Crossmodal Displays: Coordinated Crossmodal Cues for Information Provision in Public Spaces

Thesis by

Han Cao

曹 瀚

Submitted for the Degree of Doctor of Philosophy

School of Computing Science, Newcastle University, Newcastle upon Tyne, UK

May, 2012

Abstract

This thesis explores the design of Crossmodal Display, a new kind of display-based interface that aims to help prevent information overload and support information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns. By exploiting the human multimodal perception and utilizing the synergy of both existing public displays and personal displays, crossmodal displays avoid numerous drawbacks associated with previous approaches, including a reliance on tracking technologies, weak protection for user's privacy, small user capacity and high cognitive load demands.

The review of the human multimodal perception in this thesis, especially multimodal integration and crossmodal interaction, has many design implications for the design of crossmodal displays and constitutes the foundation for our proposed conceptual model. Two types of crossmodal display prototype applications are developed: CROSSFLOW for indoor navigation and CROSSBOARD for information retrieval on high-density information display; both of these utilize coordinated crossmodal cues to guide multiple simultaneous users' attention to publicly visible information relevant to each user timely.

Most of the results of single-user and multi-user lab studies on the prototype systems we developed in this research demonstrate the effectiveness and efficiency of crossmodal displays and validate several significant advantages over the previous solutions. However, the results also reveal that more detailed usability and user experience of crossmodal displays as well as the human perception of crossmodal cues should be investigated and improved. This thesis is the first exploration into the design of crossmodal displays. A set of design suggestions and a lifecycle model of crossmodal display development have been produced, and can be used by designers or other researchers who wish to develop crossmodal displays for their applications or integrate crossmodal cues in their interfaces.

Acknowledgements

Thank God for guiding me to complete this thesis at last after such a long but short time!

Sincerely thank my supervisor, Prof. Patrick Olivier for giving me the opportunity to do this research and some initial ideas and for his all supports throughout my PhD. Thanks also to my second supervisor, Christian Kray, in particular for his supervision with experimental methodologies and other suggestions.

Thanks to my research collaborators Daniel G. Jackson for coding and implementing the first version of CROSSFLOW and Stephen W. Gilroy for coding and implementing the CROSSBOARD idea. Thanks to all current and past members in Culture Lab and School of Computing at Newcastle University including John Shearer, Philip Heslop, Ahmed N. S. Kharrufa and other friends at Newcastle University for helping me on various ways.

Thanks to my parents, all brother, sisters, friends in churches and fellowships.

Thank you, Lord Jesus Christ!

Declaration

Part of the work in Chapter 2, 3, 4, 5, 8 and 9 including Experiment 1 in Chapter 5 and Experiment 5 in Chapter 8 has appeared in the various forms below. The papers below and user studies reported within them were the product of collaborative work between the author of this thesis and the others of the papers. The rest of the work including the analysis and discussion carried out within this thesis is entirely the author's own work. All other referenced material has been given full acknowledgement in the text.

Cao, H., Olivier, P., Jackson, D.G. & Armstrong, A., 2006. Mobile Devices in Crossmodal Interfaces. In *Proceedings of Pervasive Workshop on Pervasive Mobile Interaction Devices (PERMID 06)*. Dublin, Ireland. Available at: http://www.cs.ncl.ac.uk/publications/trs/papers/1004.pdf.

Cao, H., Olivier, P. & Jackson, D.G., 2008. Enhancing Privacy in Public Spaces Through Crossmodal Displays. *Social Science Computer Review*, 26(1), pp.87–102. Available at: http://dx.doi.org/10.1177/0894439307307696.

Cao, H., Olivier, P., Jackson, D. G., Armstrong, A., & Huang, L. (2006). Enhancing Privacy in Public Spaces Through Crossmodal Displays. In *Pervasive Workshop on Privacy, Trust and Identity Issues for Ambient Intelligence* (extended version accepted for Social Science Computer Review). Dublin, Ireland.

Gilroy, S.W., Olivier, P., Cao, H., Jackson, D.G., Kray, C. & Lin, D., 2006. CROSSBOARD: Crossmodal Access of Dense Public Displays. In *Proceedings of the International Workshop on Multimodal and Pervasive Services (MAPS06)*. Lyon, France. Available at: http://www.cs.ncl.ac.uk/publications/trs/papers/1005.pdf.

Olivier, P., Cao, H., Gilroy, S.W. & Jackson, D.G., 2007. Crossmodal Ambient Displays. In N. Bryan-Kinns et al., eds. *Proceedings of HCI 2006 People and Computers XX - Engage*. London, UK: Springer London, pp. 3–16. Available at: http://dx.doi.org/10.1007/978-1-84628-664-3_1.

Olivier, P., Gilroy, S.W., Cao, H., Jackson, D.G. & Kray, C., 2006. Crossmodal Attention in Public-Private Displays. In *Proceedings of ACS/IEEE International Conference on Pervasive Services (ICPS06)*. Lyon, France: IEEE, pp. 13–18. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1652201.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Thesis Aims, Research Questions and Research Methodology

Information overload and information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns (defined as *multi-user* problem in this thesis), have become two significant problems (for individuals and society) and challenges for the research and development of display-based technologies, since the vision of information being displayed everywhere all the time and people accessing information and displays anytime anywhere has become the reality in physical public spaces in modern cities.

Researches into display-based technologies including Peripheral Displays (Matthews et al. 2004) and Single Display Privacyware (SDG) (Shoemaker & Inkpen 2001) either only deal with one of the problems above or have one or more drawbacks, including a reliance on tracking technologies, weak protection for the user's privacy, small user capacity and high cognitive load demands. Findings of the research into human multimodal perception, especially multimodal integration and crossmodal interaction have received little attention and were not fully exploited (Sarter 2006); nevertheless, many design implications for the design of new interface or displays could be drawn from the relevant literature in neuroscience and cognitive and behavioural psychology. Though findings relating to human auditory-visual crossmodal interaction have resulted in several effective technologies in computer graphics (e.g. Mastoropoulou et al. 2005; Hulusic et al. 2008), display-based technologies which exploit human multimodal perception and aim to help prevent information overload and support information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns, are relatively unexplored.

1.1.1 *Thesis aims*

This thesis aims to explore the design of a new kind of display-based interface, named Crossmodal Displays to help tackle both information overload and *multi-user* problem and avoid the drawbacks associated with previous approaches. In particular, it aims to develop a systematic approach by exploiting human multimodal perception especially multimodal integration and crossmodal interaction and utilizing the synergy of both existing public displays and personal displays. It also aims to investigate the applicability of the approach in specific application areas including indoor navigation and information retrieval on dense public displays. Furthermore, the thesis aims to understand the effectiveness, efficiency, limitations and other possible effects of Crossmodal Displays through the investigation of the usability and the user experience afforded by the displays and to draw design suggestions for the future development of the displays.

1.1.2 *Research questions*

The research questions proposed for thesis are listed below. How these questions are addressed in the chapters is summarized in Section 9.2, Chapter 9.

RQ1: What are the significant problems in physical public spaces in the age of ubiquitous information and displays?

RQ2: Could Crossmodal Displays, the display-based interfaces that exploit people's multimodal perception and utilize the synergy of both existing public displays and personal displays help to tackle the problems in physical public spaces in the age of ubiquitous information and displays and avoid the drawbacks associated with previous solutions, including a reliance on tracking technologies, weak protection for user's privacy, small user capacity and high cognitive load demands?

RQ3: What are the usability and the user experience afforded by Crossmodal Displays that are developed for specific applications?

1.1.3 *Overview of research methodology*

In this thesis, we mainly use mixed methods research approach (Meissner et al. 2011) which involves the intentional collection of both quantitative and qualitative data and the combination of the strengths of each to answer RQ3, as we consider that the quantitative and qualitative approach alone is inadequate to develop multiple perspectives and a complete understanding about the usability and the user experience afforded by Crossmodal Displays that are developed for specific applications. The mixed methods are used in most of the pilot user studies and all formal experiments including Experiment 1, 2, 4 and 5, except Experiment 3 in which only quantitative approach is used.

Regarding the mixed methods designs, we adopt embedded designs, that is, to use quantitative and qualitative approaches in tandem and to embed one in the other to provide new insights or more refined thinking (Meissner et al. 2011). Different forms of data are integrated in the following ways. In Experiment 1, 2 and 5, the quantitative data collection and analysis are emphasized and the supplemental qualitative data collection (through the interview and the observation of the participants in Experiment 1, the survey and the observation of the participants in Experiment 2, the interview in Experiment 5) and analysis are embedded within the primary design. In Experiment 4, the qualitative data collection (through the survey, the interview and the observation of the participants) and analysis are emphasized and the quantitative data collection and analysis are embedded within the primary design.

1.2 Overview of Thesis

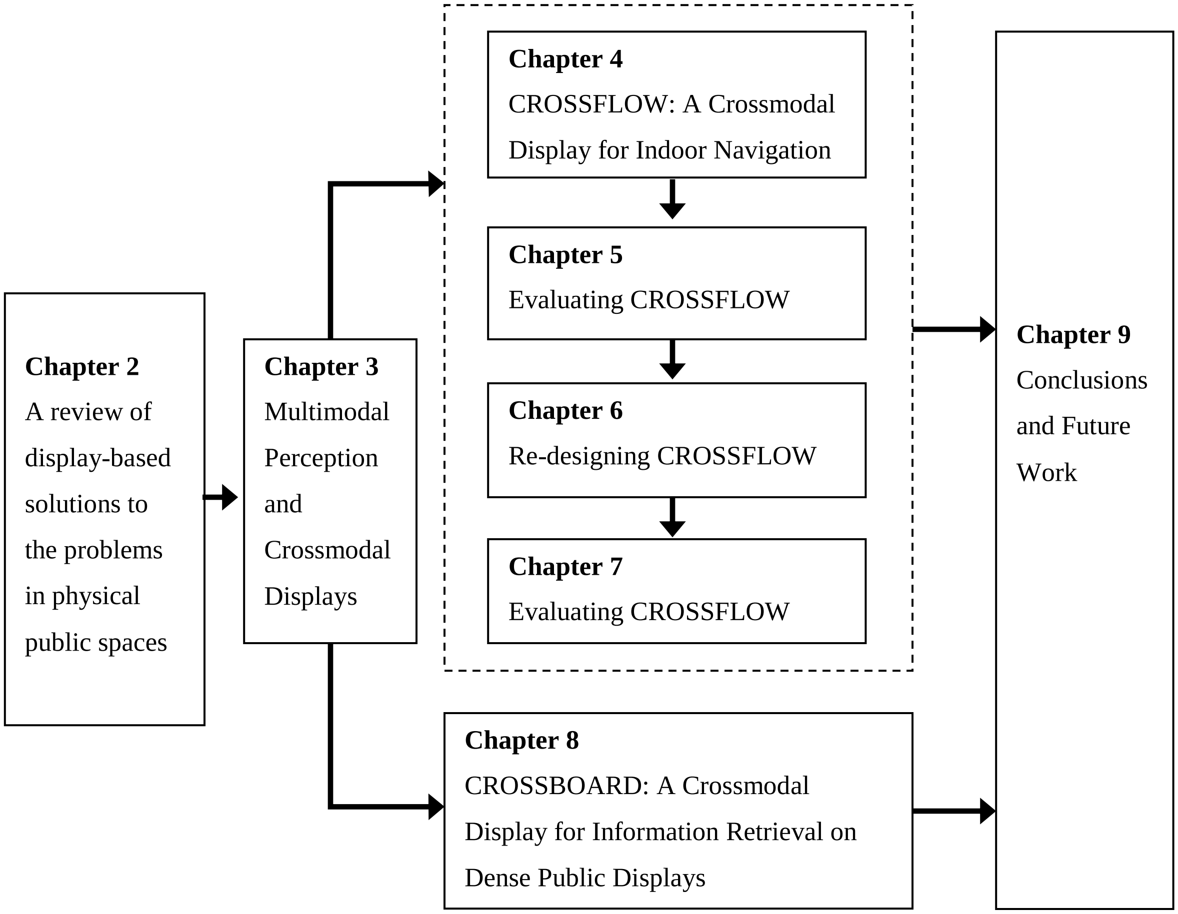

Figure 1.1 Thesis Structure

Chapter 2 identifies two significant problems relevant to the existing information displays in physical public spaces in the age of ubiquitous information and displays. A review of display-based technologies to the two problems is presented and their drawbacks are discussed, which leads to the identification of the research gaps and the three research questions that this thesis aims to address.

Chapter 3 reviews the neuroscientific literature on multimodal integration and crossmodal interaction in humans, and discusses the implications of these findings in relation to the problem of multimodal (and crossmodal) interaction and interface design. The implications constitute the foundation for a proposed conceptual model of Crossmodal Display that aims to address RQ2. The model is described so as to underpin the specific designs of Crossmodal Displays in the next five chapters.

Chapter 4 reports the exploration of the design of CROSSFLOW, a type of Crossmodal Displays for indoor navigation aiming to facilitate indoor navigation of multiple simultaneous users within an unfamiliar complex indoor environment. It details the approach used and the procedure of the development including users' task analysis, conceptual design and prototyping. In the design exploration, different design possibilities and selection criteria are presented. The first version of CROSSFLOW prototype including the system components, how they can be configured and used in practice is also described.

Chapter 5 reports a single-user and a multi-user lab study on CROSSFLOW Version 1 (Experiment 1 and 2), in which the usability and the user experience afforded by CROSSFLOW V1 are investigated in comparison with a traditional handheld paper map so as to answer RQ1 and RQ2. The quantitative and qualitative results validate the effectiveness of CROSSFLOW as a solution to help to solve the problems identified in Chapter 2 in indoor navigation application domain, demonstrate a set of distinct advantages over traditional navigation tools and existing technologies reviewed in Chapter 4, and reveal the aspects that need to be improved.

Chapter 6 presents the redesign process carried out for CROSSFLOW. The implications for redesign are drawn from the analysis of findings of Experiment 1 and 2 and the implications of multimodal perception for interface design in Chapter 3. Based on the implications, a new version of CROSSFLOW prototype system, CROSSFLOW Version 2 is developed. The new design of the system components including the grass-like pattern used for public directional information and the wind-like sound used as private cue and the corresponding configuration methods in the system is presented.

Chapter 7 reported two user studies (Experiment 3 and 4) on CROSSFLOW V2. Experiment 3 investigates appropriate parameters of components of CROSSFLOW V2. Experiment 4 investigates the usability and the user experience afforded by CROSSFLOW V2 as a type of aid helping user to identify a target (e.g. an object, location) from a large amount of visually similar objects or spatial structures within unfamiliar complex indoor environments. The results of both experiments contribute to the general design and evaluation suggestions for CROSSFLOW in Chapter 9.

Chapter 8 has reported an exploration of the design and an initial evaluation of CROSSBOARD, a type of Crossmodal Displays that combines hierarchical spatial multiplexed visual public cues with coordinated hierarchical temporal multiplexed audio and/or vibrotactile private cues to highlight the locations of information items relating to multiple simultaneous users on a large high-density information display. It demonstrates the effectiveness and efficiency of this solution through the results of a single-user lab study in which the potential of CROSSBOARD for improving the retrieval of un-indexed information from dense information displays is clearly shown.

Chapter 9 summarizes the work presented in this thesis, and relates back to the research questions outlined in Chapter 1. It also suggests a lifecycle model of Crossmodal Display development and provides a set of design suggestions for designers or researchers who wish to develop Crossmodal Displays for their applications or research. The limitations of this research including the limitations of the infrastructures of Crossmodal Displays, the design and evaluation of CROSSFLOW and CROSSBOARD are discussed, and suggestions for future work based on the issues raised by the thesis are proposed.

1.3 Main contributions of this thesis

This thesis opens a new design space for the design of displays, interfaces, interactions or information system outputs by exploiting human multimodal perception, especially multimodal integration and crossmodal interaction and utilizing the synergy of both existing public displays and personal displays.

It identifies several potentially useful design implications for the design of Crossmodal Displays as well as other multimodal interfaces from the findings of the research into human multimodal perception (see Chapter 3 for details).

This thesis also provides a set of design suggestions and a lifecycle model of Crossmodal Display development which can be used by designers or other researchers who wish to develop Crossmodal Displays for their applications or integrate crossmodal cues in their interfaces (see Chapter 9 for details).

Chapter 2. A review of display-based solutions to the problems in physical public spaces

2.1 Introduction

The purpose of this chapter is to provide the background for this research and identify the research gaps so as to inform the direction for the chapters that follow.

The chapter begins by discussing the changes that have happened in physical public spaces in the age of ubiquitous information and displays, then two problems relevant to displays and their detrimental impacts on individual and society are identified to be important in need of solutions. The remainder of the chapter reviews the solution spaces to each problem, and discusses the related works including the features and drawbacks associated with the previous display-based solutions so that the research gaps could be identified. Finally, based on the discussion of solutions, the research gaps are identified. The relation between the research gaps and the research questions that this thesis aims to address is indicated and the methods of studying these research questions are described. The main outcomes of this chapter are summarized in the conclusions section.

2.2 The significant problems in physical public spaces in the age of ubiquitous information and displays

2.2.1 *The age of ubiquitous information and displays*

Since Ubiquitous Computing (Weiser 1991), a notion that in the near future our everyday environment will be embedded with many thousands of sensors and computational devices was proposed, the quantity, ubiquity, density, heterogeneity of information and displays have continuously grown in physical public spaces due to the rapid development of information and communication technologies over the past two decades.

As a result, physical public spaces, physical spaces that are normally shared by more than one person, where traditionally a large amount of information is congregated, exchanged and communicated, become even busier nowadays (Ollson 2007; Mitchell 2000). Vast of information is presented via public displays which are situated in physical public space displaying information publicly and used separately or concurrently by multiple simultaneous users such as projection, screens and speakers, and personal displays which are worn or carried close to users' bodies such as mobile phone and Personal Digital Assistant (PDA), on which information is presented only to the owner of the display. The information transactions between devices and interactions between people and information systems could happen anytime anywhere in the spaces with the support of innovations, e.g. broadband, wireless technology, Ambient Intelligence (Aarts & Roovers 2003; Ollson 2007). The vision that information is displayed everywhere all the time and people could access information and displays anytime anywhere almost have become the reality in many places such as New York City's Time Square in modern cities (McQuire 2006). Such change has brought a number of benefits, such as that information scarcity is not a problem in public spaces any longer, and accessing services or interacting with information is not confined by locations and time. However, it also brought problems or aggravated the detrimental impacts of the existing problems.

2.2.2 *Information Overload*

Information Overload is defined as an overwhelming feeling of an individual facing an amount of information or options which is more than the individual's capability for processing them a short time (Heylighen 2004a; Jones et al. 2004). There are a wide range of research and reports (e.g. Heylighen 2004b; Advertiser 2010) about this problem and its impacts on the individual and society. The negative results such as stress, anxiety, confusion, anger, feelings of helplessness on the individual's physical and psychological aspects are mentioned frequently. From a social aspect, the results include low productivity or work efficiency, loss of control and delayed or wrong decisions or judgments.

Information overload has become very common in physical public spaces in modern cities. On the one hand, increased ubiquity and amount of public and personal displays create significant clutter in physical public spaces, which easily results in higher demand on users' limited attention and mental resources. On the other hand, increased density and amount of information presented by various displays (e.g. audio or vibration notifications in mobile phones, advertisements on outdoor plasma screens) in short amounts of time causes strong feelings of information overload. The feelings could be intensified if the information involves multimedia. Both causes and effects can be found in modern city public spaces such as Trafalgar Square in London (Figure 2.1) and New York City's Time Square (Intille 2002).

Figure 2.1 Public displays in Trafalgar Square in London

The problem also becomes more and more prominent when people attempt to search for information in physical public spaces. Retrieving relevant information on a public display (e.g. a passenger's flight information on an electronic information board), locating a relevant object (e.g. a piece of commodity on a shelf in a supermarket) and searching a sign or landmark are common tasks that people often perform in a public space. Performing these tasks using a traditional way such as visually scanning the whole display or space to locate or retrieve relevant information or object becomes more and more difficult and inefficient for people because of the increased visual clusters and heterogeneity of public displays and information. In a public space (e.g. an airport), all information that potentially interests (e.g. advertisements) or is relevant to the occupants of the space (e.g. flight numbers and corresponding destinations) usually is shown (e.g. hundreds of flight information are densely clustered and shown on a large wall-size display or several smaller-size displays). The amount, manner and content of the information so presented are unlike people normally experience in their home environment (Intille 2002). In fact, most of the displays and information are irrelevant to a user. Therefore, increased public displays and information probably impose more mental load on the user.

Research in psychology (e.g. Koch & Ullman 1985) indicates that selective attention gives human an advantage to manage attention by being able to concentrating on specific information or stimuli while ignoring the rest in an environment. However, too much input of information still causes an information overload problem. More issues about information retrieval on dense public displays are discussed in Section 8.3 in Chapter 8.

2.2.3 *Information presentation for multiple simultaneous people*

Figure 2.2 Situated information boards showing information for multiple simultaneous people in a railway station

Many public displays in physical public spaces are not designed to effectively satisfy different information needs of multiple simultaneous users and preserve their privacy and anonymity. When multiple simultaneous people in an unfamiliar physical public space who are strangers to each other, researched by Paulos and Goodman (2004), have different information needs as well as privacy and anonymity concerns, a traditional public display hardly satisfies their needs. A typical situation is that passengers in an airport or railway station view the information items about their flights or trains on an information board simultaneously (see Figure 2.2 and 8.1).

Regarding different information needs, traditional public displays usually only deliver generic information. They do not support personalization, that is, information content, presentation, interface on the displays are not tailored for different need, identity, time visited, preference etc. of an individual user or user group. Such displays include both dynamic displays (e.g. the examples shown in Figure 2.1 and 2.2) which update information from time to time or show animated information and static displays (e.g. the example shown in Figure 2.3) which seldom update information.

Regarding privacy and anonymity in a communal space, people have been familiar with each other or are encouraged to know each other thus they have less concern about privacy when accessing personal information on a shared display or in a shared physical space. However, Brignull (2005) found that most of people would feel awkward or uncomfortable when they see that their personal private information is presented on a public display unexpectedly even in a community environment. A person in an unfamiliar public environment, such as a passenger in an airport or a shopper in a shopping mall lacking the trust of bystanders or onlookers, is more likely to prefer his/her privacy (e.g. destinations, interests) and anonymity to be preserved while accessing the information relevant to him/her. However, most of public displays are designed without concern for different relationships among users. When multiple simultaneous collocated users view or interact with a single traditional public display showing information relevant to them, where both the information content relevant to a user and the relation between the information viewed and the user are exposed to public view, the user's privacy and anonymity are weakly protected and could easily be undermined, which may cause the user's negative feelings such as awkwardness, shame or other social problems. Two real world scenarios are that a group of students check their exam scores on a public announcement board, and several shoppers stand in front of a shelf in a chemist's shop to buy medicine. A more comprehensive review on the issue of information displays relating to privacy and anonymity can be found in Cao et al. (2008).

Figure 2.3 A static map posted in an airport

Drawn from above, we identify two problems that are significant in physical public spaces in the age of ubiquitous information and displays: 1) information overload problem and 2) information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and concerns for privacy, defined as *multi-user problem* in this thesis, and their detrimental impacts on individuals and society. These problems have also been the focuses of the research and development of many display-based technologies in the past two decades.

2.3 Solution Space

In response to the significance of the problems identified above in Section 2.2, various solutions, research areas, technologies have been proposed or explored. However, most of them only focus on one of the problems, and few of them deal with the all problems in a comprehensive way.

We reviewed some representative display-based technologies and systems and divided them into the two solution spaces shown below according to their main design goal. This is not intended to be an exhaustive review of all research relevant to the two problems we identified in Section 2.2, but rather to provide an overview of these display-based technologies so as to identify research gaps which requires further investigation.

2.3.1 *Solution space to information overload problem*

Intille (2002) indicated that one of the approaches to prevent information overload is to present information to the periphery of the user instead of requiring his/her focus of attention. In the last twenty years, many display-based technologies have been investigated which exploit the users's information processing capability in periphery sensations, and they form several important subgroups of ubiquitous systems including peripheral displays, ambient displays, calm technology, informative art (Matthews et al. 2003; Laakso 2004; Ljungblad 2003). These subgroups research areas are reviewed in the following subsections. Within these areas, the projects or prototypes in table 2.1 are particularly surveyed. It is found that utilizing the user's periphery to deliver all information or partial information so as to help to prevent information overload is a common feature of these displays.

A. Peripheral Displays

Matthews et al. (2003) categorized attention into four main zones: preattention, inattention, divided attention and focused attention. According to their categories, the displays that do not utilize focused attention are peripheral displays. They include most of the displays we surveyed in table 2.1 except Interactive Public Ambient Displays, Notification Collage and Semi-Public Displays.

Peripheral displays aim to deliver information in an unobtrusive way, and minimize a user's attentional and cognitive effort to process information. They convey information to a user in his/her periphery either without demanding their full attention or requiring minimal attention and cognitive effort (Matthews et al. 2003). In order to achieve the goal, the users' peripheral sensations are utilized. For example, the designers of peripheral displays such as Bus Mobile (Matthews et al. 2003) take the advantages of *peripheral vision* (Thibos & Bradley 1991; Reynolds et al. 2010) to deliver information regarding body attitude, self-motion through the environment and moving objects. Because information is delivered via users' peripheral sense, therefore peripheral displays normally just convey non-critical information, i.e. peripheral information (Maglio & Campbell 2000). Although Mynatt et al. (1998), Wisneski et al. (1998), Pedersen and Sokoler (1997) and others have explored auditory and/or haptic displays (see Table 2.1), most of the research into peripheral displays focuses on visual displays because of the superiority of human visual perception. Visual peripheral displays developers have a sophisticated understanding of the nature of visual cognition and aim to present information in a timely manner which appropriately matches the time-sharing strategies when users are performing two related tasks simultaneously. However, when the dual or multiple tasks become too demanding, the visual channel is easy overloaded and errors increase.

B. Ambient Displays

Ambient displays are a subset of peripheral displays with greater emphasize on the integration of information and display hardware in the environment. The popularity of ambient displays reflects a rise in interest in notions of user experience, and in particular, the aesthetics of the user interface and the links between user interface design and traditional art and design. In part they have been the result of a growing need and desire to situate computer interfaces in the everyday environments of users. The developers of ambient displays utilize the users' divided attention (Matthews et al. 2003).

Ambient displays extend traditional desktop configuration of peripheral displays such as Agent Tickers (McCrickard & Zhao 2000) by embedding displays in a user's everyday environment and utilizing highly aesthetically pleasing, implicit, indirect and essentially peripheral representations, either through specialized display technologies (e.g. floor lights), projection, or by dynamically controlling properties of familiar objects such as fountains and mobiles to present information without distracting or burdening the users in the environment (Laakso 2004; Mankoff et al. 2003). A number of ambient display prototypes have been developed. The Dangling String (Weiser & J. S. Brown 1996), InfoCanvas (Miller & Stasko 2001), Informative Art (Holmquist & Skog 2003) and AROMA (Pedersen & Sokoler 1997), Ambient Fixtures including Water Lamp and Pinwheels (Wisneski et al. 1998) and ambientROOM (Ishii et al. 1998) are representative examples.

The Dangling String was one of the first and simplest ambient displays as well as the representative example of calm technologies which emphasize the calm sense of information presentation in an environment (Weiser & J. S. Brown 1996). Created by the artist Natalie Jeremijenko, it was an 8-foot piece of plastic spaghetti that hangs from a small electric motor mounted on the ceiling and the sound from the installation was mapped to the change of bits of information in the network going past the Ethernet cable (Laakso

2004). Besides the general features of ambient displays, calm technologies have the following features summarized by Ljungblad (2003): (1) Information is delivered in the background and periphery of the user until the user wishes to attend; (2) Engage both the centre and the periphery of the user's attention and move back and forth between the two (Weiser & J. S. Brown 1996); (3) Do not involve subconscious awareness of the user.

Informative art is an approach exploring calm technologies, which emphasizes aesthetic quality of information presentation, and a type of ambient displays that can be used for visualizing information (Ljungblad 2003). Informative art differs from ambientROOM and Ambient Fixture in the way that it relies on the inspiration from both the appearance and the function of traditional art (e.g. painting, drawing) to convey information calmly (Holmquist & Skog 2003). Several prototypes of informative art were built to explore this idea. For example, an Andy Warhol inspired informative art is an egg clock, where time gradually replaces asparagus cans by tomato soup cans, a Bridget Riley and Piet Mondrian inspired information art show the activity in a room and email traffic where each square represent one person's mailbox or the weather in a city respectively (Holmquist & Skog 2003).

Jansen (2009) defined the purposes for which ambient displays were created and grouped them into three distinct goals in an increasing order of information demands, that is, entertaining, informing and supporting.

The Information Percolator (Heiner et al. 1999) using controlled rising air bubbles in tubes filled with water to display small and black/white image is a good example of ambient displays for entertainment, as it is designed to be highly aesthetically pleasing and not necessarily helpful for the user to complete a task.

A large portion of ambient displays are designed for informing. They keep their users informed using symbols, metaphors and abstractions so that large amounts of information can be conveyed in a less demanding way. For example, InfoCanvas (Miller & Stasko 2001), Informative Art (Holmquist & Skog 2003) and AROMA (Pedersen & Sokoler 1997), each of which incorporates abstract design elements, motivated by different styles of visual art, to represent information, e.g. weather, general trends of stock values for a trader, network traffic information for a system administrator, and so on. The designers of Ambient Fixtures including Water Lamp and Pinwheels (Wisneski et al. 1998), ambientROOM (Ishii et al. 1998), AmbiX (i.e. AmbiPower, AmbiTweet, Ambient

Anchorman, Ambient Timer and Ambient Reminder) (Müller et al. 2012) try to leverage aspects of everyday objects, media and artistic genres as a means of embedding dynamic information (e.g. from the Internet or any networked information sources) into our everyday environment. These displays have a common feature that the mapping from information changes to system state changes in a "calm" and "subtle" manner.

As Jansen (2009) classified, the third category of ambient displays is designed for supporting users to complete certain tasks in some context. For example, Interactive Public Ambient Displays (Vogel & Balakrishnan 2004) support multiple simultaneous users accessing both public and personal information, subtle notification, their privacy controls and self-revealing help without overloading each user through implicit interaction, that is, sensing the contextual cues of the users such as body orientation and position and user proximity to the display, and explicit interaction including the users' hand gestures and touchscreen input. Another example is AmbiGlasses (Müller et al. 2012), a mobile ambient display utilizing a pair of glasses with 12 LEDs that illuminate the periphery of the user's field of view, which supports navigation tasks through delivering directional information in an unobtrusive and intuitive way.

The goal to prevent information overload via ambient displays could fail in some circumstances. For example, when several devices in ambientROOM demand awareness such as sound, light, airflow and water movement in the same time, it would be very stressful and perhaps impossible for a user to process all the information (Wisneski et al. 1998).

C. Change blind displays

Change blind technologies aim to exploit change blindness in human attention to design the transition of information updates that do not attract a user's focus of attention. This approach is different from that using exploitation of periphery. Intille (2002) argued that the detection of change itself may be the main factor which causes information overload. Based on the research literature about change blindness (e.g. James et al. 1999; Henderson 1997), he proposed a set of detailed techniques for change blind displays including blanking an image, changing views, displaying "Mud slashes", changing information slowly, exploiting eye blinks or saccades and using occlusion. The most distinct feature of change blind displays that is different from ambient displays is utilizing human inattention. We surveyed Agent Tickers, AROMA, Digital Family Portrait, InfoCanvas, Information Percolator, Bus Mobile, Stock-News Displays and Guitar (see Table 2.1). These are the displays making use of human inattention to prevent information overload. As one of the examples, the Agent Tickers display changes unnoticeably by fading text so that the user can maintain awareness of changing information without being distracted (McCrickard & Zhao 2000).

Displays	**Reference**	**Attention zone**	**Supported users' sensory modality**
Agent Tickers	(McCrickard & Zhao 2000)	inattention and divided attention	visual
ambientROOM	(Ishii et al. 1998)	divided attention	visual and auditory
Ambient Fixtures (Water Lamp and Pinwheels)	(Wisneski et al. 1998)	divided attention	visual
AROMA	(Pedersen & Sokoler 1997)	inattention and divided attention	visual, auditory and haptic
Audio Aura	(Mynatt et al. 1998)	divided attention	auditory
Dangling String	(Weiser & J. S. Brown 1996)	divided attention	visual and auditory
Digital Family Portrait	(Mynatt et al. 2001)	inattention	visual
Informative Art	(Holmquist & Skog 2003)	divided attention	visual
Information Percolator	(Heiner et al. 1999)	inattention and divided attention	visual
InfoCanvas	(Miller & Stasko 2001)	inattention and divided attention	visual
Kimura	(MacIntyre et al. 2001)	divided attention	visual
Lumitouch	(Chang et al. 2001)	divided attention	visual
Sideshow	(Cadiz et al. 2002)	divided attention	visual

Displays	Reference	Attention zone	Supported users' sensory modality
UniCast	(McCarthy et al. 2001)	divided attention	visual
GroupCast		divided attention	visual
OutCast		divided attention	visual
Bus Mobile	(Matthews et al. 2003)	divided attention	visual
Bus LED		divided attention	visual
Stock-News Displays		Inattention and divided attention	visual
Ring Ticker		divided attention	visual
Guitar		divided attention	auditory
AmbiPower	(Müller et al. 2012)	divided attention	visual
AmbiTweet		divided attention	visual
Ambient Anchorman		divided attention	visual
Ambient Timer		divided attention	visual
Ambient Reminder		divided attention	visual
AmbiGlasses		divided attention	visual
Daylight Display	(Mankoff et al. 2003)	divided attention	visual
Weathermobile.	(Laakso 2004)	divided attention	visual
Interactive Public Ambient Displays	(Vogel & Balakrishnan 2004)	divided attention and focused attention	visual
Notification Collage	(Greenberg & Rounding 2001)	divided attention and focused attention	visual
Semi-Public Displays	(Huang & Mynatt 2003)	divided attention and focused attention	visual

Table 2.1 A survey of the users' attention zone and sensory modality utilized by the displays

2.3.2 *Solution space to multi-user problem*

The relationships among people who share a physical space or a situated interface are very different. Three types of relationship should be considered in information provision and presentation:

1) If people are familiar with each other, then information should be presented to support and encourage collaboration.
2) If people are unknown to each other but intend to know each other, then information should be presented to support socialization.
3) If people are unknown to each other and no intention to know each other, then information should be presented to individuals privately and the information presentation for each person should not interfere with one another.

The first type is researched and supported by Computer-Supported Cooperative Work (CSCW) (Bannon 1993) systems. The second has been investigated by Paulos and Goodman (2004) and supported by Community Displays (Brignull 2005). In the physical public spaces we investigated in this thesis such as airports, shopping malls, the relationship among most of people is the third type, therefore the interface of information systems used in such kind of environment should be designed to effectively satisfy different information needs of multiple simultaneous people who share a physical space or situated interface and preserve their privacy and/or anonymity. The representative approaches and detailed systems that can realize this are reviewed as follows.

A. Physically partitioning of the shared space

Physically partitioning a shared space is a common method to preserve people's privacy in modern public spaces. For example, the essence of a telephone booth is the creation of a physical private space within a public space, allowing for unhindered private information access. A public toilet allowing to be occupied by only one user at a time is another example. It is also utilized by Vogel and Balakrishnan (2004) to support the display of different public-private levels of information for multiple simultaneous collocated people to do their own thing without interfering each other. They developed a framework for sharable, interactive public ambient displays that supports the transition from implicit to explicit interaction with both public and personal information. Their system physically partitions the users' shared space through the utilization of a marker-based optical motion tracking system and provides four continuous phases with fluid inter-phase transitions. The shared public ambient display can present different public-private levels of information for several users each within distinct interaction phases simultaneously. Which level of information is displayed to a user depends on his location and orientation. For example, while gestures from a distance may provide an implicit cue for selecting an item or displaying more detailed descriptions of the notifications; direct touching of the screen is best suited for accurate, up-close interaction. By dividing the shared physical space into four interaction phases and using transparency, the system allows its users to reach beyond their own space to access information, i.e. see through the semi-transparent message boxes to the public information beneath.

Although the system supports presenting personal and private information by displaying a small font size and exploiting natural body occlusion when a user steps closer to the display and enters the personal interaction phase, it cannot occlude the view of the user's personal information from simultaneous bystanders. Furthermore, by the body occlusion, the user cannot possibly remain anonymous as the behaviour of interacting with private data is also exposed to other users or onlookers. Another significant drawback of their system is that the system heavily relies on sensing and tracking technologies. It is very expensive and difficult to track the user's detailed posture and gesture via IRFD tags; moreover, tags probably are intrusive for the user. In addition, because of the restriction of tracking, the number of users their system can support simultaneously is very limited.

B. Physically partitioning of the shared visual display

Besides a shared space, the way that physically partitions a shared visual display is also explored. Shen et al. (2002) partitioned the corners of a shared display as semi-private spaces for keeping an individual user's collection of photos. The main drawback of their system is that the number of users their system can support simultaneously is relatively small because limited private spaces.

Wu and Balakrishnan (2003) utilized a different approach. They project the personal or private information onto a user's tilted palm when he places his hand vertically and slightly tilted on top of a top-projected tabletop display. The main drawbacks of their system are that a reliance on tracking technologies and small user capacity.

C. Shared display combining with private channel

Another solution to the problem of multiple simultaneous users is to enable users to access personal and private information through private sensory channels.

Single Display Privacyware (SDP) (Shoemaker & Inkpen 2001), a subcategory of sharable interface, attempted to provide private visual information for each individual user within the context of a display that is shared by a group. Building on Single Display Groupware (SDG) (Stewart et al. 1999) that utilize a single, large shared display, the system of Shoemaker and Inkpen (2001) requires that pairs of users wear specially adapted glasses to view private aspects of the information shown on a shared display. Although the use of either alternating-frame shutter glasses or head-mounted displays is successful in allowing shared and private activities to take place at the same time, such configurations have drawbacks, such as limitations on the number of users (only two) able to use the shared visual display.

Morris et al. (2004) noticed the importance of multimodal SDG interfaces, and then they explored the use of a shared tabletop multimodal interface to convey personalized or private information to individual users of a group through individual audio channels. The use of single-ear headsets in the system enables both personal and shared information access and group collaboration. However, their system only supports a small number of users. The number of soundcard that the system can support restricts the number of users (only four).

D. Auxiliary displays networked with shared displays

There has been a significant amount of research into the display of private information to different users through the use of several smaller, auxiliary displays which are networked with one or more shared displays (e.g. Rekimoto 1997; Raghunath et al. 2003; Berger et al. 2005; Prante et al. 2003).

Auxiliary visual displays enable each individual user to view and manipulate private information presented on them, or to interact with the visual information presented on the shared display. However, as Shoemaker and Inkpen (2001) indicated, the requirement that the user looks back and forth between the auxiliary and the shared displays imposes a substantial cognitive load (divided attention), which can result in significant reductions in performance. Although blurring the sensitive information content can be used to preserve the private sections in symbiotic displays (Raghunath et al. 2003; Berger et al. 2005), one still has to look away from public display to examine the personal mobile device, thus revealing to other occupants of the space that he/she is examining private data, and that the display is for him (Cao et al. 2008).

Moreover, shifting attention from public displays embedded in the space encompassing the users to their personal displays might lead to a disengagement of users from their environments, which has several detrimental consequences, e.g. failing to maintain the correct balance between the protection of user privacy and social etiquette in public spaces by alienating users from other occupants of the space in which they reside.

2.3.3 *Discussion*

A. Solutions to information overload problem

All the systems or interfaces in the solution space to information overload problem help to prevent information overload through reducing cognitive or attention load, supporting attention management or facilitating information processing. The boundary between these solutions is quite blurred as they share some common features, e.g. displaying information in the background of the user, not occupying the user's full attention. The key difference between them is the way of conveying information to the user.

However, most of these display-based interfaces, except Interactive Public Ambient Displays (Vogel & Balakrishnan 2004) and Audio Aura (Mynatt et al. 1998), are not developed for supporting multiple simultaneous people who share a physical public space or a display but have different information needs and privacy concerns. They normally just convey generic public information (e.g. news, stock values, weather, traffic congestion and human activity), and have little protection for the user's privacy. Although some of them support the display of personalized or private information in public spaces by allowing the user to map the preferred information, e.g. Informative Art (Holmquist & Skog 2003) which allows a user to map information meaningful to him/her to abstract graphics, but the shared interface cannot present different personal and private information at the same time. Some displays support privacy preservation by personalized mappings but the display can only be used by a single user. For instance, InforCanvas (Miller & Stasko 2001) uses a cartoon-like beach landscape in which certain visual elements can convey personal private information that can be defined by the user such as that the latitude of a bird may indicate the activity of a particular stock which the user is interested in.

B. Solutions to *multi-user* problem

All the systems or interfaces in the solution space for *multi-user* problem support information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns. They are specialised to more than one user. The shared displays or interfaces in the systems can deliver generic information and specific information to one or a group of users determined in response to the requirements of a specific user (e.g. through option selection by the user at the display), to the identity of the user (e.g. through Radio Frequency Identification (RFID) or infrared or Bluetooth networking), to user's personal device, or to the user’s private channel.

However, most of these projects or prototypes are not developed for preventing information overload. Moreover, they have the following drawbacks: a reliance on tracking technologies, small user capacity, and/or high cognitive load demands.

The systems developed by Vogel and Balakrishnan (2004) and Wu and Balakrishnan (2003) highly rely on tracking which is relatively expensive besides high cost of supporting infrastructure. Tracking users also presents additional sets of challenges for developers and designer, both technical and ethical. Furthermore, because of the restriction of tracking, their systems only support a small number of users.

Small user capacity is one of the significant drawbacks of the systems developed by Vogel and Balakrishnan (2004), Shoemaker and Inkpen (2001), Morris et al. (2004) and Wu and Balakrishnan (2003).

As we reviewed in Section 2.3.2 D, the system that utilizes both visual auxiliary and shared displays may impose a substantial cognitive load on the users.

2.4 Research gaps and research questions

Drawn from above, we identify that a solution that is able to address both the problems, i.e. information overload and *multi-user* problem identified in Section 2.2 and avoid the drawbacks of the previous solutions reviewed in Section 2.3 is needed.

Moreover, research into human multimodal perception (especially multimodal integration and crossmodal interaction) has received little attention and was not fully exploited (Sarter 2006). Nevertheless, many design implications for the design of new interface or displays could be drawn from the literature in neuroscience, cognitive and behavioural psychology (see Chapter 3 for our detailed literature review). Though findings relating to human auditory-visual crossmodal interaction have resulted in several effective technologies in computer graphics (e.g. Mastoropoulou et al. 2005, Hulusic et al. 2008), the display-based technologies or interfaces which exploit human multimodal perception, especially multimodal integration, crossmodal interaction and crossmodal attention (Driver & Spence 1998), and aim to help to prevent information overload and support information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns, are relatively unexplored. Such technologies as solutions to our problems might have significant advantages over the solutions we have reviewed in Section 2.3. Furthermore, it is important to investigate the usability and the user experience afforded by the systems developed on the basis of such new solutions so as to understand the effectiveness, efficiency and other possible effects or limitations of the solutions.

These research gaps are filled by the answers to RQ2 and RQ3. The research questions for this thesis proposed in Chapter 1 are:

RQ1: What are the significant problems in physical public spaces in the age of ubiquitous information and displays?

RQ2: Could Crossmodal Displays, display-based interfaces that exploit people's multimodal perception and utilize the synergy of both existing public displays and personal displays help to tackle the problems (identified in Section 2.3) in physical public spaces in the age of ubiquitous information and displays and avoid the drawbacks (identified in Section 2.4) associated with previous solutions including a reliance on tracking technologies, weak protection for user's privacy, small user capacity and high cognitive load demands?

RQ3: What are the usability and the user experience afforded by Crossmodal Displays that are developed for specific applications?

This thesis addresses RQ2 and RQ3 by using a range of methods including identification of user requirements in specific contexts, task analysis, conceptual design, prototyping and lab user studies, which are adopted in many Human-Computer Interaction research methodologies. These methods are utilized in a human-centered design process or cycle similar to ISO13407 human-centered design lifecycle model (Sharp et al. 2007; p.462). In Chapter 9, all the methods used in this thesis are structured into a lifecycle model (see Figure 9.1) in order to show a clear perspective of the research methodology used in this thesis.

2.5 Conclusions

In this chapter, two significant problems relevant to traditional public displays in the age of ubiquitous information and displays in physical public spaces are identified, that is, information overload problem and information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns.

Previous solutions reviewed and discussed in this chapter either only deal with one of the problems or have one or more drawbacks such as a reliance on tracking technologies, weak protection for user's privacy, small user capacity and high cognitive load demands. With an interest in the findings of research into human multimodal perception, we found the potential to exploit human multimodal perception, especially multimodal integration and crossmodal interaction for the development of new display-based technologies that could help to tackle both the problems we identified and avoid the drawbacks associated with previous solutions.

The problems identified in Section 2.2 and the drawbacks of previous solutions identified in Section 2.3 lead to the research gaps that this thesis aims to fill. The research gaps are filled through the answers to the research questions proposed in Chapter 1. RQ1 has been answered in Section 2.2, the other two research questions are addressed in the following six chapters, and then the answers to all three questions are summarized in Chapter 9. In order to answer RQ2, in the next chapter (Chapter 3), a review of the literature of human multimodal perception is presented, which provides several design implications for the design of Crossmodal Displays, a solution proposed to help to tackle both the information overload and *multi-user* problems.

Chapter 3. Multimodal Perception and Crossmodal Displays

3.1 Introduction

Chapter 2 has identified the problems and research gaps in current research for display-based technologies. We have identified that although multimodal interfaces and multimodal information presentation have been widely applied, multisensory perception has not been fully exploited. In particular, the current understanding of crossmodal perception at the behavioural, psychophysical, and neurophysiological levels has received little consideration in multimodal display and interaction design. As indicated by Sarter (2006), this disconnection from the scientific knowledge base limits interface designers' ability to fully exploit this aspect of human perception.

