such as text during image and text tagging processing.

The original Sharepic application required modification to support the storage of metadata in each image instance. When each image object is created metadata is loaded from file into the tag component instance. If the persistence flag is enabled the
image object will load any pins and regions that are stored. A listing of all the regions is stored in a separate file that acts as a directory. This allows more than one system to store the region metadata files in a single directory and load a separate region list.

Figure 41: Architecture diagram showing metadata storage component (Tag Component) and the relationship to the basic Image component

### 9.3 Metadata storage

Each image and region contains a tag component instance which temporarily stores metadata during runtime. When the metadata is modified it is persisted to file, allowing the state to be restored when the system restarts.

Tagging an image to a map creates an entry for the tag in the metadata. If the metadata file for the image does not exist one is created. Metadata files for images contain the list of maps the image is tagged to, any regions that tags are within, relative map locations of image tags and tagged text.

```plaintext
##MAP
<map file name>
<map file name>
...

##TEXT
<region|image file name>: <text>
<region|image file name>: <text>
...

##REGIONS
<region file name>
<region file name>
...

##LOCATION
<map>: <x coordinate>, <y coordinate>, <z coordinate>
<map>: <x coordinate>, <y coordinate>, <z coordinate>
<map>: <x coordinate>, <y coordinate>, <z coordinate>
```

Figure 42: Syntax description for metadata file structure relating to images
The text and location entries in the metadata files contain identifiers. These identifiers for text state whether the text is inherited from a region or directly associated with the image. The location identifiers refer to which map the image is tagged to. Each location entry represents an image-based tag.

```plaintext
##MAP
testMap.jpg
sitBuilding.jpg

##TEXT
SITRGImage1.jpg: IT Research Group
SITRGImage1.jpg: University of Sydney
Region1.txt: Engineering Area
Region2.txt: Cleveland St

##REGIONS
Region1.txt
Region2.txt

##LOCATION
testMap.jpg: -0.200862, 0.198491, -5.0011
testMap.jpg: 1.406652, 0.228491, -5.0013
sitBuilding.jpg: -0.344322, -0.3922, -5.0012
```

Figure 43: Example of a metadata file for an image called SITRGImage1.jpg

Metadata files for regions have additional content which is not relevant for images. These contents are orientation and scale, which image tags are all represented the same. When a region is created the orientation, relative map position and scale of the region creation tool is inherited from creation tool and stored in the tag component and a sequentially named region file. The region file name is placed in the region list file. Region metadata files have only one image tag entry, since the region represents the tag to a particular location. Therefore there is only a single entry for rotation, scale, and map sections since a region can only be attached to a single map location. Regions can have multiple text entries, except regions do not inherit text tags from other regions. This means that the identifier for all text entries is the name of the region.
9.4 Processing of Metadata

Metadata is internally processed at runtime by instances of images and regions. Contents of metadata files are read at system start up and updated each time a new tag or region is created, modified or deleted. Processing of metadata involves checking for the presence of text stored with images, finding images tagged within regions and locating regions.

When an image tag is made on a map the relative map position, map name and any regions it is located in are added to the metadata collection. For each of the pieces of text that are not tagged with the image, an entry is made with an identifier and the text being copied to the object’s metadata text collection.

Figure 44: Syntax description for metadata file structure relating to regions

The metadata files for images are stored in the same directory as the actual image files, and the region metadata files in a designated region folder.
9.5 Summary

This chapter introduced the use of metadata for file persistence, caching and tag management. The use of the tag component as a metadata cache improves the performance by minimising I/O operations, while any changes to the metadata are persisted to file. MyPlanner provides the functionality for storage and processing of the necessary metadata representing tags such as text, orientation, scale, and location data.
Image tags rely on relative position to the attached object to maintain their position during runtime. If the system were to fail, tags could be reloaded at the point of the last successful update.
This chapter outlines the design of the evaluation process used to evaluate MyPlanner in terms of its supports tabletop interaction for organisation of content using tagging and maps. Evaluations are based upon studies with a small sample of users and evaluate how they perform a set of tasks. In the scop of this thesis, it was infeasible to study a larger sample of users. Moreover at this stage of the work this qualitative small scale evaluation is appropriate to inform future work. Therefore evaluations attempt to gain high quality, rich data based on observations, the user’s attitudes and feedback. To gain the necessary understanding and feedback on the prototype, we designed a set tasks attempt to make the users focus on the aspects of the interface requiring evaluation. A questionnaire that was be issued to users after a usability study, was designed to gain an accurate understanding of the user’s attitude towards the experience and how the features allowed them to perform their tasks perceived (ease of use). Figure 46 shows a video capture representing the process of evaluation collaboratively between a pair of participants.

Figure 46: Video capture of a pair of participants using the MyPlanner system
The evaluation process focuses on the tagging, interactive maps, virtual museums on tabletops, and support for collaboration that the tabletop provides for the organisation of content. The evaluation tasks consist of tagging an image to a location on the map, tagging text, creating regions and collaboration while planning.

The evaluations of the prototype were in the context of the Nicholson Museum as the scenario for creating a museum that people can visit. Tagging functionality is compared against the “TeamTag” [29] implementation where pre-defined text tags are tapped to tag text to an image. Evaluations against Team Tag focus purely on the way textual tags are attached and stored on images in the system.

10.1 Objectives

The goals of the evaluation can be divided into four categories; performance, usability, collocated collaboration and short-term memory load. User evaluations, observations and the questionnaire are used to establish understanding relating to these four aspects.

10.1.1 User Performance

The user evaluations require assessing whether users from a variety of backgrounds are able to complete a set of tasks. Performance is measured by how users are able to complete the two separate design tasks using the MyPlanner system. These design tasks are; tagging specific pieces of metadata to images; designing a museum without limitations.

Evaluations aim to collect data relating to how users are able to perform text and image tagging in the context of the museum scenario. Performance also includes error rates when attempting to perform tagging.

10.1.2 Other Usability Aspects

Beyond just success in performing tasks, we wanted to consider other usability aspects: error rates, error recovery, and learnability of features. The number of
attempts a user takes to complete a tag successfully is a key factor in establishing the usability of tagging, reflecting the learnability of system elements. If a tagging task consistently takes more than once to perform, this indicates that the participant is having difficulties performing the function and the gesture was not learnt adequately.

10.1.3 Collocated Collaboration

Tabletops offer real promise as a highly collaborative interface. The evaluations are required to indicate the level of collaboration observed as users perform tasks using the user interface elements available. For collaborative planning applications to be suited to tabletops they are required to maintain user participation and communication while completing the tasks.

Evaluations of MyPlanner on a tabletop need to assess the level of collocated collaboration that occurs while using the framework on a tabletop. To gather data for this objective, observations focus on the participation of each user as they collaboratively discuss and complete the design task.

10.1.4 Short-term memory load

In the design of MyPlanner, there was a tension between various goals. In particular confusion is reduced by sacrificing short-term memory load with respect to tagging. Evaluations are used to indicate the effects of minimising confusion at the cost of memory load by observing participants interacting with objects and how many times pieces of data require checking.

Images and regions are required to be flipped to view and attach text, with the image itself being replaced by a semi-transparent white background. The back of an image could have been represented by a semi-transparent copy of the image texture. However this would have caused confusion since the semi-transparent texture is used when the image is located above a map object to minimise obscuring the map while performing image tags.
In particular the observations note whether users have to consistently check objects to remember meanings and tags. Memory load relates to what meaning is given to objects through text tags, image tags and regions.