One of our primary research aims is to better exploit people's multimodal and crossmodal information processing capabilities in the development of novel interfaces and displays, thereby addressing the problems identified in Chapter 2. To do this, we first review the neuroscientific literature on multimodal integration and crossmodal interaction in humans. We then discuss the implications of these findings in relation to the problem of multimodal (and crossmodal) interaction and interface design. Based on our review of traditional public and personal displays in Chapter 2, we present a conceptual model of the *Crossmodal Display* that underpins the specific designs that we describe and evaluate in the chapter that follow.

3.2 Multimodal integration and crossmodal interaction in humans

It is well-known that we have sensory modalities including vision, hearing, touch, proprioception, smell, and taste. It is natural, and very important, that we can sense our surroundings through several different sensory modalities simultaneously and integrate multimodal information into a perceptual experience that is coherent and unified (Spence & Driver 2004). In the perception research literature, this ability, and effects in the process, of coordinating (or 'matching') the information received through multiple perceptual modalities, is referred to as multimodal or multisensory integration. Our perception of sensory stimuli and the impact of this on our behaviour vary according to the task and sensory modality, and a number of relevant phenomena have been explored. For example, Bertelson and Aschersleben (2003) conducted a study in which participants were required to tap their finger in the presence to two stimuli of different frequency and modality – an auditory cue in the form of a ‘clicking’ sound and visual cue in the form of a ‘flashing’ light. A bias towards the use of audition, over vision, for temporal resolution was reflected in the participants’ tendency to tap in synchrony with the auditory stimuli.

Although observations about crossmodal interactions appeared at the dawn of modern psychological science – Johannes Muller refers to ventriloquist effect in his 1839 presentation of his ‘law of specific energies of nerves’ – crossmodal interaction is a relatively new area of investigation for multisensory integration research and explores the influence that distinct sensory modalities have on each other. Researches in cognitive psychology, neuroscience and psychophysics have given rise to the identification of various crossmodal interaction phenomena. The best known of these are the *McGurk effect* (McGurk & MacDonald 1976; Green & Kuhl 1989) and the *ventriloquism effect* (Hairston et al. 2003), and the *bounce-inducing effect* (R. Sekuler et al. 1997).

The ventriloquism effect describes people’s tendency to spatially associate auditory and visual stimuli where they exhibit synchronicity, whereas the McGurk effect relates to an interaction between auditory and visual speech recognition. McGurk and MacDonald (1976) showed participants videos of phonemes being spoken over which they dubbed sound-recordings of different phonemes. For many subjects the perceived phoneme was a third, different phoneme, which was in some way intermediate to the actual and dubbed phonemes. Similar identification interactions occur outside the linguistic domain. A phenomenon was investigated by Bertenthal et al. (1993) that when two small identical visual targets towards (and over) each other on a two-dimensional display, they targets could be perceived as either passing through or bouncing off each other. R. Sekuler et al. (1997) added a loud click at the time of coincidence of two visual targets and found a significant increase in the frequency of perception of the targets bouncing off each other. This phenomenon is now referred as bounce-inducing effect.

3.3 Studies of multimodal integration and crossmodal interaction and their implications for interaction design

There are numerous empirical studies of behavioural, psychophysical, and neurophysiological aspects of multimodal integration and crossmodal interaction have implications that are potentially very useful for the designers of multimodal interfaces and displays.

3.3.1 *Crossmodal enhancement*

In visual-auditory crossmodal interaction research, Storms (1998) found that high-quality auditory displays coupled with high-quality visual displays increase the perceived quality of the visual displays in virtual environment. When investigating the factors which affect the evaluation of audio-visual quality, Winkler and Faller (2005) found that both audio and video quality contribute significantly to the perceived audio-visual quality. Analogous results have been discovered for visual-haptic crossmodal interactions. Srinivasan et al. (1996) studied the impact of visually presented spatial cues on the perception of stiffness in virtual environments and found that graphically manipulated visual information gave rise to impactful haptic illusions about the perceived stiffness of objects in Virtual Environments. Extending this to audio-haptic interactions, DiFranco et al. (1997) also examined the influence of sound on the haptic perception of stiffness. Based on their psychophysical experimental results, they proposed the use of auditory cues to overcome limitations in the representation of rigidity with force feedback devices such as the PHANToM (DiFranco et al. 1997).

Implication 1: Such studies reveal that when visual-auditory, visual-haptic or auditory-haptic crossmodal interactions occur, perception through one or more sensory modalities could be enhanced. Whilst such potential benefits have been recognised in human-computer interaction (HCI), such effects have not been fully exploited and few interaction designers seek to deploy the recent findings of multimodal integration in a controlled manner. It has been already widely accepted in HCI that multimodal information presentation has a number of distinct advantages over unimodal presentation. Multimodal presentations can increase the information delivery bandwidth, heighten perceptual awareness, lower reaction time, or reduce cognitive load especially in multi-tasking situation. Some interfaces purposely present redundant multimodal information. However, there are a number of aspects of crossmodal interaction which could enhance multimodal interface design that have not to date been fully exploited.

3.3.2 *Multimodal interaction*

As already discussed R. Sekuler et al. (1997) examined, in a lab environment, bounce-inducing effect in which a brief simultaneous sound presented at the moment that two targets coincide could resolve perceived visual ambiguity. Watanabe and Shimojo (2001) elaborated on the nature of this phenomenon, the two visual targets need to be small and the motion needs to be smooth so that observers cannot use the other cues to interpret the visual event. Watanabe (2001) employed this phenomenon as the basic paradigm to investigate various factors involved in crossmodal perception, such as the temporal window, auditory context, modality combination and intensity. The temporal windows of the audio-visual and tactile-visual interactions were investigated by Watanabe (2001) who identified that R. Sekuler et al. (1997) crossmodal interaction results depend on the relative timing between auditory and visual stimuli. Watanabe (2001) found that the temporal windows during which a sensory transient influences visual motion perception are about 400-500 milliseconds (audio-visual interaction) and 700 milliseconds (visual-tactile interaction).

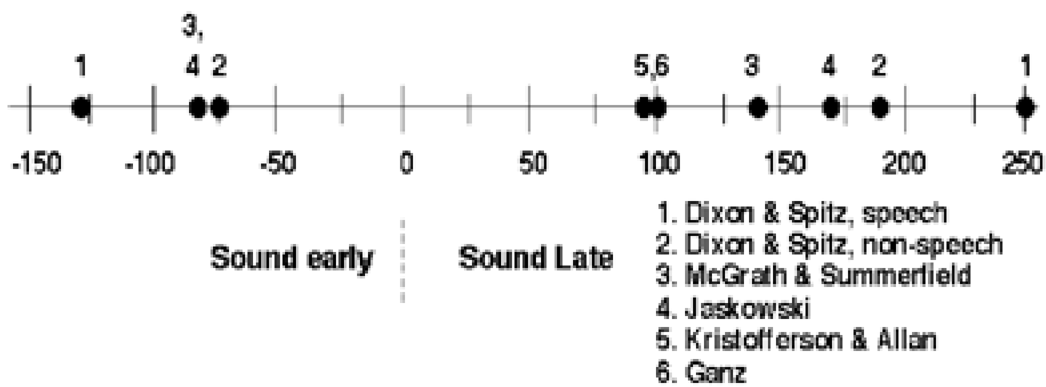

Figure 3.1 Estimates of the auditory-visual simultaneity threshold from six studies (Levitin et al. 2000)

Other researchers have found a range of temporal disparities within which humans are unable to tell that events from different modalities are asynchronous. Levitin et al. (2000) quantified the simultaneity thresholds (or temporal windows) in six (then) recent studies (Dixon & Spitz 1980; McGrath & Summerfield 1985; Jaskowski 1996; Allan & Kristofferson 1974; Ganz 1975) and arranged them in a time line (see Figure 3.1). They also reported a smaller simultaneity threshold (-41 and +45 milliseconds) for input from visual and auditory sensory modalities for crossmodal simultaneity tasks designed to have greater ecological validity.

The benefits of appropriate synchronization of sensory stimuli is further supported by the findings of Meredith et al. (1987) that auditory and visual signals arriving together at multisensory neurons, often interact to produce enhanced responses. In contrast, asynchronous multimodal information presentation can result in loss or distortion of information content, give an unnatural impression, cause communication disruption between a user and an interface, and undermine task performances.

Implication 2: The discovery of R. Sekuler et al. (1997) of the bounce-inducing effect suggests a possibility of using auditory or tactile displays to complement visual displays where visual elements of the display have some ambiguity. The results from the work of Watanabe and Shimojo (2001) imply that to achieve the desired percept, the parameters of single modality information needs to be carefully controlled (such as the size of visual objects). This can already be seen in informal recommendations as to the aspect of interface design, for example, that "content and physical properties such as level (intensity) or frequency characteristics of each item of single modality information have influence on the overall perception of the multimodal event." (Proctor & Vu 2005; page 151). The asynchrony between auditory display and visual display should not exceed the empirically established thresholds. Perceptual thresholds for auditory-haptic asynchrony are 50 milliseconds for audio lag and 25 milliseconds for audio lead (Altinsoy 2003). For auditory-whole body vibration asynchrony the perceptual thresholds are 39 milliseconds for audio lag and 35 milliseconds for audio lead (Altinsoy et al. 2002). Note that the values of these temporal windows depend on the modality combination, the modality to which attention is directed, the kind of stimuli used in the experiment, and the psychometric methods employed (Proctor & Vu 2005).

3.3.3 *Crossmodal spatial attention*

Proctor et al. (2005) reviewed a number of studies on the contribution of spatial attention to stimulus identification. Most studies have compared the difference in reaction times between the valid (when the cue and target are spatially collocated) and invalid cuing conditions (when the cue and target are not spatially collocated) and provide "a measure of the extent to which the presentation of stimuli in one sensory modality can direct, or capture spatial attention in another modality" (Proctor et al. 2005). Based on the studies on cueing effects, they concluded that audition and touch can more effectively direct visual attention, than vision can direct auditory or haptic attention. They also indicated that collocation is important in studies of crossmodal spatial attention, as many studies demonstrate that humans tend to respond quicker if information coming from more than one sensory modality is presented from approximately the same external location than presented at a different spatial location.

Implication 3: When multimodal information is presented, the spatial location of different sensory information or media should be carefully controlled. The spatial location of informative or uninformative cues and the upcoming target information should be as close together as possible so as to make users to response to the target information quicker and more accurate. The modality to which attention is directed and the modality used to direct attention should be carefully chosen and paired. This not only applies for single modalities used as alerts, but for other tasks and modality combinations. It is also suggested that audition and touch are more effective modalities for directing attention than vision (Proctor et al. 2005). The "modality appropriateness hypothesis" (Bertelson & Aschersleben 2003) also suggests that mapping auditory display to temporal tasks and visual display to spatial tasks would be more efficient and effective than vice versa. Based on 43 studies, Burke et al. (2006) evaluated the aggregate performance of visual (V), visual+audio (VA), and visual+tactile (VT) interfaces. Their results showed that visual-tactile feedback performed significantly better in alerts, warnings and interruptions.

3.3.4 *Semantic aspects of crossmodal interactions*

Stevens and Marks (1965) studied the impact of an informationally irrelevant sound on the perceived brightness of a light and found that responses to a dim light were faster and more accurate when it was accompanied by a soft or low pitched sound, and that responses to a bright light were be faster and more accurate when it was accompanied by a loud or high-pitched sound. This finding indicates that enhanced multimodal perception is contributed to by both low-level sensory correspondence and higher-level semantic link. (P Bertelson and De Gelder (2004) argued that semantic contributions from familiar bimodal contexts are not a necessary precondition for the generation of crossmodal interaction, but recognised that several empirical studies support the observation that top-down effects from familiar contexts can nevertheless enhance the demonstrated bottom-up effects of sensory factors.

Implication 4: Despite the debate on semantic contributions, current findings suggest that semantic links between different modalities have potentials for the improved design of multimodal interface. Even where user performance may not directly benefit from the effect of semantic factors, a semantically appropriate coordination of modalities may improve the user experience. For example, the semantic link between auditory and visual information may make the presented audio-visual information appear natural.

3.3.5 *Crossmodal illusions*

A final issue that we do not intend to explore in depth, is the potential for crossmodal illusions to be exploited by designers who want to induce particular cognitive effects in highly controlled environments such as virtual environment. For example, the sound-induced flash illusion (Shams et al. 2002) demonstrates that sound can alter the visual percept qualitatively even when there is no ambiguity in the visual stimulus.

Implication 5: This finding suggests that a single gunshot visual feedback accompanied by multiple gunshot sound might be perceived as multiple gunshots in a first-person shooting computer game. Thus the mechanism of such illusions should be exploited to reduce the computational costs where appropriate. In practice there is no clear example of cases where crossmodal illusions are exploited in multimodal interface design or multimodal information presentation.

3.4 Crossmodal Displays

3.4.1 *Towards a new class of displays*

In order to answer RQ2, one of the principal goals of this thesis is to understand how the implications of multimodal perception and crossmodal integration for interaction design (see Section 3.2) can be exploited to address the different categories of information presentation problem in physical public spaces characterised in Chapter 2. Crossmodal interaction effects, and the spatial location and timing factors upon which they depend, are the motivating phenomena for our interaction design proposals – these may be used to enhance or complement perceived information, or cause users to integrate information perceived via crossmodal perception into unified and meaningful new percepts (e.g. ventriloquism effect (Hairston et al. 2003)).

As suggested in Chapter 2 combining two types of displays to present information may produce complementary and/or supplementary usability and user experience, not achievable through the single type of display. One technical configuration might involve the presentation of multimodal information through a combination of a personal display, related to the body of a user, and a public display situated in the environment encompassing the user. For example, two of the interface designs we propose (see Chapter 4) are based on visual-vibrotactile information presented using a combination of the vibration motor of a mobile phone (in a user's pocket) and a large (public) visual projection.

There has been much recent interest and research as to how personal mobile displays and situated public displays can be combined (e.g. Prante et al. 2004, Eriksson et al. 2007), but most accounts focus on technical aspects of the design space and fail to recognise potential complications arising from the nature of multimodal perception and crossmodal integration. Little work so far has explored information presentation by utilizing the combination of public display and public display and exploiting crossmodal interaction effects to solve the problems identified in Chapter 2.

3.4.2 *A conceptual model of Crossmodal Display*

We define a *Crossmodal Display* as a set of public (e.g. large plasma screens or public projections) and personal displays (e.g. mobile phone or PDA) which exploit multimodal perception and crossmodal integration of stimuli that are distributed across the two classes of displays. In doing this, Crossmodal Displays mainly aim to

- prevent information overload;
- effectively satisfy different information needs of multiple simultaneous users who share a physical space or situated interface;
- preserve users' privacy and anonymity in physical public spaces;
- employ minimum sensing or tracking.

Through achieving the above, we hope Crossmodal Displays could address the shortcomings of traditional public displays, which to date have not been capable of supporting tasks which require both highly situated and personalized information provision (such as indoor navigation), through their integration with personal displays. At the same time, the divide between the developments in personal displays in the form of mobile devices and in situated public displays could also be bridged.

Figure 3.2 presents a diagrammatic representation of our conceptual model of a Crossmodal Display. In this model, the key elements are: public displays, personal displays and crossmodal cues which include public cues and private cues. As reviewed in Chapter 2, *public displays* are a class of display which is normally situated in public space and used separately or concurrently by multiple users and rendered using technologies such as projection, screens and speakers. Public displays generally present information publicly. Within our conceptual model, public displays are used to present *public cues* which are potentially perceivable to all occupants of a public space (including multiple simultaneous co-located users of the Crossmodal Display).

Personal displays are a class of display which normally worn or carried close to a user's body, on which information is presented only to the owner of the display. Personal displays allow the presentation of information privately to a user, as allowed by mobile phones and PDAs. Within our conceptual model, personal displays are used to present different *private cues* to different users or groups of users of Crossmodal Displays. We use the notion of a central server to describe the centralised process of coordinating multiple unrelated personal displays with one or more public displays so as to realise the mapping required between public and private cues. Our use of the term "cue" is borrowed from in psychology and is defined as "the part of any sensory pattern that is identified as the signal for a response" (Dictionary.com 2012). The design of public and private cues depends on the information presentation requirements and the context of use. In combination, such cues can either drive a user's attention to specific spatial locations, that is, *where* to attend; or drive temporal aspects of a user's attention, that is, *when* to attend. *Crossmodal Displays must use coordinated public and private cues to exploit their user's abilities to integrate these cues crossmodally*.

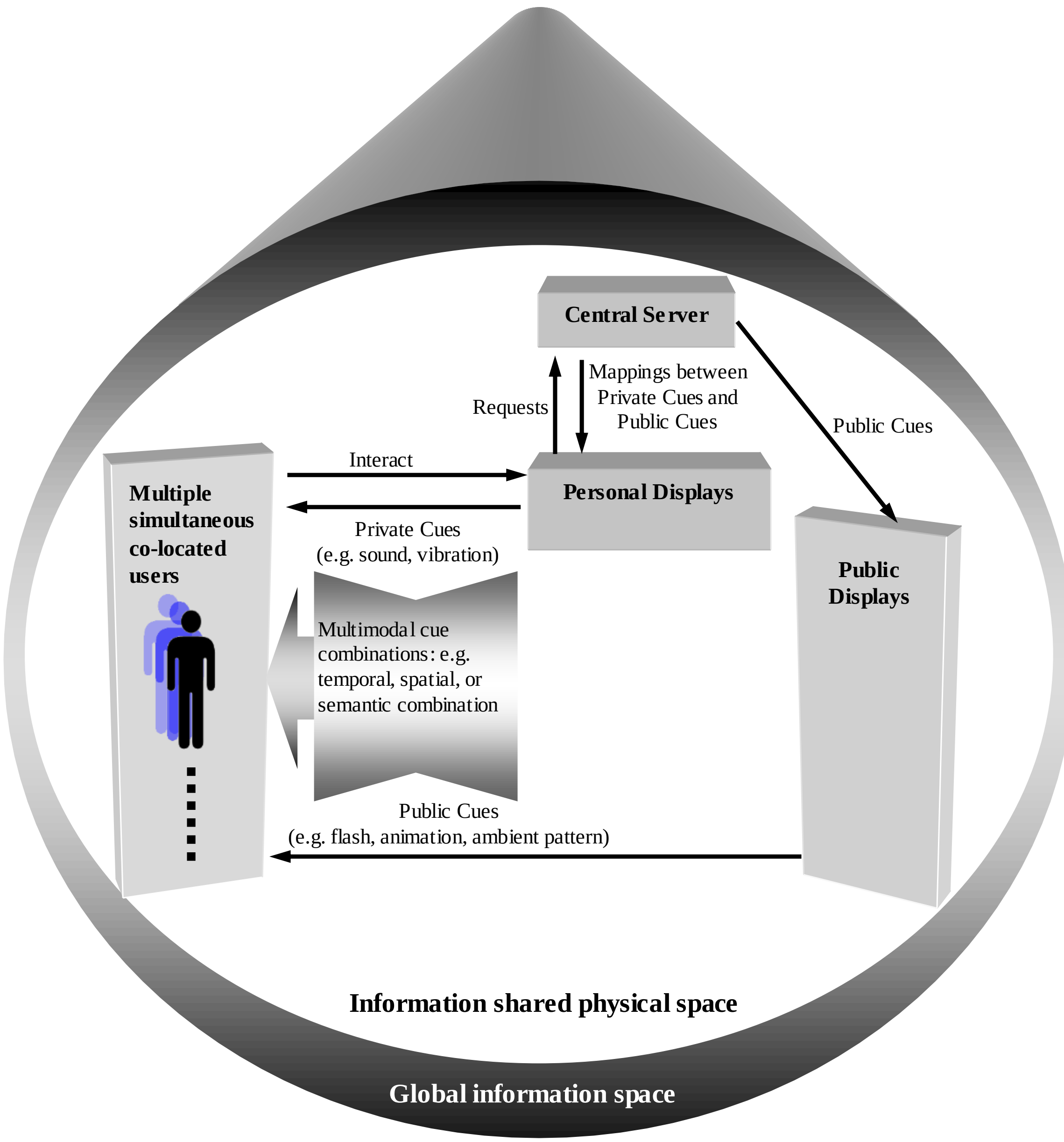

Figure 3.2 Diagrammatic representation of the conceptual model of a Crossmodal Display

In the displays that we propose in Chapter 4, 5, 6, 7 and 8, public cues are rendered as elements of the output of public displays (typically as a visual augmentation of the display's content) and are apparent to all occupants of the public space, including users and non-users of Crossmodal Display. Such public cues can be any stimulus that can be temporally sequenced, including, but not limited to, images, text, sound, and even vibration (although this is difficult to render publicly). For example, a public cue would include a flash of an LED light attached to an information board to indicate the spatial location of information that relates to a particular user or user group (or the elements of the information themselves might flash).

Private cues use a different modality (via a private sensory channel) to the public cues, and direct users' attention to the public cues relevant to them. In our example, the private cue must direct users' attention to the public cue relevant to them, that is, the particular flashing LED and associated location on the information board. This might be achieved through an auditory private cue so designed to be synchronised with the relevant element of the public cue (the flashing LED) and asynchronous with the other elements. Private cues are presented using a personal display and thus can be assumed to be received by an individual, or a specific group of users (who have the same information requirements) privately. As these are rendered on personal displays, users are allowed to control over the private cues, for example, they can turn these on or off, control their amplitude, or even choose between the modality to be used.

The selection and coordination of public and private cues are to some extent determined by the information content of the public display itself (that is, display's principal content). For example, dense information displays (mentioned in Chapter 8), where only small sub-elements of the display relevant to individual users or user groups are likely to use very different cues and cue combination to low-density information displays. Exploring this design space is a pressing but significant challenge.

3.5 Conclusions

This chapter has presented a review of the research findings of multimodal integration and crossmodal interaction relevant to the multimodal interface design, and we have outlined our notion of a Crossmodal Display with which we will seek to tackle the problems identified in Chapter 2. The review not only provides the basis for Crossmodal Display approach, but also suggests the factors which could be manipulated to design the multimodal cue combinations to achieve the desired usability and user experience goals for Crossmodal Displays.

The studies reviewed in this chapter (e.g. Spence & Driver 2004; Proctor et al. 2005) describe the perceptual and behavioural consequences of multimodal integration: decreasing sensory uncertainty (or disambiguation), decreasing reaction time, enhancing perception, increasing event detection, increasing event discrimination and illusion induction. These studies also reveal that when humans process multimodal information, one or more these factors may affect the perception: modalities; timing (or simultaneity threshold, temporal window) of multimodal stimuli, spatial locations (or spatial disparity) of multimodal stimuli, and semantic links of multimodal stimuli.

It has also been demonstrated that these factors have different significances in different contexts. For instance, Stein et al. (1996) found that the perceived luminance of an LED was significantly enhanced by a brief, broad-band auditory stimulus regardless of their relative spatial locations. Here audio-visual interaction enhances the perception of visual information, which depends on the timing of audio-visual stimuli rather than spatial disparity and semantic links. The perceptual and behavioural consequences of multimodal integration and factors involved in multimodal information processing identified in this chapter are of the upmost importance to the design of multimodal interfaces. Based on this knowledge, multimodal systems designers must take stock of, and manipulate, these factors in the fulfilment of their usability (e.g. efficiency, low cognitive load) and user experience goals (e.g. natural, aesthetically pleasing, fun) whilst avoiding the common pitfalls (e.g. conflicts of mental representation of information, mask of information, discomfort, attention overload).

Although we have at the point established our conceptual model of a Crossmodal Display, a number of issues remain to be explored. We must explore both how to design new displays, and the nature of the design space itself. We must also examine how our conceptual model can be deployed to solve the specific problems or facilitate specific tasks identified in Chapter 2 to answer RQ2. Finally, we need to explore the limits of the usability and the nature the user experience afforded by Crossmodal Displays to answer RQ3. These are the foci of Chapter 4, 5, 6, 7 and 8.

Chapter 4. CROSSFLOW: A Crossmodal Display for Indoor Navigation

4.1 Introduction

Having introduced the notion of a Crossmodal Display, the goal of this chapter is to operationalize our conceptual model in a real-world interaction design problem for public space. By doing so we hope to both demonstrate the sufficiency of Crossmodal Displays in providing personalised information in public display contexts, as well as documenting (and, in Chapter 5, evaluating) a concrete example of a Crossmodal Display. This chapter describes the motivation and design of CROSSFLOW, our first low-cost Crossmodal Display system that seeks to address the problem of providing multiple simultaneous users with personalized indoor navigation guidance through a low-cognitive load combination of public and personal displays.

The problems of negotiating large complex public spaces in modern society such as airports, hospitals and shopping malls has resulted in the indoor navigation problem becoming the prototypical application domain for ubiquitous computing and mobile interaction researchers, with typical approaches centre around the provision of expensive and novel spatial localisation technologies. Navigation or wayfinding in unfamiliar indoor environments become more of a challenge for people. Dogu and Erkip (2000) documented the many costs associated with having wayfinding problems such as loss of time, decreased safety, stress, and discomfort.

There are three aspects to the indoor navigation problem: the user, environment, and task. In our scenario we consider there to be multiple nomadic users, that is, users who have no previous knowledge of the spatial environment. The indoor environment we are concerned with will typically be large, complex and crowed spaces, which may have structural similarities to functionally related environments. Experience has shown that in many situations, users struggle to find/locate appropriate navigational information using traditional signage and maps (especially when placed under time pressure), which gives rise to the need to design more effective navigational aids. Even where navigational information is provided by an environment, it is not personalized, it is not efficient (e.g. suggest generally applicable paths) and it usually places a significant cognitive load on the user.

4.2 Existing Approaches of Supporting Indoor Navigation

Navigation in indoor environments is currently supported by the conventional methods (architectural design, signage, maps, and verbal and written instruction). In recent years a number of mobile and pervasive computing technologies have been proposed to support indoor and outdoor navigation (i.e. location-based guidance systems). In establishing the requirements for our first crossmodal prototype we must first review these existing approaches to the support of navigation.

4.2.1 *Maps, signs and architecture*

Research in architecture suggests that good architectural design gives rise to easy and error free navigation within a building (Dogu & Erkip 2000; Werner & Long 2003). But finding a particular destination can be difficult in modern complex buildings, where the corridors on different floors often look very much alike (Wright et al. 1993; Dogu & Erkip 2000). Traditional navigation aids, such as signs, and maps are enumerated by Arthur and Passini (1992) and are used by architectural designers as complement information to supporting wayfinding in addition to general building legibility. However, wayfinding signs have a number of well know limitations. Dogu and Erkip (2000) identify that although putting up signs is a universally acknowledged approach to prevent people from getting lost, the language or pictographs used in the signs are not always well understood. Locating wayfinding signs in itself can be difficult for people in modern commercial public spaces where there are high levels of visual noise, such as advertisements in the form of posters or digital displays. As a result traditionally static wayfinding signs are hard to distinguish from the environment.

Traditional maps also have a number of well understood limitations. Conventional handheld paper maps, as well as stationary embedded representations of the environment (e.g. poster maps or physical models), require users to identify their locations and their locations of their destinations in order to formulate navigation plans. Furthermore, people have significant (and well documented) problems understanding spatial layout and wayfinding performance decreases with increases in floor plan complexity (O’Neill 1991). Map also imposes heavy demands of a user's spatial cognition (including spatial memory). Navigators have to learn and memorize the sign and landmarks in the environment, on the map. Furthermore, paper maps (handheld or embedded) are essentially public information, and are not personalized to an individual user's requirements. To use a map a user has to actively undertake visual search to understand the correspondence between the representation and the environment. Of course, when people have difficulty to understand the spatial information provided by sign or map, they often use social navigation and simply ask for directions.

4.2.2 *Mobile and pervasive guidance systems*

The provision navigational support (and other location-based services) for visitors to an unfamiliar indoor space is a prototypical application for research in mobile and pervasive computing. Solutions generally seek to present maps and spatial information using multiple modalities, and in a manner sensitive to a user's needs and spatial context. However, such systems generally suffer one or more the following shortcomings:

- **Attention and cognitive overload problems**: presenting information on a personal display requires users to divide their attention between the displays and the environment.
- **Sensing and tracking technology problems**: indoor tracking technologies are both expensive and unreliable.
- **Multi-tasking problem**: Few systems are designed to support users undertaking multiple concurrent tasks while navigating.
- **Multiple simultaneous users problem**: personalised navigational information is difficult to present using traditional public displays.
- **Privacy problem**: presenting dynamic guidance information through publically is likely to violate an individual's privacy.

Many map-based mobile guidance systems, as reviewed by (Baus et al. 2005), require users to attend to either a personal display (typically a small screen on a mobile device) or to a set of verbal instructions. When using an electronic map on a personal display to navigate, a user must continuously verify both his/her current location and destination both on the electronic map and in the physical environment time to time. A significant cognitive load could be placed on the user during navigation.

Providing automated guidance requires the capability to track the location of a user and such location tracking is a standard component of nearly all ubiquitous computing designs that aim to provide spatially contextualised information. Most of navigation guidance systems rely on a range of sensing and tracking technologies to determine the location, including the usage of WLAN, Bluetooth, infrared beacons, GSM, camera and GPS (Global Positioning System) (Ciavarella & Patern 2004; Wan Bejuri et al. 2011). Wan Bejuri et al. (2011) reviewed the literature and recent development about location determination on mobile navigation systems. In particular, he categorized the location estimation methods used for predicting locations into three categories, i.e. radio frequency positioning based, camera navigation based and GPS based. As Liu et al. (2007) indicated, a radio frequency positioning system is composed with at least a signal transmitter and a measuring unit, two separate hardware component, which means the system has to utilize sensing and measuring to estimate location. Location estimation for camera navigation relies on cameras that must be able to capture image information, detect landmarks or observe environments. Standalone embedded GPS on mobile phone is used in location determination by integrating with other device or sensor in order to make it survive in unstructured environments, and such technique is called as map matching algorithm (Wan Bejuri et al. 2011).

In order to provide users with personalized navigational information, the continuous localization of each of users (and usually a wireless connection between mobile devices and servers) is required. Many navigational systems that rely on localisation therefore require an expensive (and often relatively complex) infrastructure. For example, the radio frequency positioning such as WLAN, Bluetooth and GSM, one of the two types of positioning system architecture, i.e. infrastructure-based systems have to rely on a pre-installed powered infrastructure of base stations (Wan Bejuri et al. 2011).

Furthermore, in real-world navigation tasks, people often have multiple tasks to undertake whilst they are navigating (e.g. taking care of children, transporting luggage, making a phone call, or simply just thinking about a problem as they move from one location to another). Few existing navigation systems are explicitly designed to support multiple-task handling during navigation. Moreover, very few navigation guidance systems consider how to use public displays to provide personalized navigational information to more than one user. For example, the GAUDI system (Kray et al. 2005) supported a user navigating indoor spaces by showing arrows on displays spread throughout a building. These arrows point to the target location that the user has previously selected using a display located near the entrance of a building. Although this work suggests that public display and personal display can be combined to provide navigational information, it only supported only one user at a time. Furthermore, as for all public display systems, GAUDI raises questions as to the user's privacy (i.e. personal directions displayed publicly).

4.3 Designing a Crossmodal Display for Indoor Navigation

Although multiple modalities have been adopted in many navigation systems use of coordinated multimodal cue combinations to present personalized navigational information has not been systematically explored. Sarter (2006) indicated that many works involving multimodal system outputs, including navigation systems, were dissociated from a knowledge of how the human perceptual system works and do not consider the existing knowledge base on the neurophysiology of crossmodal information processing and its implications on for human performance (e.g. Driver & Spence 2004). Drawing upon our analysis of existing navigation approaches, systems and technologies (in Section 4.2) and the implications of multimodal perception and crossmodal integration for interface design (in Section 3.2) our next step is to design a Crossmodal Display that addresses these concerns.

4.3.1 *Goals for CROSSFLOW*

CROSSFLOW has been designed as a navigation aid to facilitate indoor navigation of multiple simultaneous users within an unfamiliar complex indoor environment. Within the scope of the main aims of Crossmodal Displays, we set the following design goals for CROSSFLOW on the basis of the literature review of solutions in Chapter 2, the implications of multimodal integration and crossmodal interaction for interface design in Section 3.2, and the discussion of navigation aids in Section 4.2:

- Facilitate indoor navigation;
- Impose lower mental workload demands on users than map-based navigation aids;
- Support multiple simultaneous users accessing different navigational information without resorting to location tracking.

4.3.2 *CROSSFLOW: conceptual design*

Our conceptual design for CROSSFLOW has three sources of inspiration: (1) research findings relating to human navigation and wayfinding; (2) an analysis of a simple navigation scenario in a public space; and (3) our conceptual model of Crossmodal Displays (Section 3.3).

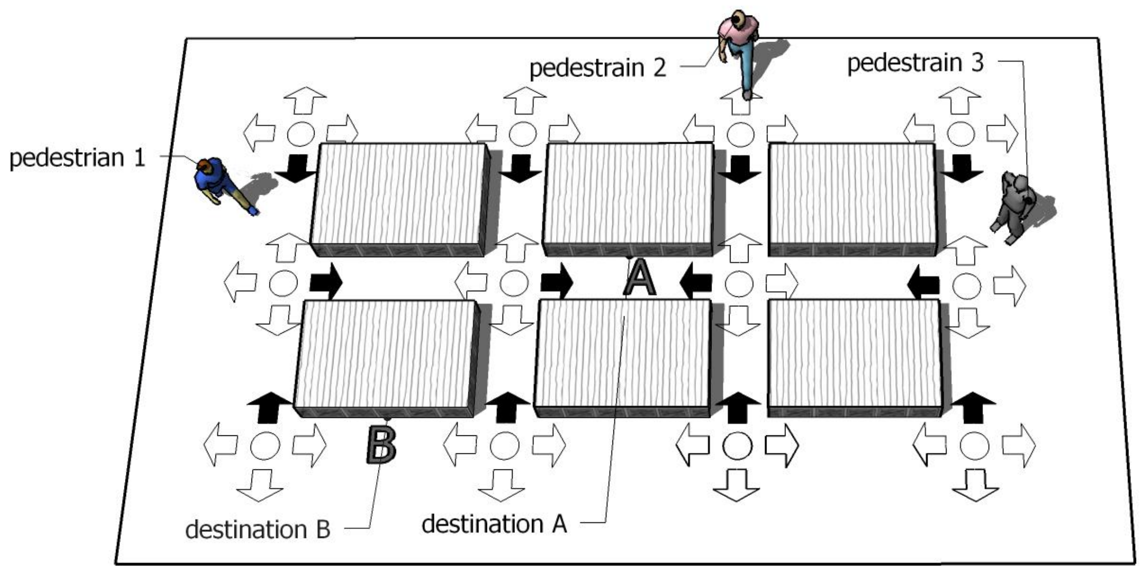

Figure 4.1 Multiple simultaneous pedestrians navigating within a physical space, with directional information corresponding to destination A.

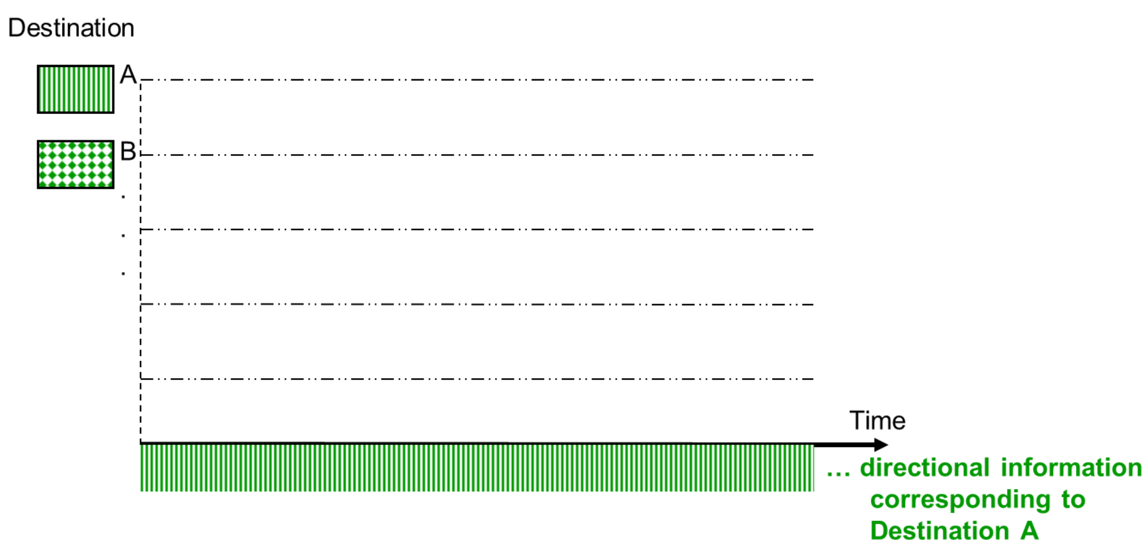

Figure 4.2 Continuous public display of directional information to destination A.

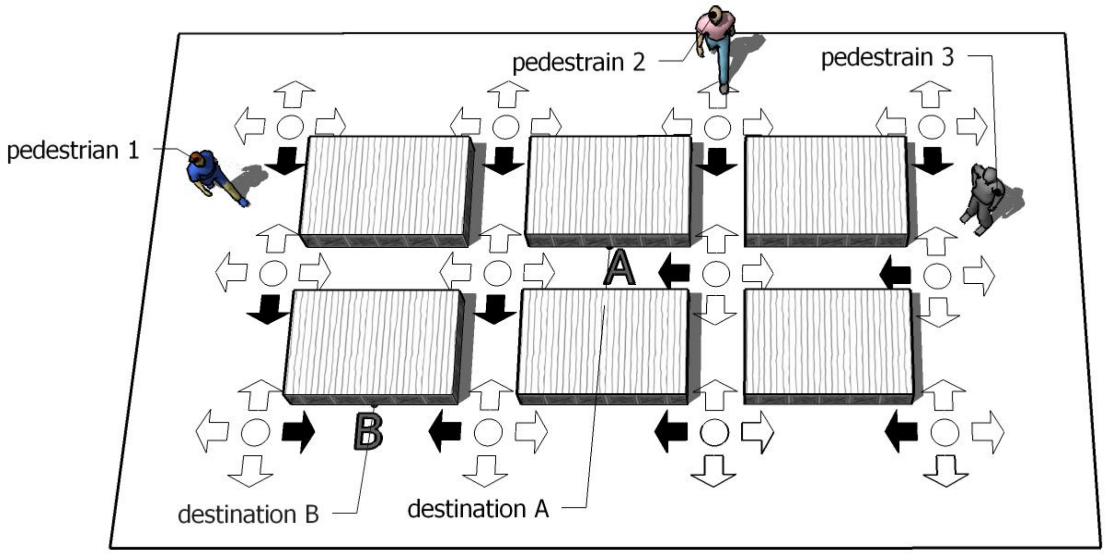

Figure 4.3 Multiple simultaneous pedestrians navigating within a physical space, with directional information corresponding to destination B.

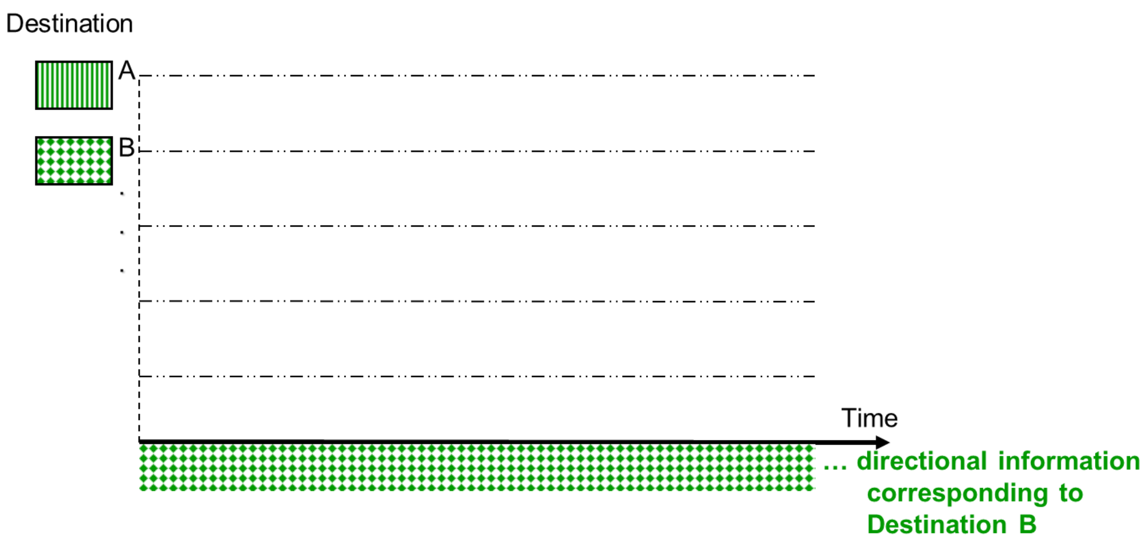

Figure 4.4 Continuous public display of directional information to destination B.

A. Human navigation and wayfinding

Wayfinding tasks can be categorized according to their functional goals as: (a) travel with the goal of reaching a familiar destination; (b) exploratory travel with the goal of returning to a familiar point of origin; and (c) travel with the goal of reaching a novel destination (Allen 1999). Our primary interest is in the task of finding one's way to a novel destination in an unfamiliar public space. In such situations, Raubal (2001) indicated that people have to rely on symbolic spatial information communicated to them through the environment. Baskaya et al. (2004) suggested that wayfinding decisions must be based on environmental information that is readily accessible so that the indoor visitors can proceed from decision point to decision point, which means a wayfinding or a navigation procedure follows the sequence: Start location ⇒ Decision point 1 ⇒ … ⇒ Decision point n ⇒ Destination. These suggest that an indoor navigation system should provide symbolic spatial information at each decision point for indoor visitors to support their navigation.

B. An indoor navigation scenario

One form of direct symbolic spatial information could be explicit directions, presented as arrows, indicating the directions to decision points or destinations. Figure 4.1 or 4.3 illustrates such an approach for an indoor navigation scenario. As shown in figure 4.1, in order to guide both pedestrian P1 and P2 to their destination A, public displays (see Chapter 2 and 3 for the definition) can be utilized to show directional information corresponding to destination A (black arrows) at all decision points in the environment all the time. Similarly, in figure 4.3, the public displays show directions to destination B (black arrows) at all decision points all the time. The relationship between time and directional information corresponding to a destination is shown in figure 4.2 or 4.4, in which directional information corresponding to destination A and B is represented as a stripe-like pattern and a spot-like pattern respectively. The scenario here is similar to an emergency evacuation scenario where all occupants of a building in a flat are guided to an emergency exit. However, such navigational information presentation is not appropriate for multiple simultaneous navigators whose destinations are different. For example, pedestrian P1 and P2 are looking for destination A while P3 stands at a decision point and requires directional information corresponding to destination B. If public displays only show the directions at all decision points to one destination all the time instead of personalized directional information corresponding to different destinations, they would fail to support multiple simultaneous navigators to reach their destinations. One way to solve this problem is to temporally multiplex the directional information corresponding to different destinations. The details are discussed as follows.

C. Proposal of Crossmodal Displays

In our proposal relating to Crossmodal Displays (Section 3.3), the public cues are apparent to all occupants of the public space, including users and non-users of a Crossmodal Display. By contrast, the private cues use a private sensory channel such as the vibrotactile or auditory channel, or a combination thereof, and these direct users' attention to one of the public cues that they index. Extending this to our application scenario, CROSSFLOW could show multiplexed visual directional information (public cues) publicly corresponding to all the destinations at each decision point using public displays at these locations, and use buzzes and/or beeps (private cues) via visitors' personal displays to indicate the particular directional information that corresponds to their destinations.

Here, visual public information and auditory/vibrotactile private cues are used because CROSSFLOW is designed for multiple simultaneous users. For this reason disturbance to other occupants of the public space, including those who not using the system, should be kept to a minimum. Private cues presented through the vibrotactile or auditory (through headphones) modalities are essentially private to the users. A buzz or beep as a private cue can resolve public cues, the ambiguity designed into the public display. This way of using auditory or tactile cues to resolve ambiguity in a visual display is closely related to the observations and experimental design of R. Sekuler et al. (1997) and the bounce-inducing effect discussed in Section 3.2.

The implications for multimodal interface design that we enumerated in Section 3.2 indicate that when the spatially incongruent multimodal cues from the spatial incongruent displays are temporally combined and presented in parallel, a user may be able to integrate the spatially incongruent but temporally synchronized multimodal information into a meaningful new percept, i.e. an intuitive directional information indication relevant to their destinations. Moreover, the perception of visual public information may be enhanced by a simultaneous auditory and/or vibrotactile private cue regardless of the spatial disparity between the cues. Therefore, instead of having to comprehend the configuration of a public space (as when using a map), a visitor to a public space can enter his destination on his personal device, and in response, his device receives a time slot schedule (cue mapping) which defines the time slot when the visitor's directional information corresponding to his destination would be shown in the space. Whenever the directional information is "valid", the personal device indicates this using the private cue.

4.3.3 *Cue mapping design*

We define that cue mapping as the spatio-temporal relation between the public cues and private cues. The design of cue mapping can be motivated by the different paradigms and techniques for the integration of information across the sensory modalities which have been reviewed in Chapter 3. Based on our conceptual design for CROSSFLOW, two types of cue mappings have been created, a *regular cue mapping* and a *random cue mapping*.