### 10.2 Scenario

There are many possible applications for the functionality provided by MyPlanner such as planning a zoo, brochures, planning museums, holiday reflection and planning building layouts. The evaluations focus on virtual museum planning, where users are required to design a virtual museum using MyPlanner. Participants were presented with the following scenario:

*The Nicholson Museum at the University of Sydney is looking to reorganise their exhibits to accommodate the Egyptian exhibit that will be on display during 2007. There are currently a small amount of exhibits that need to be rearranged to allow for a new section.*

*They have requested two volunteers to reorganise the exhibits into appropriate regions such as Roman, Greek and Egyptian ancient history. The designers are required to give equal amount of display space for each of the areas since the previous exhibits remain popular with visitors.*

*The current museum curators would like the museum divided into three sections; Greek history, Roman History and Egyptian History. A list of keywords has been provided to be associated with the exhibits, assisting in the organisation and relationship between each of the exhibits.*

### 10.3 Methodology

Researchers exploring applications and the tabletop interface have used various methodologies in evaluating the systems and achieving their results. These methodologies include: observations [33, 35], ethnographical studies [15] and surveys/questionaries [28, 29].
10.3 Methodology

The evaluation of the MyPlanner project is based on a qualitative methodology for evaluations. Evaluation of the thesis involves; prototyping, user evaluations, and questionnaire. These aim at gathering data relating to performance, usability, collocated collaboration and user memory load from the participant’s experiences.

10.3.1 Prototyping

This thesis focuses on tagging and interactive mapping on tabletop interfaces. To evaluate the effectiveness of these on tabletops in a collaborative environment beyond the regular personal computer screen, a prototype is required. The prototype builds off the Cruiser Sharepic photo sharing application to provide the basic components.

Prototyping of an existing system for comparison is also required. The TeamTag [30] system provides Text tagging, which will be the basis of a general comparison for an alternative method of text-based tagging. TeamTag was chosen because it is the only published work on tagging at tabletops. As described in Chapter 2, the published report had a focus on location of controls and duplication of them. Nonetheless, it is suitable as a comparison approach because it is a related approach and it is the only other approach in the literature.

Elements of the prototype that are to be evaluated include:

- Interactive Mapping as described in Chapter 5
- Map Replication as described in Chapter 5
- Region identification and interaction as described in Chapter 6
- Image tagging as described in Chapter 7
- Text tagging as described in Chapter 8

A prototype of the TeamTag [30] system (TeamTag-control) is used that only replicates the method of text-based tagging presented in the original system. The key element of methods for tagging text was produced to be the foundation of comparisons between MyTags (text-based tagging subsystem identified in Chapter 8) and TeamTag-control in respect to methods of tagging text. The User Interface resembles MyPlanner, with the methods of interaction being modified to resemble
10.3 Methodology

TeamTag-control based on consultation with the researchers [30] who produced TeamTag.

10.3.2 User Evaluations

This thesis involves evaluation of how groups of users can make use of particular elements to assist in collaborative organisation of content. To gain understanding of how things such as replicated maps maintain collaboration between participants and validating MyPlanner, observations of users are required.

The user evaluations are qualitative using a small sample of users to establish a general understanding of how users from various backgrounds can interact with elements of the prototype.

User evaluations involved observations made as the participants completed the tasks required and from video recordings made of the process. These observations focus on gaining understanding of participant comfort, collaboration level, interaction tendencies and how MyPlanner is used as a social tool as well as content organisation application.

10.3.3 Questionnaire

The objective of the questionnaire following the user evaluations is to gain understanding of how participants felt towards the elements being evaluated. The questionnaire aims to capturing user experiences with the tagging functionality and interactive maps.

Research relating to tabletop interfaces uses questionnaires to establish participant preferences in addition to establishing participants’ general comfort. The questionnaire establishes participant preference between techniques of MyTags and TeamTag-control methods of text-tagging, identifying the method that is more user-friendly, suitable for organisation of content and convenient to users.
10.4 TeamTag-control

10.4.1 Overview

TeamTag [30] is an implementation of two alternative methods of tagging text to images using a tabletop interface. Details for TeamTag for personal communication for pairs, multiple text tagging using piles and evaluation task specifics were not in the published papers². The alternative methods are centralised controls and replicated controls (Figure 47) containing the metadata that users may want to tag to an image. The method of tagging text to an image is done by sequential "tapping" on objects i.e. the user taps on an image to select it. Subsequent taps within any of the metadata controls then associates the metadata with the selected image.

Assigning of metadata can also be achieved through the use of piles. Users can pile images and select a special pile icon to assign the metadata to all images in that pile. The gesture for assigning to piles is the same as single tagging, except selecting the pile icon instead of the image.

The prototype for TeamTag (TeamTag-control) involves evaluation of how people collaborate while tagging pieces of metadata to images using centralised versus replicate metadata controls. These alternative prototypes are shown in Figure 47. The controls containing metadata are predefined lists stored in files that allow users to label images. These pieces of metadata can be viewed at later stages through various gestures and reloaded on the system at start up. Each of the pieces of metadata is categorised and placed in either a circular widget (centralised controls) or column (replicated controls) that is designated to the related category.

10.4.2 Reason for Comparison

TeamTag-control provides a foundation for comparison with respect to the tagging of metadata to images. Tagging of text to objects provides context to image tags and organisation. Special text objects (Chapter 8) in MyPlanner allow users to better categorise metadata.

MyPlanner provides a variation of text-based tagging (MyTags), therefore requiring a point of comparison to establish an understanding of which method users prefer when tagging text. TeamTag-control provides the ideal comparison for interaction methods for association through simple gestures compared to MyPlanner, which attempts to use more user-friendly gestures for text-based tagging.

10.5 User Evaluations

10.5.1 Overview

MyPlanner consists of the MyTags (text-based tagging components) and MyMaps (interactive mapping components) subsystems. The tasks for user evaluations are based around the scenario of users being requested to design their own museum, organising digital images and text on a map. Tasks are performed using a double-cross over method, with half the pairs using MyPlanner followed by the TeamTag-control to perform the tasks and the other half meeting the systems in the opposite order. The cross over is used to compare the text-based tagging component (MyTags) with TeamTag-control to assess results collected.
The process requires that participants are paired. Each pair is provided with a tutorial in using the different systems, familiarising the participants with the interactive mechanisms available. The task that the pairs are required to complete is common for both systems to establish a basis for comparison. The comparison between the two systems is to establish a basic understanding of different methods of text-based tagging and user interaction preferences.

10.5.2 Objectives

The high level objectives of user evaluations focus on usability, performance, short-term memory load, user preferences, and collocated collaboration. These objectives are:

- Evaluate participant preference of text tagging techniques;
- Evaluate participant tendencies relating interactive mapping;
- Evaluate participant preferences of replicated maps versus shared maps;
- Evaluate participant ability of performing image-based tagging;
- Evaluate collaboration levels between participants;
- Evaluate level of natural interaction provided by MyPlanner;
- Evaluate effects of participant technological skill levels on interaction;
- Identify performance and usability issues;
- Identify collocated collaboration levels of participants interactively organising content;
- Identify issues relating to short-term memory load.

10.5.3 Procedure

The procedure for the user evaluations are as follows:

Introduction

The introduction involves explaining to participants the procedure and tasks that were required to be carried out during the evaluation process, requesting permission to record the evaluation session and any questions the participants have regarding the
evaluations occur at this stage (see for information sheet provided to participants Appendix A).