A. Regular cue mapping

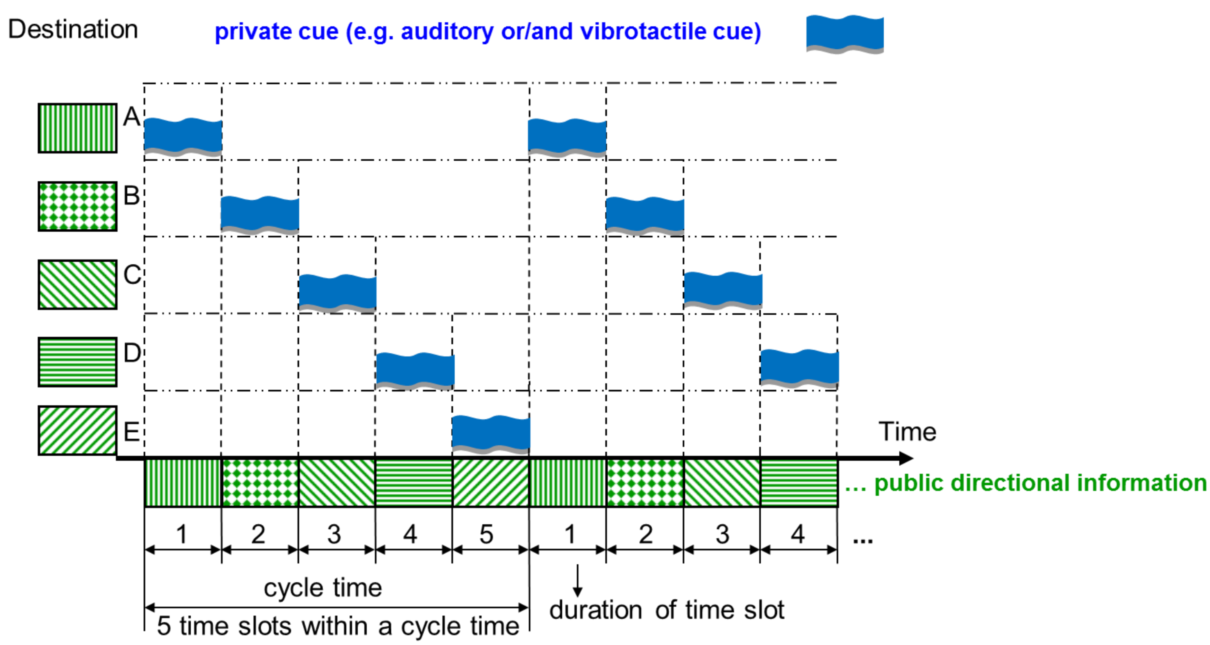

Figure 4.5 Regular cue mapping for the five-destination scenario.

Consider the problem of supporting multiple simultaneous indoor visitors navigating to five different destinations (A, B, C, D and E). At each decision point in the environment, directional information corresponding to five different destinations (i.e. five different public cues) can be displayed in a loop (fixed time cycle and order). A private cue synchronized with one of the public cues of this cycle indicates to a user which directional information relates to his/her destination. This is illustrated in figure 4.5, in time slot 1, the directional information corresponding to destination A is shown; in time slot 2, the directional information corresponding to destination B is shown, and so on. After the directional information corresponding to all destinations is shown, the cycle is repeated. A CROSSFLOW user simply needs to be provided with a mechanism to identify which time-slot in the cycle the directional information being shown corresponds to his/her destination. This is achieved using private cues (e.g. a short audio beep or vibration) presented on his personal display. Because the public cues and the private cues are combined with a fixed period and order, we define such a cue mapping is a *regular cue mapping*.

How many private cues should be presented within a cycle may depend on the navigation context, for example, a private cue may be shown to a user in two different time-slots within a cycle indicating that two public cues are relevant to him/her.

B. Random cue mapping

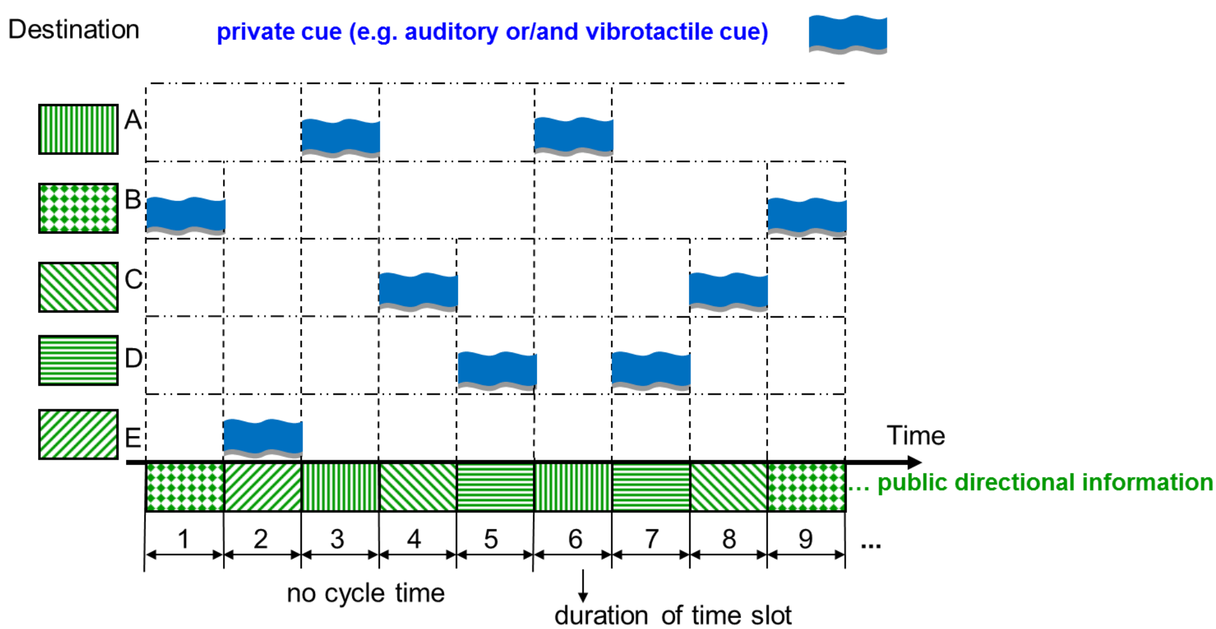

Figure 4.6 Random cue mapping for the five-destination scenario.

As shown in figure 4.6, in a random cue mapping, the public cues are not presented in a fixed time slot and order in a loop, that is, a public cue and the corresponding private cue synchronized with it are presented in a randomly selected time-slot, though within the time of two cycles between sequential presentations. Such presentation uncertainty can have both positive and negative impact on users.

The potential positive effects of a *random cue mapping* (as compared to a *regular cue mapping*) are:

- **Reduced cognitive load:** In the case of a regular cue mapping, once a user learns the mapping, he has a tendency to anticipate the appearance of a private cue in the cycle and the successive cycles, which is likely to increase cognitive load.
- **Increased continuity:** When destinations are added or removed from the destination list dynamically in the CROSSFLOW system, users are less likely to be aware of overall changes in the cues (where new cues can be unrelated to their navigational task).

Alternatively, the potential negative effects are:

- **Increased crossmodal attentional demands:** Because the timing of the presentation of the public cue and the corresponding private cue for a user is uncertain, a user cannot learn or remember the cue mapping, thus (as is intended) he/she has to rely on low-level crossmodal attention to capture his cue and then identify the directional information relevant to his/her destination.
- **Unexpected long waiting times:** Because of the non-deterministic scheduling of the cue, a user may have to wait an inconveniently long period of time for the private cue for him/her to appear at first time (though within the time of one cycle) or reappear (though within the time of two cycles).

As we are interested in indoor environments in which the number of destinations is fixed and users can learn and remember cue mappings easily, we therefore concentrate on regular cue mapping only.

4.3.4 *Public directional information (public cue) design*

Based on the regular cue mapping, the public display must provide CROSSFLOW users with visual directional information corresponding to all possible destinations in a navigational environment publicly. Methods of representing directions can take many forms. During the design of CROSSFLOW, several distinct visual design alternatives of directional information were explored including the rotating arrows shown at all decision points, highlighted markers showing all possible routes, and abstract ambient patterns covering the complete floor of the indoor environment. The detailed designs and informal evaluation results are described and discussed in the following subsections before we make our decision which design to study more rigorously.

A. Design I: Pointers

We can identify two features of traditional directional signage for public space: (1) the explicit representation of directions; (2) that directional signs are typically sited nearby or at decision points for the user (see Figure 4.7 for an example).

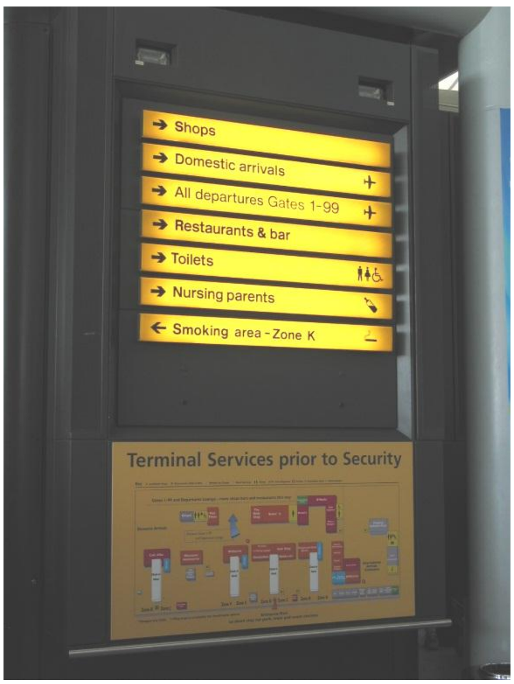

Figure 4.7 A sign used at a junction in an airport of France.

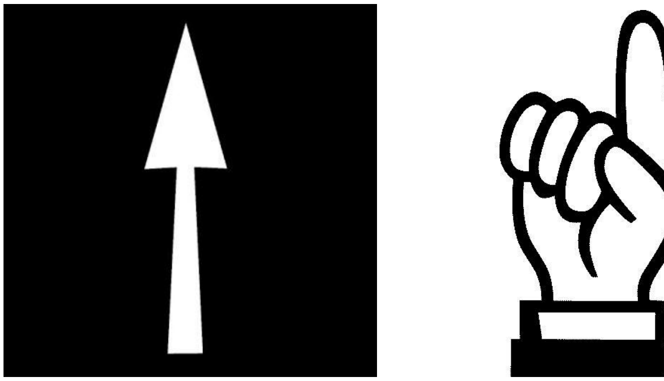

Figure 4.8 Design I: A big arrow and a big pointing finger

Figure 4.9 Projected big arrows and big pointing fingers

Figures 4.8 shows two designs, adapted for CROSSFLOW, which incorporate these features. When deployed, it is anticipated that they might be displayed at each decision point in a public space of a building and configured to indicate the direction of destinations using an animation cycle.

The size and spatial density of the pointers should have a significant impact on users' access and identification of the directional information being provided. The appropriate spatial density of symbols depends on the spatial character of the environment to be navigated. CROSSFLOW users will require a density of pointers that is sensitive to environmental factors such number of obstacles and the distances between them (including occupants of the environment). Furthermore, large numbers of densely distributed large graphical symbols would appear intrusive in a public space. The *Pointers* were informally tested in the atrium of a building as shown in figure 4.9.

B. Design II: Floor lights

Figure 4.10 Floor lights for inspiration of Design II.

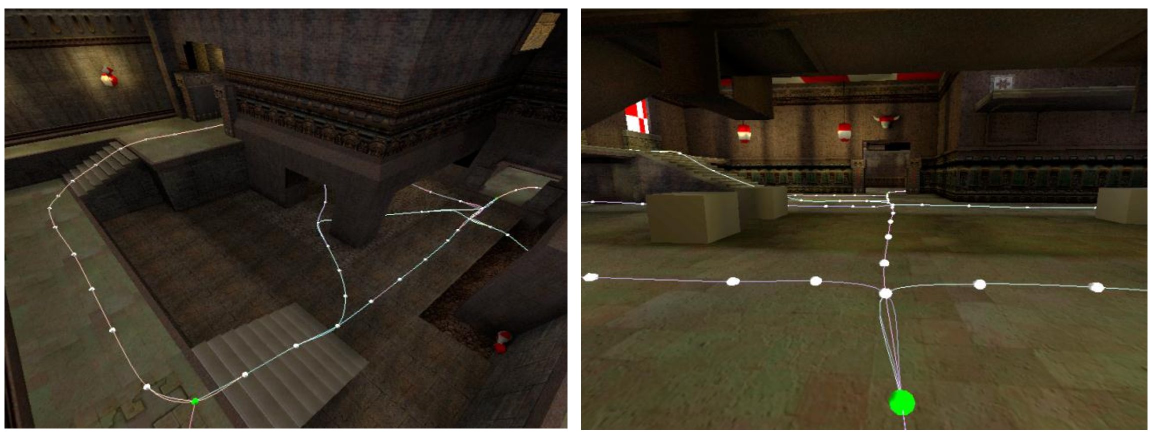

Figure 4.11 Design II in a 3D game environment.

Wayfinding was characterized by Klippel et al. (2004) as following a route segment up to a decision point, making a directional choice, following the next route segment up to the next decision point, making a directional choice, and so on. This suggests that both routes and directions at decision points towards destinations might be addressed in CROSSFLOW to support users in a wayfinding procedure. This design was inspired by instances of floor lights widely used in modern urban cities (see Figure 4.10). A type of graphical pattern was designed in a Quake Arena III model, a 3D game environment, in which, large amount of ball-like markers were densely distributed at decision points and along all route segments. In this design, the ball-like markers are highlighted by their colour, one by one, to indicate both routes and directions to all the possible destinations in the 3D environment. As shown in figure 4.11, the highlighted marker is green, while others are white. Other parameters including the size of the marker, the time duration of highlighting and the quantity and density of the markers on the route segments can be varied in this design.

The informal user feedback showed that the directional information delivered by such type of graphical pattern was not easy to be identified. For example, in some places obstacles such as columns blocked the users' perspective on the markers in an open area.

C. Design III: Ambient patterns

Designs I (Pointers) and II (Floor lights) and the corresponding informal feedback suggested that users' access and identification of directional information corresponding to a location or object is not very efficient when the graphical elements of the public cues are only displayed at decision points and on route segments. As already discussed, a dense spatial distribution of symbols, such as pointers in an environment, is also somewhat intrusive; however, this does not preclude the option of using a more aesthetically appropriate ambient representation of directional information. If such graphical elements were densely presented (in a non-intrusive manner) and distributed in large areas such as the floors, walls, ceilings, even the whole architectural space, CROSSFLOW users' navigation or wayfinding tasks may be significantly facilitated. In such a design, the whole architectural space becomes the interface between CROSSFLOW users and the digital navigational information. This is wholly consistent with the notion of an Ambient Display (Wisneski et al. 1998) as discussed in Chapter 2 that the architectural space becomes the interface between users and digital information.

The design of the ambient patterns, must be aesthetically appropriate, non-intrusive, and yet capable of expressing time varying directional information (as CROSSFLOW requires). Two such kinds of patterns were initially considered: an animated "fish-flow" pattern and an "arrow-like" pattern (see Figures 4.12 and 4.13). Both kinds of patterns are composed of a large amount of evenly-distributed graphical elements. Figure 4.12 and 4.13 show close-ups of the projection of each kind of pattern on the wooden floor of our test building (the open atrium space). The projected individual graphical element of the design (an individual fish or arrow) is approximately the size of an adult's hand (0.1 meter long and 0.1 wide) and both kinds of patterns are given a visual intensity that integrates with the floor. In the case of the "fish-flow" pattern this provides the impression of a "sparkling carpet".

The ambient patterns furnish navigational information corresponding to different destinations in a space as follows: both the fish-flow pattern and the arrow-like pattern orient their graphical elements towards different destinations, and in the case of the fish-flow pattern, move in this direction. For example, in Experiment 1 (see Chapter 5), when the fish-flow pattern shows directional information corresponding to 5 different spatial locations, the positions and the directions of the flow of the graphical elements in the pattern change 5 times within 4 seconds and once per 800 milliseconds.

Based on the regular cue mapping shown in figure 4.5, the graphical elements cycle through the schedule of directions and the collective movement is coordinated with an auditory and/or vibrotactile private cue. That is, in the first time-slot, the graphical elements (e.g. the "fish") "swim" or "point" towards destination A from every other location in the navigational environment; in the second time-slot, the elements undergo an animated transition and configure themselves in relation to the next destination B, and so on. After all directional information is shown, the cycle is repeated.

The graphical elements in the fish-flow pattern can be configured as streams to form paths (around any attendant obstacles) towards different destinations. The "head" of the "fish" can be configured toward the direction that users should move to. We expect that a user can reach his/her destination by following either one or multiple flow streams projected in the space. The way that the arrow-like pattern represents directional information of different destinations is similar to the fish-flow pattern. The main difference is that during a time slot the entire arrow set points along the designed paths to a particular destination. That is, each graphical element, a small arrow, points along a specific direction to the destination at a specific angle; the degree of the angle depends on the position of that arrow projected into the physical space. We expect that a user can reach his/her destination by following the local pointing directions of either one arrow or the collection of arrows.

Figure 4.12 Design III: A "fish-flow" pattern screenshot and its projection on a wooden floor in the test environment.

Figure 4.13 Design III: An "arrow-like" pattern screenshot and its projection on a wooden floor in the test environment.

The informal user feedback showed that the navigational information corresponding to different destinations in a physical space shown by the arrow-like pattern could not be identified as easily as the navigational information shown by the fish-flow pattern. Furthermore, the fish-flow pattern appealed to users as significantly more aesthetically pleasing. In the next stage of our research, we sought to explore the usability goals and the nature of the user experience of CROSSFLOW, and for these inquiries we selected the fish-flow pattern of Design III (reported in Chapter 5).

4.3.5 *Private cue design*

For a regular cue mapping, the private cue should induce/invoke a subtle switch of a CROSSFLOW user's attention to the corresponding public directional information that is (concurrently) displayed. Private cues in CROSSFLOW could be abstract signals or explicit information issued by a personal display such as a mobile phone or PDA, for example, a vibration coordinated with the onset of a time slot; or, an audible high pitch sound (i.e. a "beep") coordinated with the onset of a time slot.

For our initial design (see Experiment 1 in Chapter 5) we explored the combination of sound and vibration, displayed using an Orange SPVE200 smartphone, having not formally evaluated different combinations and configurations of cues yet.

4.3.6 *Duration of cue and time slot*

In a regular cue mapping, the duration of all time slots is equal (e.g. see Figure 4.5). For a private cue that includes a vibration the range of this duration (from 0.1 to 2.0 seconds) can be determined according to the usable range of vibrotactile stimuli in tactile interfaces (Geldard 1960; L. M. Brown 2007, p.14). As Gunther and O'Modhrain (2003) suggested, vibrations in tactile interface that are less than 0.1 second can provoke feelings of nudges or pokes, which may be undesirable for CROSSFLOW users; vibration of over two seconds would probably lead to unacceptable low efficiency of the CROSSFLOW system.

The value of duration was determined in an informal pilot study. A set of *influential arrows*, the design tool used to specify the fish-flow directions in the fish-flow pattern, were cyclically presented on a desktop PC monitor. As shown in figure 4.14, every 800 milliseconds the location that the arrows oriented themselves towards changes. An audio beep and a vibration were separately synchronized with one of the time slots and presented to the participants using a SPV E200 smartphone. For a configuration of an 800ms time slot, a vibration of 800ms (see Figure 4.15), and an auditory beep sound of 150ms (see Figure 4.16), the participants reported a convincing synchronization of the arrows pointing to a location and the audio beep or vibration displayed within the same time slot – they could quickly and accurately identify the corresponding target location among the many alternatives. This configuration was therefore used in the formal user studies of CROSSFLOW Version 1 (reported in Chapter 5).

Figure 4.14 An informal pilot study of CROSSFLOW using a desktop PC monitor as a public display and a SPV E200 smartphone as a personal display.

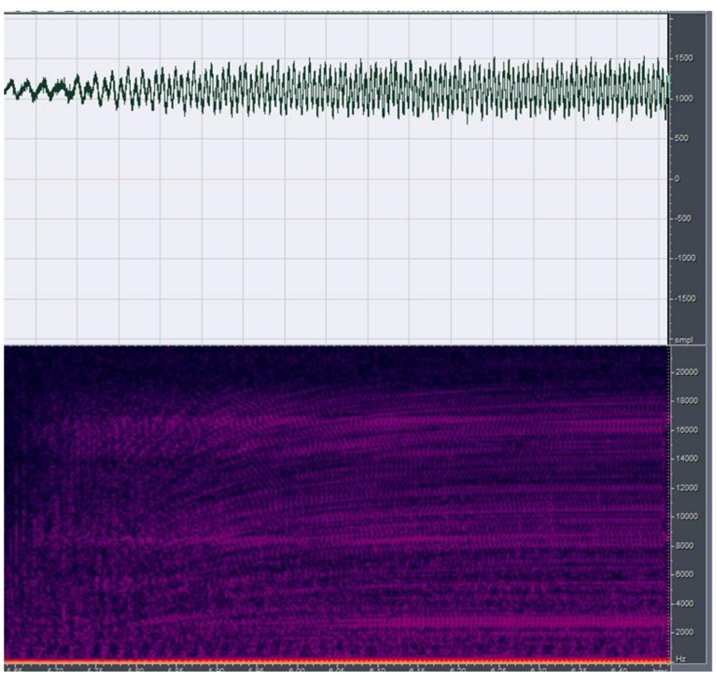

Figure 4.15 Waveform and spectral view of the output vibration from a SPV E200 smartphone (800ms length)

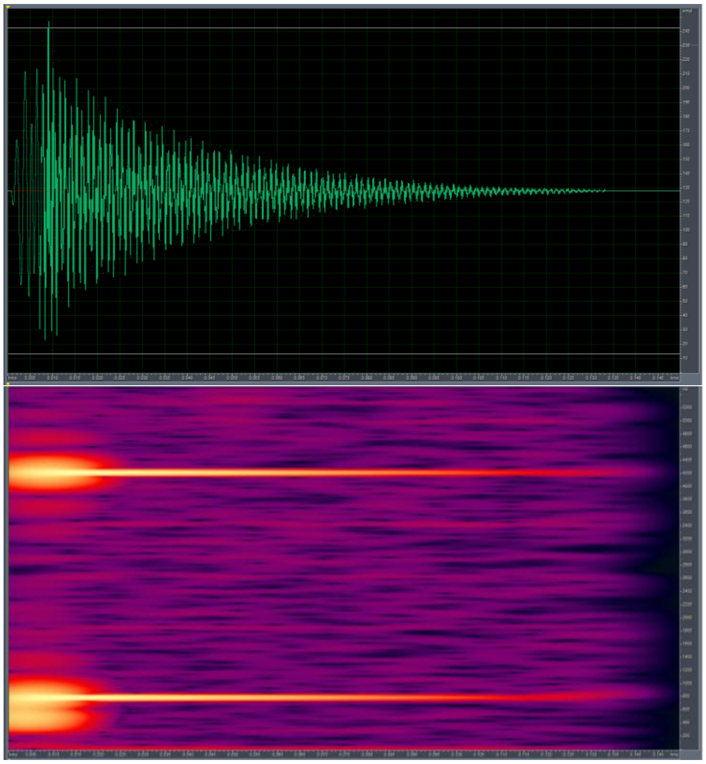

Figure 4.16 Waveform and spectral view of the output beep sound from a SPV E200 smartphone (150ms length)

4.4 CROSSFLOW V1 Prototype System Description

A prototype of CROSSFLOW indoor navigation system (CROSSFLOW Version 1, Figure 4.17 shows the schematic diagram of the system) was developed which embodied the conceptual design of CROSSFLOW by combining and integrating a regular cue mapping, visual public directional information and audio/vibrotactile private cues (see Figure 4.15 and 4.16). For brevity, in this chapter, we refer to “CROSSFLOW V1” simply as “CROSSFLOW”.

4.4.1 *System components and functions*

A. Public displays

The public display used in CROSSFLOW is a projection displayed by digital projectors connected to a networked PC server. Several different designs for the public directions (including the fish-flow pattern and the arrow-like pattern) can be displayed on the floor of an indoor environment.

B. Personal displays

CROSSFLOW uses Orange SPVE200 smartphones (including earphones) as personal displays through which sound and vibration can be displayed privately.

C. Configuration tool for visual public directional information

CROSSFLOW includes a software tool, as a component of the central server, to support the design and configuration of the visual public directional information. Figure 4.18 and figure 4.19 show a set of influential arrows which are manipulated by a designer specifying the flow directions towards the location of a white disc placed on the floor in a building. The software tool enables a designer to graphically and interactively configure the directions of graphical elements of visual public directional information corresponding to different destinations (see Figure 4.20) and specify the parameters that control the size, quantity, density, distribution areas (see Figure 4.21) and dynamic properties of the individual graphical elements (rate of movement and visual persistence), as well as the duration and number of time slot within a cycle. The tool in turn generates a configuration file saved on the central server. For example, when configuring the fish-flow pattern using this tool, a designer can steer the flow tendencies around obstacles and away from sites that are not intended to lie on the path to a destination through rotating a set of influential arrows (see Figure 4.22), and can also add or scale the influential arrows to attain the desired patterns of flow (see Figure 4.23).

D. Private cue configuration tool (auditory and vibrotactile)

CROSSFLOW also includes a software component installed on the smartphone, with which a user can configure the private cue. For example, the user can select a destination from a destination list, change the mode of private cue (vibration, sound, both or none) and the sound volume, and switch on/off private cues.

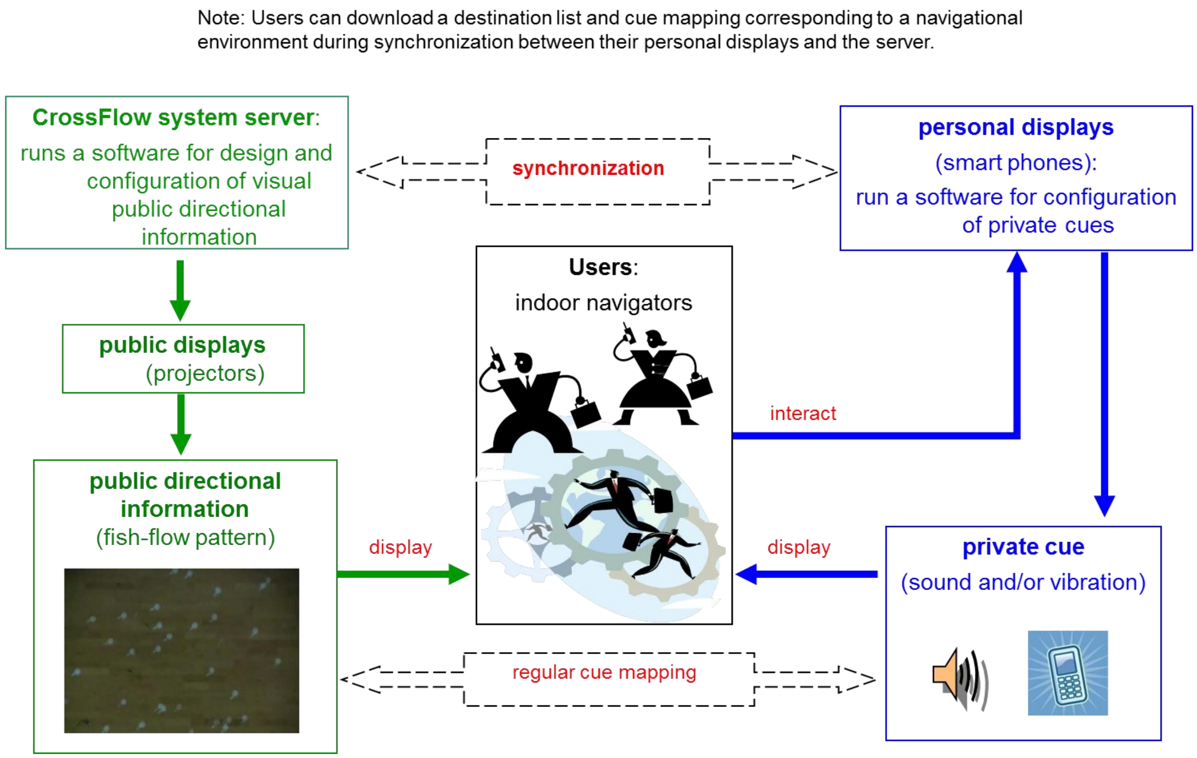

Figure 4.17 Schematic diagram of CROSSFLOW Version 1.

4.4.2 *How a user navigates using CROSSFLOW*

A user simply selects a destination in the destination list on the SPV E200 smartphone by selecting the numerical identifier of the destination. The smartphone may either be pre-configured to synchronise with the CROSSFLOW central server, or it may communicate with the server via a wireless connection to receive the cue mapping and to synchronize the timing of private cues (sound or/and vibration). There is no need for the user's personal mobile device to communicate with the central server once this synchronisation has been achieved.

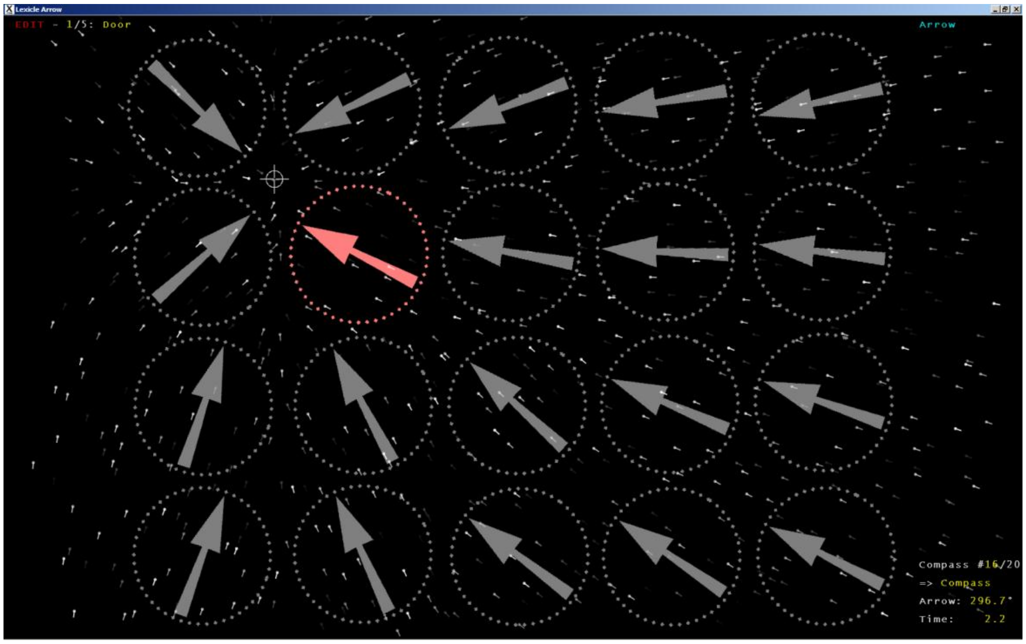

Figure 4.18 Influential arrows for configuring the fish-flow pattern on a desktop display

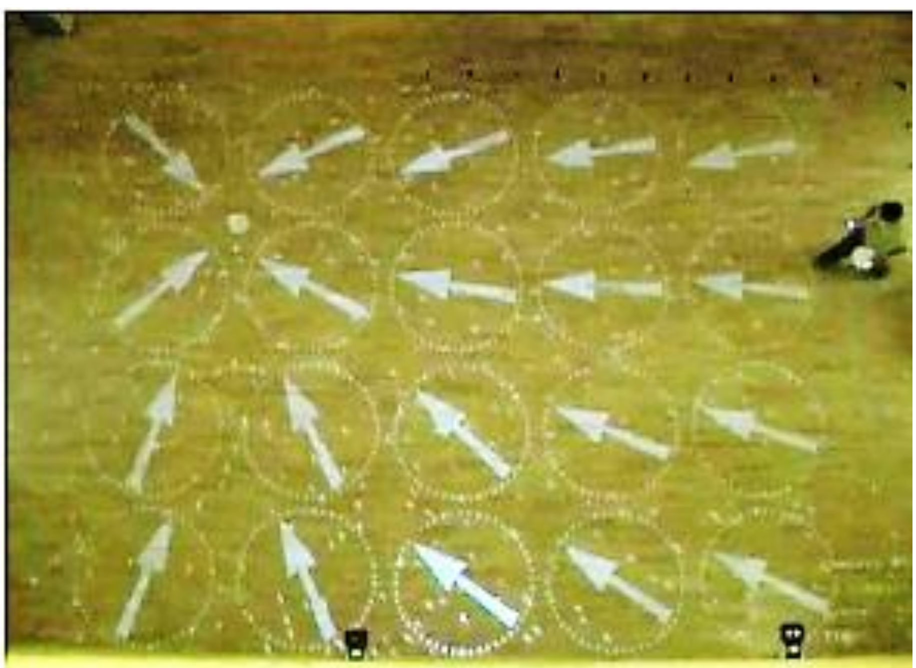

Figure 4.19 The corresponding influential arrows projected on the floor in a building

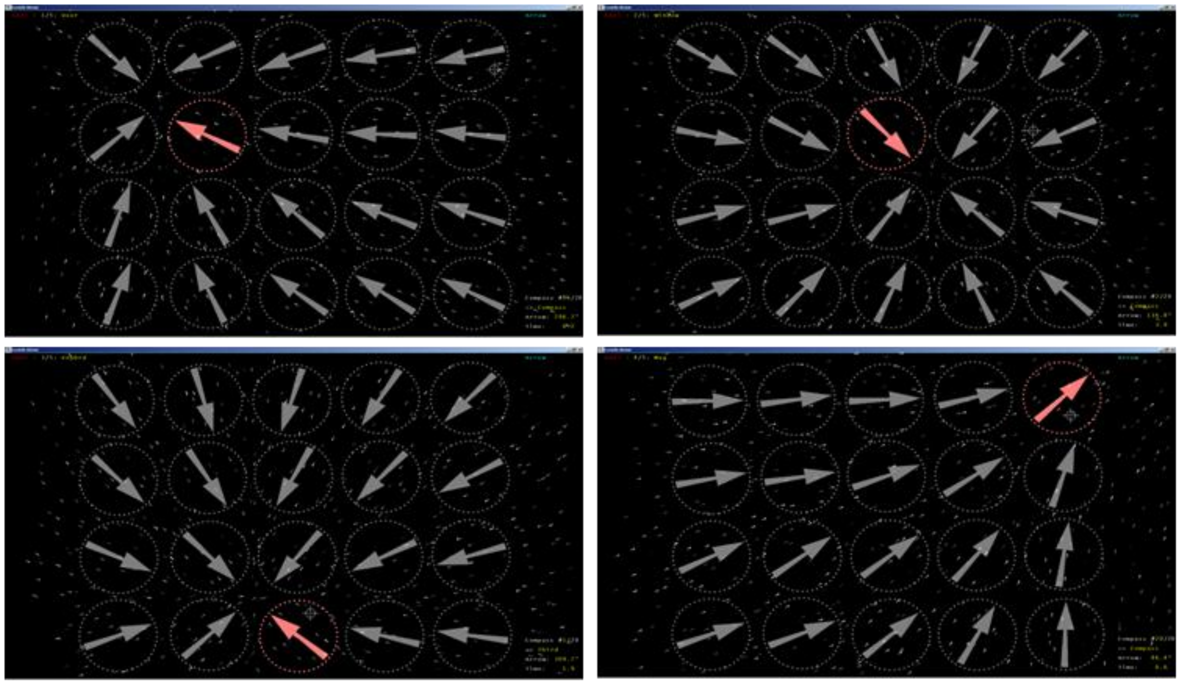

Figure 4.20 Influential arrows for configuring the directions of graphical elements of visual public directional information corresponding to different destinations

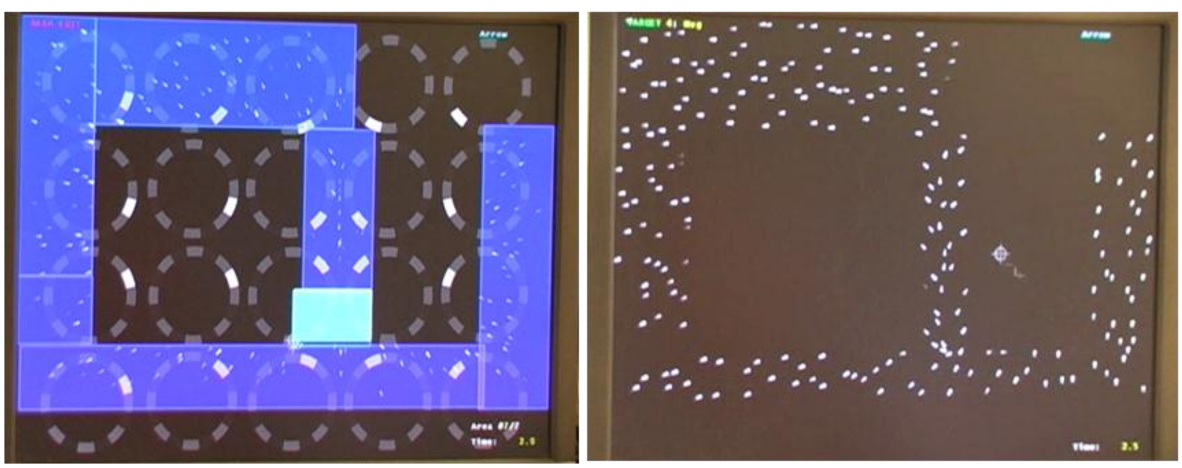

Figure 4.21 A example of how to configure the distribution areas of the fish-flow pattern

Figure 4.22 An example of how a designer steers the flow tendencies around the central area of the screen

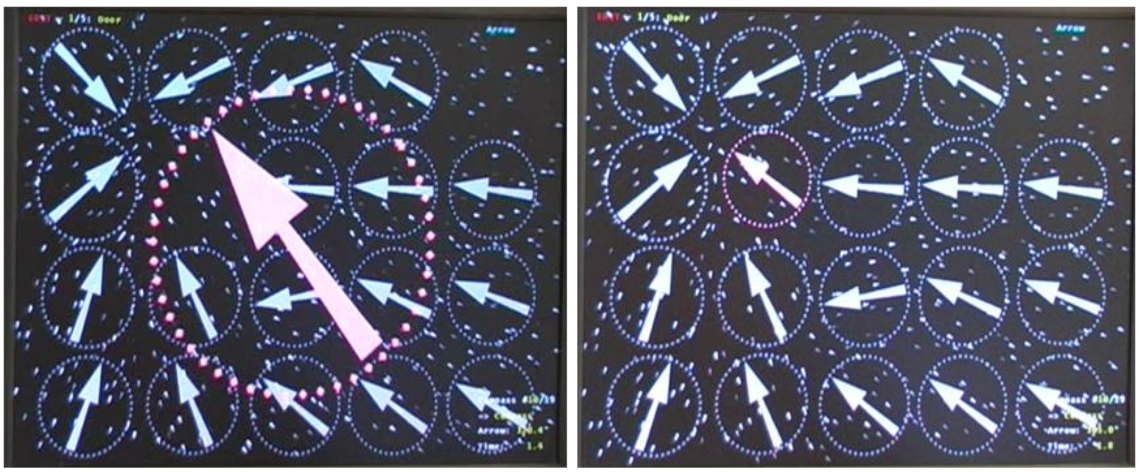

Figure 4.23 An influential arrow (in red) added and scaled down for configuring the fish-flow pattern

4.5 Conclusions

This chapter has reported our exploration of the design of CROSSFLOW, a Crossmodal Display for indoor navigation aiming to facilitate indoor navigation of multiple simultaneous users within an unfamiliar complex indoor environment. The approach taken in this exploration is to identify the design goals based on the problems identified and literature review of the corresponding solutions in Chapter 2, the design implications of multimodal integration and crossmodal interaction for interface design in Chapter 3 and the analysis of existing navigation approaches, systems and technologies, establish conceptual design for CROSSFLOW from the research findings relating to human navigation and wayfinding, an analysis of a simple navigation scenario in a public space and our conceptual model of Crossmodal Displays in Chapter 3, and then explore the design of the components of CROSSFLOW including cue mappings, public cues and private cues, and finally embody the concept design as a lab prototype (CROSSFLOW V1) which provides the combination of aesthetically pleasing visual ambient patterns projected on the floor of an indoor environment and sound or/and vibration displayed by one or more smartphones. In Chapter 5, two user studies of CROSSFLOW V1 (Experiment 1 and 2) are reported, in which we investigate the usability and the user experience afforded by CROSSFLOW V1 so as to answer relevant research questions proposed in Chapter 2.

Chapter 5. Evaluating CROSSFLOW Version 1

5.1 Introduction

Having designed and built the first CROSSFLOW prototype, a Crossmodal Display for indoor navigation, it is important and interesting to see the usability and the user experience of CROSSFLOW in the context of navigating in a real physical indoor environment in comparison with traditional navigation aid such as a hand-held paper map. If it were possible to use CROSSFLOW for real indoor navigation, this would make the use of Crossmodal Display such kind of interface more commercially viable, as the projectors and mobile phones, PDAs and smartphones are common devices and featured in many common public spaces such as shopping malls and art galleries.

As there is no previous approach for evaluating Crossmodal Displays, we performed two different evaluations that referred to most of the techniques found in previous works (e.g. Dix et al. 2004; Kim & Rieh 2005; Maglio & Campbell 2000). The first user study (Experiment 1) was a dual-task lab study, the methods of which included objective records of performance on the primary and secondary task; and subjective measurement of task workload (NASA-TLX) (Hart & Staveland 1988; NASA 1988). The second user study (Experiment 2) included objective records of performance on the navigation tasks; subjective measurement of task workload (NASA-TLX); and qualitative information on questionnaires.

Because the dual-task method is widely used in experimental psychology and human factors research (Kim & Rieh 2005) and proved to be valid to assess mental effort in Human-Computer Interaction (HCI) field (e.g. Maglio & Campbell 2000), it is probably a valid method to find out whether the users invest less mental effort using CROSSFLOW than traditional paper-map-based navigation aid.

NASA-TLX (NASA 1988) is often used where performance measures cannot be measured directly, although this is not the case for both Experiments 1 and 2, we use it as a subjective validation of the dual task performance measures. Furthermore, NASA-TLX is one of the best-established subjective workload measures with the advantages including a sound basis, relatively high reliability and sensitivity and low intrusiveness and cost (Farmer & Adam Brownson 2003).

We use qualitative measures through questionnaires mainly to elicit the users' specific subjective experience associated with the navigation systems that the performance measures and the subjective workload measures are unlikely to reveal.

5.2 Experiment 1: Evaluating CROSSFLOW V1 in Single User Setting

In this formal user study, individual participant was given dual tasks to perform using either CROSSFLOW V1 prototype or hand-held paper map. The performances of the participants using different navigation aids were compared. The results of this empirical study would determine whether CROSSFLOW could offer more benefits on indoor navigation tasks than hand-held paper map in dual-task situation.

The main aim of this experiment was to investigate CROSSFLOW on these aspects:

- **Suitability of Components:** to observe the suitability of the components in CROSSFLOW including the cue mapping, public directional information and private cue and the effectiveness of CROSSFLOW 's user-interface and interaction model.
- **Utility and Usability:** to observe the utility and usability of CROSSFLOW as a type of indoor navigation aid that helps user to reach his/her target object among many visually similar objects in an unfamiliar open indoor environment.
- **CROSSFLOW in comparison with hand-held paper map:** to determine whether CROSSFLOW is a better navigation aid in comparison with hand-held paper map in terms of navigation task performance, primary task performance and subjective workload in dual-task condition. There are two dual-task conditions in this experiment which are referred to as the CROSSFLOW condition and the MAP condition.

5.2.1 *Experimental Settings of Experiment 1*

A. Experimental area

The experiment was implemented on the atrium of a building of Newcastle University (Figure 5.1). The size of the experimental area was approximately 10 meters wide by 6.5 meters long. Within the experimental area which is empty and clean on the floor, only 15 white containers with the same size and shape were positioned at different locations in each condition. 5 out of the 15 containers held navigational information for the user. The other 10 containers that were not valid destinations contained the statement “this is the incorrect location”. The use of these “distractor” containers was with a view to adding a degree of real world complexity and ambiguity for navigation tasks. As the experiment was conducted in a relatively small space in comparison with a physical public indoor environment such as an airport, it was necessary to have a significantly denser array of locations for the user to navigate between.

Figure 5.1 Experimental area of Experiment 1 and 2

B. Public display and public directional information

A laptop was used as the server of CROSSFLOW system. A SXGA projector was connected to the laptop. The laptop and the projector were both positioned at the third floor of the building. When CROSSFLOW V1 was in use, the projected fish-flow pattern covered the whole experimental area, and was configured to flow to the 5 target containers in the CROSSFLOW condition at 5 different time slots within 4 seconds cycle time. The lighting of the building was also lessened to make the fish-flow pattern more visually accessible.

C. Personal display and private cue

A SPV E200 smartphone was synchronized with the server of CROSSFLOW before they were used in the experiment. The beep sound (see Figure 4.16) and the vibration (see Figure 4.15) used in the previous informal user study were used in this experiment as private cues and displayed simultaneously to each participant via the SPV E200 smartphone.

Although we had not yet empirically studied the impact of different type of private cue (auditory cue, vibrotactile cue or the combination) on the usability and the user experience of CROSSFLOW, as indicated in Section 3.3.1 in Chapter 3, either the sound or the vibration could be used as a redundant cue which might lead to increased confidence or reliability of identification of public directional information relevant to users.

D. Cue mapping

The cue mapping used in this experiment is shown in figure 5.2, in which there are five 800-millisecond time slots within a-four-second cycle time corresponding to the locations of 5 target containers in the CROSSFLOW condition.

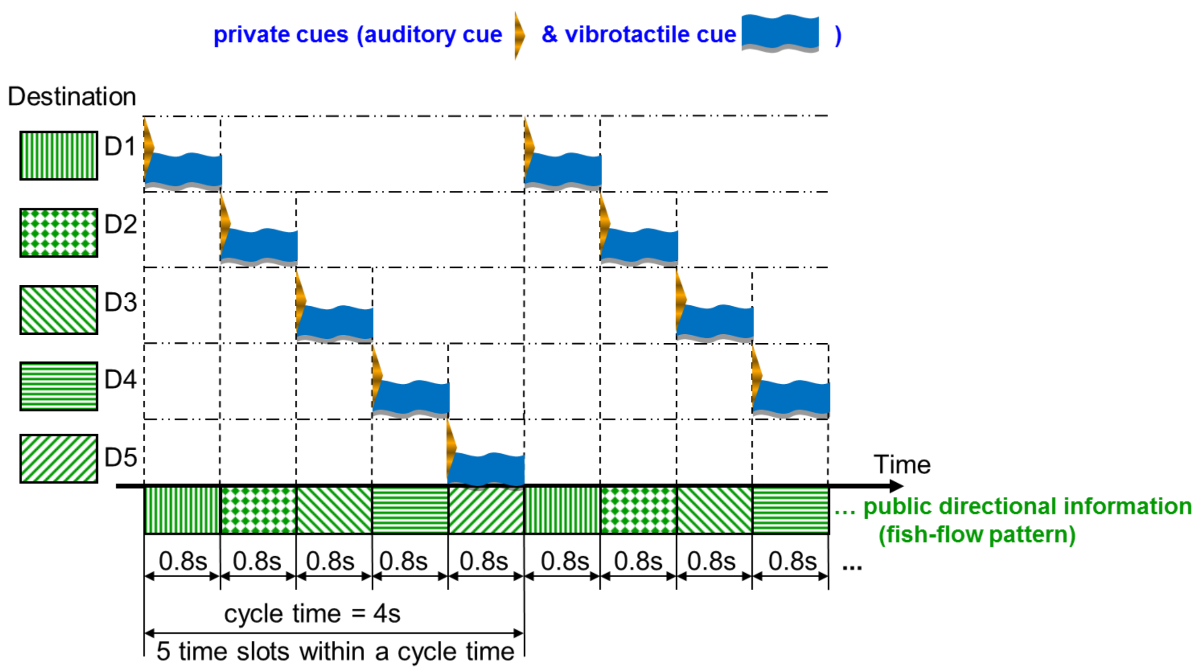

Figure 5.2 Regular cue mapping for Experiment 1

E. Video camera

Two digital video cameras were positioned at 2nd floor of the building. Through them we captured and recorded the activities of the participants in the whole experiment with a bird-eye view.

5.2.2 *Experimental Method of Experiment 1*

A. Experimental Design

The study used a within-subject design (Dyer 1995). We utilized one independent variable, type of navigation aid, with two levels of treatment: (1) navigation using a hand-held paper map in dual-task condition (answering arithmetic questions and navigating), (2) navigation using CROSSFLOW in dual-task condition (answering arithmetic questions and navigating). Such a design has advantages over a between subject design including lowering the possibility of individual differences skewing the results. With the within-subjects design, the conditions are always exactly equivalent with respect to individual difference variables since the participants are the same in the different conditions (Hall 1998; Shuttleworth 2009). So, in our experiment, any factor that may affect performance on the independent variable (type of navigation aid) such as navigational skills and arithmetic skill will be exactly the same for the two conditions, because they are the exact same group of people in the two conditions. Therefore, it is no need to benchmark the navigational and arithmetic skills of the subjects.

B. Participants

Nine participants (5 males, 4 females) with no prior experience of using CROSSFLOW took part in this experiment; all were students of Newcastle University and Northumbria University, aged between 20 and 30 years old and declared themselves free of visual, auditory, tactile disability and physical impairment. The content and the method of the experiment were agreed by the participants and approved by the supervisor of this research who is in the School of Computing Science at Newcastle University.

C. Task

The primary task involved answering a set of arithmetic questions posed one-by-one by the experimenter whilst each participant undertook the navigation task. For these reasons, answering arithmetic questions was chosen as the primary task: (a) it is relatively easy to design such kind of task; (b) it is relatively easy to control the task demands; (c) the low intrusiveness for the participants and (d) low cost.

The secondary task for each participant was to find 5 target containers out of the 15 containers within the experimental area with the aid of either a hand-held paper map or CROSSFLOW.

D. Testing Procedure

Each participant was presented with all three conditions during the experiment. The order of the conditions was randomized for every participant. The three conditions are:

- **Single-task condition**: answering arithmetic questions without navigating.
- **MAP dual-task condition**: answering arithmetic questions and navigating with a hand-held paper map.
- **CROSSFLOW dual-task condition**: answering arithmetic questions and navigating with CROSSFLOW.

Before the experiment, each participant was given a brief introduction of tasks and demonstration on the use of CROSSFLOW and the paper map. In addition, they were told that in the dual-task conditions answering the arithmetic questions was the primary task which would not stop until all designated 5 targets were found, and they would be no time limitation on each condition. The importance of performing both tasks to the best of their ability was emphasized.

In the single task condition, each participant answered 18 arithmetic questions. Participants were asked to try their best to answer all the arithmetic questions correctly, and no time limit was enforced. The mathematical ability of each participant was elicited through a screening questionnaire. The evaluation of performance on arithmetic questions yielded a mean time per question of 4.0 seconds and a mean accuracy of 97%.

In the dual-task conditions, each participant navigated in the experimental area to find 5 target containers out of the 15 containers. The set of targets used in dual-task conditions was randomized for each participant to avoid the spatial learning effect. Figure

5.3 shows the experimental area with a sample distribution of the targets used in the dual-task conditions.

In the MAP dual-task condition (Figure 5.4), each participant navigated in the experimental area with the aid of a hand-held paper map on which the locations of 5 target containers were marked, and was asked to give spoken answers to the arithmetic questions posed by the experimenter. When he/she started to navigate, the name of the first target container was given, and subsequent target containers contained the name of the next target. On finding a target, the participant read the name of the next target from the container and continued. In the CROSSFLOW dual-task condition (Figure 5.5), the primary task was the same and navigation task was completed with the aid of CROSSFLOW. Each participant held a SPV E200 smart phone in hand to navigate. When he/she started to navigate, the number of the first target container was given, and subsequent target containers contained the number of the next target. On discovery of a target, the participant entered the number corresponding to the next target on the smartphone keypad.

After each participant completed all tasks in three conditions, he/she was asked to complete a questionnaire based on NASA-TLX (Task Load Index) to assess his/her subjective workload of the tasks. No time limit was enforced for completing this questionnaire.

All the activities of each participant were recorded on digital videotapes. The responses of participants to the arithmetic questions were recorded on paper by the experimenters during the experiment.

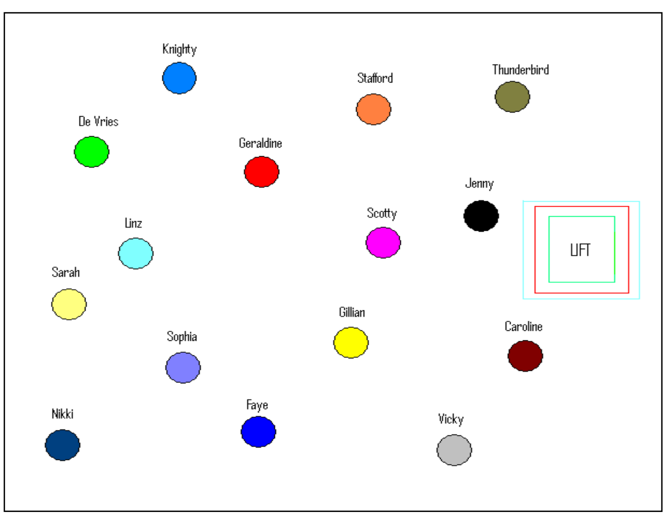

Figure 5.3 Experiment 1: Illustration of the map given to subjects in MAP condition

Figure 5.4 Experiment 1: MAP condition

Figure 5.5 Experiment 1: CROSSFLOW condition

E. Measurements

The following measures of performance were defined and collected; they are also served as the dependent variables in the statistical analysis. Performance on the answering arithmetic questions alone was measured and served as the baseline to compare to arithmetic questions answering performance in the two dual-task conditions.

- **Primary task reaction (time taken per question)**: The time taken to answer an arithmetic question, measured from when a question is asked to when the participant answers.
- **Primary task accuracy (% Correct)**: The percentage of the correctly answered arithmetic questions.
- **Navigation task success rate**: The number of successful navigation tasks versus the number of navigation tasks. Successful navigation task was defined as the participant reaching the target container regardless of time taken and navigation errors.
- **Navigation error**: The number of navigation errors in discovering the 5 targets. Navigation errors were recorded formally when the participants addressed the wrong targets during the navigation task. For example, in the course of searching 5 target containers in the MAP condition, if a participant incorrectly identified a distractor container as the next destination and also returned to a previous destination in order to ascertain the location of the next destination, two navigation errors would be recorded.
- **Total time**: The total amount of time participants spent on finding 5 target containers, and answering the arithmetic questions.
- **Subjective task workload**: The subjective measurement of task load using the NASA-TLX software (NASA 1988).

5.2.3 *Results of Experiment 1*

Condition	Time taken per question (second)	% Correct	Total time (second)	Number of navigation errors	NASA TLX score
MAP	8.47 (1.88)	84.13 (16.75)	133.04 (56.37)	1.22 (2.22)	78.41 (7.30)
CROSSFLOW	6.12 (1.49)	98.13 (3.72)	79.23 (28.08)	0.44 (0.73)	52.85 (11.93)
SINGLE-TASK	3.99 (0.64)	96.90 (5.65)	79.67 (12.78)	N/A	N/A

Table 5.1 Experiment 1: Mean measures (with standard deviations) for the MAP and CROSSFLOW dual-task conditions and single-task condition.

Table 5.1 presents the mean measures for the MAP and CROSSFLOW dual-task conditions and single-task condition, and standard deviations in brackets.

Before performing statistical analysis the data were tested using Kolmogorov-Smirnov normality test with a Lilliefors significance correction and Shapiro-Wilk normality test (Park 2008) to see if they were normally distributed. As the data of the percentage of correctly answered questions in each condition and the number of navigation errors in the MAP and CROSSFLOW condition did not follow a normal distribution, non-parametric analyses were carried out using Friedman test and Wilcoxon signed-rank test (Leeper 2000) for within-participant comparisons. For the data which followed a normal distribution, one-way repeated measures ANOVA and two-tailed paired samples t-tests (Dyer 1995) were used for within-subjects comparisons.

A. Comparison of primary task performance

The performance for the primary task was compared across the single briefing task, and the MAP and CROSSFLOW dual-task conditions. Two aspects of the performance on the primary task were considered: primary task reaction (time taken per question) and primary task accuracy (% Correct).

The single-task condition had the time taken per question measured in seconds (Mean = 3.99, Standard Deviation = 0.64), followed by the CROSSFLOW condition (Mean = 6.12, Standard Deviation = 1.49) and the MAP condition (Mean = 8.47, Standard Deviation = 1.88). Figure 5.6 shows the mean time taken per question with standard deviations for the MAP and CROSSFLOW conditions.

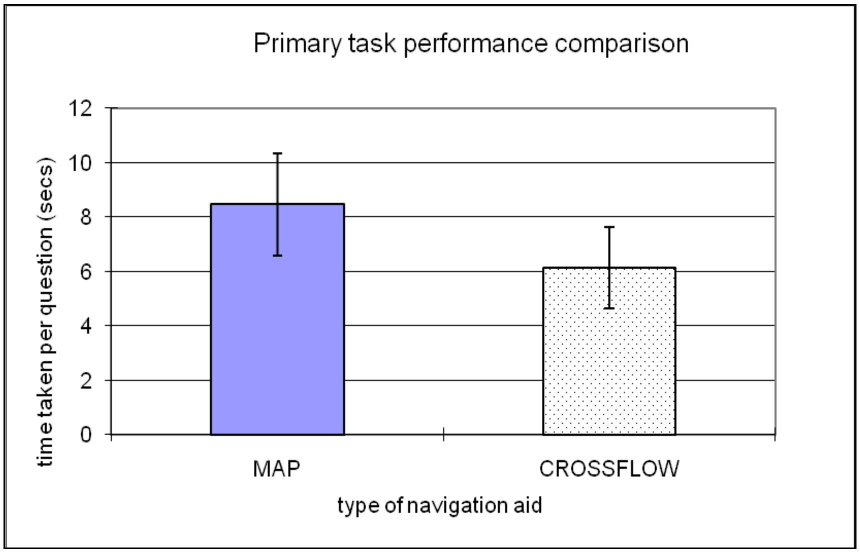

Figure 5.6 Experiment 1: Mean time taken per question (with standard deviations) for the MAP and CROSSFLOW conditions.

With respect the time taken per question, a one-way repeated measures ANOVA was calculated across the three conditions, showing a significant differences between performance without and with the navigation task ($F(2,16) = 42.28$; $df = 2,16$; $p < 0.001$). The *post-hoc* two-tailed paired samples t-tests showed that differences in the times taken to answer an arithmetic question between the single-task condition and the CROSSFLOW dual-task condition ($t(8) = 4.32$, $p = 0.003$), and the single-task condition and the MAP dual-task condition ($t(8) = 7.66$, $p < 0.0005$) were statistically significant. The difference between the CROSSFLOW dual-task condition and the MAP dual-task condition was also statistically significant ($t(8) = 6.60$, $p < 0.0005$), with the mean time taken per question for the CROSSFLOW dual-task condition being 38.4% less than which for the MAP dual-task condition.

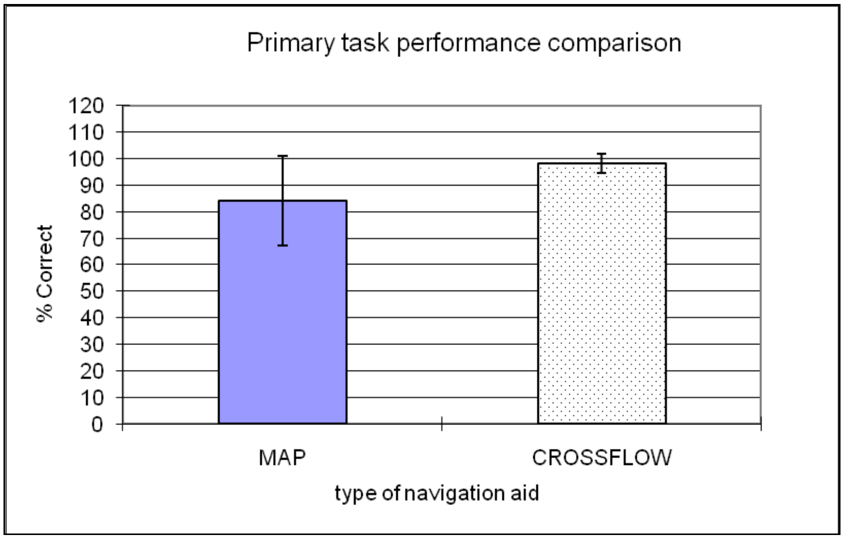

Figure 5.7 Experiment 1: Mean percentage of correctly answered questions (with standard deviations) for the MAP and CROSSFLOW conditions.

To compare the percentage of correctly answered questions, a Friedman test revealed that the difference was not statistically significant across the three conditions, p = 0.082. Figure 5.7 shows the mean percentage of correctly answered questions with standard deviations for the MAP and CROSSFLOW conditions. The mean percentage of correctly answered questions for the CROSSFLOW condition was 16.7% higher than which for the MAP condition.

B. Comparison of secondary task performance

The performance for the secondary (navigation) task was compared across the MAP and CROSSFLOW condition according to three criteria: navigation task success rate, navigation error and total time.

All participants found the 5 targets in each condition, resulting in 100% navigation task success rate. Thus there was no statistically significant difference between the two conditions.

In the MAP condition subjects averaged 1.2 navigational errors and for CROSSFLOW the average was 0.4. A Wilcoxon signed-rank test was conducted and the result showed that the difference of the number of navigation errors was not statistically significant between the MAP and CROSSFLOW condition, p = 0.157.

As for total time taken, the mean total time for the CROSSFLOW dual-task condition was 67.9% less than which for the MAP dual-task condition. Figure 5.8 shows the mean time spent on finding 5 targets with standard deviations for the MAP and CROSSFLOW conditions. A paired samples t-test further shows that the total time decreased significantly from the MAP condition to the CROSSFLOW condition, $t(8) = 3.46$, $p = 0.009$.

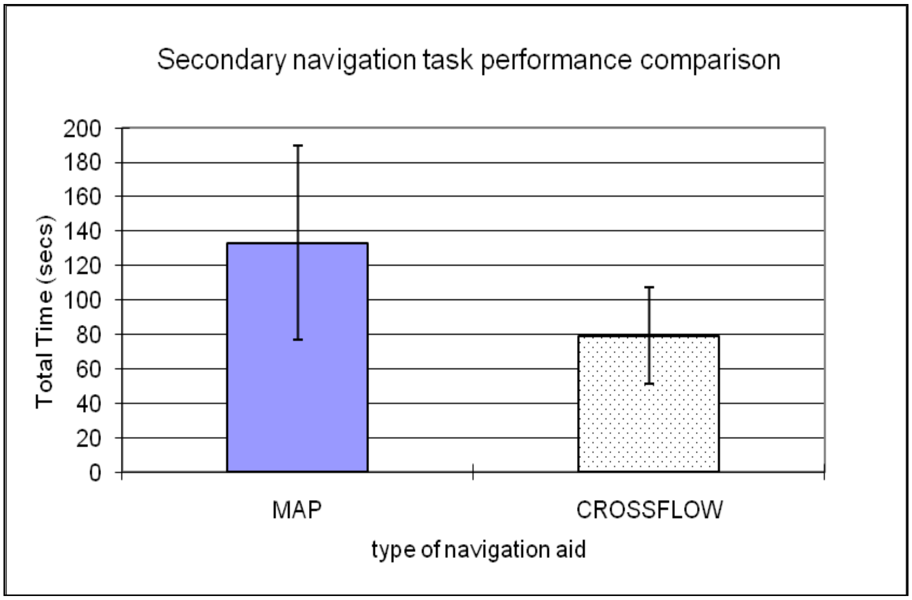

Figure 5.8 Experiment 1: Mean time spent on finding 5 targets (with standard deviations) for the MAP and CROSSFLOW conditions.

C. Comparison of subjective workload

Figure 5.9 shows the mean NASA TLX score with standard deviations for the MAP and CROSSFLOW conditions. A paired samples t-test was conducted to compare the scores of NASA-TLX subjective workload of the two conditions. The result showed a significant reduction in the perceived task workload when using CROSSFLOW as compared to MAP, $t(8) = 6.24$, $p < 0.001$.

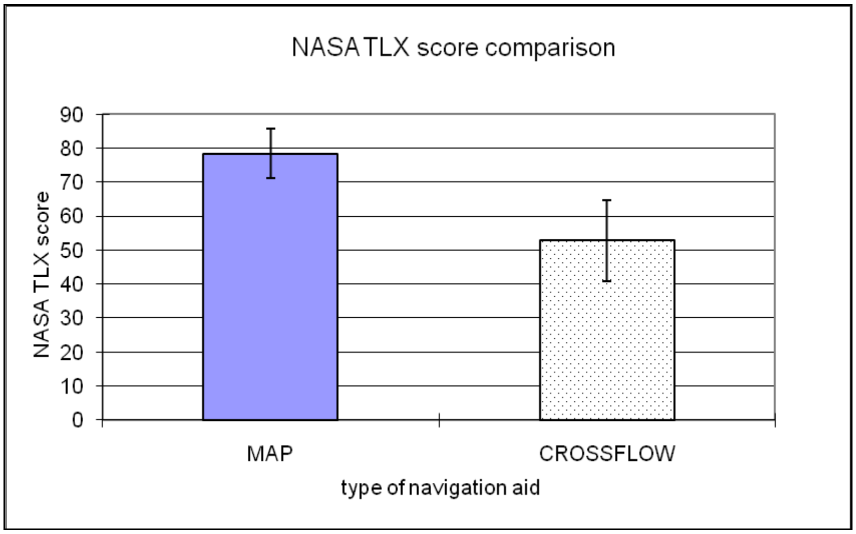

Figure 5.9 Experiment 1: Mean NASA TLX score (with standard deviations) for the MAP and CROSSFLOW conditions.

D. The user experience of CROSSFLOW

The user experience afforded by CROSSFLOW was not formally addressed in this user study, though informal feedbacks received after the study showed that the participants felt at ease with CROSSFLOW and most found it fun to use and helpful.

E. Observations of participants' behaviour

From the observation of the participants' activities recorded on the videos (analyzed by the experimenter frame by frame in some parts), the behaviour of the participants appeared differently in the two conditions. In the MAP condition, the participants held the paper map while navigating, continuously glancing back and forth between the surrounded environment and the map while walking around. Once reached a target container, they tended to pause more time then started a new navigation task, which is probably because they needed more time to identify their local position and next target on the map and in the environment and plan a route.

In the CROSSFLOW condition, the participants paid attention to the pattern shown in the experimental environment most of the time and looked at the interface on the smartphone occasionally. It seemed that they started a new navigation task immediately after reaching a target container. Some participants tended to pause more in order to gradually hone in on the correct target. One of the subjects paused for a particularly long time as he struggled to find the next target, which was actually right beneath his feet. This situation may mainly be attributed to the artificially small spatial scales over which the participants have to move.

5.2.4 *Discussion of Experiment 1*

The results from this study of the CROSSFLOW lab prototype are promising.

A. Utility and usability of CROSSFLOW

In the CROSSFLOW condition, the participants successfully found target objects among many visually similar objects in a physical open indoor environment using CROSSFLOW with 100% success rate regardless navigation errors. This result demonstrates the success of CROSSFLOW's user-interface and interaction model and indicates that CROSSFLOW can be used as an indoor navigation aid. The result of the navigation task success rate in CROSSFLOW condition also indicates that the designed components of CROSSFLOW including regular cue mapping, abstract public directional information (the projected fish-flow pattern in this experiment) and private cues (the beeps and vibrations from smartphone in this experiment) can be used for personalized directional information presentation, and that users are able to identify the correlation between private cues and public directional information corresponding to a target location or object with less than five minutes brief demonstration, and then reach it successfully.

B. CROSSFLOW in comparison with MAP

The participants' primary task performances in the CROSSFLOW dual-task condition are significantly better than which in the MAP dual-task condition in terms of the time taken to answer each question. These results are consistent with the result of the comparison of subjective workload that the scores of NASA-TLX for the CROSSFLOW condition is significantly lower than the MAP condition. The results of the comparison of primary task performance and subjective workload suggest that CROSSFLOW imposes less attentional and cognitive load on users on indoor navigation tasks than hand-held map.

This might be because that the participants invest different amounts of mental effort into the navigation tasks when using different navigation aids. Using hand-held map requires a user to identify his/her current location and the target location both on the map and in the physical environment and plan a route, while using CROSSFLOW to navigate only needs a user to capture and identify the public directional information corresponding to the target location shown by public displays. A user may invest relatively less mental effort into the navigation tasks and allocate more attentional and cognitive resources to the primary task in the CROSSFLOW condition, thus resulting in significantly better primary task performance and lower subjective workload.

The results of the comparison of navigation success rate and navigation error between the CROSSFLOW and MAP condition show a similar level of navigation performance as no statistically robust conclusions were able to be drawn from the data, which suggest that using CROSSFLOW to find a object or location in a indoor environment is as successful and accurate as using a hand-held map.

The participants averagely spent less time to find 5 target containers within the experimental area in the CROSSFLOW condition in comparison with the MAP condition. Moreover, the difference of time spent was statistically significant between the two conditions. The results together with the other results above suggest that CROSSFLOW is a better type of navigation aid than hand-held paper map in terms of multitasking, workload and navigation efficiency.

5.3 Experiment 2: Evaluating CROSSFLOW V1 in Multiple Simultaneous Users Setting

In this user study, two groups of participants performed several navigation tasks respectively and simultaneously within a maze-like complex indoor environment using either CROSSFLOW V1 prototype or hand-held paper map. In order to simulate a scenario in real world that a group of people were strangers to each other and busy with their own tasks respectively in a public space, the participants were required not to communicate with each other during the experiment. The usability and the user experience of CROSSFLOW and hand-held paper map were investigated and compared. The results of this empirical study would enable the more design issues about CROSSFLOW to be dealt with before approaching the important research questions identified in Chapter 2.

The main aim of this study was to inform our in-depth understanding of CROSSFLOW in multi-user situation, in particular on these aspects:

- **Suitability of Components:** to observe the suitability of the system components of CROSSFLOW including the cue mapping, public directional information and private cue and the effectiveness of CROSSFLOW's user-interface and interaction model in multiple simultaneous users setting.
- **Utility and Usability:** to observe the utility and usability of CROSSFLOW as a type of indoor navigation aid for multiple simultaneous users in an unfamiliar complex indoor environment.
- **User Experience:** to observe the user experience afforded by CROSSFLOW.
- **CROSSFLOW in comparison with hand-held paper map:** to determine whether CROSSFLOW is a better type of navigation aid in comparison with hand-held paper map in an unfamiliar complex indoor environment in terms of navigation task performance, subjective workload and user experience. There are two conditions in this experiment which are referred to as the CROSSFLOW condition and the MAP condition.

5.3.1 *Experimental Settings of Experiment 2*

A. Experimental area

This experiment was still implemented on the atrium of a building of Newcastle University. The size of experimental area was approximately 11 meters wide by 7 meters long, which was slightly larger than the previous. In order to make the navigation tasks more demanding, the complexity of the floor plan within the experimental area was increased significantly. As shown in figure 5.10, 28 polystyrene foam walls (1.2 x 2.4 x 0.3 meters each) were distributed in the experimental area and configured a maze-like structure. 16 out of 28 walls were possible target locations where 16 envelopes and pieces of paper with information were attached to each of the foam walls (12 pieces of paper had numbers written on them, and 4 pieces of paper had the statement "This is a wrong target."). The use of these “distractor” walls was again for the purpose of adding a degree of real world complexity and ambiguity for navigation tasks.

The floor plan layout of the "maze" is symmetrical, so as the distribution of the 16 possible target walls. One purpose of this setting was to increase the spatial complexity of navigation tasks, because a symmetrical layout with repetition and few differentiation can increase the difficulty of wayfinding (Baskaya et al. 2004). Another purpose was to keep the spatial complexity identical and consistent across the MAP and the CROSSFLOW condition.

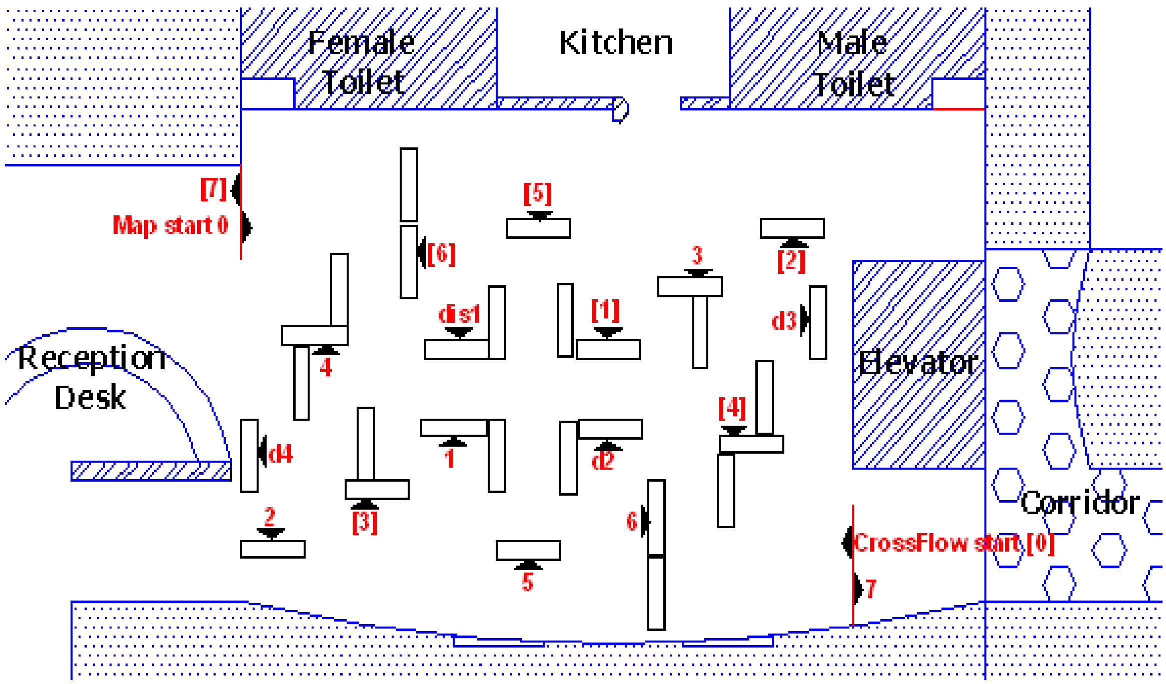

Figure 5.10 Experiment 2: Floor plan depicting the spatial configuration of the experimental area.

B. Public display and public directional information

The settings of public display in this experiment are similar to which in Experiment 1. The main difference is that additional laptop connecting with a SXGA projector was employed to project the fish-flow pattern covering the shadow area caused by the foam walls. Moreover, the pattern on the two laptops was synchronized and configured to flow to the 7 target locations in the CROSSFLOW condition.

C. Personal display and private cue

Several SPV E200 smartphones were synchronized with the server of CROSSFLOW before they were given to the participants in the CROSSFLOW condition. The same 800-milliseconds-long vibration used in Experiment 1 was used in this experiment as the private cue.

D. Cue mapping

The cue mapping used in this experiment is shown in figure 5.11, in which there are seven 800-millisecond time slots within the 5.6 seconds cycle time corresponding to the 7 target locations in the CROSSFLOW condition.

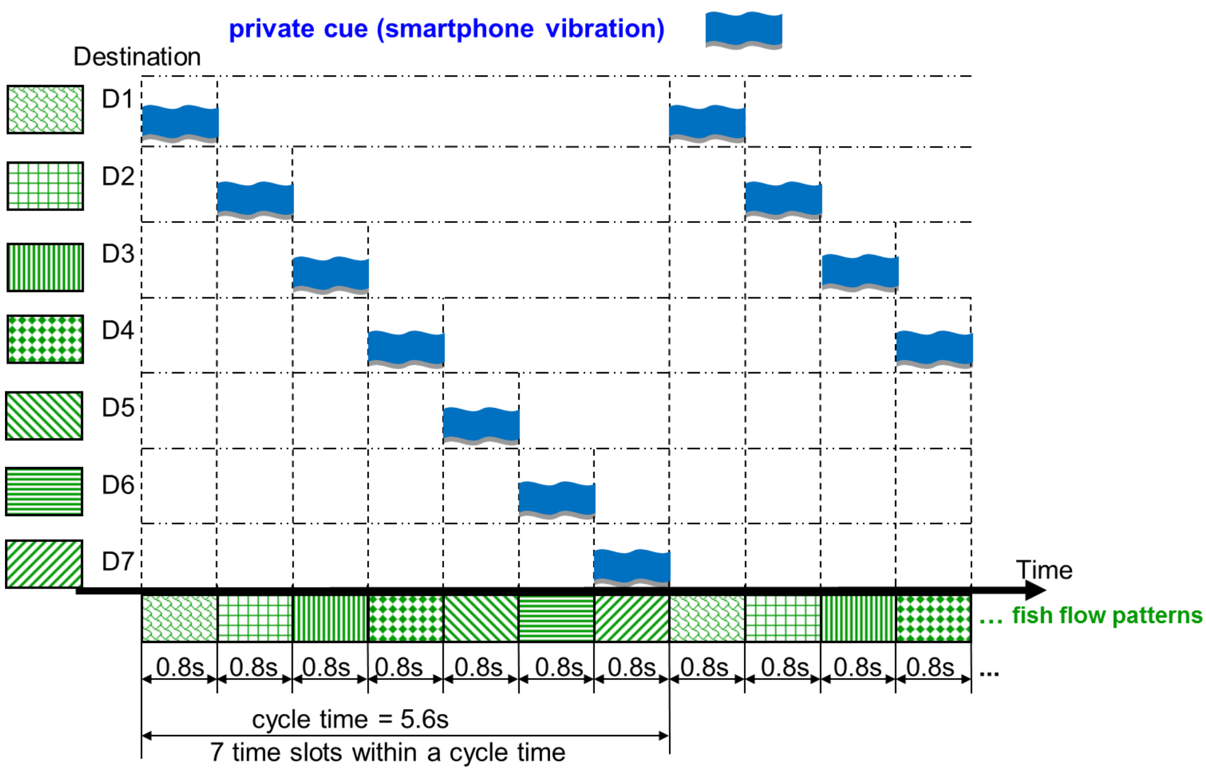

Figure 5.11 The regular cue mapping for Experiment 2

E. Video camera

The setting of video cameras in this experiment was the same as which in the Experiment 1.

5.3.2 *Experimental Method of Experiment 2*

A. Experimental Design

The study used a within subject design. We utilized one independent variable, type of navigation tool with two levels of treatment: (1) navigation using a hand-held bird-eye view paper map; (2) navigation using CROSSFLOW. The rationale of why such experimental design was adopted is the same as which is discussed in Section 5.2.2 A.

B. Participants

Ten participants (4 males, 6 females) with no prior experience of the CROSSFLOW took part in this experiment; all were students of Newcastle University and the Northumbria University, aged between 20 and 30 years old and declared themselves free of visual, auditory, tactile disability and physical impairment. They were divided into two groups (Group 1 and 2). Two males and three females were randomly assigned to each group. Each participant was paid £20. The content and the method of the experiment were agreed by the participants and approved by the supervisor of this research who is in the School of Computing Science at Newcastle University.

C. Task

Each participant was asked to perform 7 navigation tasks in each condition in the experiment, finding the 6 target walls and the end location within the experimental area with the aid of either a paper map or CROSSFLOW. The order of searching the 6 target walls is randomized and different from participant to participant. The detailed order for each participant is shown in table 5.2.

The locations of 6 target walls and the ending location in the MAP condition were totally different from which in the CROSSFLOW condition, but they are symmetrical (see Figure 5.10). Moreover, the order of navigation tasks in the two conditions is the same as shown in table 5.2. Thus the level of difficulty of navigation tasks in the two conditions was equivalent. As a result, the task searching the location "2" in the MAP condition and the task searching the location "[2]" in the CROSSFLOW condition were logged as the same "task II"; other tasks were logged in the same way.

Group 1 participant code	the order of searching locations for the participants in the **MAP condition**								the order of searching locations for the participants in the **CROSSFLOW condition**							
A	start location 0	1	4	2	6	5	3	end location 7	start location [0]	[1]	[4]	[2]	[6]	[5]	[3]	end location [7]
B	start location 0	6	3	1	5	2	4	end location 7	start location [0]	[6]	[3]	[1]	[5]	[2]	[4]	end location [7]
C	start location 0	4	2	5	3	1	6	end location 7	start location [0]	[4]	[2]	[5]	[3]	[1]	[6]	end location [7]
D	start location 0	3	5	4	2	6	1	end location 7	start location [0]	[3]	[5]	[4]	[2]	[6]	[1]	end location [7]
E	start location 0	5	1	6	4	3	2	end location 7	start location [0]	[5]	[1]	[6]	[4]	[3]	[2]	end location [7]
Group 2 participant code	the order of searching locations for the participants in the **CROSSFLOW condition**								the order of searching locations for the participants in the **MAP condition**							
G	start location [0]	[1]	[4]	[2]	[6]	[5]	[3]	end location [7]	start location 0	1	4	2	6	5	3	end location 7
H	start location [0]	[6]	[3]	[1]	[5]	[2]	[4]	end location [7]	start location 0	6	3	1	5	2	4	end location 7
I	start location [0]	[4]	[2]	[5]	[3]	[1]	[6]	end location [7]	start location 0	4	2	5	3	1	6	end location 7
J	start location [0]	[3]	[5]	[4]	[2]	[6]	[1]	end location [7]	start location 0	3	5	4	2	6	1	end location 7
K	start location [0]	[5]	[1]	[6]	[4]	[3]	[2]	end location [7]	start location 0	5	1	6	4	3	2	end location 7

Table 5.2 Experiment 2: The order of searching target locations for the participants

D. Training Procedure

The purpose of the training section was to make sure the participants understand how to use their navigation aids and tasks. Therefore we did not measure their performance. Trials proceeded in the following way. The first group of five participants was guided to a training space in a building at Newcastle University. They were given a brief introduction of the procedure of the whole experiment and demonstration on the use of CROSSFLOW and a paper map of the training space. Printed instructions were also provided. Then they received the training one by one. Each participant in the group was asked to perform two training tasks, finding a target out of seven possible targets in the training space using the paper map and finding another target using CROSSFLOW. He/she was told whether the target was right or not after he/she finished each task. After the first group completed the training section and the formal experiment section, the second group was guided to the training space and received the trainings. Two groups of participants received the similar training; the only difference is that the order of training tasks was counterbalanced across the two groups. In the training section, each participant spent less than five minutes in completing the training tasks.

E. Testing Procedure

In the formal experiment section, each group of participants was presented with two conditions. The order of the two conditions was counterbalanced across two groups to prevent any carryover effects. The two conditions are:

- **MAP condition**: navigating with a hand-held paper map.
- **CROSSFLOW condition**: navigating with CROSSFLOW.

After a group of participant completed their training, they were guided to the formal experimental area, the atrium of a building at Newcastle University.

Before the formal experiment started, the participants were asked to wear the hats given by the experimenter. The use of the 5 hats with different colours is to label the participants so that the activity of each participant could be easily analysed from the observation. The participants were also told that each of them should complete his/her own tasks as quickly and accurately as possible in his/her normal pace and would be limited to 30 minutes in each condition. In addition, they were asked to not compete and communicate with each other during the formal experiment.

In the MAP condition (Figure 5.12), firstly, each participant in a group was given an A4-size birds-eye view paper map (Figure 5.13) on which the locations of 6 targeted walls, the starting and ending locations were marked and an A5-size task book (on each page a number corresponding to a target wall in the MAP condition was printed). Then the participants were asked to stand in a line at the location "Map start 0". When the experimenter said "start", they started their navigation tasks immediately and simultaneously. When a participant reached a wall or the ending location, he/she uncovered the paper on the wall or table to reveal the information. On discovery of a right location, the participant ripped off the page corresponding to the target of the task book and put it in the envelope at that location. On discovery a wrong location, he/she carried on to find the right one. In this condition, an airport-public-announcement-like sound file was played and broadcasted for the participants in the formal experimental area with the purpose of imitating a real world scenario.

In the CROSSFLOW condition (see Figure 5.14 and 5.15), firstly each participant in a group was given a SPV E200 smart phone on which the vibrations was synchronized with the fish-flow pattern projected on the experimental area, and a new A5-size task book (on each page a number corresponding to a target location in the CROSSFLOW condition was shown). Then the participants in a group were asked to stand in a line at the location "CROSSFLOW start [0]". When the experimenter said "start", they started their navigation tasks immediately and simultaneously. When a participant reached a wall or the end location, he/she uncovered the paper on the wall or table to reveal the information. On discovery of a right location, the participant ripped off the page corresponding to the target of the task book and put it in the envelope at that location, then entered the number corresponding to the next target on the smartphone keypad. On discovery a wrong location, he/she carried on to find the right one. In this condition, a different airport-public-announcement-like sound file was played and broadcasted for the participants in the formal experimental area. The activities of all the participants in both MAP and CROSSFLOW conditions were recorded on digital videotapes.

At the end of the formal experiment, the participants were asked to complete a questionnaire based on NASA-TLX assessing their subjective workload of the tasks. They were also asked to complete a post-experiment questionnaire (see Appendix) which was designed to assess their experience on the CROSSFLOW system that included the questions about the difficulty they experienced during their navigation, and open-ended questions about what factors they thought contributed to successful and unsuccessful navigation tasks in each condition and the suggestions about improvement on the CROSSFLOW system. Their answers were summarized and analysed.

Figure 5.12 Experiment 2: MAP condition

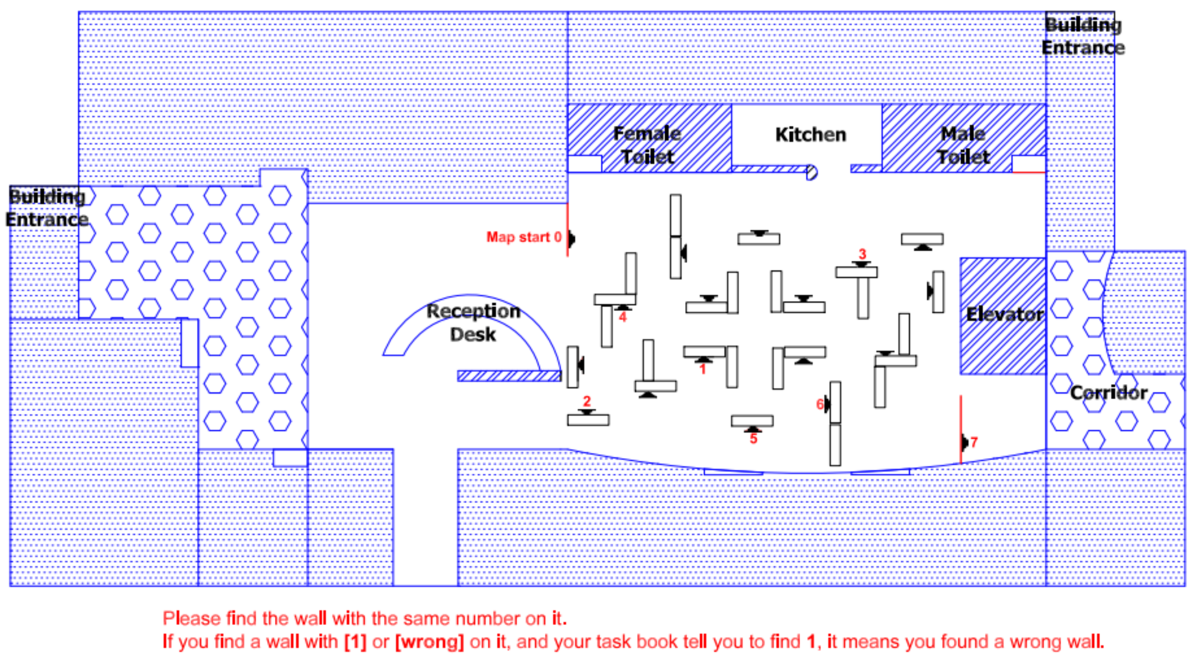

Figure 5.13 The paper map given to participants in the MAP condition.

Figure 5.14 Experiment 2: CROSSFLOW condition

Figure 5.15 Experiment 2: A participant performing navigation tasks using CROSSFLOW

F. Measurements

The following measures of performance were defined and collected; they are also served as the dependent variables in the statistical analysis:

- **Navigation task success rate**: The number of successful navigation tasks versus the number of navigation tasks. Successful navigation task completion was defined as the participant reaching the target location, regardless of time taken and navigation errors.
- **Navigation task time**: The completion times of successfully completed navigation tasks (successful task completion time)
- **Navigation error**: The number of errors of successfully completed navigation tasks. Navigation errors were defined as the times of the participants reaching the wrong locations before successfully reaching the right locations. For example, if a participant reached two distracter walls before identified a right target wall [2] in the CROSSFLOW condition, his/her navigation error performance on task II in the CROSSFLOW condition were logged as 2 times of navigation errors.
- **Subjective task workload**: The score of subjective measurement of task workload using the NASA TLX software.

5.3.3 *Results of Experiment 2*

A. Quantitative results

Statistical analysis was performed on the results for all tasks to understand whether the participants' performances using different navigation aids were significant or not. Before performing statistical analysis the data were tested using Kolmogorov-Smirnov normality test with a Lilliefors significance correction and Shapiro-Wilk normality test to see if they were normally distributed. As the data (the completion times of finding location 2, 3, 4, 5 in the MAP condition and location [3], [6] in the CROSSFLOW condition, the number of errors of finding location 2, 4, 5 in the MAP condition and location [2], [3], [4], [5], [6] in the CROSSFLOW condition) did not follow a normal distribution, non-parametric analysis was carried out using Wilcoxon signed-rank test for within-subjects comparisons. For the data which followed a normal distribution, two-tailed paired samples t-tests were used for within-participants comparisons.

1) **Comparison of navigation task performance**

Three kinds of measures of performance on the navigation tasks, navigation task success rate, navigation task time and navigation error, were compared across the MAP and CROSSFLOW condition.

The performances on some tasks were logged as invalid tasks performances and eliminated from the evaluation as some target locations were forgotten to be found by some participants or not be found because of the shadow that covered the projection of fish-flow pattern, and some participants confused the location number in the MAP condition with the number in the CROSSFLOW condition, accidentally found a target location without using any navigation aid, or misunderstood the experimental setting (e.g. assumed that the ending location was a foam wall instead of a table.) Therefore the results are only reported for the performances on task II, III, IV, V and VI.

Condition	task II	task III	task IV	task V	task VI	Overall
MAP	location 2	location 3	location 4	location 5	location 6	
	9/9 (100%)	10/10 (100%)	8/8 (100%)	9/9 (100%)	9/9 (100%)	45/45 (100%)
CROSSFLOW	location[2]	location[3]	location[4]	location[5]	location[6]	
	10/10 (100%)	8/8 (100%)	9/9 (100%)	9/9 (100%)	10/10 (100%)	46/46 (100%)

Table 5.3 Experiment 2: Navigation task success rates for the MAP and CROSSFLOW conditions.

The navigation task success rates gathered from the experimental video are shown in table 5.3. As the success rate was the same across different tasks in the MAP and CROSSFLOW conditions, no significant statistical difference between the conditions is evident.

Condition	task II	task III	task IV	task V	task VI
MAP	location 2	location 3	location 4	location 5	location 6
	11.08	22.78	26.78	9.68	8.80
CROSSFLOW	location[2]	location[3]	location[4]	location[5]	location[6]
	46.72	29.22	49.92	31.28	22.38

Table 5.4 Experiment 2: Median task completion time in seconds for the MAP and CROSSFLOW conditions.

According to Robson (1994), the data (the completion times of finding location 2, 3, 4, 5 in the MAP condition and location [3], [6] in the CROSSFLOW condition) did not follow a normal distribution, thus the median completion time of successfully completed navigation tasks in seconds for the MAP and CROSSFLOW conditions are shown in table 5.4 instead of the means and the standard deviations.

As shown in table 5.4, the MAP cases managed to a consistently faster time than CROSSFLOW cases across task II, III, IV, V, VI. Wilcoxon signed-rank tests were carried out to investigate the differences of the completion time of the tasks II, III, IV, V, VI between the MAP and CROSSFLOW condition. For the task III ($p = 0.161 > 0.05$) and IV ($p = 0.176 > 0.05$), the pair-wise tests revealed no significant differences between the conditions, which indicates that the participants spent a similar amount of time to complete task III and IV in the MAP condition and CROSSFLOW condition. For the task II ($p = 0.011 < 0.05$), V ($p = 0.012 < 0.05$) and VI ($p = 0.021 < 0.05$), there were significant differences between the conditions, which indicates that the participants used much less time reaching the target location 2, 5 and 6 in the MAP condition than reaching the target locations [2], [5] and [6] in the CROSSFLOW condition.

Condition	task II	task III	task IV	task V	task VI
MAP	location 2	location 3	location 4	location 5	location 6
	0	1.00	0.50	0	0
CROSSFLOW	location[2]	location[3]	location[4]	location[5]	location[6]
	1.00	0	0	0	0

Table 5.5 Experiment 2: Median number of navigation errors for the MAP and CROSSFLOW conditions.