**System 1 Tutorial**

The first system as part of a double-cross over technique is not the same for all pairs of participants. For half of the evaluations MyPlanner was presented as the first system, and the other half TeamTag-control was the first system.

Tutorials of the systems involve presenting the prototype to the participants and running through each of the elements they are required to use. The participants were demonstrated each feature, explaining the purpose and functionality triggered by certain gestures. After the demonstration participants were given a short period of time to practice using elements and gestures (see Appendix A). The elements included:

- The Region Creation Tool
- The Blackhole
- The Map objects
- Text objects
- Image objects

The tutorials also involved introducing the participants to the available gestures that trigger events in the prototyped systems, as part of the explanation of user interface elements (see Appendix A). These gestures included:

- Movement
- Dwell
- Swipe
- Select
- Flip
- Rotate
- Resize
- Deletion
Participants were provided time to practice the elements and gestures to establish a basic understanding.

**Task 1 – Design Museum with 10 images**

The task required that participants tag 10 images to the map, associating a selection of metadata with each image. The layout of image-based tags and regions on the map was at the participants’ discretion. However, metadata for each image was outlined on personalised information sheets (see Appendix B and Appendix C for participant crib sheets containing metadata for each image).

The map requires three designated regions representing Greek, Roman and Egyptian exhibit collections. These regions can be positioned at the participants’ discretion with an example provided in Figure 48. The coloured rectangles in Figure 48 indicate three numbered regions on the map (represented by the black box around the three regions) where the overlapping indicates an image tagged to the map in this area would inherit tags associated with both regions. For example the region 1 (red) is assigned to Greek, region 2 (green) assigned to Roman, and region 3 (blue) assigned to Egyptian history.

![Figure 48: Example of a layout of 3 regions on a map object that participants should try create during the user evaluation process](image)

Each image has an average of 5 text required to be attached. These pieces of metadata can be tagged using any of the methods illustrated in the system tutorial.
**Task 2 – Free design task**

This task is a free form where participants are allowed to use the interface to organise content to design their own museum without design constraints. This task provides insight into the way participants use the system elements without constraint.

**Cross-over**

A tutorial is provided for the cross-over system i.e. if MyPlanner was the first system, this tutorial and the following tasks will be using TeamTag-control and vice versa. The tutorial for the second system follows the same steps as the first system, where participants receive an explanation of the elements and gestures.

After the tutorial is completed the participants repeat Task 1 and 2 using the new system. This is part of the double-crossover method to establish a basis for comparison minimising first system effect.

**Questionnaire**

Participants are requested to complete a short questionnaire that relate to their experiences with the alternative prototypes and preferences (Appendix A and Appendix D).

**10.6 Questionnaire**

**10.6.1 Overview**

User evaluations and observations provide some of the data for participant comfort with elements of prototypes and their understanding of what each element represents. The use of questionnaires is a method of reflection to document participant preferences and allow user feedback [31]

There are many aspects that can result in participants appearing to be uncomfortable during the evaluation processes. Some of these aspects may include, the environment that the evaluations take place is foreign with participants might have an obligation to feel they have to perform differently than they naturally would under evaluation
conditions. The questionnaire is present to improve the qualitative collection of data from participants.

The questionnaire also provides the opportunity for participant feedback and identification of certain methods they liked and disliked. This feedback gives valuable data in determining the viability of some of the objectives for this thesis such as how natural is image tagging and interactive mapping are when implemented on tabletop interfaces.

10.6.2 Objectives

The questionnaire is a part of the evaluations to assist in understanding the comfort and preference of participants clarifying observations made during the user evaluation process. The main objectives are:

- Identify participant text-tag method preferences;
- Identify preferences relating to replicated versus shared interactive maps;
- Understand the influence social background between participants had on the experiences completing the tasks;
- Identify any confusion relating to elements of the user interface and interaction methods that participants experienced after system cross-over;
- Identify participants comfort in performing particular gestures;
- Clarify performance and usability issues;
- Clarify collocated collaboration levels between users.

10.6.3 Questions

The questions that are provided for each participant to answer at the end of performing the tasks associated with user evaluations are:

1. What previous social contact have you had with the person you were paired with?
2. How comfortable were you in communicating with your pair?
3. Which alternative system did you prefer and why?
4. How convenient was using the regions to copy text to images as a form of assigning text in bulk?
5. How convenient was using piles to assign text to images in bulk?
6. Which method of assigning text in bulk did you prefer (regions or piles) and why?
7. How well did you remember the things taught in the tutorials for each system being presented?
8. During the free-form task for each system which one was easier to use once there was no limitations on design?
9. During the evaluations were there any difficulties in remembering gestures and how to perform particular tasks such as attaching text and if so what were they?
10. After switching between TeamTag-control and MyPlanner, what were the difficulties in relation to remembering the parts that were in TeamTag-control but not in MyPlanner and vice-versa when trying to perform tasks, if any?
11. How did you feel comfortable with how to attach images to the map (Creating the pins on the map indicating that an image was attached there)?
12. How did you like the idea of having piles of text each so that you can both attach text to images?
13. How did replicating interactions between maps help in achieving the tasks set?
14. How did you feel about having the text objects able to be moved around, and would you prefer the text objects be fixed in an arrangement directly in front of each participant?

Questions 1 and 2 are designed to document the social background and what influence this had on collaboration between the participants. They confirm observations made during the user evaluations and reviewing of video recorded relating to user comfort and communication level.

Questions 3 to 8 are structured to gain feedback on the user experience using the alternative systems creating a foundation for preferences. Preferences are obtained by focusing on the types of tasks (specific design task and free-form design task). These questions involve the comparison between the MyPlanner and TeamTag-control methods of text tagging.
The remaining questions are structured to identify usability feedback. The usability issues aimed at clarifying is the use of replicated versus shared maps, and the use of floating text objects versus text objects with a fixed location, which are two of the key uncertainties during the design process. Questions 9 to 11 focus on the learnability of MyPlanner. A system that is difficult to learn is more likely to be undesirable to be accepted by users.

10.7 Summary

This chapter introduced the design for user evaluations of the MyPlanner prototype to collect evidence about the four key objectives; performance, usability, collocated collaboration and short-term memory load.

The procedure involves a double-cross over between MyPlanner and TeamTag-control, where each pair of users is required to perform two separate tasks under each system. At the end, participants are asked to complete an open-ended questionnaire aims at identifying user preference, opinions and reinforcing observations.

Results from the evaluation process are discussed in Chapter 11. These relate directly to the objectives, tasks and procedure that were observed for user evaluations.
Chapter 11 RESULTS

This chapter reports the results of the major evaluation of MyPlanner outlined in Chapter 10. Results are structured into the four key areas of user performance, other usability, collocated collaboration, and user short-term memory load. The key aspects of image-based tagging, interactive mapping, and text-based tagging are elaborated on from results gained from user evaluations and questionnaires.

The evaluations were conducted with 12 participants (4 female, 8 male), with different backgrounds ranging from Office Clerks to PhD Computer Scientists. Six participants had a non-computing background.

The participants worked in pairs, with two of the pairs consisting of participants who had no prior contact. Each evaluation took approximately 2 hours to complete the tasks for both TeamTag-control and MyPlanner. The user evaluation process was video taped and reviewed for further observation data collection.

The questionnaire provided participants the chance to give their opinions allowing comparison and clarification of observations (see Appendix D). Questions were structured to encourage the participants’ to elaborate their opinions.