As the data (the number of errors of finding location 2, 4, 5 in the MAP condition and location [2], [3], [4], [5], [6] in the CROSSFLOW condition) did not follow a normal distribution, the median errors of successfully completed navigation tasks for the MAP and CROSSFLOW conditions are shown in table 5.5 instead of the means and the standard deviations.

Wilcoxon signed-rank tests were carried out to investigate the differences of the number of errors of the successfully completed navigation tasks between the two conditions. For the task II ($p = 0.066$), III ($p = 0.18$), IV ($p = 0.915$), V ($p = 0.317$) and VI ($p = 0.18$), the pair-wise tests revealed no significant differences between the MAP and CROSSFLOW conditions, which indicates the similar level of navigation errors in the MAP and CROSSFLOW condition.

2) Comparison of subjective task workload

The mean NASA-TLX scores with standard deviations for the MAP and CROSSFLOW conditions are 64.90 (Standard Deviation = 14.42) and 62.40 (Standard Deviation = 14.70) respectively. The score of the MAP condition is slightly higher than the CROSSFLOW condition. However, a paired samples t-test was conducted and the result showed that the difference between the two conditions was not statistically significant, $t(9) = 0.428$, $p = 0.679$.

B. Qualitative results

The qualitative results from the post-experiment questionnaires and interviews are summarized as follows.

1) Comments about the components of the CROSSFLOW system

a) The fish-flow pattern:

Participants were asked whether they could identify the public directional information corresponding to different targets shown by the fish-flow pattern.

Six out of the ten participants confirmed that they could, whereas the other four said that the directions of flows to some target locations were difficult to be identified. One of the four participants explained that it was because the individual "fish" was too small and not noticeable; another two explained that it was because the presentation time of visual pattern corresponding to a target (i.e. duration of time slot) was too short; and the last one expressed the both opinions.

Four participants suggested that direction indication by visual pattern should be more explicit, and one of them said that arrows may be better than small fish.

Four participants suggested that the presentation time of visual pattern corresponding to a target (i.e. duration of time slot) should be longer.

Five out of the ten participants made positive comments on their overall perception of the fish-flow pattern, and they said that the fish-flow pattern was "beautiful", "attractive" or "vivid".

b) Correlation between private cue and visual pattern:

The participants were asked whether they could identify the correlation between vibration cue from the smartphones and visual pattern corresponding to different targets.

Seven of the ten participants answered "Yes". Of the rest three participants did not make this answer, two said that sometimes they could, while sometimes they could not integrate vibration and synchronised visual pattern into a unified stimulus within an exact time slot. The last one commented when he was alerted by the private cue, he was often too late to identify the directional information from the visual pattern displayed simultaneously with the private cue, in which case he had to wait the next cycle to identify it and even wait additional cycle to confirm his identification. He suggested that duration of time slot should be longer or private cue should precede the visual pattern corresponding to a user’s target.

Another participant suggested that the presentation time of visual pattern corresponding to a user's target should be longer than private cue, that is, the private cue and the corresponding visual pattern should start at the same time, after the private cue stops for a short period of time then the pattern changes into the successive pattern.

c) Interface for personal display:

All participants commented that the interface for the smartphone was easy to be understood and interacted with.

2) CROSSFLOW in comparison with MAP: the usability aspects

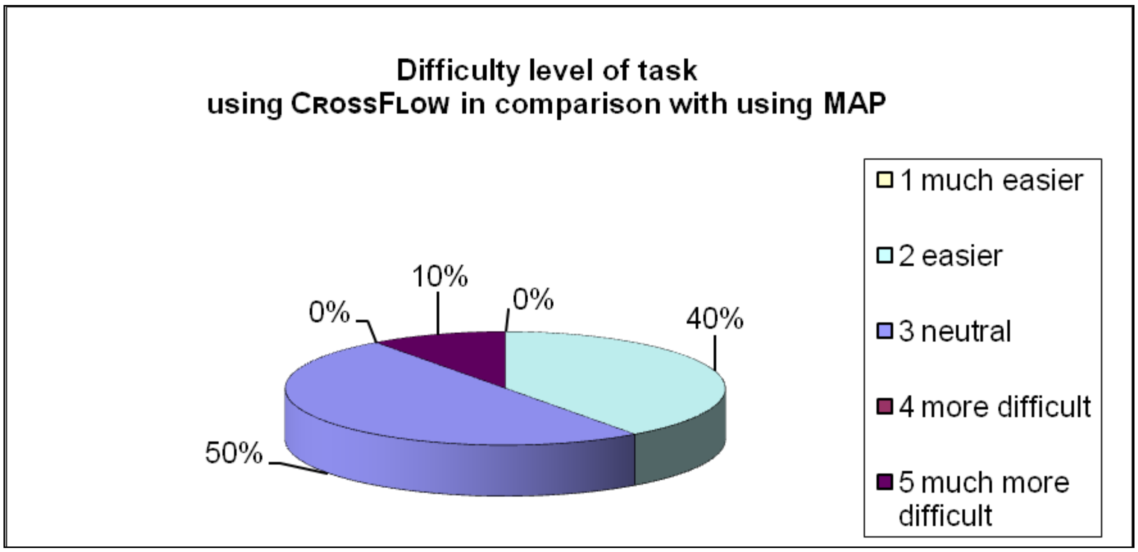

Figure 5.17 Experiment 2: Subjective rating of the difficulty level of navigation tasks using CROSSFLOW in comparison with MAP

A main question in the questionnaire asked participants to rate the difficulty level of navigation tasks in the CROSSFLOW condition in comparison with which in the MAP condition using a 5-point Likert Scale (Sharp et al. 2007) and give explanations. As shown in figure 5.17, "1" represents that using CROSSFLOW is much easier than using hand-held paper map and "5" represents that using CROSSFLOW condition is much more difficult than using map, 50% of responses are "3" which indicates that five out of the ten participants felt that the task difficulty level was similar in the CROSSFLOW condition and the MAP condition; 40% of responses are "2" which indicates that four out of the ten participants felt that the task difficulty level in the CROSSFLOW condition was lower than which in the MAP condition; and 10% of responses are “5” which indicates that only one of the ten participants felt that using CROSSFLOW was much more difficult than using map.

The comments of the participants regarding the usability of CROSSFLOW in comparison with MAP are grouped as follows:

a) The shadows of people covered the flowing fish thus affected viewing the visual pattern.

→ This was indicated by the only participant who rated 5 as the task difficulty level. She complained that when she navigated in the experimental environment, the shadows of other people sometimes covered the projected visual pattern thus affected her task performance.

b) The current system just works in dark environment instead of a normal bright environment.

→ This was also indicated by the same participant who rated 5 as the task the difficulty level.

c) Time-consuming at each decision point.

→ This opinion was expressed by six participants. Five of them said that the presentation time of visual pattern corresponding to a target (i.e. duration of time slot) was so short that they often needed two or three cycles to identify and confirm a direction at each crossway. The last one said that the cycle time was too long, and she felt that she had to wait a few seconds for a vibration.

d) CROSSFLOW did not work well in such kind of environment. However using CROSSFLOW might be easier than using hand-held map in larger or open spaces without so many partitioned spaces.

→ This opinion was expressed by three participants.

e) CROSSFLOW should be used to support navigation in the situation where there is a little or no time pressure, such as navigating in museums.

→ This opinion was expressed by one participant.

The comments of the participants which are evidently positive about the CROSSFLOW system are grouped as follows:

a) There was no need to learn and remember the environment using CROSSFLOW thus saved the time for navigation, while using map to navigate required navigators to learn and remember the spatial information from both map and environment.

→ This was indicated by most of the participants. The representative comments are:

"You can ignore the environment surrounding you."

"It is easy to identify the direction, no need to identify the landmarks in an environment."

b) CROSSFLOW helps users to identify directions to their destinations and also indicates best routes. Users just follow the direction of a flow at each decision point to the destination or observe several flow directions in the open areas to get an approximate direction to their destination. The navigation aid is easy to use and better than using hand-held map.

→ This was indicated by most of the participants. The representative comments are:

"No matter where I was, CROSSFLOW helped me with the direction."

"It seems to be much more direct and easier to understand than a map. If people get familiar with CROSSFLOW, it should be very useful."

"I have no sense of direction, using map is more difficult. CROSSFLOW is easier and more distinct."

"It is a new way to find the accurate target location. No need to understand map or people's instruction."

"Using CROSSFLOW to navigate is not only better than using a map, but also better than following signs or landmarks and speech guidance."

c) CROSSFLOW is a personal navigation aid and preserves certain level of privacy and anonymity in public space.

→ This was explicitly indicated by one of the participants on the questionnaire, which was also considered by her as one of the most distinct advantages of CROSSFLOW. She commented *"It is good for personal use, everyone has a private guider."* In the open interview after the formal experiment, many participants expressed the similar opinion and gave the reason that others could not perceive the vibrations issued from their smartphones on their bodies. However they also said that where they would go could be identified if someone actively observed their behaviours.

3) CROSSFLOW in comparison with MAP: the user experience aspects

Participants were asked to rate the following user experience aspects of CROSSFLOW in comparison with MAP.

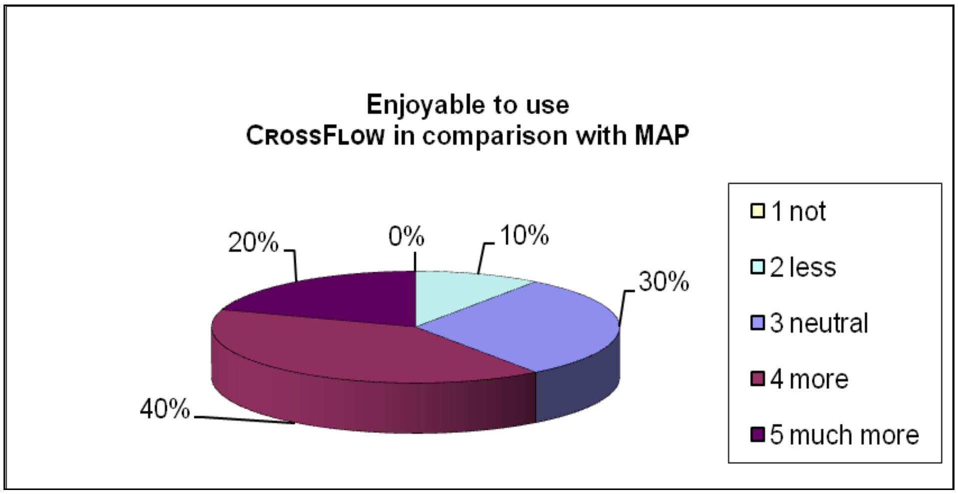

Figure 5.18 Experiment 2: Subjective rating of enjoyable to use (CROSSFLOW in comparison with MAP)

As shown in figure 5.18, “1” represents that using CROSSFLOW is totally not enjoyable than using MAP and “5” represents that using CROSSFLOW is much more enjoyable than using MAP, 20% of responses are “5” which indicates that two out of the ten participants felt that using CROSSFLOW was much more enjoyable than using hand-held paper map; 40% of responses are “4” which indicates that four out of the ten participants felt that using CROSSFLOW was more enjoyable than using hand-held paper map; 30% of responses are “3” which indicates that three out of the ten participants felt that using CROSSFLOW and using hand-held map were similarly enjoyable; and 10% of responses are “2” which indicate that only one of the ten participants felt using CROSSFLOW was less enjoyable than using hand-held paper map.

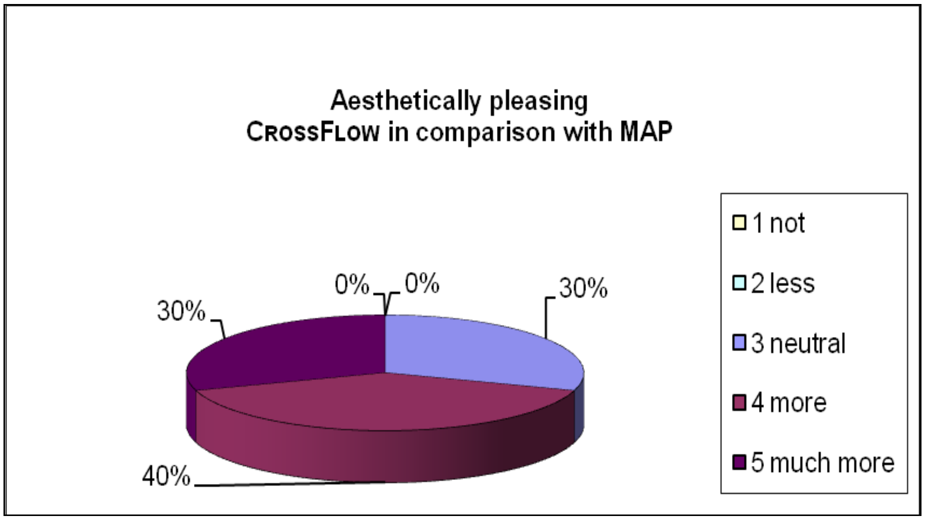

Figure 5.19 Experiment 2: Subjective rating of aesthetically pleasing (CROSSFLOW in comparison with MAP)

As shown in figure 5.19, "1" represents that CROSSFLOW is totally not aesthetically pleasing than paper map and "5" represents that CROSSFLOW is much more aesthetically pleasing than paper map, 30% of responses are "5" which indicates that three out of the ten participants felt that CROSSFLOW was much more aesthetically pleasing than paper map; 40% of responses are "4" which indicates that four out of the ten participants felt that CROSSFLOW was more aesthetically pleasing than paper map; and 30% of responses are "3" which indicates that three out of the ten participants felt that CROSSFLOW and paper map were similarly aesthetically pleasing.

As shown in figure 5.20, "1" represents that navigational information delivered by CROSSFLOW is not comprehensible than paper map and "5" represents that navigational information delivered by CROSSFLOW is much more comprehensible than MAP, 70% of responses are "4" which indicates that seven out of the ten participants felt that navigational information delivered by CROSSFLOW was more comprehensible than MAP; 20% of responses are "3" which indicates that two out of the ten participants felt that comprehensibility of navigational information delivered by CROSSFLOW and MAP was at the similar level; and 10% of responses are "2" which indicates that only one out of the ten participants felt that navigational information delivered by CROSSFLOW was less comprehensible than MAP.

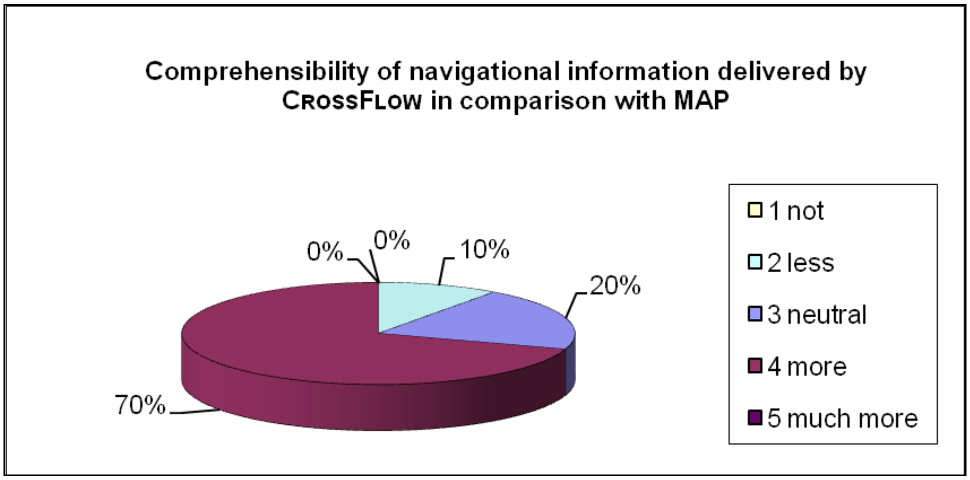

Figure 5.20 Experiment 2: Subjective rating of comprehensibility of navigational information delivered by CROSSFLOW in comparison with MAP

4) Overall user experience of CROSSFLOW

Participants were also asked open-ended questions about the overall experience of CROSSFLOW in the experiment. Most of the participants liked the idea of using CROSSFLOW to support indoor navigation and their comments were generally positive. In particular, three participants commented that using CROSSFLOW to navigate was fun and interesting.

5) Other suggestions from the participants

Some suggested improvements in the design and evaluation of CROSSFLOW system from the participants have been integrated into the above sections; the rest is summarized as follows:

a) Should indicate distance from the location of users to their destinations.
b) Should use audio private cue instead of vibration.
c) Should compare CROSSFLOW with map posted on the wall instead of hand-held map (suggested by one of the participants, and his original suggestion is *"It was not really comparable between CROSSFLOW and hand-held map, and it would be more comparable between CROSSFLOW and maps posted on the wall."*).

C. Observations of participants' behaviour

From the observation of the participants' activities recorded on the videos (analysed by the experimenter frame by frame in some parts), the participants tended to pause and walk more within the experimental area in the CROSSFLOW condition than in the MAP condition. In the CROSSFLOW condition, the participants sometimes walked to the border of experimental area and paused there for a while in order to gain the perspective on the projected visual pattern, then kept approaching a potential target location. When approaching a potential target wall, instead of going straight to uncover the number on it, some participants stepped away from it and paused for a while observing the flow pattern and waiting their private cues to confirm their decisions. Some of them had such behaviour several times before they reached a wall and uncovered the number on it. Such behaviour may be because they need to get a perspective of more flows towards that location from different directions and collect more directional information as the evidence of a target location. However, in the MAP condition the behaviours above were not very explicit, when approaching a potential target wall, the participants tended to go straight to uncover the number on it. The participant often paused in the MAP condition as well, which may be because they often needed to stop to identify the position of themselves and a target on the map and plan a route.

The participants were also observed that they tended to mistake a distractor wall nearby a target wall for the target location in both the CROSSFLOW and MAP conditions.

5.3.4 *Discussion of Experiment 2*

A. Utility and usability of CROSSFLOW

Excluding the invalid tasks, in the CROSSFLOW condition the participants found the target foam walls among many visually similar walls in a complex indoor environment which contains many small and partitioned spaces using CROSSFLOW with 100% success rate regardless navigation errors. Although they only used vibration cue combing with the fish-flow visual pattern, they were still able to complete the tasks within a reasonable period of time. This result proves the success of CROSSFLOW's user-interface and interaction model and supports our hypothesis that CROSSFLOW could support multiple simultaneous users to perform navigation tasks in an unfamiliar complex indoor environment. The result of the navigation task success rate in the CROSSFLOW condition is also a evidence that the designed components of CROSSFLOW including regular cue mapping, abstract public directional information (the projected fish-flow pattern) and private cues (the vibrations from smartphone) can be used for personalized directional information presentation for multiple simultaneous users, and that users are able to identify the correlation between private cues and public directional information corresponding to a target location or object with less than five minutes training, and then reach it successfully.

B. CROSSFLOW in comparison with MAP

Looking at CROSSFLOW in comparison with hand-held paper map case indicates a similar level of navigation task success rate, navigation error, subjective task workload and completion time of two successfully completed navigation tasks (tasks III and IV), because no statistically robust difference was able to be drawn from the data. However, the mean navigation task time in the CROSSFLOW condition is more than which in the MAP condition on all the five navigation tasks, moreover, the participants spent significantly more time in the CROSSFLOW condition than in the MAP condition on completing three out of five navigation tasks. Such results are inconsistent with the results of the Experiment 1 that the task completion time in the CROSSFLOW condition was significantly shorter than which in the MAP condition. This may be on account of the following factors, which were also confirmed in the participants' qualitative feedbacks.

1) Increased density of decision points in the experimental area.

Although the size of the experimental area in Experiment 2 is similar to which in Experiment 1, in order to make the navigation tasks more demanding than which in Experiment 1, 28 foam walls partitioned the experimental environment into many dense and small spaces. As a result, in comparison with the environmental area of Experiment 1, much more decision points were densely distributed in this small scale area (at least four decision points between two target locations). In contrast, there were much fewer decision points in Experiment 1 as there were only targets and distractor targets within the experimental area and no obstacles.

When using CROSSFLOW to navigate, the participants have to rely on the directional information provided by CROSSFLOW at each decision point, where they need to wait for a private cue so as to identify a direction to a target location, which may take them a whole cycle time. Sometimes they may wait one more cycle time to confirm the identified direction. This was confirmed by six out of the ten participants on the post-experiment questionnaire. Thus more decision points on the way to a target location means more cycle time the participants have to spend.

In comparison with Experiment 1, more and denser decision points were distributed in the navigational environment, thus the participants may also take longer time to learn the map and the environment in the MAP condition, but they may not need to stop or spend time as long as in the CROSSFLOW condition at each decision point during navigation.

Drawing from the analysis above, it can be concluded that the density of decision points in such a small scale experiment area may have a greater impact on the navigation task completion time in the CROSSFLOW condition than in the MAP condition. This inference is supported by the observations of the participants' behaviour in the two conditions. When the participants encountered a decision point after walking every two or three steps, they apparently paused longer and more frequently in the CROSSFLOW condition than in the MAP condition.

Therefore, the worse time performance in the CROSSFLOW condition might be attributed to the increased density of the decision points in the experimental area, which might also be the cause of a complaint about the duration of cycle time of CROSSFLOW. One of the participants complained that the duration of cycle time of CROSSFLOW was too long, which is probably because she stopped and waited for her private cues too frequently.

2) **Shadows of the obstacles and participants in the CROSSFLOW condition.**

In the CROSSFLOW condition, although two projectors had been employed to project the fish-flow pattern to avoid shadows caused by the foam walls and other obstacles within the experimental area. The shadows near some target locations still existed and blocked the projected visual pattern. Moreover, the shadows of the participants themselves could block their views of the projected pattern as well, which may also impair their navigation task performances. The effect of shadows on the participants' task performances was confirmed by one of the ten participants on the questionnaire and other participants in the interview. Although the data relevant to the shadowed target locations was removed from the statistical analysis (as mentioned in Section 5.3.3 A. 1)), the participants' task performances may still be impaired due to the shadows.

3) The fish-flow pattern

The worse time performance in the CROSSFLOW condition may also be attributed to the design and configuration of the visual pattern.

The identification of public directional information corresponding to a target is highly related to navigation task performance, which may involve two processes. The first is the integration of the private cue and the visual pattern corresponding to a target; the second is the identification of the public directional information from the visual pattern corresponding to a target. The design and configuration of visual pattern play important role in the second process. If the design and configuration of visual pattern was not satisfactory for indication of different locations or objects in an indoor environment which contains many small and partitioned spaces, the identification of public directional information from the visual pattern would be difficult. As a result, the navigation performance using CROSSFLOW would be worse than using hand-held map in that kind of environment.

The fish-flow pattern was clear and explicit for indication of locations or objects in an open navigational environment like the environment in Experiment 1, which resulted in better performances in the CROSSFLOW condition than the MAP condition on almost all measures. However, on the post-experiment questionnaire of Experiment 2, four of the ten participants (40% of responses) commented that directional information relevant to them was difficult to be identified from the fish-flow pattern, thus the pattern should be re-designed; and four participants (40% of responses) indicated that the indication of different target locations of the fish-flow pattern should be more explicit. These results and the observation of participants' behaviour in Experiment 2 suggested that the fish-flow pattern was not very satisfactory for indication of several different locations or directions in an indoor environment which contains many small and partitioned spaces, in detail, the flowing directions near in time and the convergences of flows indicating target locations near in space were not very clear and explicit.

The results of the comparison of navigation task success rate and navigation error are consistent with the results of Experiment 1, which suggests that using CROSSFLOW to perform relative more demanding navigation tasks is still as successful and accurate as using a hand-held map.

The difference of the score of NASA-TLX subjective workload for the CROSSFLOW condition and the MAP condition is not significant (Subjective workload: CROSSFLOW = MAP). This result is inconsistent with the result of Experiment 1 that the scores of NASA-TLX subjective workload for the CROSSFLOW condition is significantly lower than the MAP condition (Subjective workload: CROSSFLOW << MAP). This result may due to the same factors that discussed in the analysis of the results of navigation task time performance above.

5.4 Discussion of Both Experiment 1 and 2

The quantitative and qualitative results of Experiment 1 and 2 validated these distinct advantages of CROSSFLOW:

A. Facilitate navigation in certain type of indoor environment and support multi-tasking while navigating.

Wright et al. (1993) stated that finding a particular destination could be difficult in many modern building complexes, where there were many visually similar spatial structures such as corridors, doorways and visually similar information or objects. Among those buildings, some of them contain relatively less and sparser decision points and there are many possible target objects, directions or destinations at each decision point. Some of the indoor environments contain relatively more and denser decision points (i.e. the distances between decision points are small) and there are less than four possible target objects, directions or destinations at each decision point.

The results of Experiment 1 and 2 imply that when a user uses an aid to navigate or search physical objects, in comparison with hand-held paper map, CROSSFLOW probably is more efficient and impose much lower workload on users in the former type of indoor environment, while not so efficient and impose a similar level of workload on users in the latter type of indoor environment. Moreover, CROSSFLOW supports multi-tasking while navigation and is as effective and successful as hand-held paper map in both types of indoor environment.

B. Support personalized navigational information presentation for multiple simultaneous users without resorting to location tracking.

In contrast to the GAUDI system (Kray et al. 2005) reviewed in Section 4.2.2, the results of Experiment 2 show that CROSSFLOW scales well with multiple simultaneous users. All users successfully completed their navigation tasks. Moreover, an increased number of users would not likely result in any deterioration of system performance.

Technically, when using CROSSFLOW to navigate, there is no need for tracking individual users or for network connectivity beyond the initial registration, which is particularly difficult for indoor environments. Beyond selecting the target location and synchronizing the time, no further interaction is needed between the server of CROSSFLOW and user's mobile device. Although in Experiment 1 and 2, target locations were predefined and devices were synchronized before being given to the participants, it is straightforward to realize these steps using only unidirectional transmissions (e.g. from the server to the mobile device). Not only does this reduce the cost of sensors and the requirements imposed on the mobile device (e.g. it only needs to be able to receive information but not to send it after initial registration), but also contribute to the increased degree of privacy and anonymity when people navigating.

C. Preserve certain level of privacy and anonymity in physical public spaces.

As reviewed in Chapter 2, a person's privacy and anonymity in public space would be impaired when the relationship between the person and the information he/she owned or interacted with and the information content were exposed to occupants in the same space. For example, when a person constantly interacts with his private information on his mobile phone screen in a shopping mall or a railway station, his privacy and anonymity in that space will be impaired as his acts of interaction and information content on his mobile phone are exposed to others by shoulder surveillance.

CROSSFLOW only need a user to interact with his mobile device once to download the schedule of the time slots (i.e. the cue mapping) for an environment. No further interaction is needed during the navigation process unless the user needs to select a new destination. During navigation, the relationship between the user and the information he owned (a specific pattern relevant to the user in this case) is hardly identified by other occupants in the space. Furthermore, the content of the information relevant to a user (the directional information in this case) is represented by abstract ambient patterns and can only be decoded in conjunction with the user's private cues (i.e. the vibrations from the user's mobile phone) that can hardly be detected by other occupants of the space. Therefore, a certain level of privacy and anonymity of the user could be preserved without hindering the user’s interaction with, and access to, public space, though there is a possibility to undermine user’s privacy and anonymity by hacking the server and the user’s personal mobile device or constantly observing and following the user.

The feedbacks of the participants of Experiment 2 (i.e. Section 5.3.3 B. (2) (c) in the comments of the participants which are evidently positive about the CROSSFLOW system) have proven that preserving user's privacy and anonymity in public space is a distinct advantage of CROSSFLOW over traditional map-based and tracking-relied navigation aids or approaches.

D. Aesthetically pleasing and fun to play with.

This advantage has been demonstrated in the qualitative results of both Experiment 1 and 2.

5.5 Conclusions

This chapter has reported the evaluation of CROSSFLOW V1, the first version of Crossmodal Display for indoor navigation. The evaluation contains two user studies which investigated the usability and the user experience of CROSSFLOW as an indoor navigation aid in comparison with traditional hand-held paper map.

In the first user study (Experiment 1), individual participant was given dual tasks to perform using either CROSSFLOW V1 prototype or hand-held paper map. In the second user study (Experiment 2), multiple simultaneous users navigated within a maze-like complex indoor environment using either CROSSFLOW V1 prototype or hand-held paper map. The quantitative and qualitative results of two experiments have proved the success of the design of CROSSFLOW and our conceptual model of Crossmodal Display, and validated the effectiveness of CROSSFLOW as a solution to help to solve the problems identified in Chapter 2 in indoor navigation application domain, in particular, validated these distinct advantages over traditional navigation tools and existing technologies we reviewed:

- Facilitate indoor navigation and support multi-tasking while navigating.
- Support personalized navigational information presentation for multiple simultaneous users without resorting to location tracking.
- Preserve certain level of privacy and anonymity in physical public space.
- Not use or only use limited sensing or tracking.
- Aesthetically pleasing and fun to play with.

The results of also showed that the quantity and density of decision points in the indoor environment have greater impact on the navigation efficiency and workload on users when using CROSSFLOW than using traditional hand-held paper map. Moreover, the usability and the user experience of CROSSFLOW were impaired not only because of environmental factors such as the quantity and density of decision points and the shadows of the obstacles and the occupants of the environment, but also the components designed for CROSSFLOW V1 including the fish-flow pattern and short beeps and vibrations.

In order to improve the usability and the user experience of CROSSFLOW, we developed a new version of CROSSFLOW (CROSSFLOW Version 2) based on further analysis of findings of Experiment 1 and 2 and implications from them, the process of which is reported in Chapter 6. The evaluation of CROSSFLOW V2 is reported in Chapter 7.

Chapter 6. Redesigning CROSSFLOW

6.1 Introduction

Chapters 4 and 5 have reported our initial exploration of the design and evaluation of a Crossmodal Display for indoor navigation, CROSSFLOW V1. The results of evaluation not only validated the success and advantages of CROSSFLOW, but also revealed several aspects for improvement including cue mapping, public cue and private cue.

In order to improve those aspects to achieve better usability and user experience of CROSSFLOW, we further explore the design space of CROSSFLOW and report the design iteration in this chapter. First, further analysis of the findings of the previous user studies (Experiments 1 and 2) together with the implications of multimodal perception and crossmodal integration for interface design in Chapter 3, are drawn upon to derive several implications for CROSSFLOW re-design. Then these implications are referred in the re-design process, including public directional information re-design and private cue re-design. Finally, the chapter describes prototype system components and functions developed for CROSSFLOW Version 2.

6.2 Analysis of Findings of Experiment 1 and 2 & Implications

6.2.1 *Implications 1: Visual pattern should be changed.*

The results and the observations of participants' behaviour in both Experiments 1 and 2 suggest that the fish-flow pattern is not satisfactory. Furthermore, we identified that the identification of public directional information from the visual pattern corresponding to a target is related to several factors, including the size and shape of individual graphical element, the animation of the visual pattern and the projection. We analyse these factors in detail below so as to draw implications for our public directional information re-design.

A. Size and shape of individual graphical element

The shape of the fish is represented by a circle connecting to a short bar, which is not as directional as the shape of an arrow. Moreover, the size of the fish was criticized that it was too small. If the size of individual graphical element was increased and the shape was more directional, when CROSSFLOW users are navigating, they could see individual graphical element and the direction indicated by individual graphical element more clearly, furthermore, they may be able to identify a destination indicated by multiple graphical elements and distinguish two destinations nearby more easily.

B. Animation of the visual pattern

In the experiments, the participants have difficulty to distinguish two successive flow directions in the fish-flow pattern and they tend to mistake the public directional information presented in the former or later time slot as the correct public directional information corresponding to their private cue. This might be because they are distracted by the transition animation between two successive flow directions. In the animation, the fish in different projection areas rotate different angles in order to orient themselves to a successive flow direction.

Thus removing the transition animation might be better. Moreover, using fish flow motion to indicate different directions may not be as good as fixing the location of the graphical element to point to them.

C. Projection

Using projection as public display has two limitations. One is that the darker lighting is required in the indoor environment where CROSSFLOW is deployed as the brightness degree of projection is relative low. The other is that the shadows of obstacles and the occupants of the indoor environment caused by the projection can block CROSSFLOW users' view of the projected visual pattern. These limitations might be able to be solved by projecting the visual pattern on the walls and ceilings or using other types of public displays such as floor lights, large plasma screens or LED displays which can be embedded in the indoor environment.

D. Distance information

It was suggested in the participants' feedback in Experiment 2 that the distance information, e.g. how far from the location of a user to his/her destination is, should be shown in the visual pattern. As the common indoor environments are much smaller in comparison with large scale outdoor environments, the distance information may be not so important for indoor navigation; furthermore, adding distance information to the visual pattern of the CROSSFLOW system may impose higher attention and cognitive load on CROSSFLOW users.

6.2.2 *Implications 2: The most appropriate duration of time slot should be identified.*

We have indicated in Chapter 5 that the identification of public directional information corresponding to a target of CROSSFLOW users may involve two processes: the process of the integration of temporally congruent but spatially incongruent private cue and visual pattern corresponding to a target, and the process of the identification of public directional information from the visual pattern.

If the duration of time slot (which equals to the presentation time of each visual pattern in CROSSFLOW V1) is too short, users could not complete the two processes during the first presentation of the visual pattern corresponding to their target. Furthermore, they had to wait the next presentation to complete the processes and even wait additional presentation to confirm their identification of the public directional information corresponding to their target, which means that they may need several full *cycle time* at each decision point in an indoor environment. This would result in poor navigation efficiency. The same result would happen if the duration of time slot is too long and the number of time slots in a cycle (which equals to the number of destination) is certain, as the length of cycle time is calculated by the duration of time slot multiplies the number of time slots in a cycle, the longer duration of time slot is, the longer cycle time will be.

Although only the former case and relevant complaints were found in the findings of previous experiments, the analysis above implies that the appropriate duration of time slot, which is as short as possible (to avoid long cycle time) while long enough for CROSSFLOW users to complete the two processes during the first presentation of the visual pattern corresponding to a target, should be investigated. The investigation was done in a user study (Experiment 3) reported in Chapter 7, in which we compared the effect of different duration of time slot and cycle time on users' task performance.

6.2.3 *Implications 3: Users need more training and the system lag of synchronization should be controlled.*

Two of the ten participants in Experiment 2 reported that sometimes they could not integrate vibration from their smartphones and visual pattern corresponding to their target into one exact time slot.

There are two possible causes for this. Firstly, it might be because that the participants were unfamiliar with the integration of temporally congruent but spatially incongruent multimodal stimuli, i.e. the vibrotactile private cue and the synchronized visual pattern. Depending on individual difference in multisensory integration capability, some people may have difficulty to maintain the simultaneity across modalities. The crossmodal simultaneity would be more difficult to be maintained if a user just paid attention to the visual pattern corresponding to his target far away from his/her local position during the smartphone vibrated. If this is the cause, then more guidance and training on using CROSSFLOW would be necessary.

Secondly, although we have tried to make the synchronization of private cue and each visual pattern displayed within the same time slot as precise as possible, when synchronizing the clocks on smartphone with the server, there might still be a system lag between them which was not compensated by the software. If the system lags was larger than the simultaneity threshold for input from visual and vibrotactile sensory modalities, the user might mistake the public directional information shown in the successive time slot for the directional information relating to his/her target. Although the system lag could not be avoided completely, it should not exceed visual-auditory and visual-haptic simultaneity threshold reviewed in Section 3.3.2.

6.2.4 *Implication 4: Private cue should be presented earlier than visual pattern corresponding to a target*

It was suggested in the findings that a private cue should precede the visual pattern corresponding to a target. A range of psychological evidences have demonstrated that an irrelevant sound can augment the perception of both concurrent and subsequent visual event (McDonald et al. 2000; Vroomen & Gelder 2000; see Driver & Spence 1998 for a review). More specifically, Sheth and Shimojo (2004) indicated that an auditory cue maximally enhanced visual salience if the sound preceded the visual target by 50~100 milliseconds. Therefore, to present private cue earlier than the corresponding visual pattern indicating a target may facilitate CROSSFLOW users to capture the visual pattern relevant to them and identify their target via the visual pattern. In an ideal situation, the users may be able to identify their target when the visual pattern relevant to them is shown to them at the first time.

6.2.5 *Implication 5: There should be a semantic correlation between private cue and corresponding visual pattern.*

In Experiment 1 and 2, private cue and visual pattern corresponding to different targets were only temporally correlated, and they have no semantic link between each other.

According to the implication in Section 3.3.4 that a semantically appropriate coordination of modalities may improve the user experience although where user performance may not directly benefit from the effect of semantic factors, a semantically appropriate coordination between private cue and visual pattern may not increase the efficiency of CROSSFLOW directly, but may be able to enhance the salience of visual pattern corresponding to each target and improve the user experience afforded by CROSSFLOW. For example, if natural everyday sounds were used as private cue which have a semantic link to visual pattern animation that simulates natural phenomenon, they would make more sense for a user than abstract beep sound. This implies that the design space of semantic correlated and aesthetic pleasing visual-audio and visual-tactile information combinations should be explored in the re-design of private cue and public directional information.

6.3 CROSSFLOW V2 Prototype System Description

CROSSFLOW Version 2 (CROSSFLOW V2) was developed based on the analysis of findings of Experiment 1 and 2 and implications in Section 6.2. The system components and functions of public displays, personal displays, and private cue configuration tool installed on the smartphone were not changed. We mainly developed a tool on the server supporting design and configuration of the grass-like pattern and the auditory private cue and synchronization. The tool was developed in *Pure Data* (Puckette 1996) environment on a PC. How a user navigates using CROSSFLOW V2 is the same to using CROSSFLOW V1.

6.3.1 *Public directional information re-design*

Based on the previous public directional information design in Section 4.3.4, the analysis of findings of Experiment 1 and 2 and implication 1, and inspired from a grassland in a natural environment (see Figure 6.1) and a natural phenomenon, grass blowing in wind in wild, a grass-like ambient pattern was created in Pure Data (with Gem library) for CROSSFLOW V2. The pattern is composed of a large amount of evenly-distributed grass-like graphical elements (see Figure 6.2 and 6.3). When it is projected on the floor, it can be given a visual intensity that integrates with the floor, which provides the impression of "green grassland" (see Figure 6.4 and 6.5). The way that the grass-like pattern represents directional information corresponding to different destinations is similar to the arrow-like pattern but much more clearly. During a time slot the entire blades of grass point towards a particular destination. Each blade of grass points in a specific direction depending on its projected position in relation to that destination in the physical space (see Figure 6.3 and 6.4). We expect that a user can reach his/her destination by following the local pointing directions of either one blade or the collection of blades.

In comparison with the fish-flow pattern and arrow-like pattern, the following changes and features were introduced into the grass-like pattern:

A. Shape of individual graphical element

Based on the analysis in Section 6.2.1 A., the shape of individual graphical element was designed like the blade of grass and the apex of the blade could be configured to point to different directions when the pattern is projected on a two-dimensional surface. Such shape appears more directional and clearer than the shape of fish in the fish-flow pattern and aesthetically pleasing than the shape of an arrow.

Figure 6.1 A grassland in a natural environment for the inspiration of the "grass-like" pattern

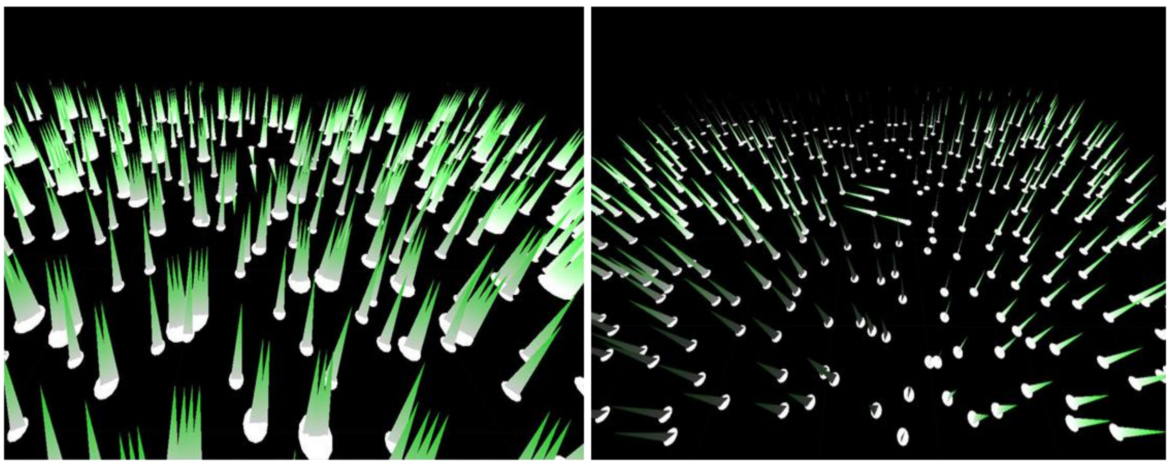

Figure 6.2 Screenshots of the grass-like pattern

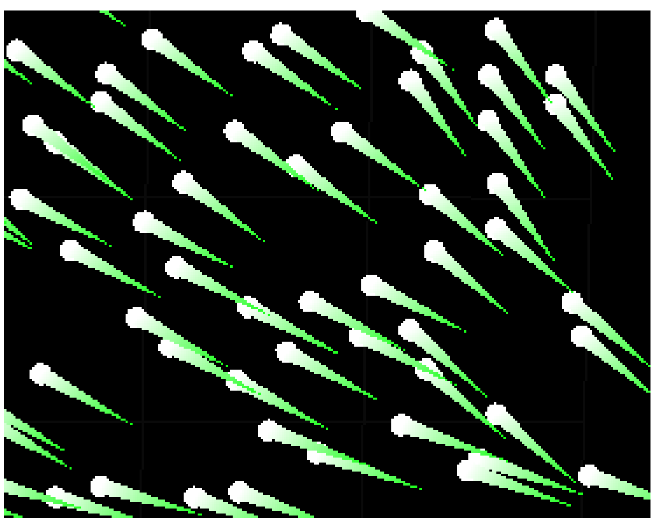

Figure 6.3 A screenshot of the grass-like pattern: each blade of grass points in a specific direction to the destination

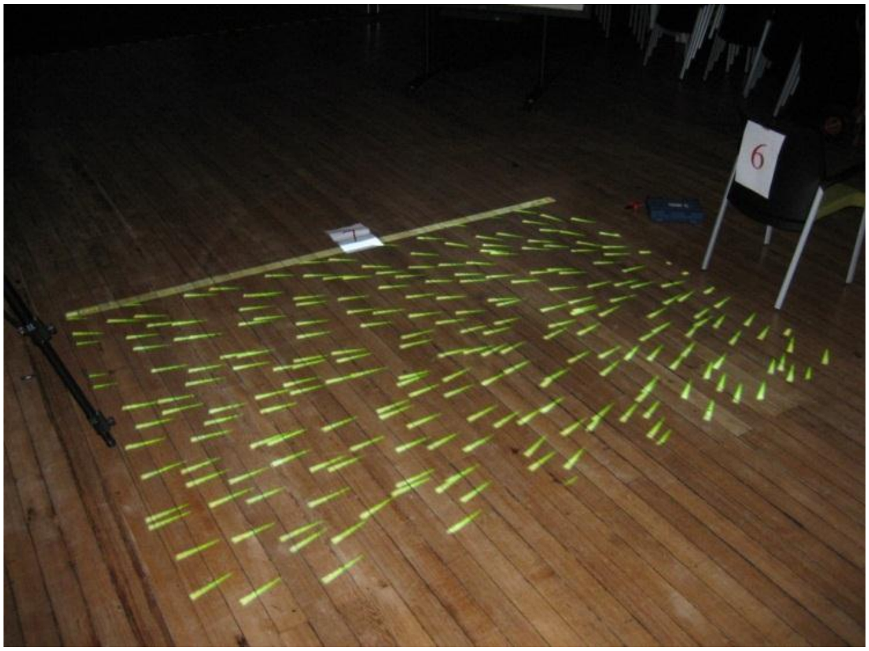

Figure 6.4 **A grass-like pattern projected on a wooden floor in the test environment indicating a specific location or object within a designated time lot.**

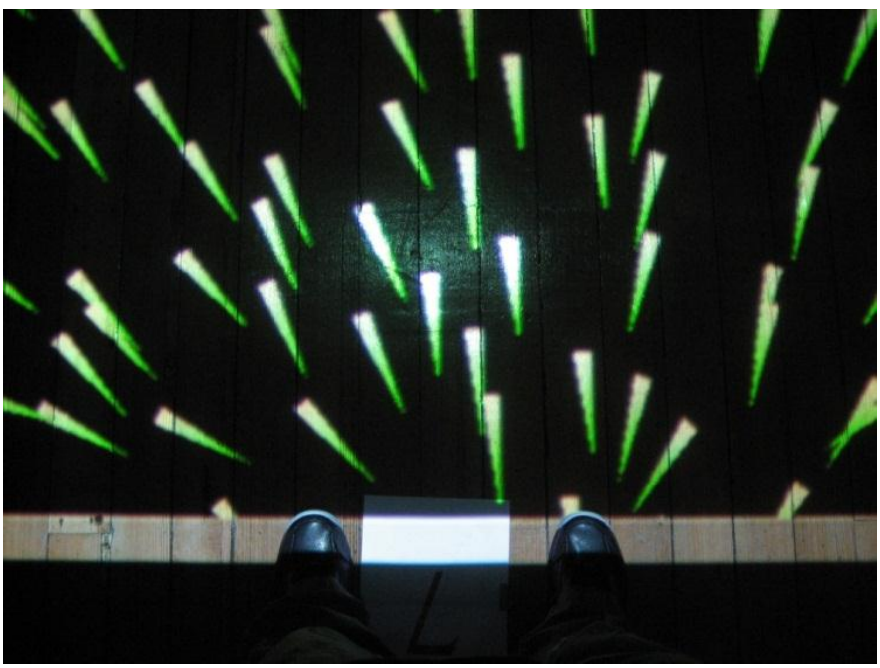

Figure 6.5 **A close-up of the grass-like pattern projected on a wooden floor in the test environment.**

B. Size of individual graphical element

The individual blade of grass was approximately 20-centimeter-long when the static grass-like pattern was projected on a floor (see Figure 6.5), thus it appears much bigger than projected individual fish or arrow. The reason of such change is also based on the analysis in Section 6.2.1 A.

C. Animation of the pattern

The presentation of the grass-like pattern contains a “flickering animation”, in which, all blades of grass lean towards the direction corresponding to the destination and then lean back, and the length of the blade of grass changes five times within a time slot to simulate the natural phenomenon that the blades of grass flick in a direction for a period of time in wind. This animation was made to not only make the pattern appear natural, vivid and aesthetical pleasing for viewers (both CROSSFLOW users and non-users), but also cause the CROSSFLOW users to pay attentions to the visual pattern relevant to them easily. When the flickering animation corresponding to one destination is completed, the blades of grass immediately start to "grow" (i.e. change length and lean) towards the direction corresponding to the next destination. Unlike the transition animation in the fish-flow pattern and the arrow-like pattern in which the graphical elements undergoes an animated transition in configuring themselves for the next destination, no transition animation between any two pointing directions near in time in the grass-like pattern. Such change was made in order to avoid distracting and confusing users so that the users could distinguish different directions more easily.

6.3.2 *Configuration component for the grass-like pattern*

On the server, one of the components in the configuration tool enables a designer to interactively configure the pointing directions of the blades of grass corresponding to different destinations sequentially (see Figure 6.6) and specify the parameters that control the size, quantity, density, distribution areas (see Figure 6.7) and dynamic properties (i.e. animation) of the individual blade of grass, as well as the duration and the quantity of time slot within a cycle.

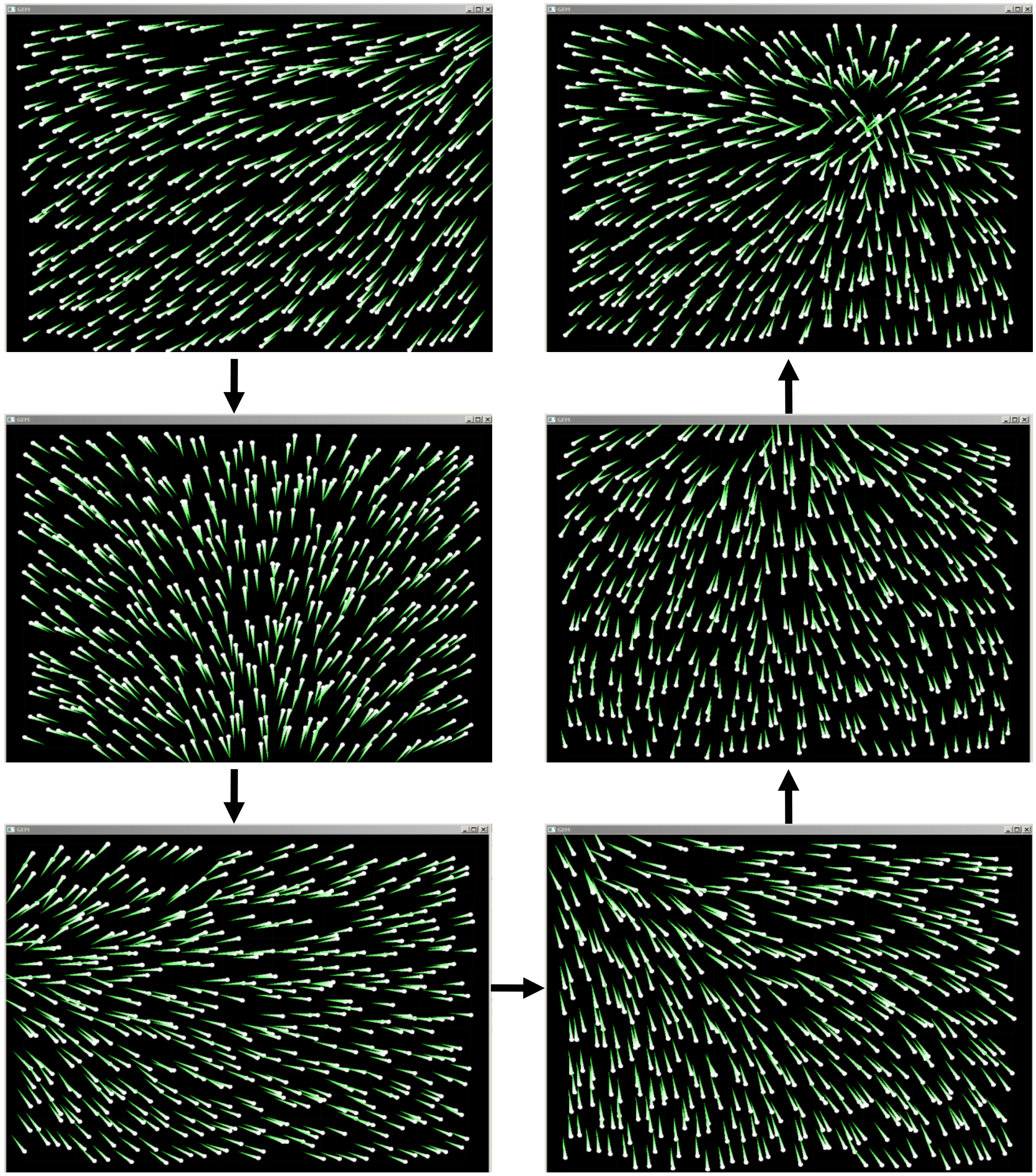

Figure 6.6 Screenshots of the grass-like patterns corresponding to different destinations

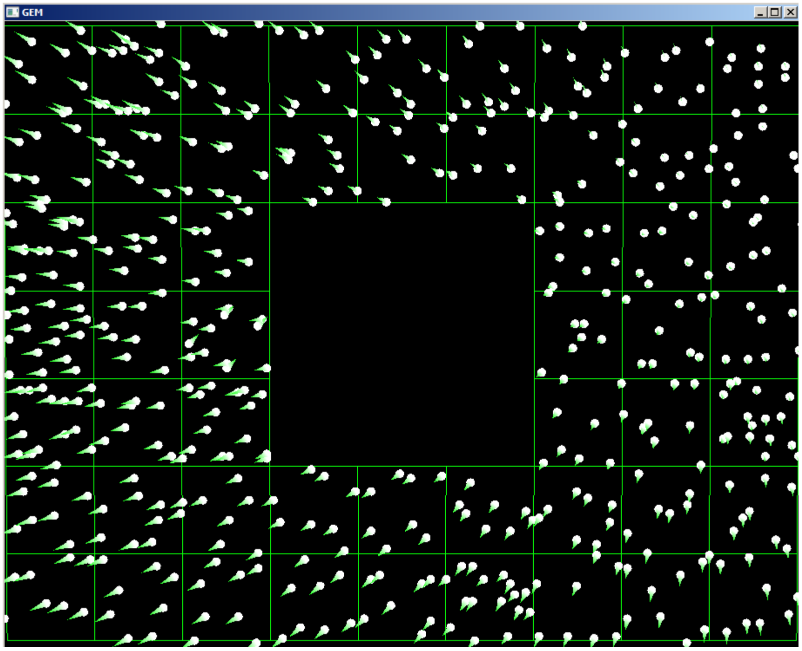

Figure 6.7 An example of how the blades of grass are distributed around the central area of the screen (the green grids is made for the designer's convenience for viewing the distribution area of the grass)

6.3.3 *Private cue re-design*

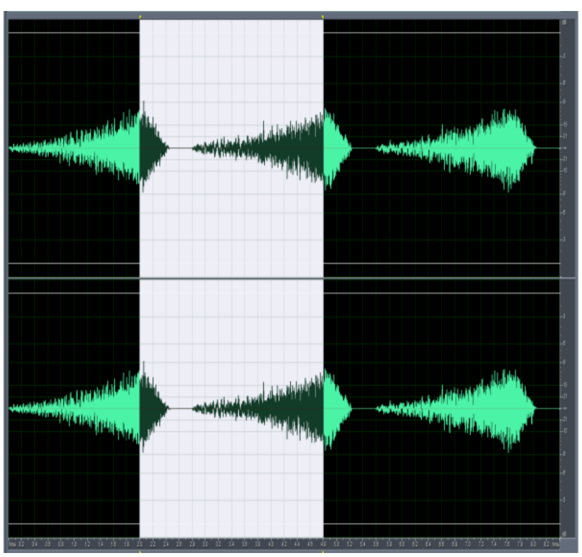

Figure 6.8 The waveform of the 400ms-time-slot wind-like sound (a 2800ms cycle time is highlighted in white part)

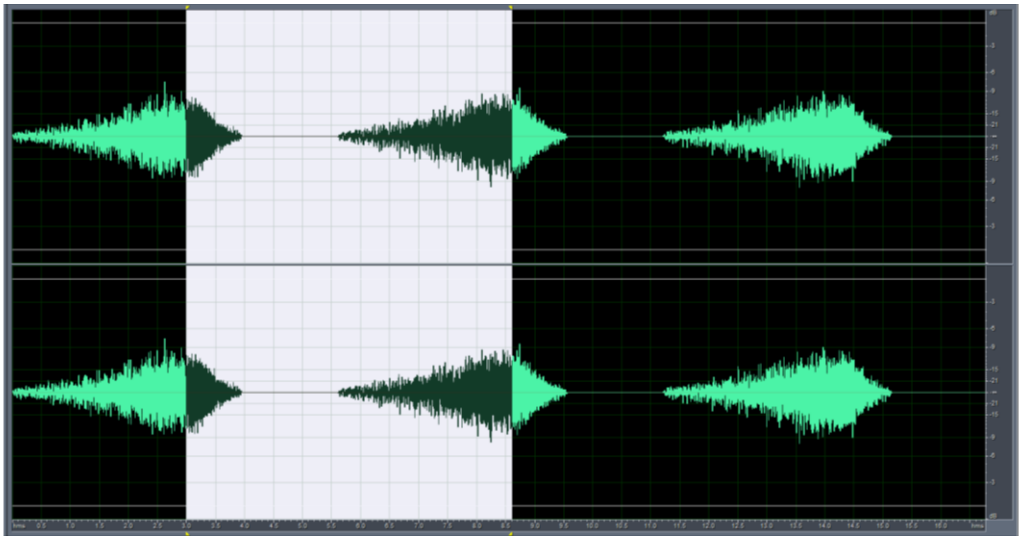

Figure 6.9 The waveform of the 800ms-time-slot wind-like sound (a 5600ms cycle time is highlighted in white part)

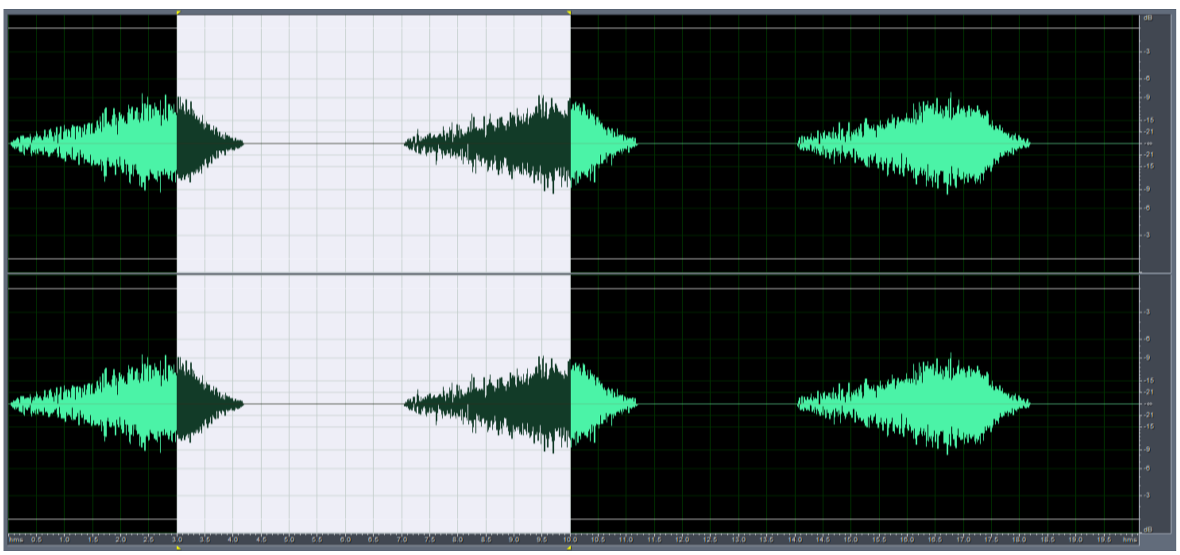

Figure 6.10 The waveform of the 1000ms-time-slot wind-like sound (a 7000ms cycle time is highlighted in white part)

Based on the analysis of findings of Experiment 1 and 2 and implications, and inspired from a natural phenomenon grass blowing in wind in wild, we explored the wind-like sound as the new type of auditory private cue.

Such type of auditory private cue can be produced by a noise module which outputs unit-amplitude white noise and modulated in a *Pure Data* environment.

For a regular cue mapping in which a cycle time contains 7 time slots, three pieces of wind-like sound, 400ms-time-slot, 800ms-time-slot and 1000ms-time-slot wind-like sound were created. The wave form of the 400ms-time-slot wind-like sound is shown in figure 6.8, in which the sound fades in and reaches the maximum amplitude within 2 seconds, and then fades out within 400ms. In the case of the 800ms-time-slot wind-like sound (Figure 6.9) and the 1000ms-time-slot wind-like sound (Figure 6.10), the sound fades in and reaches the maximum amplitude within 3 seconds, and then fades out within 800ms and 1000ms respectively.

The presentation of the wind-like sound is coordinated with the changes of the grass-like patterns in this way, the fade-out part of the sound is presented simultaneously with the flickering animation of the corresponding visual pattern, that is, the onset of one of the time slots in a cycle time; the fade-in part of the sound is presented immediately before the fade-out part accompanying the other changes of the grass-like pattern.

Based on the implication 5 in Section 6.2.5, the fade-out part of the wind-like sound which has a semantic correlation with the flickering animation was created so as to enhance CROSSFLOW users' perception of the corresponding visual pattern and enable users to integrate the sound and the corresponding visual pattern into a unified and meaningful percept easily. Moreover, the fade-in part of the sound was designed according to the implication 4 in Section 6.2.4 in order to alert users to pay attention to the subsequent corresponding visual pattern so that the users could identify the directional information relevant to them easily and quickly.

Musical sound can also be used as auditory private cue, in which each visual pattern is mapped to a unique timbre or pitch, or the corresponding visual pattern is mapped to a unique timbre and pitch and other visual patterns are mapped to the same timbre and pitch. However, this type of auditory private cue may demand significant cognitive loads on a user as the mappings between the sound and the visual patterns of which is much more complex than the mapping between the wind-like sound and the visual patterns.

6.3.4 *Configuration component for auditory private cue*

Besides the configuration component for the grass-like pattern in the software tool on the server, another the component enables a designer to specify the parameters that control the amplitude, frequency, the duration of fade-in and fade-out part of the wind-like sound. Moreover, the wind-like sound that could satisfy our requirements of usability and user experience of CROSSFLOW can be exported from *Pure Data* as a wave file and imported into a smartphone. Moreover, the clock of auditory cue is precisely synchronized with the clock of the grass-like pattern.

6.4 Conclusions

This chapter has presented the re-design process carried out for CROSSFLOW. The analysis of findings of Experiment 1 and 2 and implications of multimodal perception and crossmodal integration for interface design suggested that the size and shape of individual graphical element, the animation of the visual pattern should be changed, the appropriate duration of time slot should be found, users need more training and the system lag of synchronization should be controlled, a private cue should be presented earlier than the corresponding visual pattern, and the private cue and the corresponding visual pattern should have semantic correlation. Most of suggestions were adopted in the re-design process.

The re-design was embodied in CROSSFLOW V2 lab prototype in which the grass-like pattern and the wind-like sound are used as the public directional information and the private cue in the CROSSFLOW system. Although the parameters including the duration of time slot and the duration of fade-in and fade-out part of the wind-like sound have not been decided in this chapter yet, the appropriate parameters of components of CROSSFLOW V2 are investigated in a preliminary study of CROSSFLOW V2 (Experiment 3) which is reported in Chapter 7. In the same chapter, the investigation of the usability and the user experience of CROSSFLOW V2 (Experiment 4) is also reported, which reveals more interesting insights into the advantages and limits of CROSSFLOW.

Chapter 7. Evaluating CROSSFLOW Version 2

7.1 Introduction

The analyses of findings of Experiment 1 and 2 have suggested that the appropriate parameters of the components of CROSSFLOW including the parameters of private cue and the duration of time slot should be invested before the evaluation of the overall system. Thus in this chapter we compare three regular cue mappings corresponding to three combinations of wind-like sound (the 400ms-time-slot, 800ms-time-slot and 1000ms-time-slot wind-like sound), grass-like pattern and duration of time slot (400ms, 800ms, and 1000ms) in Experiment 3 to determine which one is most appropriate for use in the CROSSFLOW V2 system in terms of time taken to identify a target, target identification accuracy and subjective preference level. Furthermore, we applied the most appropriate parameters in CROSSFLOW V2 and evaluated the system in Experiment 4 so as to see whether users' perception and experience of the components and overall usability and user experience of CROSSFLOW are significantly improved or not and inform our in-depth understanding of the applicability of CROSSFLOW and the aspects that need to be further improved.

7.2 Experiment 3: Investigating Appropriate Parameters for CROSSFLOW V2

In this user study, individual participant was given target identification tasks to perform using CROSSFLOW V2 prototype. The main aim of this experiment was to investigate CROSSFLOW V2 on these aspects:

- **Suitability of Components:** to observe the suitability of the component of CROSSFLOW V2 including public directional information (the grass-like pattern) and auditory private cue (the wind-like sound).
- **Utility and Usability:** to observe the effectiveness and the efficiency of CROSSFLOW V2 as a type of navigation aid at a decision point in an indoor environment.
- **Appropriate regular cue mapping:** to compare three regular cue mappings corresponding to three combinations of the wind-like sound (the 400ms-time-slot, 800ms-time-slot and 1000ms-time-slot wind-like sound), the grass-like pattern and

the duration of time slot (400ms, 800ms, and 1000ms) to determine which one is most appropriate for use in the CROSSFLOW V2 system in terms of time taken to identify a target, target identification accuracy and subjective preference level. There are three conditions in this experiment which are referred as the 400ms, 800ms and 1000ms condition.

7.2.1 *Experimental Settings of Experiment 3*

Figure 7.1 Experiment 3: Lab settings

A. Experimental area

The experiment was implemented in a hall-like space in the building of Culture Lab of Newcastle University. The size of the experimental area is approximately 4 meters wide by 6 meters long. Six targets were positioned at different locations around the experimental area representing possible directions around a CROSSFLOW V2 user (see Figure 7.1). One target (number “7”) was positioned at the center of the experimental area representing the user's local position. Each target is an A4-size white paper showing a unique number on it. The purpose of this setting was to simulate a situation that, a user has already received the cue mapping and the private cue corresponding to his/her goal (e.g. a direction, location or object) from the server of CROSSFLOW V2, he/she who is standing at a decision point within an open area in an indoor environment needs to identify his/her target out of several visually similar objects or spatial structure (e.g. gates, corridors or doorways) as accurately and quickly as possible.

B. Public display and public directional information

A desktop PC was used as the server of the CROSSFLOW V2 system. The PC was connected with a SXGA projector to display the grass-like pattern. As shown in figure 7.2, the desktop PC and the projector with tripod were positioned by the side of the experimental area. The projected grass-like pattern covered the whole experimental area and the blades of grass were configured to point to the 7 target numbers at 7 different time slots within a cycle time. In order to make the pattern as visually prominent as possible, the lighting of the hall was switched off and the curtains were drawn.

C. Personal display and private cue

In order to avoid the system lag when synchronizing the clocks on smartphone and in the server, smartphone was not used in this experiment. Two sound speakers were connected to the server directly so as to play the wind-like sound of CROSSFLOW V2. The stereo speakers were positioned above the experimental area.

The 400ms-time-slot, 800ms-time-slot and 1000ms-time-slot wind-like sound (the waveforms are shown in Figures 6.8, 6.9 and 6.10) were used as private cues for comparison in this experiment (see Figure 7.3 for a clearer comparison).

D. Cue mapping

The three regular cue mappings corresponding to the three combinations of wind-like sound (the 400ms-time-slot, the 800ms-time-slot and the 1000ms-time-slot wind-like sound), the grass-like pattern and the duration of time slot (400ms, 800ms, and 1000ms) for comparison in this experiment are shown in figures 7.4, 7.5 and 7.6, in which there are seven time slots within a cycle time corresponding to the 7 target numbers.

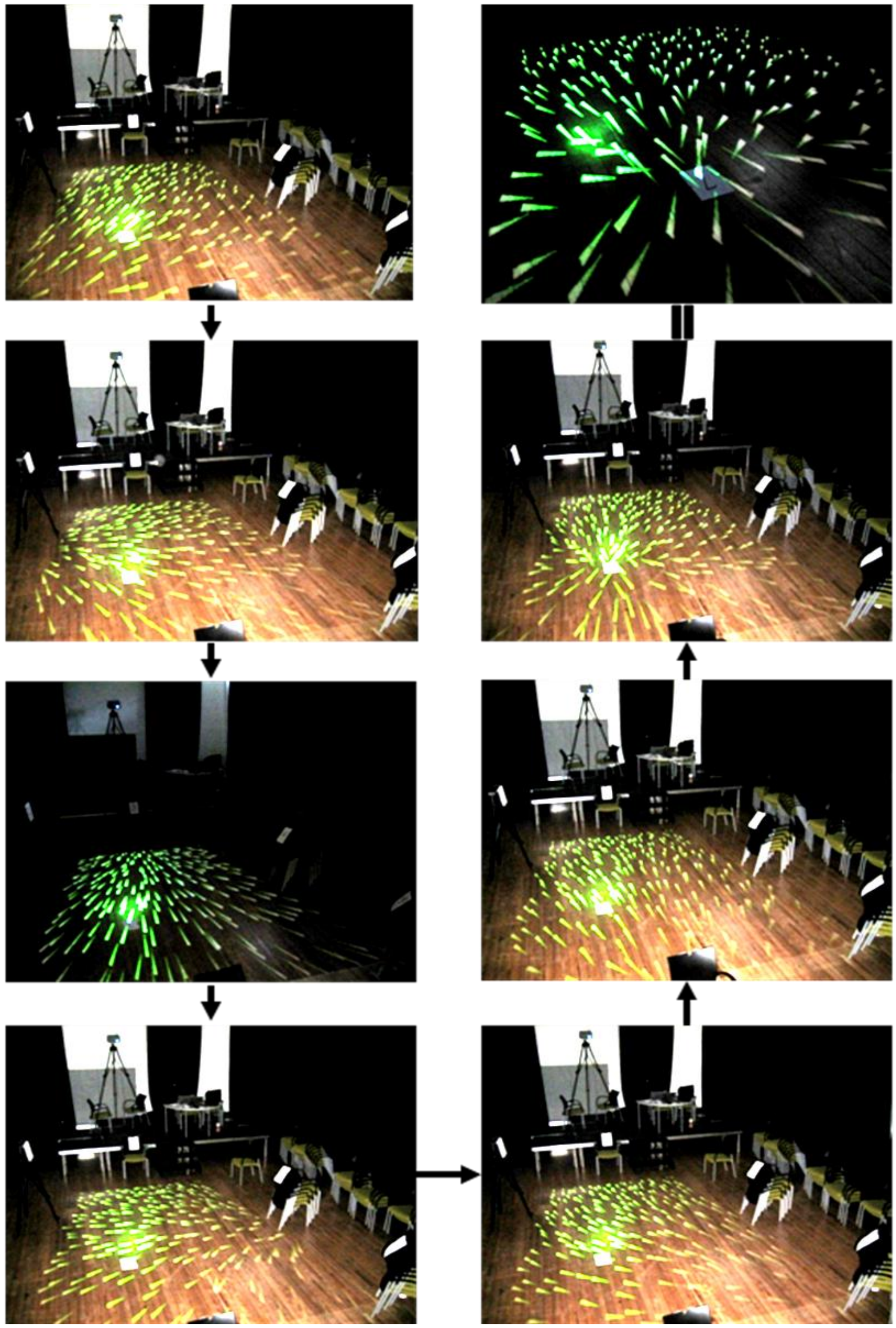

Figure 7.2 Experiment 3: Projections of the grass-like pattern corresponding to 7 targets in the experimental area (the upper right one is the close-up the grass-like pattern corresponding to the target number "7").

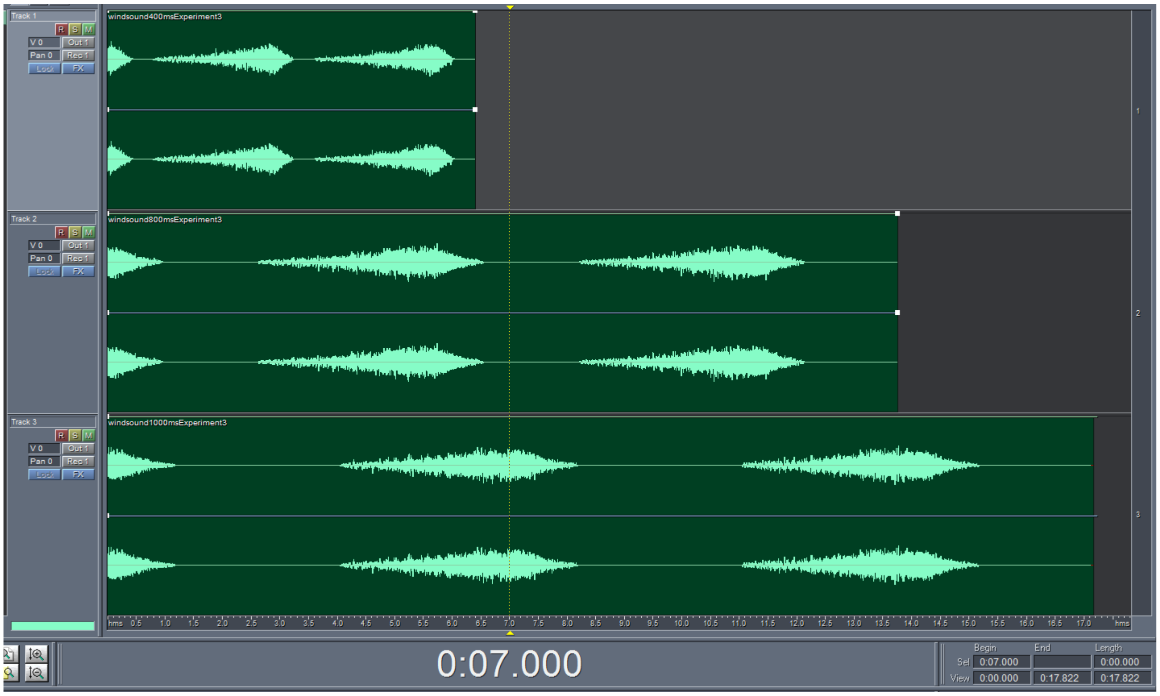

Figure 7.3 The waveforms of the 400ms-time-slot, the 800ms-time-slot and the 1000ms-time-slot wind-like sound (The upper sound track: 400ms time slot, 2800ms cycle time; the middle sound track: 800ms time slot, 5600ms cycle time; the lower sound track: 1000ms time slot, 7000ms cycle time).

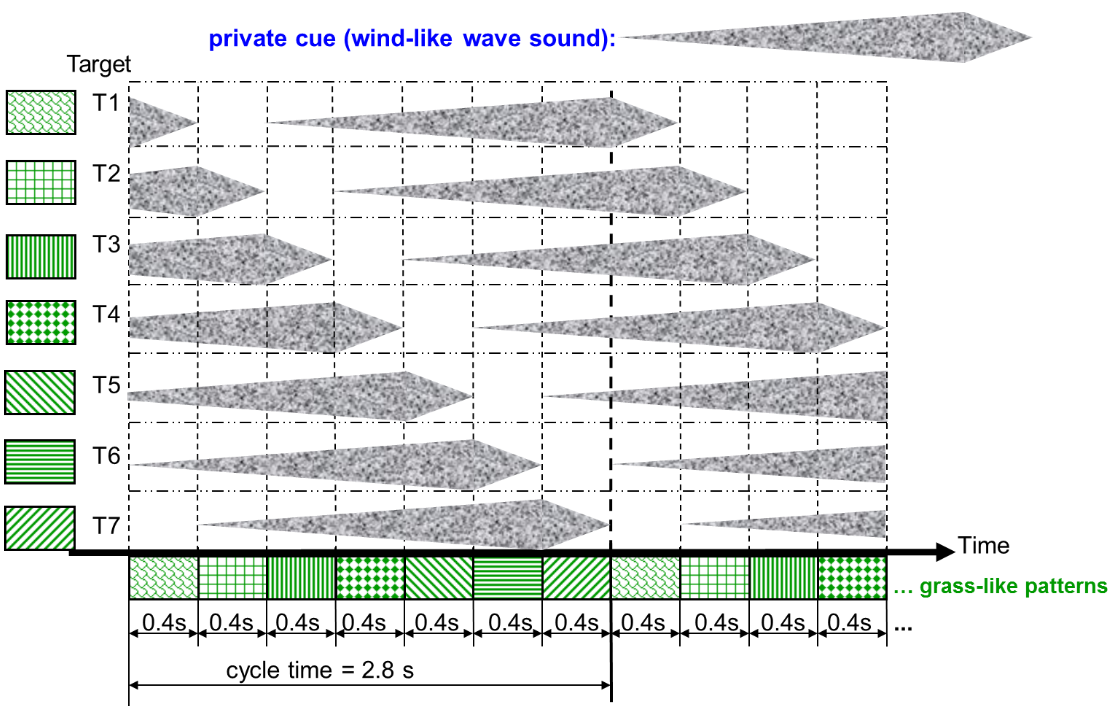

Figure 7.4 Regular cue mapping corresponding to the 400ms-time-slot wind-like sound.

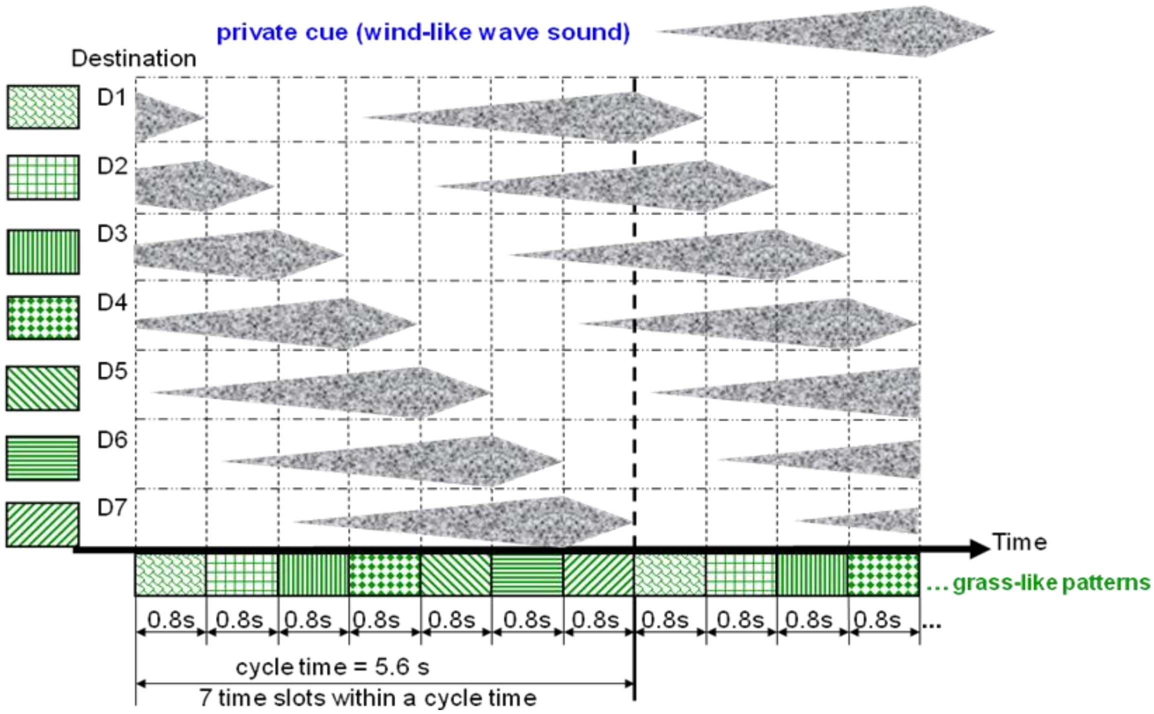

Figure 7.5 Regular cue mapping corresponding to the 800ms-time-slot wind-like sound.

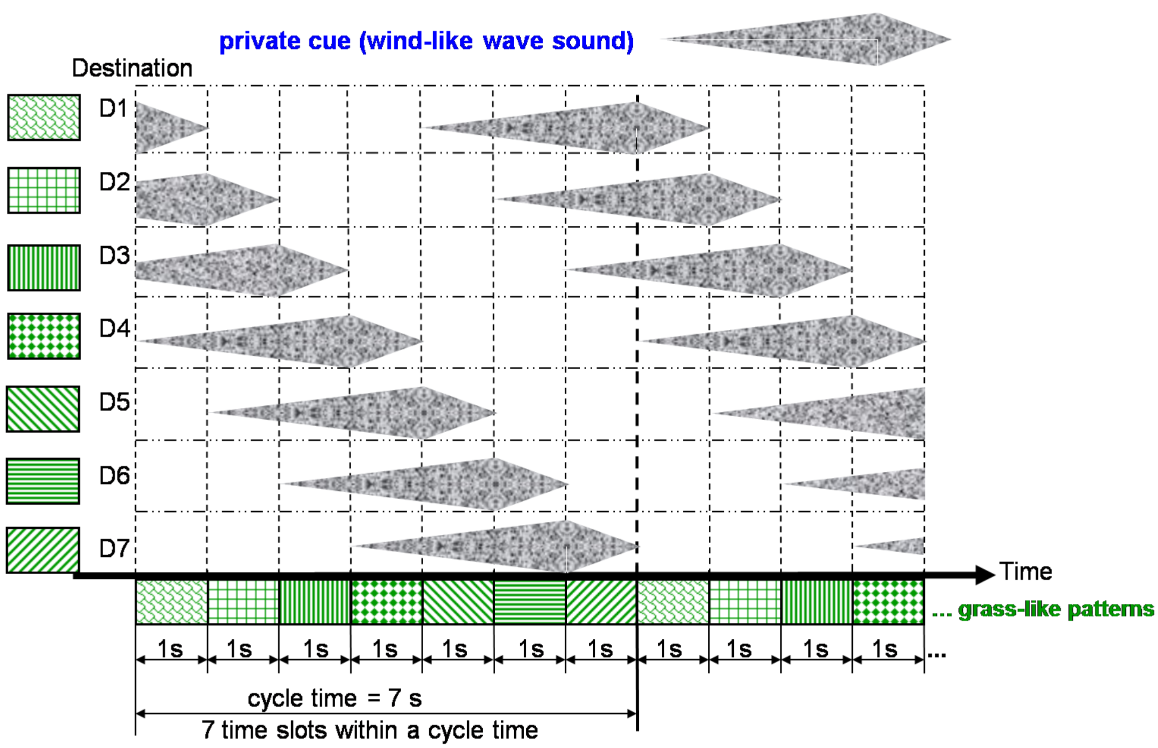

Figure 7.6 Regular cue mapping corresponding to the 1000ms-time-slot wind-like sound.

E. Video camera

A digital video camera was used to capture and record the activities of the participants in the whole experiment.

7.2.2 *Experimental Method of Experiment 3*

A. Experimental Design

The study used a within-subject design. We utilized one independent variable, kind of regular cue mapping, with three levels of treatment: (1) the regular cue mapping corresponding to the 400-time-slot wind-like sound and grass-like pattern animation (referred as the 400ms condition); (2) the regular cue mapping corresponding to the 800-time-slot wind-like sound and grass-like pattern animation (referred as the 800ms condition); (3) the regular cue mapping corresponding to the 1000-time-slot wind-like sound and grass-like pattern animation (referred as the 1000ms condition). The rationale of why such experimental design was adopted is the same as which is discussed in Section 5.2.2 A.

B. Participants

Fifteen participants (12 males, 3 females) with no prior experience of using CROSSFLOW took part in this experiment; all were students and staff of Newcastle University, aged between 20 and 40 years old and declared themselves free of visual, auditory disability and physical impairment. The content and the method of the experiment were agreed by the participants and approved by the supervisor of this research who is in the School of Computing Science at Newcastle University and the building manager.

C. Procedure

Each participant took part in the experiment that lasted approximately 30 minutes and consisted of two parts (see Figure 7.6): a training session (3 tasks); the formal experiment (30 tasks). The purpose of the training session was to make the participants understand how to use CROSSFLOW V2 to complete tasks successfully and provide feedbacks to experimenter after each task. The training performances were not measured.

Before the training session, the experimenter introduced the CROSSFLOW system and demonstrated how to use it to identify a target number within the experimental area for a participant. The importance of performing tasks correctly and quickly to the best of his/her ability was emphasized.

Then a participant was asked to complete three identification tasks. The participant stood within the experimental area and was told that he/she was free to walk within the area to observe the pattern. Once the participant informed the experimenter that he/she was ready to do the tasks, one of the three wind-like sound with the corresponding grass-like pattern animation was presented within the experimental area. The presentation did not stop unless the participant identified a target and said "OK". After the participant informed the experimenter the target number he/she just identified, he/she also need to rate the preference level (1~5 from the lowest to the highest preference level) about the combination of visual pattern and sound just presented. Whether the number that the participant reported was correct or not was indicated by the experimenter. Each participant was presented with all three conditions in the training session. The order of the conditions was randomized for every participant.

The formal experiment consisted of thirty tasks, consisting of ten presentations of each combination of visual pattern and sound, with the order randomised for every participant. The procedure of each formal task was similar to that of the training task. The only difference is that no feedback was given to the participants after each task was completed.

The time taken on each task was recorded via the tool on the server of CROSSFLOW V2. The number of successfully completed tasks and the ratings of the preference level were recorded on paper by the experimenter during the experiment. The activities of each participant were recorded on digital videotapes.

D. Measurements

The following measures of performance were defined and collected; they are also served as the dependent variables in the statistical analysis.

- **Identification Time (IT)**: The time taken to identify the target number from 7 numbers, measured from when the tool on the server of CROSSFLOW V2 started to render the combination of visual pattern and sound to when the participant identified a target number by saying "OK" to the experimenter.

- **Subjective Preference Level (PL)**: The preference level was defined as how the participant felt about the parameters of cue mapping, visual pattern and sound in terms of comfort and preference. The participants were asked to rate the preference level from 1 to 5, the mark 1 represents the lowest level of comfort and preference and mark 5 represents the highest level.

7.2.3 *Results of Experiment 3*

Table 7.1 presents the mean measures of Identification Time (IT) and Subjective Preference Level (PL) for the 400ms, 800ms and 1000ms conditions, and standard deviations in brackets (see Figures 7.7 and 7.8 for the charts).

Before performing statistical analysis the data were tested using Kolmogorov-Smirnov normality test with a Lilliefors significance correction and Shapiro-Wilk normality test (Park 2008) to see if they were normally distributed.

With respect Identification Time (IT), for the data followed a normal distribution, a one-way repeated measures ANOVA was carried out and the result showed that there was a significant differences across the three within-participants conditions ($F(2, 28) = 4.78$; df = 2, 28; $p = 0.016$). As shown in figure 7.7, the mean time to identify a target number in the 400ms condition is shortest, then follows the 800ms condition and the 1000ms condition, i.e. IT1000 > IT800 > IT400. The *post-hoc* two-tailed paired samples t-tests showed that the difference in the times taken to identify the target number between the 400ms condition and the 800ms condition ($t(14) = 0.88$, $p = 0.395$) was not statistically significant. This means that the participants used similar length of time to identify the target number in the 800ms condition and the 400ms condition, i.e. IT800 >= IT400. However, the differences between the 800ms condition and the 1000ms condition ($t(14) = 2.87$, $p = 0.012$), and the 400ms condition and the 1000ms condition ($t(14) = 2.43$, $p = 0.029$) were statistically significant, which means that the participants used significantly less time to identify the target number in the 800ms condition than in the 1000ms condition as well as in the 400ms condition than in the 1000ms condition, i.e. IT1000 >> IT800; IT1000 >> IT400. Moreover, the mean time taken to identify the target number for the 800ms condition is 15.7% less than which for the 1000ms condition. The mean time taken to identify the target number for the 400ms condition is 23.2% less than which for the 1000ms condition.

Therefore, the results of the comparison of Identification Time across the three conditions can be summarized as: IT1000 >> IT800 >= IT400 and IT1000 >> IT400.

As the data of Subjective Preference Level (PL) represented measures on an ordinal scale (Dyer 1995) and the data of PL in the 1000ms condition did not follow a normal distribution, non-parametric tests were carried out.

To compare PL, a Friedman test revealed that the difference was statistically significant across the three conditions, $p = 0.003$. As shown in figure 7.8, the mean Subjective Preference Level for the 800ms condition is the highest, and then follows the 1000ms and 400ms condition, i.e. PL800 > PL1000 > PL400. The *post-hoc* paired wise comparison were carried out using Wilcoxon signed-rank tests, the results showed that the difference on PL between the 400ms condition and the 800ms condition ($p = 0.033$) was statistically significant. This means that the participants felt much more comfort and preferable about the parameters used in the 800ms condition than the 400ms condition, i.e. PL800 >> PL400. Moreover, the mean PL for the 800ms condition is 16.5% higher than which for the 400ms condition. While the difference in the subjective preference level between 800ms condition and the 1000ms condition ($p = 0.208$), and the 400ms condition and the 1000ms condition ($p = 0.078$) were not statistically significant. This means that the participants felt equally comfort and preferable about the parameters used in the 800ms condition in comparison with the 1000ms condition as well as the 400ms condition in comparison with the 1000ms condition, i.e. PL800 >= PL1000; PL1000 >= PL400.

Therefore, the results of the comparison of Subjective Preference Level across the three conditions can be summarized as: PL800 >= PL1000 >= PL400 and PL800 >> PL400.

Condition	Mean Identification Time (ms)	Mean Preference Level
400ms	5754.10 (2101.27)	3.39 (0.74)
800ms	6125.85 (3297.77)	3.95 (0.42)
1000ms	7089.30 (4019.04)	3.80 (0.72)

Table 7.1 Experiment 3: Mean measures (with standard deviations) for the 400ms, 800ms and 1000ms conditions.

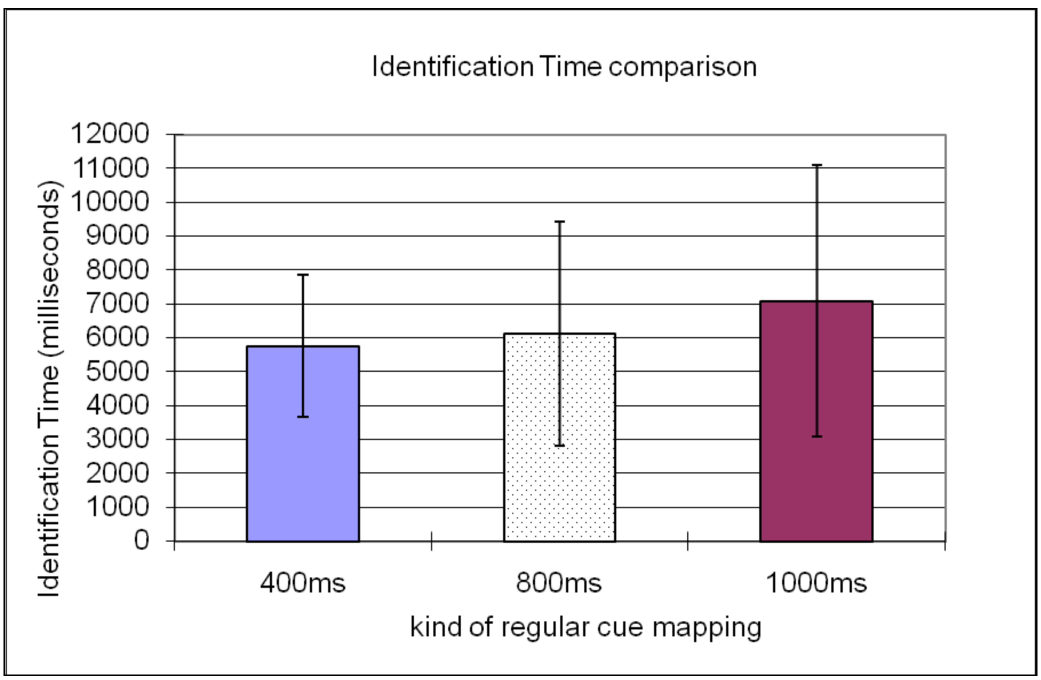

Figure 7.7 Experiment 3: Mean Identification Time (with standard deviations) for the three conditions.

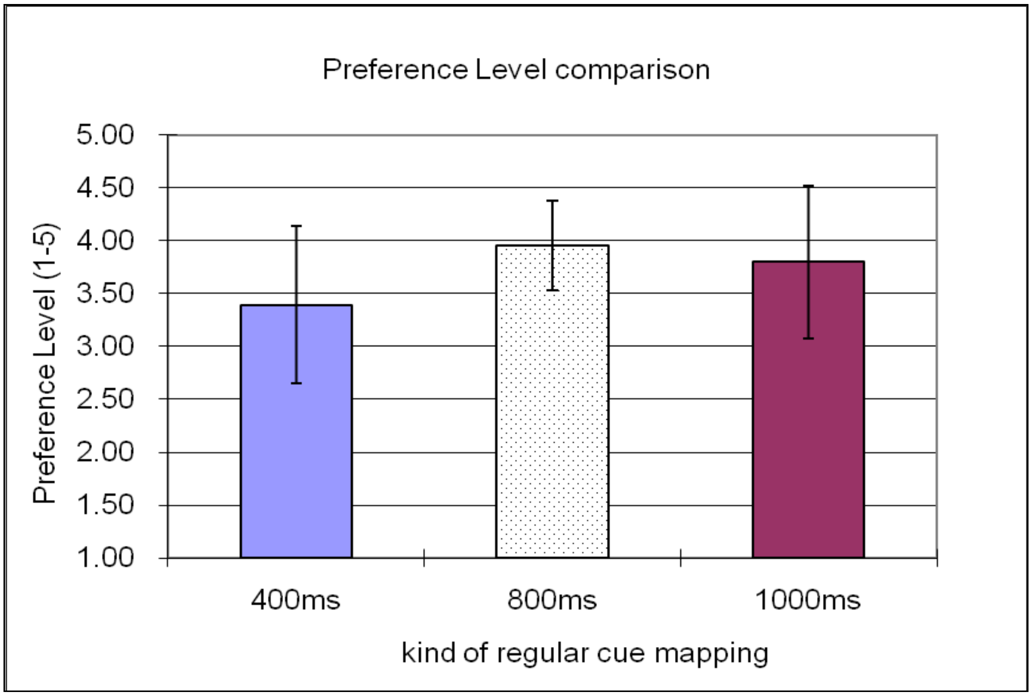

Figure 7.8 Experiment 3: Mean Subjective Preference Level (with standard deviations) for the three conditions.