11.1 Results relating to Evaluation Goals

11.1.1 User Performance

The evaluations illustrated how participants are able to effectively complete a set of design tasks using the elements of tagging and interactive mapping under both
TeamTag-control and MyPlanner systems. Users were able to successfully tag pieces of text to images and regions with minimal errors.

<table>
<thead>
<tr>
<th>Participant Pair</th>
<th>TeamTag-control: Time (minutes)</th>
<th>MyPlanner: Time (minutes)</th>
<th>Notes Relating to Participant Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (A, B)</td>
<td>27</td>
<td>33</td>
<td>Minimal Computing background</td>
</tr>
<tr>
<td>2 (A, B)</td>
<td>14</td>
<td>19</td>
<td>No prior contact with each other</td>
</tr>
<tr>
<td>3 (A, B)</td>
<td>13</td>
<td>19</td>
<td>Experienced Tabletop Users</td>
</tr>
<tr>
<td>4 (A, B)</td>
<td>16</td>
<td>12</td>
<td>Minimal Computing background</td>
</tr>
<tr>
<td>5 (A, B)</td>
<td>45</td>
<td>14</td>
<td>No prior contact with each user affected early communication. Mixture of participant with minimal computing background and experienced computer user</td>
</tr>
<tr>
<td>6 (A, B)</td>
<td>41</td>
<td>21</td>
<td>Mixture of participant with minimal computing background and experienced computer user</td>
</tr>
</tbody>
</table>

Table 3: User evaluation times taken by evaluators to complete the task 1 using each of the systems. Pairs 1-3 used TeamTag-control first and 4-6 used MyPlanner first.

<table>
<thead>
<tr>
<th>Average of System being used First (minutes)</th>
<th>Average of System being used Second (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TeamTag-control</td>
<td>MyPlanner</td>
</tr>
<tr>
<td>34</td>
<td>23</td>
</tr>
<tr>
<td>TeamTag-control</td>
<td>MyPlanner</td>
</tr>
<tr>
<td>18</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 4: Average times taken per pair using each of the systems when each system was used first and second

The errors that occurred during the evaluations were a result of the participants not planning enough in advance. For example, a pair would tag an image to the Greek region and realise that the image also belongs in the Roman region, requiring the participants to remove and recreate the tag in the more suitable location. These errors were able to be corrected by the participants through recreation of tags meeting the new user requirements.
11.1 Results relating to Evaluation Goals

11.1.2 Other Usability

During the evaluation process there were several noticeable aspects of usability that were experienced by participants. The most prominent aspect was the use of a swipe gesture to tag text in the MyPlanner system that was not originally in the design. This gesture is where users move a text tag across the back of an image to perform a text tag.

Participants were able to easily learn gestures required for performing tagging and interaction on the MyPlanner system compared to TeamTag-control. The implementation of TeamTag-control was difficult for participants to learn the appropriate timing to signify a tap on the table using the pen stylus devices. Participants that had not used the tabletop interface before tended to dwell on the pens to begin with until they realised that this signifies different gestures and triggering events unintentionally. The MyPlanner system had minimal dependency on time based gestures such as dwell and tap. The dwell is the only time-based gesture used by MyPlanner and was performed comfortably and efficiently when performing image-based tagging.

![Figure 49: Video capture of evaluation after a user had just performed an image-based tag (new image tag/pin located towards centre of semi-transparent image)](image)

The general cause of errors when using MyPlanner was relating to the participants not planning sufficiently in advance. For example, on several occasions participants realised that an image tag should be placed in two of the regions instead of one, after they had already tagged it to the map, which participants removed, recreating the
image tag in the more suitable location. There was the occasional error made in text tagging, in the MyPlanner system deletion of mistaken tags was performed simple and efficiently correcting the mistakes that were made.

Primary text objects were duplicated with each participant having a pile of text. Users were observed to initially focus on a single pile, until certain tags were unable to be located on the interface. When unable to locate, the participants searched the other pile of text. The positioning of the piles was observed to be a determining factor with participants focusing on the pile closest to both users. The main issue observed with the duplicate piles of text was clutter that resulted from having a large number of text objects on the interface. Two pairs used the blackhole to control the clutter by removing a copy of the text objects once the majority of the tagging tasks were completed.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TeamTag-control</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>MyPlanner</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Participant (e.g. 1:A means pair 1 user A) preferences relating to TeamTag-control and MyPlanner with pairs 1-3 using the TeamTag-control system first and pairs 4-6 using MyPlanner first

In user evaluations participants provided preferences on their preference between MyPlanner and TeamTag-control. Preferences were based on usability and comfort experienced during the tasks. Ten of the twelve preferred the system that they used first with two participants preferring MyPlanner despite the order (Table 5). General comments from participants stated that MyPlanner was easier to use to tag text, and TeamTag-control was more convenient.

### 11.1.3 Collocated Collaboration

The evaluation process examined collocated collaboration when using a design framework such as MyPlanner. One of the main elements of MyPlanner that required understanding was how participants interact with maps when they are duplicated for each user. The evaluations were required to show whether users preferred using individual copies of map objects or a single shared map.
In the user evaluations, it was observed that participants preferred using a single map for planning. Participants collaboratively interacted with the map taking turns to designing the locations and performing the image tags. Two pairs made use of the second map as a thumbnail view of the shared map to assist in viewing where tags were located. The use of a single shared map over replicated maps seems partially due to a hardware limitation. The questionnaire identified that participants felt limited by the hardware only allowing single interactions and this influenced the majority of the participants to focus on using a single map. Another single pair determined that the replicated maps would have been more desirable to use if the data presented on each map contained more personalised information specific to each user.

Participants shared and passed objects between each other to complete text and image tagging tasks using the elements provided by the MyPlanner system. Each pair collaborated to constructively complete the design tasks, with several groups being social and relaxed discussing background they knew beyond the information that was presented on the information sheets provided. Pairs of varying background and social relationship completed the task collaboratively and successfully.

One other interesting observation was that the hardware limitation forced participants to be more aware of each other’s activities on the interface. Some aspects of the planning task were designed so that participants could perform them simultaneously without the hardware limitation. However, when users learnt that they could not interact simultaneously, user interaction was slowed to monitor what each participant was doing when attempting to complete a task.

### 11.1.4 Effects of Short-term memory load

User short-term memory load is a key aspect of MyPlanner that required evaluation. MyPlanner minimises confusion at the cost of memory load on the participants during use.

In the evaluations it was observed that users had to check what an image was or what text was tagged. However, the load was minimal since users could see what text was
associated to an image while tagging new text, compared to TeamTag-control where participants had to constantly check what has been attached since the image did not require to be flipped while tagging text. Participants commented that the reason they preferred to use MyPlanner was because it minimised the amount of checking of tags when attached compared to TeamTag-control.

Image-based tagging had memory load for the reference/pins on the maps. Evaluations presented only a few occasions when users had to check what pins referred to certain images. The majority of the time users remembered the image tags which could be directly due to the small sample of images used in the evaluations.

Users demonstrated that MyPlanner is slightly more efficient than TeamTag-control in attaching text with users having to check what text has been tagged. The small sample of text and images lowered the short-term memory load. However, users were still able to remember which tags referenced certain images.

### 11.2 Observation Results

#### 11.2.1 Interactive Mapping

The evaluation procedure presented participants with a scenario for creating a virtual museum using interactive maps. Participant interaction and use of these maps was observed and video taped during the evaluation process.