7.2.4 *Discussion of Experiment 3*

In this experiment, each participant successfully completed the 30 tasks and identified the target number in each task correctly. Such high identification accuracy indicates that CROSSFLOW V2 can successfully be used as navigation aid at a decision point in an indoor environment, and also is an evidence that the projected grass-like pattern (public cue), wind-like sound from smartphone (private cue) and the corresponding regular cue mapping can be used to deliver directional information, and that users are able to understand the correlation between the sound and the grass-like pattern animation corresponding to a direction or object with a brief demonstration and short training and then identify it correctly.

The statistical analysis shows that the mean PL for the 800ms condition is the highest among the three conditions, which indicates that CROSSFLOW V2 users felt most comfortable and preferable about the parameters used in the 800ms condition. Although PL for the 800ms condition is not significantly higher than which for the 1000ms condition, it is significantly higher than which for the 400ms condition. The reason for the latter result may be because the participants felt capturing and identifying the grass-like pattern corresponding to a target number within in the first cycle in the 800ms condition was much easier than in the 400ms condition.

The statistical analysis also shows that the mean IT for the 400ms condition is the lowest among the three conditions. Looking at the 800ms condition in comparison with the 400ms condition indicates a similar level of IT, because no statistically robust difference is able to be drawn from the data of the two conditions. These results indicate that, the participants correctly identified a target number in the 800ms condition as quickly as in the 400ms condition.

Although the mean IT for the 400ms condition is the lowest among the three conditions, most of the participants did not identify the grass-like pattern corresponding to a target number in the first cycle in the 400ms condition. They did not respond until several cycles of time was passed. The main cause of the lowest mean PL for the 400ms condition may be that the repetition rate of the visual pattern animation and the wind-like sound was too high.

IT for the 1000ms condition is significantly longer than which for the 800ms and the 400ms condition, that is because some participants were observed in the 1000ms condition that they did not capture the grass-like pattern corresponding to a target number in the first cycle and they had to wait 7 seconds time for that pattern to reappear. The mean PL for the 1000ms condition is less than the 800ms condition may be because the participants felt bored about waiting too long time.

Therefore, the results of the comparison of both Identification Time and Subjective Preference Level across the three conditions suggest that the parameters of the cue mapping, the grass-like pattern animation and the wind-like sound used in the 800ms condition is most appropriate for the CROSSFLOW V2 system. They were used in the formal user study of CROSSFLOW V2 reported in the next section.

7.3 Experiment 4: Evaluating CROSSFLOW V2 Prototype

The main aim of this lab-based user study was to inform our in-depth understanding of CROSSFLOW V2 on these aspects:

- **Suitability of Components:** to observe the suitability and improvement of the components of CROSSFLOW including cue mapping, private cue and public directional information.
- **Utility and Usability:** to observe the utility and the usability of CROSSFLOW as a type of aid helping a user to identify his/her target (e.g. a direction, location or object) from large amount of visually similar objects or spatial structures within unfamiliar complex indoor environments.
- **User Experience:** to observe the user experience of CROSSFLOW V2 in comparison with that of CROSSFLOW V1.

7.3.1 *Experimental Settings of Experiment 4*

This experiment was still implemented in a hall-like space in the building of Culture Lab of Newcastle University. There are two areas (Area A and B) within the space were used in the formal experiment. The sizes of Area A and B are 4 meters wide by 7 meters long (see Figure 7.9) and 4 meters wide by 3 meters long (see Figure 7.11) respectively.

The 800ms-time-slot wind-like sound (see Figure 6.9) and the corresponding grass-like pattern animation, the regular cue mapping (see Figure 7.5) were used as the private cue, public directional information and cue mapping of CROSSFLOW V2 in this experiment.

Within each area, the grass-like pattern was configured to point to a target (see Figures 7.10, 7.12) which was required to be identified by a participant in one task and the targets in other areas sequentially. A target is a piece of A4-size paper showing a unique number covered by an A4-size envelope. Besides the target, six to twenty A4-size envelopes were distributed within the same area and used as distractors for the purpose of adding a degree of complexity and ambiguity for the target identification task.

In each experimental area, all the other settings were identical with which in Experiment 3.

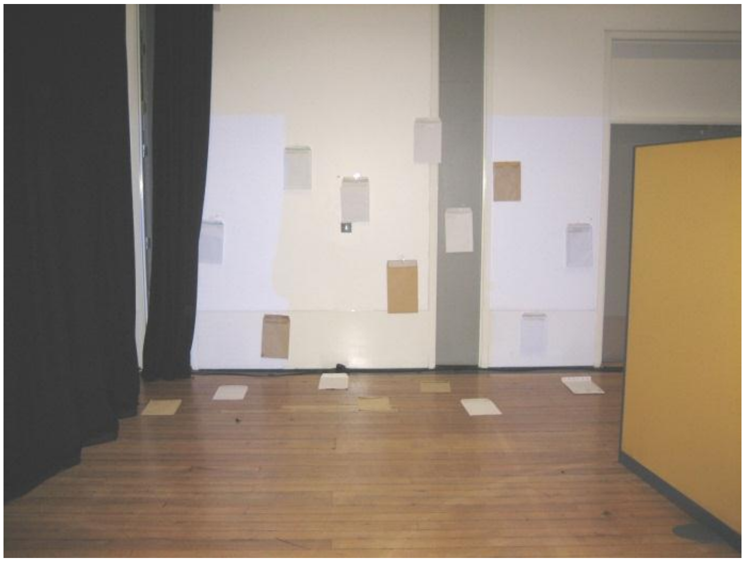

Figure 7.9 Experiment 4: Setting in Area A.

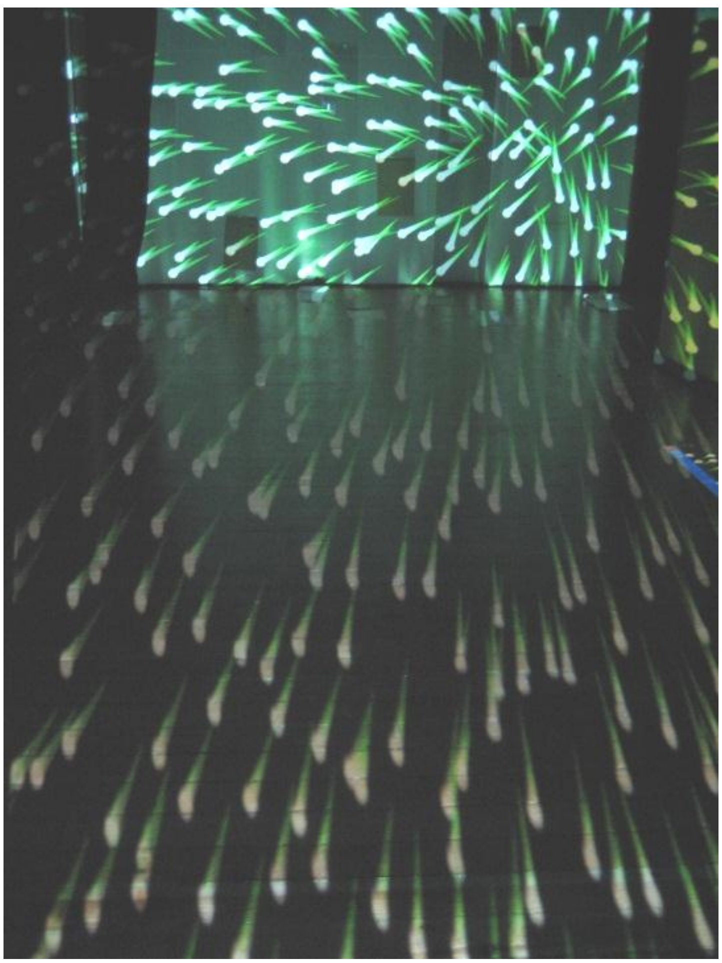

Figure 7.10 Experiment 4: Projection of the grass-like pattern corresponding to Target A in Area A.

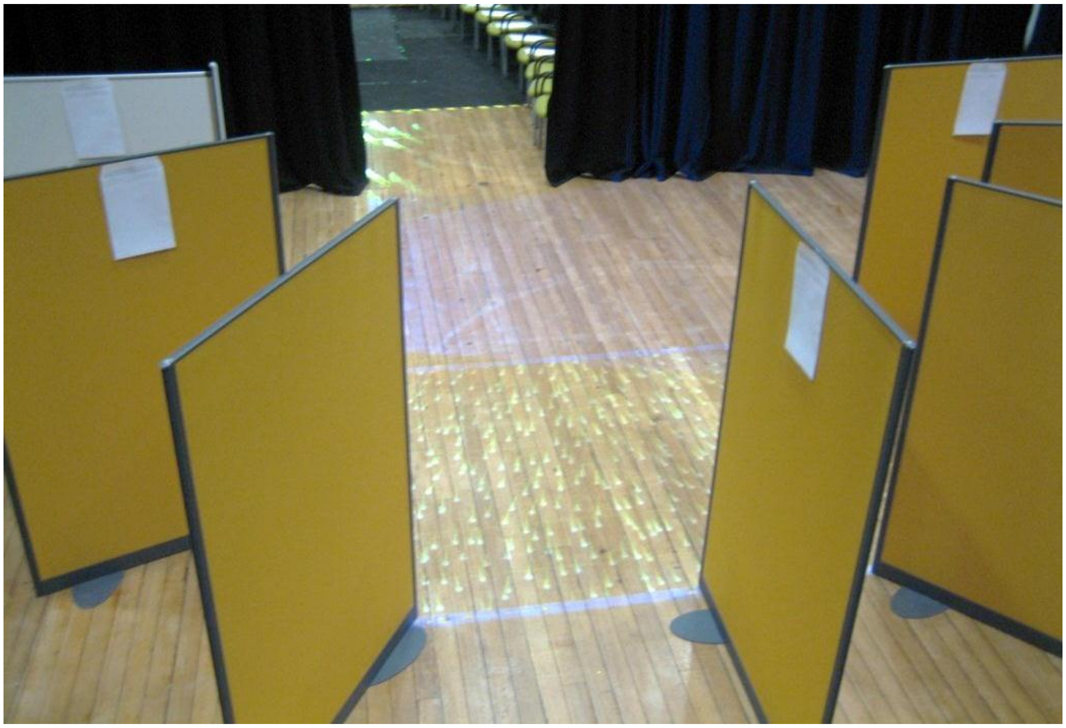

Figure 7.11 Experiment 4: Setting in Area B.

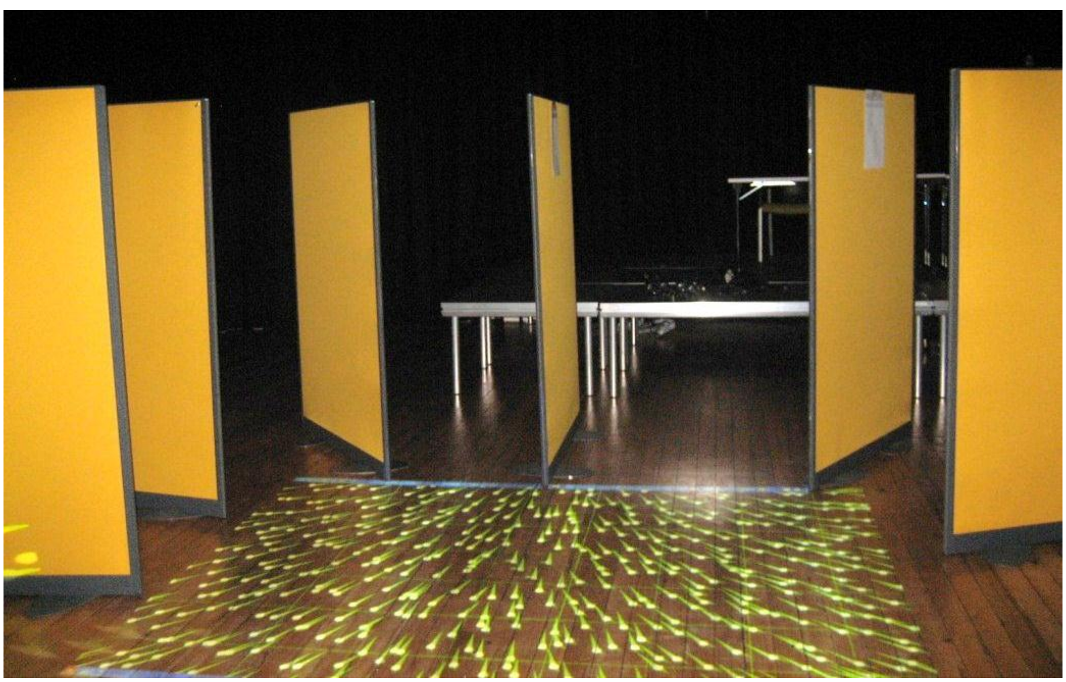

Figure 7.12 Experiment 4: Projection of the grass-like pattern corresponding to Target B in Area B.

7.3.2 *Experimental Method of Experiment 4*

A. Participants

Twelve participants (8 males, 4 females) with no prior experience of CROSSFLOW V2 took part in this experiment; all were students of Newcastle University, aged between 21 and 44 years old and declared themselves free of visual, auditory disability and physical impairment. Each participant was paid £20. The content and the method of the experiment were agreed by the participants and approved by the supervisor of this research who is in the School of Computing Science at Newcastle University and the building manager.

B. Procedure

Each participant took part in the experiment that lasted approximately 30 minutes and consisted of three parts: a training session; the formal experiment (Task A and B); and a post-experiment interview and questionnaire.

The purpose of the training session was to make sure the participants understand how to use CROSSFLOW V2 to complete a task successfully. The training performances were not measured.

Before the training session, the experimenter introduced the CROSSFLOW system and demonstrated how to use it to identify a target within an experimental area for a participant. The importance of performing tasks correctly and quickly to the best of his/her ability was emphasized.

Then the participant was asked to complete two identification tasks. At first, the participant stood at a fixed location within a training area and faced a designated direction. After the 800ms-time-slot wind-like sound with the corresponding grass-like pattern animation was presented within the area, the participant was free to walk within the area to observe the pattern. The presentation did not stop unless the participant identified a target by uncovering the envelope that covered a target number and informing the experimenter the target number beneath the envelope. The feedback was given by the experimenter after each training task was completed.

In the formal experiment, the participant asked to identify Target A in Area A and Target B in Area B respectively. The order randomized for every participant. The procedure of each task was similar to that of the training task. The difference was that no feedback was given to the participant after each task was completed.

The time taken on each task was recorded via the tool on the server of CROSSFLOW V2. The number of successfully completed tasks was recorded on paper by the experimenter during the experiment.

Having completed the formal experiment, the participant was interviewed for 10 minutes by the experimenter and asked to complete a post-experiment questionnaire (see Appendix). Their answers were summarized and analysed.

C. Measurements

The following measures of performance were defined and collected.

- **Task success rate:** The number of successfully completed tasks versus the number of tasks. Successful task completion was defined as the participant identifying the target successfully, regardless of time taken and errors.
- **Task completion time**: The time taken to identify Target A or B, measured from when the tool on the server of CROSSFLOW V2 started to render the combination of visual pattern and sound to when the participant reached Target A or B and uncovered its envelope.

7.3.3 *Results of Experiment 4*

Task	Mean task completion time (ms)	Expert's task completion time (ms)	Mean task completion time excluding Expert's time performance (ms)	Task success rate
Task A	21473 ms ≈ 3.8 cycle time (13487 ms ≈ 2.4 cycle time)	6542 ms ≈1.2 cycle time	23132 ms ≈ 4.1 cycle time (13179 ms ≈ 2.4 cycle time)	10/10 (100%)
Task B	11939 ms ≈ 2.1 cycle time (5908 ms ≈ 1.1 cycle time)	7902 ms ≈ 1.4 cycle time	12343 ms ≈ 2.2 cycle time (6065 ms ≈ 1.1 cycle time)	11/11 (100%)

Table 7.2 Experiment 4: Measures for Task A and B.

Table 7.2 shows the mean time taken to complete the task A and B with standard deviations in brackets. Putting this in perspective, it is worth noting that the "expert", one of the participants who already had been familiar with CROSSFLOW V1 completed the Task A and Task B in 6542 ms and 7902 ms respectively, which is 72% and 36% quicker than the mean task completion time of all the other participants. Table 7.2 also shows the task success rate, how successful users were able to complete the Task A and B. Of the 21 tasks, 100% were completed. Three tasks were not logged due to experiment setting failure.

The main findings from the post-experiment questionnaires and interviews are summarized as follows.

A. Comments about the components of CROSSFLOW V2

1) The grass-like pattern

The participants were asked several questions about the grass-like pattern of CROSSFLOW V2.

a) Indication by the grass-like pattern

In one of the questions, the participants were asked whether they were satisfied with the indication of the targets by the grass-like pattern or not.

Four of the twelve participants were very satisfied with the indication by the grass-like pattern. The representative comments are "*I can easily identify a location through the pointing directions of lots of grass or through the junction of the grass*" and "*It is obvious where is being indicated*".

The rest of eight participants answered "Yes" and confirmed that they could identify the targets indicated by the grass-like pattern. Their comments also suggested that the number, density and distribution of graphical elements and the projection angle of the grass-like pattern should be improved.

Five participants commented that the number of graphical elements distributed in the experimental area was excessive. The representative comment is "*It was not easy to tell exactly where the arrows were pointing, because there were too many arrows to look at.*" Two of them suggested that the number of graphical elements should had a trend towards the target, that is, the system should present more and denser grass far away from the target and present fewer and sparser grass near the target so that users could locate the target more easily.

Six participants commented that a target (e.g. Target B) was easier to be identified or located where the distractor targets were far away from it, while a target (e.g. Target A) was more difficult to be identified or located where the distractor targets were close to it or where the target and the distractor targets were distributed in a three-dimensional space like Area A. The representative comments are "*It is easy to get a rough direction, but accurately locating one of several envelopes that were closed to each other was not easy.*", and "*The indication of a location or object in a 3D space (Area A) by the pattern is not very clear, while it is clearer in a 2D surface (Area B).*"

Two participants mentioned the projection of visual pattern. The representative comment is "*It was much easier to identify in the overhead projection than the other projection areas, because you were not 'dazzled' by the projector. Also there were fewer shadows produced than with other angles of projection.*"

b) Overall perception and experience

The participants were also asked to report their overall perception and experience of the grass-like pattern.

Seven of the twelve participants made positive comments, they said that they liked the pattern and described their impressions of the grass-like pattern as "comfortable", "good", "interesting", "beautiful", "eye-catching", "appealing" or "impressive".

Two participants commented that the graphical elements in the pattern were like arrows or daggers rather than grass.

Two participants suggested that the graphical elements should be designed and animated like real grass, flowers or pointing fingers.

Three participants said that they felt uncomfortable after watching the visual pattern animation for a while, and concerned that watching the flickering light of the pattern for a long time may cause certain people ill (e.g. epilepsy). One of them also commented "*I did not find the pattern to be ambient, I found it very busy, and it looked quite technological*".

2) **The wind-like sound**

Participants were also asked several questions about the wind-like sound of CROSSFLOW V2.

a) **Fade-in part of the wind-like sound**

The participants were asked whether the fade-in part of the wind-like sound has alerted them to pay attention to the subsequent visual pattern which was presented simultaneously with the fade-out part of the sound or not.

A related question asked if the fade-in part of the sound was long enough to alert them.

One participant said that he did not pay attention to any visual pattern animation during the fade-in part of the wind-like sound was played, he just paid attention directly to the visual pattern which was presented simultaneously with the fade-out part of the sound.

All the other eleven participants answered "*Yes*" for both questions. In addition, one of them said that she was a bit confused as she assumed that after the fade-in part of the wind-like sound played, the sound would stop immediately or else; another two participants said that they felt the fade-in part of the sound was a bit too long.

b) **Fade-out part of the wind-like sound**

The participants were also asked whether the fade-out part of the wind-like sound has helped them to capture the grass-like pattern relevant to them and identify the target or not.

A related question asked if the fade-out part of the wind-like sound was long enough for them to capture the grass-like pattern corresponding to their target within the first cycle time.

Ten out of the twelve participants answered "*Yes*". Of the participants who answered "*Yes*", one said "*After a while the sound could be possible to come out of focus and drift into a daze as the pattern is quite repetitive. The decreased sound breaks one out of the daze!*"; another said "*Yes, the cut off in the sound was easy to associate with the direction of the arrows at that moment.*"; two felt the fade-out part of sound was too short and one of them suggested that it would be better if the fade-out part of sound was a bit longer than the grass-like pattern corresponding to their target; another three participants commented that although the fade-out part of the sound was long enough for them to identify directional information corresponding to their target within the first cycle time, they still needed one or more additional cycles to confirm their identification. It seems that two participants who did not answer "*Yes*" did not notice or understand that the fade-out part of the sound was presented to indicate the grass-like pattern corresponding to their target and enhance their perception of it. One of them said that she thought the fade-out part of the sound was just the echo of the fade-in part of the sound.

c) Semantic correlation between sound and visual pattern

The participants were asked whether they could identify the semantic correlation between the fade-out part of the sound and the concurrent grass-like pattern animation.

Nine out of twelve participants confirmed that they could. The representative comment is "*Yes, when the visual pattern was sharp, and the sound was also sharp and abrupt.*" Of the participants who did not make this comment, one said that sometimes she could associate the sound with the pattern animation semantically, sometimes not; two did not notice or misunderstood the semantic correlation.

d) Overall perception and experience

The participants were asked about their overall perception and experience of the wind-like sound.

They confirmed that the sound was clear and facilitated their task. Three of them did not like the sound while one expressed the opposite opinion. The representative negative comment is "*The sound was slightly annoying; however it did make me focus on the task.*" The positive comment is "*The wave sound was quite nice... the wave sound had the makings (the visual pattern) ambient.*" In addition, two participants suggested "*A faster moving pattern which was accompanied by music would be good. Different users would hear different music, the rhythm of the music would emphasize the correct direction to take.*"; and "*Perhaps replace the noise with a calm friendly voice saying 'Please follow the arrow'*".

3) Duration of time slot and cycle time

Four out of the twelve participants in total concerned about the duration of time slot and/or cycle time.

Two of them raised the concern when commenting the grass-like pattern. They said that the presentation time of the relevant visual pattern was too short to identify directional information in the given length of cycle time and suggested that it should be extended.

Two of them raised the concern when commenting the wind-like sound. They said that the fade-out part of the wind-like sound was too short; and one of them suggested that the sound should be a bit longer than the visual pattern corresponding to the target.

Three of them complained that cycle time was too long and suggested that it should be shorter so that they could confirm their identified target more quickly. The representative comment is "*The sound may cycle faster so that no need to wait too long when I did not capture it (directional information corresponding to the target) at the first time.*"

B. Overall usability of CROSSFLOW V2

Most of the participants felt that the tasks were not very difficult. The comments of the participants regarding the overall usability of CROSSFLOW V2 are grouped as follows:

a) Require short memory and fast identification

The only participant who felt the tasks were difficult and rated 4 was the participant who was 44 years old and much older than the other participants (20~31 years old). He explained "*As the arrows disappeared or changed direction once the sound was over, it was difficult to be sure where they were pointing as you needed to rely on memory to do this. ...There was only a short time to judge which way they were pointing as the pattern were not steady. ...There was an expectation to respond quickly which was not always easy.*"

b) Accuracy of indication is relatively lower in small scale spaces and a three dimensional space.

The oldest participant also mentioned that the accuracy of indication of visual pattern is relatively lower in small scale spaces and a three dimensional space; therefore Task B was much easier than Task A because of the clearer pointing and projection.

Another two participants expressed the similar opinion, even though they repeated their comments about the grass-like pattern. The representative comment is "*Task B was done in a small region. The target (Target B) could be easily identified by observing the* *junction of much grass, but the junction of grass looked not very tidy. A target might be more easily identified if it was indicated by less grass. However a target in a larger area may not be easily identified by observing the indication of less grass, because the indication may not be very precise. Therefore in larger area more grass should be presented to indicate the target location or object.*"

c) Not better than map.

Such opinion was expressed by only one participant and the original comment is "*I did not think the system could provide more convenience than map.*"

d) Require divided attention.

This was indicated by one participant and the original comment is "*The system needed me to concentrate to detect the relevant visual pattern and sound at the same time.*"

The following groups of comments are evidently positive about the CROSSFLOW V2 system:

a) Useful in large environment, dark environment or night time.

This is indicated by two participants. Their original comments are:

"*In a large location, it would be more convenient to point my way/route.*"

"*It is helpful, especially for people find a place in night time.*"

b) Can be used by multiple users.

This advantage was identified by one of the twelve participants. His original comment is "*Multiple users can get to different destinations from the same visual stimulus.*"

c) Efficient, time-saving.

This was indicated by two participants. The representative comment is "*After 2 iterations of the pointing (two cycle time), I could find the location.*"

d) Easy to understand and use, and better than using map.

This was indicated by most of the participants. The representative comments are:

"*It is easy to find an object, direction, location using this system.*"

"*It is clear to identify the target based on the pattern and sound.*"

"*It is easier than map.*"

"*The dynamic directional information is acceptable, simple and easy to be understood. Especially it is good for the people who have difficulty to use map.*"

"*I have difficulty to use map, discriminate directions. Thus for me using this pointing system is much easier than using map.*"

"*Be able to find something without actually knowing where it is to start off with. Using this system meant that I did not have to read a map and could concentrate fully on reaching the target.*"

It is worth noting that the "expert", one of the participants who already had been familiar with CROSSFLOW V1 commented "*Through the combination of visual and auditory display, a user can reach the target without thinking, just follow the sound cue. The visual pattern indicating the approximate direction is enough for identifying the targets. The indication of a target or location is clearer than CROSSFLOW V1, thus it is easier to identify a target (using CROSSFLOW V2).*"

C. Overall user experience of CROSSFLOW V2

Participants were asked the open-ended question about the overall experience of CROSSFLOW V2.

Most of the participants liked the idea and the CROSSFLOW V2 system, and said the system or idea was "good", "nice", "useful", "helpful", "interesting", "innovative" or "great". Two participants were also interested in seeing the final product for real application.

D. Suggested improvements

It was suggested that the following components of CROSSFLOW V2 should be improved: The grass-like pattern was ranked the highest priority for improvement by four participants; the wind-like sound was ranked the highest priority for improvement by three participants; and duration of time slot was ranked the highest priority for improvement by three participants. The detailed suggestions have been integrated into the above sections.

7.3.4 Discussion of Experiment 4

A. Public directional information

Regarding users' satisfaction of indication of public directional information, four out of ten participants (40% of responses) in Experiment 2 felt that directional information was difficult to be identified shown by the fish-flow pattern and two of them explained that it was because the size of "fish" was too small.

The improvement has been shown in the findings of Experiment 4 that all participants (100% of responses) confirmed that they could identify the directions or locations shown by the grass-like pattern and no complaint about the size of the graphical element of visual pattern was found. Although the qualitative feedbacks of Experiment 4 showed that the indication precision of the grass-like pattern needed to be improved, the grass-like pattern appeared much more precise and clearer than the fish-flow pattern when indicating different locations or directions in different experimental areas in the informal test. As the results of Experiment 4 suggested, improving the number, density and distribution of grass and projection angle might be able to further improve the indication precision of the grass-like pattern.

Regarding overall users' perception and experience of public directional information of CROSSFLOW, the positive comments and the percentage of positive responses in Experiment 2 and 4 were very similar. However, three participants of Experiment 4 felt that the grass-like pattern were not comfortable or ambient. This might result from the individual difference. Some people such as those autism, epilepsy or photosensitivity are very susceptible to flickering lights or bright lights or the combination of them. For those people, the combined factors such as the frequency of the flash, brightness, contrast with background lighting, distance between the viewer and the light source may trigger the onset of seizures, headaches, nausea or dizziness. In particular, flashing lights most likely to trigger seizures are between the frequency of 5 to 30 flashes per second (Hertz) (CEA 2008).

In the setting of Experiment 4, in order to make the projected pattern as visually prominent as possible, the background lighting in the experimental indoor environment was made dark. Furthermore, the size of each "grass" was changed 6 times within each 800ms time slot to produce "flickering" animation. Thus viewers may perceive the visual pattern animation as 7.5Hz flashing lights with high brightness and high contrast with a dark background environment. Those three participants might be too susceptible to such lights. This problem may be able to be resolved by reducing the grass-size-changing time (e.g. from 6 times to 3 times) or visual pattern brightness and contrast with background environment.

The detailed suggestions about the design and the configuration of the visual public directional information of CROSSFLOW were summarized in Chapter 9.

B. Private cue & correlations between private cue and the corresponding visual pattern

The results of Experiment 4 indicate that the fade-in part of the wind-like sound can be easily identified and is useful, as most of the participants (91.7% of responses) confirmed that the fade-in part of the sound alerted them to pay attention to the subsequent visual pattern corresponding to their target. Although two participants of Experiment 4 felt the fade-in part of sound a bit long within a cycle, three-second fade-in part of auditory cue is necessary since users may be affected by the factors such as environmental noise, visual distractions and stress when using CROSSFLOW V2 in real indoor navigation situations, it may take longer time for them to be alerted.

The results of Experiment 4 also indicate that the fade-out part of the wind-like sound can be easily identified and is useful, as most of the participants (83.3% of responses) confirmed that the sound has helped them to capture the visual pattern corresponding to their target. Moreover, the semantic correlation between the fade-out part of the sound and concurrent grass-like pattern animation is identifiable and useful, as most of the participants (75% of responses) have confirmed this.

As most of the participants in Experiment 4 were satisfied with the wind-like sound in terms of usability and experience, the success of private cue re-design is proved.

Although both utility and aesthetics were considered when designing auditory private cue of CROSSFLOW V2, three participants (25% of responses) were not satisfied with the wind-like sound in terms of aesthetics. This suggests that the design of private cue which can balance both utility and aesthetics need to be further explored.

Regarding the temporal correlation between private cue and visual pattern, in Experiment 2 seven of the ten participants (70% of responses) confirmed that they could identify that private cue was synchronized with one of the visual patterns and cycled, while the rest three participants had difficulties to integrate the private cue and synchronized visual public pattern into a unified stimulus within an exact time slot. In Experiment 4, eight of the twelve participants (66.7% of responses) could easily identify the temporal correlation between private cue and visual pattern, although four participants commented that it was not easy to identify the correlation. The level of identification of temporal correlation between private cue and visual pattern in Experiment 2 and Experiment 4 is similar. This result suggests that most of CROSSFLOW users have no difficulty in the first process of identification of public directional information relevant to them, that is, the integration of private cue and visual pattern corresponding to their target.

The results of Experiment 4 show that six of the twelve participants of (50% of responses) either had difficulties to identify temporal or semantic correlation between the wind-like sound and the grass-like pattern animation, or did not notice or understand the function of the fade-in part or fade-out part of the wind-like sound. There are two possible causes for this. Firstly, it might be because the cue mapping and the private cue used in Experiment 4 were more complex than which used in Experiment 1 and 2. Secondly, it may be because the training of using CROSSFLOW V2 in Experiment 4 is too limited. Providing more training tasks before the formal experiment might have improved the users' task performance, perception and experience of the wind-like sound, the grass-like pattern and the correlations between the sound and the pattern. However, the training in this experiment was intentionally limited, since in real indoor navigation situation it is unlikely that the users would participate in extended training before using CROSSFLOW V2.

C. Duration of time slot and cycle time

Regarding the duration of the time slot and the cycle time, seven of the ten participants (70% of responses) in Experiment 2 raised the concern about the duration of time slot and/or the cycle time, in particular, six of them (60% of responses) complained that the presentation time of visual pattern corresponding to each target, i.e. the duration of time slot was too short and suggested it should be prolonged. While in Experiment 4 only four of the twelve participants (33.3% of responses) raised the concern, and only two participants (16.7% of responses) had the similar complaint and suggestion. These results show a significant improvement in users' perception of the duration of time slot. Such improvement appears to be a direct result of the improvements of private cue, public directional information and cue mapping, as the same duration of time slot (800ms) and the cycle time (5600ms) were used in Experiment 4 and Experiment 2. This implies that the way to improve users' perception of duration of time slot is to improve the design of private cue, public directional information and cue mapping instead of shortening or prolonging the duration of time slot.

D. Utility and usability of CROSSFLOW V2

Excluding the unlogged tasks, the participants identified the targets out of large amounts of visually similar objects or spatial structures within unfamiliar complex indoor environments using CROSSFLOW V2 at 100% task success rate within a reasonable period of time. The result is an evidence that the components of CROSSFLOW V2 including the projected grass-like pattern (public cue), the wind-like sound from smartphone (private cue) and the regular cue mapping can be used to deliver personalized navigational information, and that users are able to understand the correlation between auditory cue and the visual pattern corresponding to their target with a brief demonstration and very little training, and then capture and identify the directional information corresponding to a location or object and reach it successfully.

Excluding the time performance of "expert", the average completion time of Task A of the participants is 72% longer than the time taken by the "expert". The average completion time of Task B of all the other participants is 36% longer than the time taken by the "expert". Such difference implies that if the participants received more training, their performances might be much better.

The large deviations in task completion time may be explained partly by the range of normal walking spaces in the participants; however from the observations and the feedbacks of the participants, it can be concluded that this was not the major factor. Rather, firstly, two of the participants took several cycle times, nearly two standard deviations from the mean time longer to complete the tasks. Secondly, different strategies through which the participants used the system to navigate have different impacts on the task completion time. We identified two different strategies in the analysis of the behaviours of the participants. Some participants waited more than one cycle time in order to confirm their identification after identified the public directional information corresponding to their target, then walked and uncovered the envelope on the top of the target. This approach allowed for increased confidence but slower completion time. Some participants did not wait and just walked towards the envelope immediately after they identified the public directional information, which allowed for less confidence but faster completion time.

E. Overall user experience and usability of CROSSFLOW

The results of Experiment 4 and 2 suggest a similar level of overall user experience of CROSSFLOW V2 and V1, as most of the participants in both experiments provided positive feedbacks.

The results of both experiments show that the subjective task difficulty level ratings were similar, the task success rates were the same, and the feedbacks about the overall usability of the system were inter-complementary and partially repetitive. More importantly, much fewer users felt that using CROSSFLOW to complete a task is time-consuming in Experiment 4 (25% of responses) than in Experiment 2 (50% of responses), which significantly benefited from the re-design of CROSSFLOW's components including cue mapping, private cue and public directional information. Therefore, regardless the difference of the settings in the experimental areas between the two experiments, it might be concluded that the overall usability of CROSSFLOW V2 is better than CROSSFLOW V1, although further comparisons on different aspect of usability are still needed such as target identification task completion time, identification error and subjective workload.

F. Age of participant

One finding in the results of Experiment 4 was that most of feedbacks from the participant who was 44 years old and much older than the other participants (20~31 years old) were negative, especially his feedbacks about the grass-like pattern and overall usability and experience of CROSSFLOW V2. As there was only one participant is over 40-years-old in Experiment 4, a larger scale study would provide a more accurate representation of the feedbacks of older people.

A series of studies has been performed on aging and its effect on perceptual processing and working memory capacity for visual stimuli. It was found that a number of perceptual abilities, especially visual perceptual abilities that diminish with age (Faubert 2002). Controlling for the age of participants was outside the scope of this thesis, but this factor should be considered if CROSSFLOW is intended for use by older people.

7.4 Conclusions

This chapter has reported two user studies (Experiment 3 and 4) of re-designed CROSSFLOW prototype, CROSSFLOW V2. Experiment 3 investigates appropriate parameters of components of CROSSFLOW V2. The results have shown that the regular cue mappings corresponding to the combination of the 800ms-time-slot wind-like sound, the grass-like pattern and the 800ms duration of time slot and the corresponding parameters were most appropriate for CROSSFLOW V2 considering both target identification time performance and users' subjective preference level. Experiment 4 investigates the usability and the user experience afforded by CROSSFLOW V2 as a type of aid helping a user to identify his/her target (e.g. a direction, location or object) from large amount of visually similar objects or spatial structures within an unfamiliar complex indoor environment.

The main results have shown that some aspects of usability of CROSSFLOW V2 are better than that of CROSSFLOW V1, which means that the re-designed public directional information, private cue, correlation between private cue and the corresponding visual pattern might significantly improve the usability of CROSSFLOW and the perception of the duration of time slot. Although the improvements on users' perception and experience of public directional information and the overall user experience were not very significant.

The suggestions and lessons learned in the CROSSFLOW design and evaluation iteration in Chapter 6 and 7 provide contributions to the overall suggestions for CROSSFLOW design and evaluation which are summarized in Chapter 9.

Chapter 8. CROSSBOARD: A Crossmodal Display for Information Retrieval on Dense Public Displays

8.1 Introduction

As traditional public displays are highly relevant to the problems identified in Chapter 2, it is necessary to improve the displays and facilitate people's interactions with them. Augmenting traditional public displays by integrating crossmodal cue and cue mapping proposed in Chapter 3 might be an effective solution. Drawn design inspirations from CROSSFLOW, we explore a Crossmodal Display prototype application named CROSSBOARD that harnesses hierarchical crossmodal cues and cue mapping to support efficient information retrieval of multiple simultaneous users on high-density information displays. This chapter reports the design of the CROSSBOARD system through describing the design of a hierarchical crossmodal cue mapping and demonstrates the effectiveness and efficiency of this solution through the results of a single-user lab study in which the potential of CROSSBOARD for improving the retrieval of un-indexed information from dense information displays is clearly demonstrated.

8.2 Design Inspirations from CROSSFLOW

The CROSSFLOW system explored in Chapter 4 and 5 can project ambient patterns that could be configured to indicate the directions to a destination from all decision points in an environment, but only supports indicating as many destinations as the sequence of public cues allows (the sum of the time slots). In the system used in Experiment 2, the duration of time-slot for each public cue is 800ms and directions to each destination need to be repeated every 5.6 seconds, this allows the system to provide public cues for only 7 destinations. If we view each of destinations as a particular information item, we will find the amount of information item indicated by CROSSFLOW is very small. Although finding a large number of destinations can be supported by utilizing contextual sensors which have been mentioned in Chapter 3, such approach would again have to rely on sensing and tracking. Furthermore, we need to consider the situation if the number of information items (where an element is information relating to one user or user group) runs to many hundreds or thousands. Apparently, in such a situation it is not efficient to cue each information item one by one, therefore, the research challenge about how to design appropriate public cues, private cues and cue mapping of Crossmodal Displays to deal with large amount of information items arises.

Furthermore, CROSSFLOW is designed for nomadic users in an open public space, thus ambient patterns as public cues are designed and displayed in a large area such as an atrium floor of a building. As the relative smaller display platforms such as wall-size information board are more common in public space, it is necessary to explore how public cues can be designed and used on such platforms in the design and evaluation of Crossmodal Displays.

Therefore, by contrasting the features and limitations of CROSSFLOW, we can identify a new design space for Crossmodal Displays.

8.3 Information Retrieval on Dense Public Displays

It has been very common that people perform *visual search tasks* (Nowell et al. 2002; Hollands & Wickens 1999) finding the detail for one or more particular items (e.g. the flight/train they are trying to catch) which are presented on traditional public displays such as information boards in railway stations and airports (examples are shown in Figure 8.1 and 8.2). On such displays, a large amount of visually similar information items are usually densely arranged and presented on multiple collocated displays or one large size of screen, and the number of information items can run to many hundreds (see examples in Figure 8.2 and 8.3). Users regularly encounter difficulties locating (retrieving) relevant information on the displays and other impacts caused by information overload problem mentioned in Chapter 2. A typical case was reported by Fisher and Kichhanagari (2006) that drivers spend too much time glancing at the airport terminal sign relating to them and thus precipitating slowdowns or crashes.

Studies in the psychophysics have divided factors which affect the efficiency of visual search into three groups (Steinschneider 1990; p.391): (1) nature of stimulus or visual search itself (complexity of target, structure of search field); (2) specific attributes of observer (attention, memory, age, training); and (3) external conditions such as luminance and size of search field, size and contrast of target. The first and third groups of factors are mostly considered by interface designer and information architect in practice. Furthermore, the ways to decrease the time of visual search on high-density information display were proposed, such as using a *Hierarchical Set* or *Tabular Set* (Tidwell 1999).

Figure 8.1 Train information displays in a train station in London

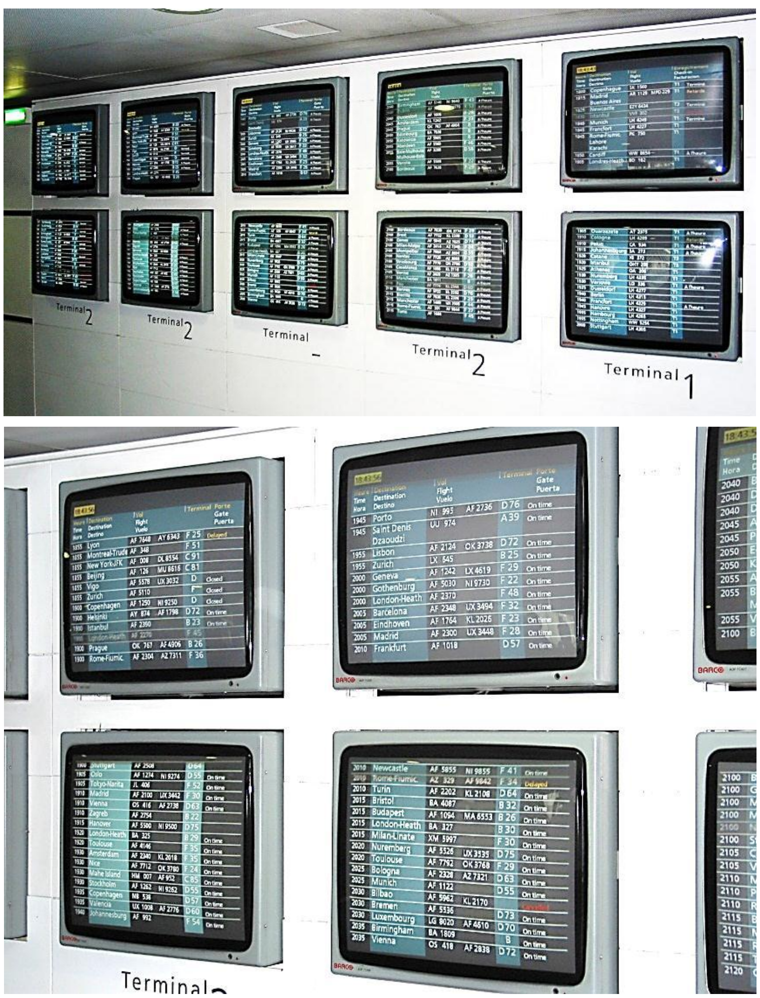

Figure 8.2 Flight information displays in the airport of Paris, France (CDG-Charles de Gaulle)

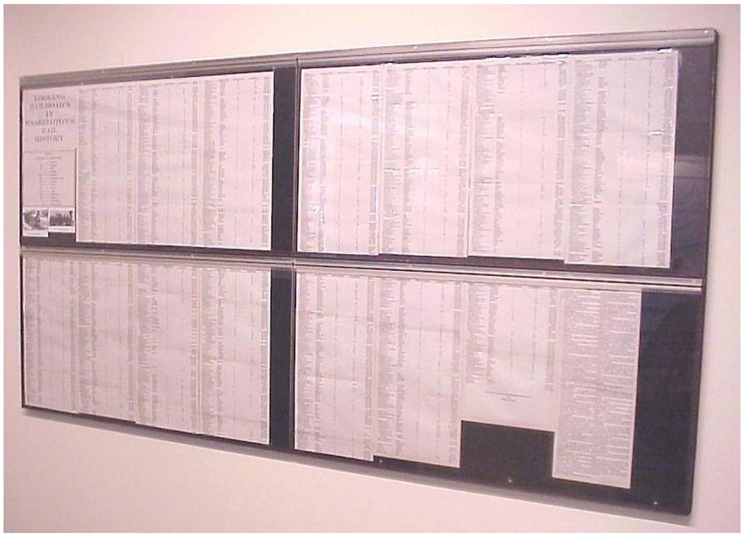

Figure 8.3 Large amount of visually similar information presented on a blackboard

However, previous researches only focus on improving visual information presentation, few of them considered to use multimodal information presentation. Most of HCI researchers try to improve desktop displays, the public displays we interested has received little attention (Shneiderman & Plaisant 2005).

Before we provide the solution to both the problems identified in Chapter 2, we need to understand how people access information on a dense public display as typified by a flight information board in an international airport.

There are a range of information retrieval tasks currently supported by traditional public displays that present high-density information. We classify them as follows:

- Retrieving indexed information
- Retrieving un-indexed information
- Relational queries (multiple retrievals)
- Monitoring

Retrieval of indexed and un-indexed information relates to the nature of the activity that a user is engaged in. For example, the principal function of an airport display board is to inform users the updates of the departure times of aircrafts and the relevant gate from which the departure will occur. Thus the flight time (and destination) is the principal index by which a tabular display of the information is organized. The layout corresponding to this indexing is conventionally (e.g., in international airports) top-down and left-to-right, both within and between multiple collocated displays. Consequently un-indexed information, such as the time and destination corresponding to a particular flight number can only be retrieved through sequential search (or other ad hoc strategies).

Where the quantity of information is particularly large, there are well-understood conventions for limited temporal multiplexing of its presentation on public displays. This can occur within, and between elements of the display. Within-element multiplexing occurs when a field sequentially displays related information, for example, the identifier of the planned departure gate for a flight, and instructions such a “wait in lounge”. Between-element multiplexing is used occurs where the amount of information is too large to show at one time on the displays available. Thus displays scroll through pages of information and include subheadings indicating the screen number currently being displayed and the total number of screens.