Some participants attempted to use the replicated maps as they were intended with each having their own copy of the map to minimise reaching and allow for correct orientation of objects. However, all participants concluded sharing a single map was easier with hardware limitations preventing simultaneous interaction.

Participants interacted successfully with the map for geographical organisation of content. During the evaluations some participants attempted to plan things on top of the map and when selecting the map all objects that were originally on top were pushed beneath requiring either the map to be moved or repeating the planning
11.2 Observation Results

process. Participants adjusted accordingly by either placing the images around the perimeter of the map or using the second map as a guide.

A second copy of the map was used occasionally as a guide for organisation and a preview of the map being shared. The second map was also used for visual feedback to whether an image based tag was performed since the pen, shadows and image obscured the location that pins were made on several occasions. Figure 50 shows video capture of two participants interacting using a copy of the map each.

![Video capture of users interacting with duplicate maps](image)

**Figure 50: Video capture of users interacting with duplicate maps**

The interactive mapping is an element of MyPlanner, made available in the TeamTag-control equivalent for providing a consistent environment for comparison. This is due to the comparison between the two systems is focused on methods of text-based tagging.

**11.2.2 Image-Based Tagging**

Image-based tagging is an element of MyPlanner that is made available in the TeamTag-control equivalent for evaluation purposes. Participants were required to tag 10 images to the maps using the tagging gestures available. Figure 51 shows a video capture of the map with a collection of image tags on it during a user evaluation.
Participants performed image-based tagging in the context of geographical organisation relatively naturally and efficiently. One of the main problems occurred when dwelling on an image while it is not over a region, pin or map: this triggered an intended audio recording event. Participants adjusted and responded appropriately when triggering the audio recording by mistake. The participants did not repeat the mistake during later stages of use.

![Figure 51: Video capture of image tags (reference points – dark dots) on the map within 2 of the 3 regions (represented by blue rectangles) circled in red](image)

Participants who were not familiar with the tabletop required the opportunity during the introduction to be familiarised with the gestures and timing for the dwell gesture. Demonstration and practice for image tagging were understood and replicated by participants during the tasks with relative ease.

### 11.2.3 Region Creation

Region creation was the main aspect of difficulty observed. Participants generally created a new region in a location that they did not desire due to the sensitivity of the dwell gesture used for creating a region using the Region Creation Tool. This error occurred on average three times during each user evaluation.

The accuracy of positioning the regions needs addressing, since it is dependent on the participants to dwell towards the centre of the tool to create a region resembling the positioning of the tool itself. Participants often worked around this by either
recreating until the region resembled what they desired, or they used the created region ignoring its positioning. The creation of region process improved after the first attempt resulting in a mistake with sensitivity of creation being the key problem in later attempts.

Region creation was more often performed by only one of the participants during the evaluation process, due to the small number of regions required. This was normally determined by who was closest to the region creation tool. In two of the pairs both participants took turns creating regions. However, in four different pairs the role of creating regions was reversed after system cross-over.

11.2.4 Text Association

Participants experienced difficulties with the TeamTag-control approach, with some tags requiring multiple attempts. This difficulty could have been a direct result from the implementation of TeamTag-control interactions from specifications provided. The Cruiser uses pen stylus devices, which are not used for the actual TeamTag system, which was implemented on a DiamondTouch interface [16, 19, 32, 33, 35, 36].

Timing was an issue that made the TeamTag-control method of text tagging difficult. Participants dwelled on objects instead of tapping, resulting in timing required to be learnt to know the stage that a select becomes a dwell gesture. Participants did adjust after some practice. The reasonable number of text tags required allowing for users to familiarise themselves with the interaction to formulate accurate opinions.

The method of assigning text in bulk in TeamTag-control relied on piling images. This resulted in tagging mistakes going unnoticed with some images unknowingly being placed in a pile, and participants had already completed the necessary tagging for those images.

A more conclusive comparison to TeamTag-control would require placing the map, region and blackhole components in the actual TeamTag-control system. This would
also require MyPlanner being implemented on a direct touch tabletop interface for a more comprehensive comparison.

The text tagging method MyPlanner was observed to be less frustrating to participants. Since objects were required to be flipped, this allowed for participants to view the current tags instead of having to check tags later. Text tagging was achieved by participants with fewer mistakes and improved feedback. Multiple text tags could be achieved by either individually moving text objects onto back of images, or using regions.

Participants used the single text tag, with only 2 pairs attempting to use the region and pile methods. The task was designed to evaluate how participants would assign text to an image using the functionality provided. Observations showed participants focused on one image at a time instead of detailed planning to identify common tags for groups of images.

11.3 Questionnaire Results

11.3.1 Interactive Mapping

The effects of replication of maps had the potential to improve collaboration by lowering the need to orientate the map to suit each participant. Questions posed related to the effectiveness of replicated maps for achieving the desired tasks.

Observations showed that participants focused more on the use of a shared map instead of the individual copies. Results from the questionnaire demonstrated that the hardware limitation of only one interaction at anytime was influencing the participants’ decisions to use only one of the maps. This was unexpected, even though 2 pairs and one other participant liked the multiple maps except for the hardware issue.

The hardware limitation was not expected to impact the participants’ use of replicated maps in the context of the organisation task. It was assumed the users would interact
taking turns and adjusts for those moments when simultaneous interaction is required. However, users initially attempted to interact simultaneously in organising the table surface, with hardware forcing them to take turns.

### 11.3.2 Image-Based Tagging

Participants had no difficulty performing image tags and the response was positive, commenting that the task of performing images was simple and they were comfortable in performing the task. The tutorial showed participants how to perform an image tag. Participants commented that once being shown how to perform a tag, the interaction was simple to perform.

### 11.3.3 Text Association

The evaluation process compared several methods of tagging. There were several questions devoted to identifying participant preference relating to methods of text tagging. The comparison was performed between a system resembling TeamTag-control and MyPlanner. Comparisons included:

- Regions (MyMaps) against Piles (TeamTag)
- Swipe/movement (MyTags) against tapping (TeamTag) shown in Figure 52

![Figure 52: Video capture of the swipe gesture. Left shows a collection of text objects that the participant wanted to associate with the flipped image. Right shows the text objects in a pile that has been formed by moving each text across the back of image to do the tag](image)

Participants were asked which method of multiple tagging they preferred and their opinions on each of the methods (Table 1). Regions could be used to performing
multiple text tags, by placing images within a region that had text tags. Four of twelve of the participants found the regions convenient for the task. The other participants felt that it was more convenient to assign text individually to images and how the regions copied text was complex.

Piles are part of the TeamTag-control system that is used to assign text to multiple images. Seven of twelve participants found the use of piles simpler and easy to understand. However, most participants preferred assigning text individually because they still had to check if the tag was made. 4 out of 12 participants commented that piles were not suited to the task since it involved having to plan ahead. One other participant (pair 5, participant A) did not recall using piles to assign text to multiple images.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Piles</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regions</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Participant (e.g. 1:A means pair 1 user A) preferences relating to the TeamTag-control piles and MyPlanner region methods of associating text to multiple images (X signifies users that liked those methods)

The participant preferences for which method of assigning multiple text tags was even with one pair liking both but preferring piles. One participant commented that the presentation of text tags on the information sheet might have suited piles more if the images were sorted and grouped by tags instead of images.

Single text tagging preference was presented through observations and participant opinions on alternative methods for text-based tagging and the fixture of text. The general outcome was that participants found that the system they used second was preferred, with only 2 people from different pairs preferring MyPlanner over TeamTag-control even when MyPlanner was used first.

The general comment when comparing to the 2 methods of single attachment was that swipe was easier and less frustrating than the TeamTag-control method of tapping. However, with the swipe method this resulted in text being sometimes misplaced. Participants (9 out of 12) preferred text being in a fixed location. TeamTag-control
allowed participants to maintain text in a fixed location and this was an influencing factor in system preferences.

### 11.4 Conclusions

The combination of results gathered from qualitative user evaluations and questionnaires provides the basis for several conclusions. Tagging images to the map was the most conclusive piece of data collected with both observations and questionnaires showing that participants were comfortable and understood the process for performing image tags on the tabletop. Participants collaborated using the data provided to decide on a desirable place for the image on the map.

Image tagging was performed in five of six pairs using a shared map. The other group attempted to use the duplicate maps to complete the design task. However, eventually, they too focused on using a single shared map. Observations relating to the use of maps presumed sharing a single map was a result of limited display space. The questionnaire indicates that the participants were using a shared map because of the hardware limit of only one interaction at a time. Five of twelve participants indicated that the duplicate maps had potential use, one participant indicating that more personalised information for each user would improve the usefulness of duplicated maps. Hardware limitations were observed to force participant collaboration. Participants had to be aware of what each user was doing when interacting with the table.

Comparisons between methods for text tagging resulted in participants slightly favouring the swipe method over tapping. However the multiple text tagging methods were evenly split depending on how participants approached the task. Future systems would benefit by providing both the piles and region alternatives for assigning text, with participants commenting that the usage depended on the context of the task.

Overall the MyPlanner system was the preferred approach with 8 out of 12 participants, first used system effect resulting in ten out of twelve participants preferring the system used first (Table 5). The combination of single and multiple text
tagging methods present that MyPlanner was liked for the design task more than TeamTag-control. Further comparisons with TeamTag-control is required since the results are based on a system that was developed using specifications provided by the researchers of TeamTag-control using available technology and implementations. Results indicate participants rarely liked both of the text tagging methods which showed that the potential to provide participants with an application allowing user choice of methods could provide an improved solution.

11.5 Summary

This chapter presented results gathered from conducting qualitative evaluations, involving user evaluations and a questionnaire. Results demonstrated that MyPlanner is an appropriate system using image-based and text-based tagging for collocated collaborative design tasks. Participants were able to successfully learn the elements of the system and complete the design tasks using image and text tagging, with the only errors being a result of the participants not planning sufficiently in advanced prior to making a tag.

Participants collaborated using data provided efficiently designing a virtual museum using both of the systems. The task was completed faster by the system used second with less planning involved and a better understanding by the participants of the task to be completed. Generally MyPlanner performed better and was learnt faster since it required minimal use of time-based gestures such as dwell and tapping, which users tended to have difficulties learning. However, region creation was a problem with sensitivity for creation.

Significantly, user evaluations showed use of a swipe gesture that was not originally designed. Participants would attach text in the MyPlanner system by moving text across an image. When the text tags were done using swipe instead of moving text onto the back, less text objects were lost during the tagging process. Eleven of the users made use of the swipe gesture to complete the text tagging process. This swipe gesture has been unreported in any literature and is a valuable contribution extracted from the evaluations.
Participant preferences favoured the system used second, with the exception of two people who preferred MyPlanner despite it being used first (Table 5). This was an indication that users felt that the second was simpler because it took less time, understanding the task and interactions better. In the context of geographical organisation of content, the results illustrate image-based tagging was a comfortable method of assigning images to a map.

Overall the performance and usability of the MyPlanner system was shown to be effective in allowing users to complete the required design tasks. The concern of short-term memory load was evaluated and some issues were experienced in remembering certain tags, but with the small sample of text and images, it was not a major problem for users.

MyPlanner has been shown to support collocated collaborative planning activities in the context of image based tagging and interactive mapping demonstrated by a high performance by users of a variety of backgrounds, observations, logs and data collected from questionnaires.
The field of research relating to tabletop interfaces is still developing with new elements and applications. MyPlanner explores image-based tagging with relation to interactive mapping and an alternate method for text tagging on tabletop interfaces, furthering research in using tabletops of collocated collaborative planning activities. However, there are still many areas that can be enhanced to further tagging and map interaction for organisation of content on tabletops. Some of the key areas of future work are:

- **Natural interaction:** further the investigation of methods for inputting text into the system using natural affordances such as handwriting recognition. Research into tagging currently depends on text being defined and stored externally to the main system.

- **Personalisation:** investigate the representation of personalised data to users during the planning process. MyPlanner presents a personalised representation of image-tags and this can be enhanced by allowing users to define what can be viewed by users on maps. There is the possibility of recommendations to users based on tags based on a user model to define what areas interest the user.

- **Interface:** further the application of tagging to include 3-Dimensional mapping. Currently the maps are represented in 2-Dimensional form in MyPlanner and tagging in 3-Dimensional space would enable expansive use in planning in the context of structural engineering.

- **Application:** investigate the use of MyPlanner for the organisation of file systems using geographical techniques. Evaluations have been performed in the context of organisation and planning of structures such as museums. Extension to include the organisation of files using image-based tagging.
would link into the research conducted on interacting with file systems on tabletops [13].

Other enhancements to the MyPlanner system would include internal aspects such as efficiency. The concept of efficiency would involve replacing the flat file storage of metadata with a database system minimising I/O operations required when performing updates. This chapter outlines only some of the key aspects that this thesis work points to as the most promising and important for future research and evaluation.
Chapter 13: CONCLUSION

The primary goal of this thesis was to develop a prototype framework that used image-based tagging and interactive mapping to support geographical organisation of digital content on a tabletop interface. There were several sub-goal, including providing alternative methods of presenting virtual museums to collocated groups of users, and providing users with a framework for collaborative planning tasks.

Image-based tagging and interactive mapping were evaluated using the MyPlanner prototype framework that is a plugin for the Cruiser Tabletop (University of Sydney) photo sharing application. The research issues addressed by MyPlanner include:

- **Improving understanding in tabletop personalisation through the enhancement of tagging in a collaborative environment.** Tagging is a method of organising digital content, identifying a level of user ownership over a piece of work. Tagging elements provided by MyPlanner were discussed in Chapter 7 and Chapter 8. These elements were assessed through user evaluations, questionnaire and observations from video reviews (Chapter 10 and Chapter 11).

- **Explore the application of a map metaphor for organisation of digital content in tabletop interfaces.** MyPlanner evaluates image-based tagging through interactive mapping outlined in Chapter 5. User evaluations were conducted to assess the effect on collaborative interaction with maps in organising digital content. The results from evaluations are outlined in Chapter 11.

- **Replication of actions between objects on tabletop interfaces.** MyPlanner explored collaborative use of replicated maps assisting organisation of digital content (Chapter 5). Sharing a map proved popular with a hardware limitation preventing simultaneous interaction. Results from the use of replicated interactive mapping are discussed in Chapter 11.
• Exploring an alternative method of tagging text to objects. MyPlanner explored the use of movement of digital objects compared to the tapping gesture used by TeamTag-control outlined in Chapter 8. Evaluations were performed through a double-cross over comparison between MyPlanner and TeamTag-control, (reported in Chapter 11) showed MyPlanner was the more preferred alternative for text-tagging with improved user performance.

• Exploring gestures for performing tagging of objects to locations and regions on other objects. MyPlanner was the vehicle for evaluating the usability of maps and tagging on tabletop interfaces built upon the Cruiser tabletop gestures for interaction. Usability as assessed in user evaluations and questionnaire feedback (Chapter 11) was that all users were comfortable in tagging images to the map after being demonstrated in the evaluation tutorial.

• Discovery of a swipe gesture. The implementation of a swipe gesture has not been previously reported in literature. This gesture resulted from participant use during evaluations and is outlined in Chapter 3 and Chapter 11.

The MyPlanner prototype framework was implemented successfully with evaluations showing users were able to perform collocated collaborative learning tasks with image-based tagging and geographical organisation. The evaluation process allowed understanding of the collaboration and usability by users of a variety of computational skill levels. Participants used MyPlanner collaboratively and comfortably when asked to design a virtual museum using the framework. MyPlanner method of tagging text to objects was compared against the TeamTag-control approach and it was found that using the pen stylus devices MyPlanner method is more user-friendly and easier for participants to learn the interactions required.

Evaluations identified that participants preferred a single shared map over individual copies, since personalised information is minimal and hardware limitations made simultaneous use of duplicate maps impractical. It was found that the hardware influenced the collaboration level between users, with each user requiring having awareness of the intentions and actions of others.
This thesis provided an initial investigation into the concept of collocated collaborative image-based tagging on tabletops. Evaluations and implementation of a prototype allowed for initial understanding in the use of image-based tagging for geographical organisation of digital artefacts using tabletop interfaces.
REFERENCES


34. Schmitz, P., Inducing Ontology from Flickr Tags. in, (2006).


36. Smeaton, A., Foley, C., Gurrin, C. and Lee, H., Collaborative Searching for Video Using the Fischlar System and a DimondTouch Table. in *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, (2006), IEEE.


42. Wu, M., Shen, C., Ryall, K., Forlines, C. and Balakrishnan, R., Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. in First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, (2006), IEEE.
Appendix A : USER EVALUATION INFORMATION SHEET

Disclaimer

These evaluations have the option of being video recorded. If you are satisfied with being video taped during the evaluation please sign below

I ____________________ agree to being able to be video taped while completing the evaluation.

Signature: __________________ Date: __________

Introduction

MyPlanner and TeamTag are two examples of systems that offer the association of objects. These systems are designed for providing collaborative applications, with the expansion towards planning applications, such as planning a museum or zoo. The tasks are based around the scenario of being requested to design a new virtual museum by designating regions of the museum to particular aspects of history, and associating particular exhibits to these regions.

The process requires that each user is paired with another based on the level of computer knowledge of the user, and social background with the person they are paired with. These pairs are established to understand how relationships between users influence design decisions and participant comfort with the person they are collaborating with influences their comfort in using the system.

Each pair is provided with a short tutorial in using the different systems, familiarising the participants with the interactive mechanisms available to complete the designated tasks. The task that the pairs are required to complete is common for both systems to establish a basis for comparison.
Scenario

Nicholson Museum at the University of Sydney is looking to reorganise their exhibits to accommodate the Egyptian exhibit that will be on display during 2007. There are currently a small amount of exhibits that need to be rearranged to allow for a new section designated to Egypt, in addition to the previous Greek and Roman exhibits that are present.

They have requested two volunteers to reorganise the exhibits into appropriate regions such as Roman, Greek and Egyptian ancient history. The designers are required to give equal amount of display space for each of the areas of history since the previous exhibits remain fairly popular with visitors.

The current museum curators would like the museum divided into three sections; Greek history, Roman History and Egyptian History. A list of keywords has been put together to be associated with the exhibits, assisting in the organisation and relationship between each of the exhibits.

The Tasks

Task 1

The task requires the tagging of 10 images to the map, associating a selection of text with each image. The layout of image-based tags and regions on the map was at the participants’ discretion. These pieces of text are provided on personalised information sheets which require collaboration amongst participants to identify the correct region of the museum they are to be assigned to.

The map requires three designated regions representing Greek, Roman and Egyptian exhibit collections. These regions can be positioned at the participants’ discretion with an example provided in picture below. The coloured rectangles in image below indicate three numbered regions on the map (represented by the black box around the three regions) where the overlapping indicates an image tagged to the map in this area would inherit tags associated with both regions. For example the region 1 (red) is assigned to Greek, region 2 (green) assigned to Roman, and region 3 (blue) assigned to Egyptian history.
Each image has an average of 4 text required to be attached. These pieces of text can be attached using any of the methods illustrated in the system tutorial provided prior to starting the task. These methods, depending on the system that the task is being performed under are:

- **TeamTag System**:
  - Assigning text individually to every image one by one
  - Assigning text required by multiple images by making piles and assigning to the pile

- **MyPlanner System**
  - Assigning text individually by flipping images over and assigning to the image
  - Assigning text through use of regions which replicates text assigned to regions to any images tagged to the map in a location that exists within the area of the region.

The pieces of text are provided on the information sheets given to each participant. Each sheet do not contain all the text required, and therefore collaboration will be required to determine where things should go based on the participants “personal” knowledge provided by the information sheets.

**Task 2**

This task is a free form where participants are allowed to use the table to design their own free-form museum using the images and text provided. This task is limited to 10 minutes to observe how the participants can use the system to plan and create museums without specified requirements to the organisation.
Task 3

Complete the Questionnaire provided for participant feedback and preferences to support what has been observed during the previous tasks.

Questionnaire

1. What previous social contact have you had with the person you were paired with?
2. How comfortable were you in communicating with your pair?
3. Which alternative system did you prefer and why?
4. How convenient was using the regions to copy text to images as a form of assigning text in bulk?
5. How convenient was using piles to assign text to images in bulk?
6. Which method of assigning text in bulk did you prefer (regions or piles) and why?
7. How well did you remember the things taught in the tutorials for each system being presented?
8. During the free-form task for each system which one was easier to use once there was no limitations on design?
9. During the evaluations were there any difficulties in remembering gestures and how to perform particular tasks such as attaching text and if so what were they?
10. After switching between TeamTag-control and MyPlanner, what were the difficulties in relation to remembering the parts that were in TeamTag-control but not in MyPlanner and vice-versa when trying to perform tasks, if any?
11. How did you feel comfortable with how to attach images to the map (Creating the pins on the map indicating that an image was attached there)?
12. How did you like the idea of having piles of text each so that you can both attach text to images?
13. How did replicating interactions between maps help in achieving the tasks set?
14. How did you feel about having the text objects able to be moved around, and would you prefer the text objects be fixed in an arrangement directly in front of each participant?
Appendix B: Participant 1 Crib Sheet

The images used and contained in this crib are a collection from Wikipedia [8] and photos taken from the Nicholson Museum.

<table>
<thead>
<tr>
<th>Image/Exhibit</th>
<th>Description</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statue of Neptune the Roman god of the Sea and has other names under other cultural mythologies.</td>
<td>God Mythology Roman</td>
</tr>
<tr>
<td></td>
<td>Pictorial representation of the Greek god Zeus.</td>
<td>God Greek Picture</td>
</tr>
<tr>
<td></td>
<td>Famous portrait painted by the famous European inventor/artist Leonardo Di Vinci the Mona Lisa is a representation of a real person that has been renowned for not smiling.</td>
<td>Person Artwork Portrait</td>
</tr>
<tr>
<td>Image</td>
<td>Description</td>
<td>Role</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>Sculpture of Rameses II who was one of the ancient world's greatest Egyptian figures</td>
<td>Sculpture</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>Claudius Caesar, Emperor of Rome and was one of the main people responsible for the expansion and establishment of the Roman Empire</td>
<td>Emperor</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>Julius Caesar, one of the most well-known of Roman's Emperors, with his assassination resulting in what was to be the end of the Roman Empire</td>
<td>Person</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td>Ancient artefact of the Egyptian god Ra</td>
<td>Person</td>
</tr>
<tr>
<td>Image</td>
<td>Description</td>
<td>Person</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td><img src="image1" alt="Sculpture of the God Nike." /></td>
<td>Sculpture of the God Nike.</td>
<td></td>
</tr>
<tr>
<td><img src="image2" alt="Sculpture of the Greek God Apollo, this god is not specifically Greek and has been known to be present in other mythologies" /></td>
<td>Sculpture of the Greek God Apollo, this god is not specifically Greek and has been known to be present in other mythologies</td>
<td></td>
</tr>
<tr>
<td><img src="image3" alt="Sculpture of Alexander the Great, leader that instigated the conquest of the Persian empire." /></td>
<td>Sculpture of Alexander the Great, leader that instigated the conquest of the Persian empire.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: PARTICIPANT 2 CRIB SHEET

The images used and contained in this crib are a collection from Wikipedia [8] and photos taken from the Nicholson Museum.

<table>
<thead>
<tr>
<th>Image/Exhibit</th>
<th>Description</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statue of Poseidon the Greek god of the Sea and has other names under other cultural mythologies.</td>
<td>God Mythology Greek Person</td>
</tr>
<tr>
<td></td>
<td>Pictorial representation of Zeus.</td>
<td>Person Greek Mythology</td>
</tr>
<tr>
<td></td>
<td>Portrait of the Mona Lisa from Italian based inventor/artist Leonardo Di Vinci.</td>
<td>Roman Portrait Picture</td>
</tr>
<tr>
<td>Appendix C : Participant 2 Crib Sheet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Sculpture of the great Pharaoh Rameses II who ruled Egypt during the 1300’s BC.** | Sculpture  
  Egyptian  
  Pharaoh |
| **Claudius Caesar, one of the most prominent Emperors of the ancient world.** | Emperor  
  Person  
  Sculpture |
| **Julius Caesar, one of the most well known Romans, with his assassination leading to the suicide of Queen Cleopatra of Egypt** | Person  
  Assassinated  
  Roman  
  Egypt |
| **Ancient artefact of the Egyptian god Ra who was a person with a creatures head similar to most of the Egyptian Gods** | Egyptian  
  God  
  Creature  
  Person |
<table>
<thead>
<tr>
<th>Sculpture</th>
<th>Description</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Sculpture of the Greek God Nike" /></td>
<td>Sculpture of the Greek God Nike.</td>
<td>Greek Sculpture</td>
</tr>
<tr>
<td><img src="image" alt="Sculpture of the Roman God Apollo" /></td>
<td>Sculpture of the Roman God Apollo, this god is not specifically Roman and has been known to be present in other mythologies.</td>
<td>Roman God Sculpture</td>
</tr>
<tr>
<td><img src="image" alt="Sculpture of Alexander the Great" /></td>
<td>Sculpture of Alexander the Great, Emperor of Ancient Greece, and another of the ancient world that was assassinated.</td>
<td>Emperor Greek Assassinated</td>
</tr>
</tbody>
</table>
The answers listed for each of the questions is a small sample of the results collected from each of the participants of user evaluations.

1. What previous social contact have you had with the person you were paired with?

   “Colleague”
   “University course friend”
   “No prior contact”

2. How comfortable were you in communicating with your pair?

   “Comfortable”

3. Which alternative system did you prefer and why?

   “TeamTag, because tapping to tag text was more convenient.”
   “TeamTag, good to be able to keep text objects in the same location”
   “MyPlanner, region replication made more sense for planning a museum”
   “MyPlanner, easier to tag images with text”
   “MyPlanner, easier to use, understand and quicker to associate words”
4. How convenient was using the regions to copy text to images as a form of assigning text in bulk?

“Not convenient, hard to find common tags”
“Was easier assigning images individually rather than by regions”
“Very convenient for task, although evaluation only required on tag per region, so wasn’t used for much work”
“Complex to understand”

5. How convenient was using piles to assign text to images in bulk?

“Required more planning ahead than regions”
“Was not very convenient”
“Very convenient, simple to understand and easy to do”
“Wasn’t very useful, if the crib sheet was grouped by tag rather than image, it may have made more sense for piles”
“Simpler for assigning images by bulk”
“Didn’t like and needed to double check after use”

6. Which method of assigning text in bulk did you prefer (regions or piles) and why?

“Piles, no need to flip images”
“Piles, quicker and easier to do”
“Regions, piles not convenient but not a huge need to assign tags by bulk in task”
“Regions, doesn’t explicitly require gathering everything into a pile to assign text”
“Piles, does not need to be tagged to map in a region”
“Piles easier, but time consuming”
“Regions, didn’t need to be as careful when assigning text”
7. How well did you remember the things taught in the tutorials for each system being presented?

“Clear and remembered well”
“Remembered first system better (MyPlanner)”
“Wasn’t clear until partner showed how to do things”
“Had habit of trying to use MyPlanner method in TeamTag (MyPlanner used first)”
“Hands on experience was better than just being shown”

8. During the free-form task for each system which one was easier to use once there was no limitations on design?

“TeamTag was easier to use with tap text and tap image”
“MyPlanner essentially easier than TeamTag”
“MyPlanner overall was easier”
“Regions, as we could see the overlap in the regions we assigned. This helped us easily sort the images into spaces assigned”

9. During the evaluations were there any difficulties in remembering gestures and how to perform particular tasks such as attaching text and if so what were they?

“Tried moving regions after creation, but everything else was straightforward”
“Remembering parts of TeamTag with touch compared to move (TeamTag was used second)”
“Remembering timing for piles, made system appear unresponsive”
10. After switching between TeamTag-control and MyPlanner, what were the difficulties in relation to remembering the parts that were in TeamTag-control but not in MyPlanner and vice-versa when trying to perform tasks, if any?

“Was trying to do drag method in TeamTag (MyPlanner used first)”
“Kept thinking that flipping was required (MyPlanner used first)”
“Hard to remember differences between methods”
“No difficulty”

11. How did you feel comfortable with how to attach images to the map (Creating the pins on the map indicating that an image was attached there)?

“Very comfortable knowing that images were already attached”
“Very comfortable and easy once shown in Tutorial”
“Simple, but dilemma when image belonged to more than one region”
“Size of pins made difficult to remove”

12. How did you like the idea of having piles of text each so that you can both attach text to images?

“Didn’t like having text started in piles – would have preferred a single pile”
“Useful and saved time”
“Worked well when used”
“Hardware limitation prevented use of both piles at once”
“Piles needed sorting”

13. How did replicating interactions between maps help in achieving the tasks set?

“Focused only on one map”
“When used had to constantly rescale to use both maps at once”
“No help, since only one person could use pen at once”
“Easier to collaborate using the same map”
14. How did you feel about having the text objects able to be moved around, and would you prefer the text objects be fixed in an arrangement directly in front of each participant?

“Liked freedom to move and organise text”
“Preferred fixed, less time consuming to find text objects”
“Preferred fixed, sometimes hard to find text on table”
“Liked not fixed, as it is easier to pass text objects to other users and movement is more convenient”