Relational queries and monitoring both require the stable and indexed display of information. For example, relational queries, such as which of two flights leaves first, involves the identification of the two flights, their relative locations on the display, and a mapping from these locations (via the indexing) to the query. Likewise, monitoring, such as checking for state change in a flight departure display, requires the identification of the spatial location of an element and easy repeated retrieval. Both tasks require that the elements displayed on the board are relatively stable, that the indexing of the information does not change, and that changes to the location of displayed information occurs in a well understood manner. In case of a departure board the entries move up, and to the left, as time passes.

The stable display of information also has an impact on privacy and collaborative use of displays allowing easier verbal and multimodal co-referencing and shoulder surfing (i.e., identifying what information a user is attending to).

8.4 Design a Crossmodal Display for Information Retrieval on High-density Information Displays

8.4.1 *Goals for CROSSBOARD*

Based on the conceptual model of Crossmodal Display proposed in Chapter 3 and visual search literature in Section 8.3, we explore a new type of Crossmodal Display, CROSSBOARD to facilitate multiple users accessing individual entries that are densely presented on large public displays such as flight departure boards in international airports.

Within the scope of the main aims of Crossmodal Display, we specify the following design goals for CROSSBOARD:

- **Display personalized information by attention guiding**

The CROSSBOARD system aims to pinpoint information relevant to a user by directing the users' attention to the portion of the screen that shows information relevant to them, instead of filtering information according to who is viewing the display.

- **For multiple simultaneous users**

The design of CROSSBOARD aims to support a large number of simultaneous co-located users whose information is all displayed on the public display of the system at once to find the information item relevant to each of them, and the system does not need to attempt to keep track of which user is attending to the public display at a particular moment.

- **Retrieve particular information item among a large amount of items**

CROSSFLOW only supports as many destinations as the sequence of crossmodal cues allows (the sum of the time-slots). With CROSSBOARD, the aim is to support multiple users retrieving their particular information items respectively among hundreds of visually similar information items which are densely presented on collocated displays or a single large screen.

8.4.2 *Hierarchical crossmodal cues and cue mapping*

CROSSBOARD augments traditional public displays by integrating hierarchical crossmodal cues and cue mapping, a type of regular cue mapping which is the core of the design of the system. Here, public cues, hierarchical spatial multiplexed highlighting of regions of the board coordinated with private cues, temporal multiplexed sequence of sound and/or vibration from personal devices, allow CROSSBOARD users to rapidly identify the region of the board that they should search for relevant information.

In order for the user to locate the item of information they are interested in efficiently, we systematically divided the large high-density information display into sections which are highlighted at the appropriate time interval that the user’s private cue is triggered. The unsystematic division will apparently take more steps and longer time to cycle through all the information items, therefore is automatically discarded. The number of levels of subdivision applied depends on the granularity to be presented to the user: a block of items, a list of items or the individual item of interest. More levels of subdivision lead the user closer to the item, but require more cycles of highlighting.

In general, the public cue can highlight all the divided areas in turn, which allows the display to be utilized by several users, whose audio and/or haptic private cues are provided at different time steps. Each user has private cues at each level of subdivision that direct them to attend to a single area. All such areas are then divided again, and the user must attend to the highlighting within the relevant division. As the density of the board increases, the number of time steps required to pick out an individual item in a single level increases. Subdivision trades off the time taken to pick out an area with the cycle time until the user can synchronize with the cues again.

A display can be subdivided in a number of different ways at different levels depending on the layout of items. For a simple grid, a row/column division can pinpoint an item in two cycles (of r and c steps respectively), but if the grid is large, it will take a long time to cycle through all of the rows or columns. Instead the display can be divided into fewer columns or rows made up of more than one item, and then each of these groups is divided on another cycle. For example, in a CROSSBOARD prototype, each screen can be divided into 16 cells, with 5 items in each cell. Subdivision is used to reveal the cell the item is in, and then a further row division is used to pick the item from a cell.

A. Area division

Area division partitions a display into a number of discrete areas. They can be any size and shape as long as they tile the area completely. CROSSBOARD commonly uses binary division, or quartering, and highlights the quarters as shown in figure 8.4.

B. Row/column division

Division by row or column is a special case of area division. A row division has m areas with a width equal to that of the whole display and a height of h/m. A column division has n areas of equal width w/n and a height equal to the height of the section.

Unless the section consists of a single row or column, a posse of rows and a posse of columns are both needed to select a cell in the grid. So, a 2x2 grid can be divided by 1 cycle of 4 square areas (1x1) or 2 cycles of rows (1x2) then columns (2x1). The total cycle time is 4 steps in either case. Selection from within these groups is performed at the next level of subdivision, or left to user scanning. This is shown in figure 8.5.

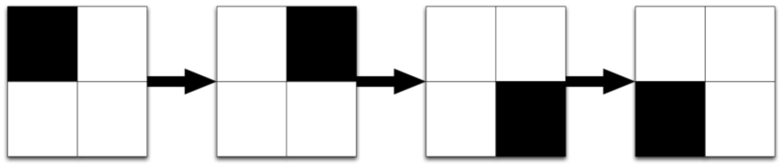

Figure 8.4 Binary area subdivision of a display.

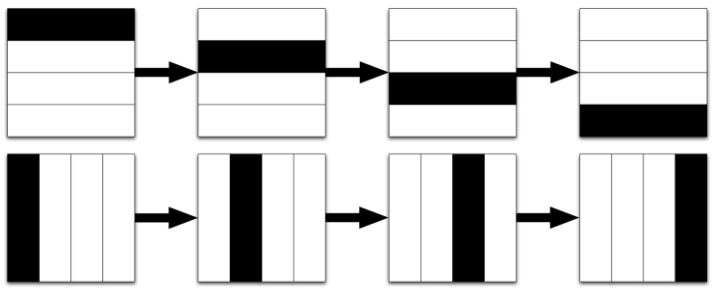

Figure 8.5 Row division (top) and column division (bottom) of a display.

C. Multiple subdivision

For larger arrangement of cells, there is a choice of how to subdivide them. As an example, consider a 4x4 arrangement of 16 cells. A single cell can be indicated by one level of division, using 4 columns and 4 rows, giving a cycle time of 8 steps.

Alternatively it can be divided by two levels of 2x2 area subdivision, of 4 steps each, again a total of 8 steps. In the one-level arrangement the user has to remember the row until the column is selected too. In the two-level, they only have to remember the current section, which is then further subdivided.

Now consider an asymmetrical arrangement of 8x4 cells. A row/column division will give 8+4 = 12 steps. A subdivision of 4 areas of 2x1 then a row division of the 2x1 areas gives 4+2 = 6 steps—twice as fast.

In general binary subdivision will use fewer timesteps than two row/column passes, which are more efficient than highlighting every element. Area and row/column divisions result in the same number of steps if a grid is square or consist of one row or column. However area divisions imply more subdivision, while row/columns require the user to attend to a previously highlighted area for longer. Subdivision of a display into quarters, then each quarter into 5 rows is shown in figure 8.6. The private cue for the selected item in each cycle is shown underneath.

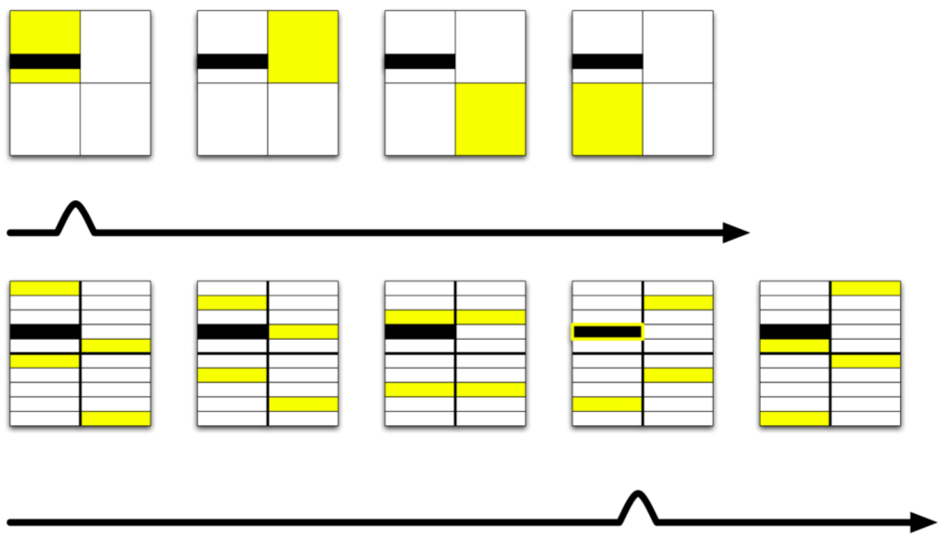

Figure 8.6 Hierarchical temporally multiplexed crossmodal cue and cue mapping in CROSSBOARD.

8.5 Experiment 5: Evaluating CROSSBOARD

A user study was conducted to investigate the effect of crossmodal cues on information retrieval from a large display, i.e. how hierarchical crossmodal cues can influence user performance on retrieval of indexed and un-indexed information.

The study used a within subject design. The rationale of why such experimental design was adopted is the same as which is discussed in Section 5.2.2 A. We utilized one independent variable, type of display, with three levels of treatment (three conditions): (1) no private or public cues (plain display condition, see Figure 8.7); (2) Private and public cues (CROSSBOARD condition, see Figure 8.6); (3) visual public cue only (flash display condition). The dependent variable measured was the time taken to find the required information item.

Each participant needs to perform two kinds of retrieval tasks: (1) retrieval of indexed information; and (2) searching for un-indexed information. All tasks are to find the remaining part of a three-part item given the two other parts. The indexed retrieval task was to find the flight number of a particular flight given the destination and departure time which was indexed on the display. Non-indexed retrieval required finding the departure time for a given flight number and destination which were un-indexed on the display.

Figure 8.7 Display setup, using a 3-screen CAVE in a "wall" arrangement.

8.5.1 *Hypotheses*

It was hypothesized that

- (H1) without the aid of public and private cues (non-CROSSBOARD condition), the participant would perform better at indexed task retrieval than un-indexed, as they can utilize the board ordering to effectively scan for the indexed item. Items that were at the beginning of the ordering would be found more quickly than those at the end.

- (H2) For the aid of public and private cues (CROSSBOARD condition), there would be less difference between the indexed and un-indexed retrieval tasks, as the cues drill down to an item irrespective of whether the indexed information is known or not. It was still expected that the indexed task will be completed in less time, as once at a suitable level of subdivision, the user can leverage the existing board ordering to find the item within a division without needing cues.
- (H3) Using public and private cues (CROSSBOARD) would be much quicker than using a static display when accessing non-indexed information.
- (H4) The highlighted board without auditory private cues (flash display) would not affect the participants' time performance, although they might distract them who do not receive the audio cues.

8.5.2 *Participants*

Eight volunteers (6 male, 2 female) with normal vision took part in this study. They were students or staff from Newcastle University, aged between 20 and 32. The participants have no previous experience in such experiment, and did not know the aim of the experiment. The content and the method of the experiment were agreed by the participants and approved by the supervisor of this research who is in the School of Computing Science at Newcastle University.

8.5.3 *Display configuration*

The test display was based on an airport departure board, with each item giving the flight number, destination and departure time. Items were indexed by departure time, as is the norm. The display consisted of 240 items, spread across three large screens which are the same size (7X8 feet each, 7X24 feet in total). The first cycle of cues selects the relevant screen. A binary area subdivision was used to partition each screen into 16 cells, and then each cell contains 5 items, highlighted by a row subdivision. This setup is shown in figure 8.7, and the sequence of subdivisions in figure 8.8.

In the first two seconds of a display sequence one audio cue will sound in one of four 500 milliseconds that four different regions of the display flash, this indicates a spatial subdivision of the display in which the user's information resides. In the next two seconds a second cue (and coordinated flash) indicates a sub-region of the first subdivision further localizing the information.

The 80 items on each screen are arranged as 20 rows and 4 columns, which can be subdivided in a number of ways ('*a*', '*r*' and '*c*' refer to area, row and column respectively):

- two levels of quartering plus row division

  (4a) + (4a) + (5r) = 3 levels, 3 cycles, 13 steps (1)

- one level of row/column division

  (20r + 4c) = 1 level, 2 cycles, 24 steps (2)

- quartering plus row/column division

  (4a) + (10r + 4c) = 2 levels, 3 cycles, 18 steps (3)

- row/column plus row division

  (4c + 4r) + (5r) = 2 levels, 3 cycles, 13 steps (4)

As can be seen (1) and (4) are equivalent in terms of number of time step and cycles, but (1) utilizes more levels of subdivision. Style (1) was chosen in order to investigate both area subdivision and row/column subdivision while minimizing the total number of steps.

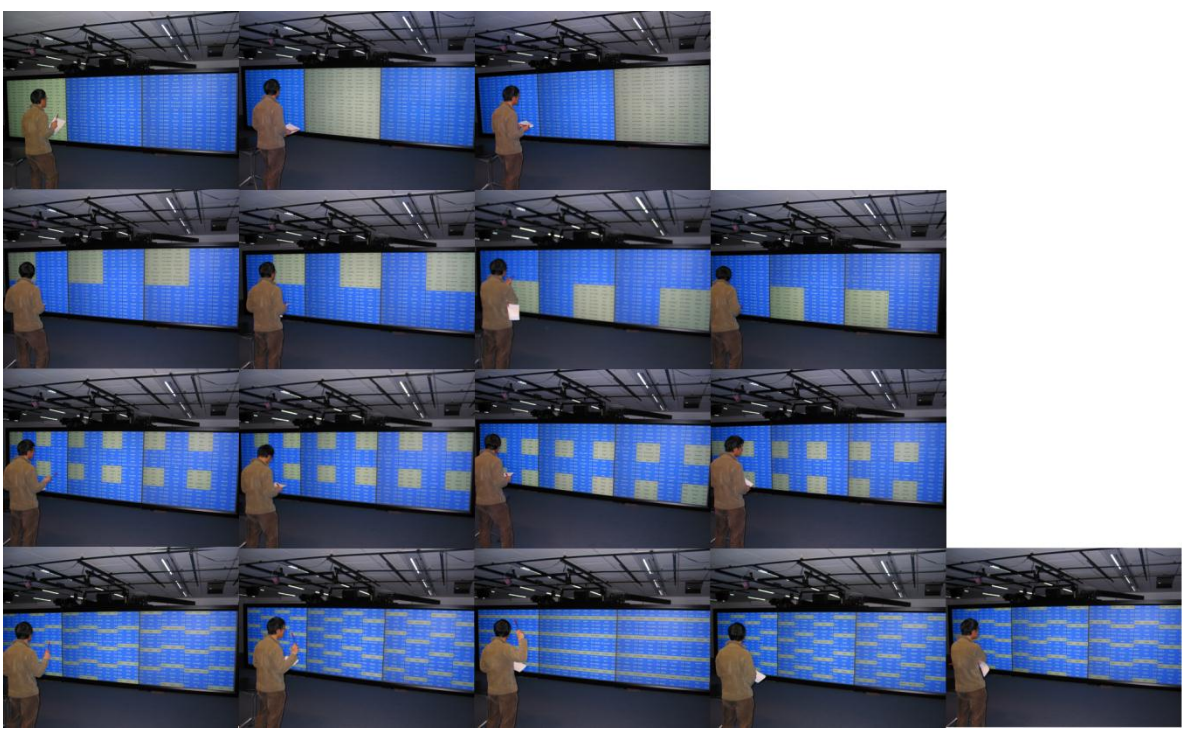

Figure 8.8 Subdivision of the display

8.5.4 *Procedure*

The experiment consisted of 30 tasks. Each participant performed 10 retrieval tasks for each display condition, in a random order one by one: 5 indexed retrievals and 5 non-indexed retrievals. Each task has two steps, in the first step, a two-part flight information item (e.g. 10:00 London) was shown on the screen, a participant was asked to write down to remember and retrieve the remaining part of the information item on the display shown in the next step as quickly as he/she could. In the second step, the visual information display was started immediately (the cycle of audio cue started simultaneously in the CROSSBOARD condition) after a participant informed the experimenter that he/she was ready. Once the participant located the target information item (e.g. 10:00 KLM161 London), he/she indicated his/her response by informing experimenter and reading the remaining part of the information item aloud, then a task ended.

The order of presentation of the displays for each subject was randomized. The flight information was also generated randomly, and there were flights to the same destination with different numbers leaving at different times, and also flights leaving at the same time, but to different destinations. A different board of flight times was presented for each task's iteration. All users received the same boards in the same order.

Before starting the experiment, each participant was asked to stand at a fixed location which was 5 meters away from the middle screen and informed that he/she was free of orientation during the experiment. Participants were not given an extensive training on using CROSSBOARD, but they were shown one example at the beginning of the experiment.

8.5.5 *Results of Experiment 5*

The time performance was calculated from being shown a display to finding the required item (time taken to read out the requested information was ignored).

Treating each task as a one-way repeated measures ANOVA revealed no significant difference across the three conditions for the indexed information retrieval ($F(2,14)=2.286$; $df=2,14$; $p=0.138$). However, the mean value and standard deviation of the Crossmodal Display is the minimum among the three conditions. The performance for un-indexed information retrieval revealed a significant difference across conditions ($F(2,14)=24.659$;

df=2,14; p<0.001). Again, the mean value and standard deviation of the Crossmodal Display is the minimum.

The descriptive statistics (Table 8.1) reveals that the performance for un-indexed tasks in CROSSBOARD condition is 448% better than in plain display condition ((Pu-Cu)/Cu = 448%); 417% better than in flash display condition ((Fu-Cu)/Cu = 417%).

The performance for indexed tasks in CROSSBOARD condition is 71% better than in plain display condition ((Pi-Ci)/Ci = 71%); 182% better than in flash display condition ((Fi-Ci)/Ci = 182%).

	N	Minimum	Maximum	Mean	Std. Deviation
Index(1)	8	6.25	67.11	21.48	21.41
Index(2)	8	4.96	19.44	12.56	4.63
Index(3)	8	7.52	120.55	35.48	38.58
Un-indexed(1)	8	27.84	82.97	53.04	19.52
Un-indexed(2)	8	5.04	18.72	9.68	4.91
Un-indexed(3)	8	27.68	69.48	50.05	16.21
Valid N (listwise)	8				

Table 8.1 Descriptive statistics of the results of Experiment 5

8.5.6 *Discussion of Experiment 5*

These results here support our original hypotheses and suggest that the Crossmodal Display is effective at reducing retrieval time for information on a high-density information display. (H3) Especially when the information is arranged in a fashion unordered by the searched value, the performance for un-indexed information retrieval was significantly better using CROSSBOARD than using a traditional high-density information display (corresponding to H3). Besides in the CROSSBOARD condition, participants performed better at indexed task retrieval than un-indexed (H1). This demonstrates that locating information by index is efficient on traditional high-density information display, while in the CROSSBOARD condition, the difference of time performance between the indexed and un-indexed retrieval tasks is relative small (H2). It was interesting that the un-indexed tasks were completed in less average time than indexed tasks. This result demonstrates that retrieving information only using crossmodal cues is more efficient than using the combination of crossmodal cues and time index. (H4) Although the time performance on a highlighting display without cues was comparable to a non-highlighted display, the participants did comment that the flashing of the highlighting was distracting (H4). All users completed all the tasks using CROSSBOARD successfully without an extensive training though they were shown one example at the start of the experiment, which demonstrates that CROSSBOARD is easy to learn and use.

8.6 Conclusions

This chapter has reported our exploration of the design and an initial evaluation of CROSSBOARD, a Crossmodal Display prototype system that combines hierarchical spatial multiplexed visual public cues with coordinated hierarchical temporal multiplexed audio and/or haptic private cues to highlight the locations of information items relating to multiple simultaneous users on a large high-density information display. The cues are associated with regions of the display that are flashed in sequence. Depending on the number of items displayed, these regions may then be divided into sub-regions, which also have crossmodal cue combination associated with them. As the cued regions become smaller, it allows a user to rapidly search the region for the required information. This hierarchical cueing is designed to not affect the use of display in a traditional manner without cues, but allows those users with cues both to narrow down a region of the display that is searchable by some indexed part of the information item (e.g., flight departure time) and also to quickly locate an item that is not searchable by an index, which would otherwise require systemic scanning of the whole display.

The initial user study on CROSSBOARD (Experiment 5) mainly investigates the usability of CROSSBOARD. The results have demonstrated the effectiveness and efficiency of CROSSBOARD as a solution to help to solve the problems identified in Chapter 2 and Section 8.3, in particular, validated the following distinct advantages over the previous solutions we reviewed:

- **Facilitate information retrieval on high-density information displays;**

  We have demonstrated that hierarchical crossmodal cues in CROSSBOARD can be utilized to provide greatly enhanced retrieval performance for un-indexed information on dense public information displays without adversely affecting the performance of conventional display usage (for the retrieval of indexed or unindexed information).

- **Not use or only use limited sensing or tracking;**

  This is one of the significant advantages of Crossmodal Displays over previous display-based systems that highly rely on tracking technologies, which is demonstrated by the CROSSBOARD system.

- **Large user capacity.**

  This is another significant advantage of Crossmodal Displays. In theory, a Crossmodal Display such as CROSSFLOW or CROSSBOARD can support a large amount of users as it is only limited by the processing capability of the server handling users' synchronization requests.

Although the following aspects are not validated directly through this evaluation, we have planned the methods to do them in future work.

- **Support multiple simultaneous users accessing different information content without resorting to tracking users;**

  A study of multiple simultaneous users can validate this, and the experimental method has been demonstrated in Chapter 5.

- **Preserve privacy and anonymity in physical public spaces;**

  The research into the user experience of multiple simultaneous users afforded by CROSSBOARD can validate this.

- **Impose lower mental workload demands on users than traditional static displays.**

  The subjective task workload on users could be compared between CROSSBOARD and a static display in a user study so as to validate this.

How this chapter provides the answers to the research questions identified in Chapter 2 as well as the contributions to the design and evaluation suggestions for CROSSBOARD and Crossmodal Displays in general are summarized and discussed in Chapter 9.

Chapter 9. Conclusions and Future Work

9.1 Introduction

This chapter summarizes the work presented in this thesis, and relates back to the research questions outlined in Chapter 1 and 2. It also suggests a lifecycle model of Crossmodal Display development and provides a set of design suggestions for designers or researchers who wish to develop Crossmodal Displays for their applications or research. Then the limitations of this research including the limitations of infrastructures of Crossmodal Displays, design and evaluation of CROSSFLOW and CROSSBOARD are discussed, and suggestions for future work based on the issues raised by the thesis are proposed.

9.2 Summary of Research Carried Out

We summarize how our research questions are addressed in this thesis as follows.

RQ1: What are the significant problems in physical public spaces in the age of ubiquitous information and displays?

RQ1 is addressed in Chapter 2. In the chapter we identified two significant problems and their impacts on individuals and societies, that is, information overload problem and information presentation for multiple simultaneous people who share a physical space or situated interface but have different information needs and privacy concerns, i.e. *multi-user* problem.

RQ2: Could Crossmodal Displays, display-based interfaces that exploit people's multimodal perception and utilize the synergy of both existing public displays and personal displays help to tackle the problems in physical public spaces in the age of ubiquitous information and displays and avoid the drawbacks associated with previous solutions including a reliance on tracking technologies, weak protection for user's privacy, small user capacity and high cognitive load demands?

RQ2 is addressed in Chapter 3, 4, 5, 6, 7 and 8. In Chapter 3, we reviewed the literature of research into human multimodal integration and crossmodal interaction and drew the design implications which could potentially be exploited for the design of multimodal interfaces or displays. Based on the design implications and existing public and personal displays, we propose a systematic solution, i.e. a conceptual model of Crossmodal Display aiming to help to solve the problems (identified in Section 2.3) and avoid the drawbacks (identified in Section 2.4) associated with previous solutions. In Chapter 4, 5, 6, 7 and 8, we developed two types of Crossmodal Display prototype applications: CROSSFLOW for indoor navigation and CROSSBOARD for information retrieval on high-density information displays. The evaluation results of the prototype systems reported in Chapter 5, 7 and 8 have demonstrated the effectiveness and efficiency of Crossmodal Displays on solving the problems identified in Chapter 2 in two specific application areas, i.e. indoor navigation and information retrieval and validated the distinct advantages over the previous solutions we reviewed.

RQ3: What are the usability and the user experience afforded by Crossmodal Displays that are developed for specific applications?

RQ3 is addressed in Chapter 5, 7 and 8. We evaluate the usability and the user experience afforded by two versions of CROSSFLOW and one version of CROSSBOARD through one multi-user (reported in Chapter 5) and four single-user (reported in Chapter 5, 7 and 8) lab studies. Most of the evaluation results are encouraging for the reason that most of the design goals of CROSSFLOW and CROSSBOARD and the aims of Crossmodal Displays are achieved and the advantages of the displays are validated (summarized in Table 9.1 below). Although partial results are not ideal, this might be result from the experimental design approaches, operational errors, inherent limitations of the systems, training or other factors. Some of the factors are discussed in Section 9.4.

Moreover, based on our experience of the exploration of Crossmodal Displays and inspired from *The Usability Engineering Lifecycle* (Mayhew 1999; Sharp et al. 2007, p.461), we propose a lifecycle model of Crossmodal Display development (Figure 9.1) in order to help interaction designers or researchers who wish to develop Crossmodal Displays for their applications or research. The model is intended to be suggestive instead of prescriptive; other ways to develop or improve Crossmodal Displays are also possible.

Problems	Drawbacks of previous solutions	Aims for Crossmodal Displays	Advantages, usability and user experience of CROSSFLOW	Advantages, usability and user experience of CROSSBOARD
P1) info overload problem	High cognitive load demands	Prevent information overload;	Facilitate navigation in certain type of indoor environment. Impose lower mental workload on users than hand-held paper map in certain type of indoor environment, and support multi-tasking while navigation (demonstrated by CROSSFLOW V1).	Facilitate information retrieval on high-density information displays. Impose lower mental workload demands on users than traditional static displays (future work to validate).
P2) multi-user problem		Effectively satisfy different information needs of multiple simultaneous users who share a physical space or situated interface;	Support personalized navigational information presentation for multiple simultaneous users without resorting to location tracking.	Support multiple simultaneous users accessing different information content without resorting to tracking users (future work to validate).
	A reliance on tracking technologies	Employ minimum sensing or tracking;	Not use or only use limited sensing or tracking.	
	Weak protection for user's privacy	Preserve users' privacy and anonymity in physical public spaces.	Preserve certain level of privacy and anonymity in physical public spaces.	Preserve privacy and anonymity in physical public spaces (future work to validate).
	Small user capacity		Large user capacity	
			Special user experience: Aesthetically pleasing and fun to play with	

Table 9.1 Summary of the main research outcomes

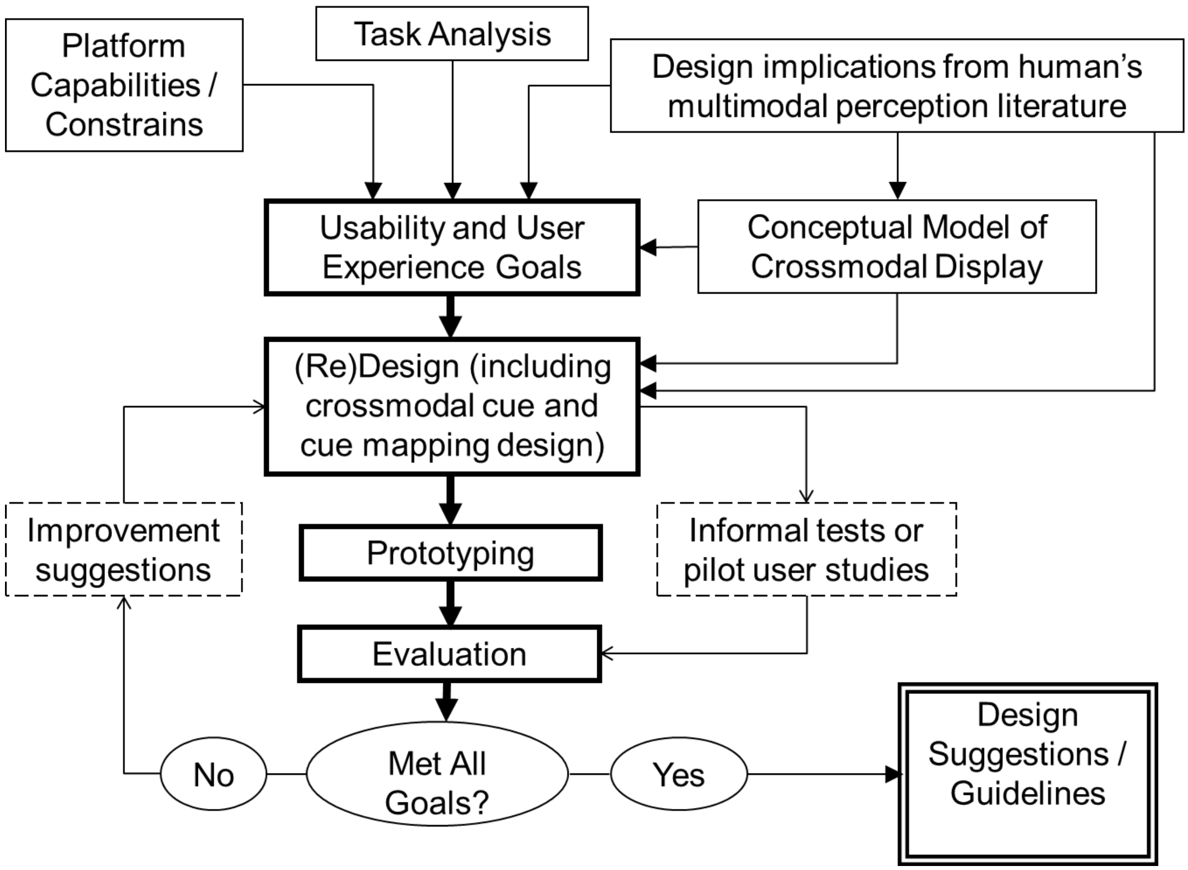

Figure 9.1 A lifecycle model of Crossmodal Display development

9.3 Design Suggestions for Crossmodal Displays

9.3.1 *Design suggestions for CROSSFLOW*

A. Design suggestions for visual public cues of CROSSFLOW

There are many visual modality parameters and properties which could be controlled in the design of visual public cues for directional information presentation, such as the size, colour, motion, quantity, density of the graphical elements. Both graphical design possibilities and the possible effects on the usability and the user experience afforded by CROSSFLOW should be considered.

Based on the literature of Ambient Displays and Peripheral Displays reviewed in Chapter 2, the design explorations in CROSSFLOW V1 and V2, and the analysis of findings of Experiment 1, 2, 3 and 4, we draw the following design suggestions for visual public cues, i.e. public directional information of CROSSFLOW. These suggestions can be used as guidelines for designing and selecting the appropriate visual public cues.

1) Individual graphical element

a) The individual graphical element of the ambient pattern should be recognizable in the environment and clear to CROSSFLOW users.

b) The size of the individual graphical element should be as large as the size of a palm of a normal adult at least.

c) One part of the individual graphical element should be fixed instead of moving.

d) The shape of the individual graphical element should be easily recognized as pointing in a direction. Furthermore, any two different directions that an individual graphical element pointing in should be easily distinguished by users.

Note: All the above suggestions are not for the graphical element in dynamic animation state but for the static state in which the element indicates a location or object in a physical space most clearly.

2) Visual directional information

Visual directional information corresponding to a location or object in a physical space should be intuitive to be understood. The indication of different directions or locations should be explicit and clear in both large-scale and small-scale spaces, which means when users are searching a destination, they should be able to intuitively distinguish different directions or locations nearby that are indicated by the visual public cues in different time slot.

3) Transition animation

No transition animation should be between any two visual public cues successive in time or make the transition animation undetectable by users.

4) Distance information

Distance information, e.g. how far from the location of a user to his/her destination is, might be not so important for indoor navigation. Furthermore, adding distance information, such as numerical data to the visual pattern of the CROSSFLOW system may impose higher attention and cognitive load demands on users. However, it might be able to be delivered in an abstract way such as by the shape or the length of the graphical element.

5) Distribution

Public directional information should be shown at all decision points in the navigational space at least. Showing directional information throughout the whole physical extent of the navigational space such as projecting the ambient pattern everywhere have both advantages such as that users can easily discover the directional information without visually searching them, which may be necessary for the space which contains relatively large open areas, and disadvantages such as that impair the visibility of other information in the environment (e.g. signs and landmarks) if the visual public cues could not be embedded in the environment naturally.

6) Duration

800 milliseconds as the duration of the time slot and the length of visual public cue is most appropriate and efficient for directional information identification, which is proven in the results of Experiment 3 in Chapter 7.

7) Intrusiveness

The visual public cues should not be intrusive, but be naturally embedded in the environment and be as peripheral as possible. This suggestion is inspired by the concept and the design principles of Ambient Displays and Peripheral Displays which are intended to make users process information delivered by the displays or systems with minimum attention. The suggestion is also consistent with the vision of ambient displays that seamlessly integrates digital information (the navigational information for CROSSFLOW users) into physical environments (the navigational environments) around users.

Although the dynamic changing visual public cues may be viewed by users as augmented and dynamic signs embedded in the navigational environment and might attract people's attention more easily than traditional static signs, a public cue should exist in the background of people's attention and only demand the full attention of CROSSFLOW users when it is enhanced by the corresponding private cue.

8) Aesthetics

To be aesthetically pleasing is one of the important features of ambient displays. It can be adopted as a design principle and selection criterion for visual public cues used in CROSSFLOW.

B. Design suggestions for auditory private cues of CROSSFLOW

1) Usability

Usability should be considered at first. Therefore, the precondition for auditory private cues should be simple and clear. Furthermore, users should be able to easily discriminate between the private cue of CROSSFLOW and common phone call signals.

2) Semantic correlation with visual public cue

Private cues should semantically relate to the corresponding visual public cue. Drawn from the multimodal perception literature in Chapter 3, if different representations of information and different modalities linked semantically well, the perception of the information would be better. Based on this, we design the private cue of CROSSFLOW V2 by simulating wind sound in the natural environment. The findings from the evaluation results demonstrated that the users can associate the wind-like sound private cue with the grass-like pattern animation easily and naturally. Therefore, we suggest that private cue and corresponding visual directional information should be both temporally and semantically correlated in order to achieve relative better user experience.

3) Timing

The private cue should be presented earlier than the corresponding visual pattern indicating the user’s target, i.e. the upcoming visual public cue relevant to the user, so as to cause the user's pre-attention to it.

4) Aesthetics

The auditory private cues should be pleasing to the ear instead of unpleasant. A cue can be a higher pitch mixed in a piece of music. Thus users may be able to choose specifically composed music as their private cues or background music of the private cues.

5) Modality comparison and selection

In the design of the private cue, sound alone, vibration cue alone and the combination of sound and vibration should all be considered and compared. The effects of these three on the usability and the user experience of Crossmodal Displays is not compared formally in this thesis, which is one part of the future work. Moreover, different types of sound as private cues should also be compared in order to identify the most appropriate type. The usability between the wind-like auditory sound used in Experiment 3 and 4 and the single beep used in Experiment 1 and 2 might not be significantly different, but the user experience might be.

9.3.2 *Design suggestions for CROSSBOARD*

Visual public cues, i.e. the public information on CROSSBOARD should not flash to highlight. The participants' complaint about the flashing suggests that the smooth and aesthetically pleasing visual transitions between any two successive highlighted information levels are needed in the visual public cues of CROSSBOARD.

9.4 Limitations of This Research and Future Work

The following subsections discuss the limitations of the research reported in this thesis and propose a number of avenues for future research in this area.

9.4.1 *Limitations and future work for the research of Crossmodal Display infrastructures*

The two prototypes (CROSSFLOW and CROSSBOARD) address two tasks – navigation and information retrieval. They can be viewed as components of a wider Crossmodal Display infrastructure that could be deployed in a physical public space to provide contextualized navigation information to users within that space. In a basic configuration, the public displays of CROSSFLOW and CROSSBOARD can be deployed in such a manner, with ambient patterns as directional information projected within the space, and dense information displays augmented with hierarchical spatial multiplexed visual public cues at appropriate locations. In a usage scenario where users have networked personal mobile devices, a user's interaction is then simply a matter of communicating with the central server of the Crossmodal Display and acquiring the correct time slot for notification. The contextualized information can be sent to these devices respectively. Such technology has distinct advantages. Firstly, the intuitive navigation directions are directly embedded in the environment without users having to direct their attention to a separate device in their hand. For information retrieval, external public displays which can be situated contextually in an environment (e.g. near a specific gate/platform) give a large area to display information. In addition, on a large display, if enough distinct items of information need to be displayed, more detail per item could be presented on a user-carried device, since seeing more information than that which is directly relevant to a user can be useful in some situation. For instance, seeing what occurs before and after the event the user is interested in, e.g. "the next available train/flight" to a destination.

A. Cue Delivery

Although neither CROSSFLOW nor CROSSBOARD utilize a user's personal computing device to deliver explicit information, it is the means by which the contextual cue is given to the user. This may be a sound, or a vibration, or some other sensory cue. If the device is networked, then it can coordinate its time value with the crossmodal system, and be told a specific time to trigger the private cue. However, one of the aims of the Crossmodal Display system is to deliver private cues independent of an external network, to support a wider range of devices and scenarios where constant network availability is not feasible (e.g., wireless dead spots, lack of GSM signal, restricted environments, cost of connectivity). In this case, the device gives a private cue at a pre-defined interval, and this interval is synchronized with the rest of the crossmodal system. This implies two main problems need to be solved for a particular crossmodal infrastructure: how to calculate appropriate time intervals, and how to synchronize the private cue device with the system.

As explained earlier, CROSSFLOW uses a regular cue mapping, where a set number of destinations are cycled through regularly, implying that for n destinations and a wait time of Δt that the interval between private cues is $n\Delta t$. There will be a lower limit on Δt by which a user is able to distinguish between different locations. They must be allowed time to react to a private cue and associate the private cue and the corresponding public cue with the current destination before the public cue changes. Long intervals between private cues will mean the user has to interrupt their activity (typically walking), wait and/or attend more closely to the public cues in order to identify the direction of the destination. In addition, given the fixed duration of time slot, the more destinations in a navigational environment, the longer time users have to wait for their private cues to appear and reappear, which might result in lower navigation task performances. This is an inherent limitation of CROSSFLOW. However, carefully designed cue mapping might be able to reduce the detrimental impact of this limitation such as allocating shorter presentation time of public directional information (visual public cue) corresponding to those less important destinations.

CROSSBOARD uses a hierarchical decomposition of a display to narrow down the portion of relevance, so for a decomposition of n levels, the user receives n private cues in a particular cycle. The user may not need all private cues to identify their information if they have other cross-indexing information. For each level, there are $n(i)$ divisions, and therefore $n(i)$ cues. If the wait between cues is again Δt then a single level will last for an amount of time equal to $n(i)\Delta t + \Delta c$ (where Δc is a wait between levels, that may be different to the wait between private cues in a single level). So, for a complete cycle, the time is $n(i)\Delta t + \Delta c + C$ (where C is the wait between cycles). For a large number of items, this is less overall cycle time compared to a linear division, but more cues. However, a user does not necessarily need all cues to locate information as all the information is displayed at once and the user can use cross-referencing information to help locating.

Synchronizing the cues with the server of a Crossmodal Display system depends on the network connectivity between the server and the personal mobile devices. As mentioned before, if a device is constantly connected to a network, the server can deliver cues synchronously to the device. However, the usual case is that the device will partially or totally be disconnected from the system's network. Therefore, the device must be able to run asynchronously. This is achieved by having the cycle times programmed into the device, along with the cycle division to be cued. The beginning of the cycle must then be synchronized with the server. Partially-connected devices can retrieve all of this information from the network when they are connected and update it whenever they re-connect to keep the cues in synch. For disconnected devices, the cycle times of the server must be entered directly on the device and synchronization performed manually. One way to do this is to have a server-generated start-of-cycle cue displayed somewhere and, when this is seen, the device’s cycle is started. Care must be taken to factor in user-reaction time as with any other cue. A semi-automatic way of recognizing the cue could be achieved by using the device’s microphone or camera to pick up the cue.

B. Coordination and Contextualization

Integrating CROSSFLOW and CROSSBOARD into a larger infrastructure requires contextual information. The system needs to know which navigation list the user is following or which CROSSBOARD they are attending to. If they switch between a CROSSBOARD and CROSSFLOW, then the cue mapping must be updated. In order to make the users distinguish the cue mappings for both CROSSFLOW and CROSSBOARD easily, different modality private cues can be used, i.e. auditory private cues for CROSSFLOW and haptic private cue for CROSSBOARD or vice versa.

By placing a CROSSBOARD at the locations in a navigation list of CROSSFLOW, the choice of a navigation destination informs the CROSSFLOW system the location of the CROSSBOARD a user is interested in. Synchronization of the user’s mobile device can be managed across CROSSBOARDs and CROSSFLOW, so all that needs to be done is to make the system know when the user reaches the CROSSBOARD. When the user informs the device he is at his destination, the cues are switched to that for the CROSSBOARD at that location or the next navigation cues. The device could be programmed with or stores the cue mappings for all appropriate CROSSBOARDs, and then the user goes through the synchronization step when he reach a CROSSBOARD, such as manually entering the CROSSBOARD ID or using a camera, microphone or network connection to decode the display details.

Once the system knows which CROSSBOARD a user is attending to, information can be further contextualized. That board can add extra details for users it knows are attending. This can be used as a limited form of privacy to ensure information is only displayed when needed. Critically private information can be stored on the user’s device which displays the augmented information when attending to a board. As well as other users, consideration must be given to non-users who still are using the board. The initial CROSSBOARD experiment showed that the presentation of visual public cues did not impair a non-cued user’s performance, though the users anecdotally complained that it was distracting.

The final problem of coordination is when a board updates timely visual information. If the visual information item is sorted on the board, this may affect the portion of the board a user's information is in, and hence the cue mapping that applies to the whole board. Thus, either the cue mapping for that user's device needs to be updated automatically which might be realized in the way that the server "pushes" the new cue mapping to the user's device or better synchronization schemes between the server of the CROSSBOARD and the mobile devices of users should be employed.

9.4.2 *Limitations and future work for the studies of CROSSFLOW*

Experiment 1, 2, 3 and 4 are lab-based user studies which were conducted in the lab environment with specifically designed physical configurations. The environment and tasks used are different from real situations. In the real world, indoor navigation commonly takes place in an environment which has signs and landmarks. When people navigate using a tool such a hand-held paper map or an interactive system, they might also use signs and landmarks at the same time. As Sharp et al. (2007; p.641) indicated that "*ecological validity concerns how the environment in which an evaluation is conducted influences or even distorts the results"*, the influences of the environment in which an evaluation is conducted should be concerned in the evaluation of an interactive system. Thus CROSSFLOW should be tested in a real navigational environment in which there are natural navigation aids such as signs and landmarks.

People do not generally navigate while performing arithmetic questions and the spaces they occupy are usually populated by other people undertaking a range of activities. In Experiment 1, the task of performing mental arithmetic is a very particular kind of cognitive task that is not necessarily representative of the kinds of tasks people would actually be engaged in navigation in the presence of CROSSFLOW. Therefore, to test cognitive load, the experimental design should refer to the real world navigation scenarios such as talking or listening to music while navigating.

In Experiment 2, all navigation tasks in the two conditions led to a high degree of completion time variability. Increasing the number of test participants or further constraining the experimental conditions may lower this.

In order to continue to improve the usability and the user experience of CROSSFLOW, formal comparison of navigation task performance and subjective user experience between CROSSFLOW V2 and V1 should be made. Moreover, further user experience afforded by CROSSFLOW including the aspect about privacy and anonymity should be investigated in future work.

9.4.3 *Limitations and future work for the studies of CROSSBOARD*

Although the results of Experiment 5 support and validate the applicability and the usability of CROSSBOARD and the proposed conceptual model of Crossmodal Display, the measures of cognitive load (e.g. NASA TLX) and the detailed user experience (e.g. the privacy and anonymity aspect, the subjective preference level about the parameters of cue mapping, the visual pattern and the sound) should be incorporated in the future iteration of user studies. The experiment just studied the single user using crossmodal cues on a high-density information display in a fixed setting, while the applicability and the user experience of multiple users simultaneously accessing personalized information on CROSSBOARD should be investigated in the future. The evaluation approach used in the multi-user study of CROSSFLOW in Experiment 2 could be used as a reference.

With a view to increasing the ecological validity of our empirical paradigm we also plan larger scale multi-user and multi-display user studies in natural environments. More ecologically natural environments will also permit the study of additional properties that are desirable in a public information displays such as *legibility at a distance*. In many spaces and contexts, multiple simultaneous users access public displays in a dynamic manner; both reading the displays while moving, and/or walking close enough so as to be able to retrieve information. One useful feature of CROSSBOARD is the location of information relevant to the users which is more legible at significantly greater distances than the specific detail of the element. For example, the location of information about a particular flight on a departure board (e.g. a highlighted large section) is apparent at distances at which the text is not readable. We anticipate that in dynamic environments users will be able to beneficially coordinate their use of public displays as a result. For example, moving towards sections of the display where their information resides, and only moving close enough to be able to read a single entry rather than a distance where they can comfortably scan all entries.

9.4.4 *Future work for the research of crossmodal cues*

We need further investigations and user studies about how the following factors or parameters, and the effects of multimodal integration and crossmodal interaction reviewed in Chapter 3 affect the usability and the user experience afforded by Crossmodal Displays in the next stage of exploration so as to answer the questions proposed here:

- Timing window： What are the lower and upper limits on time window between the public cue and the corresponding private cue within which the enhanced perception of the public cue could be generated?
- Semantic correlation: Whether the semantic correlation between public cue and the corresponding private cue can significantly improve both the usability and the user experience of Crossmodal Displays?
- Modality selection: Whether there are significant performance differences among using sound alone, vibration alone and the combination of sound and vibration as the private cue of a Crossmodal Display?
- Spatial location of private cue: Whether the location of vibrotactile private cues on a user' body affects the usability and the user experience of Crossmodal Displays?

Appendix A: Experiment 2 Files

A.1 Experiment 2: Experiment introduction for the observers (CROSSFLOW V1 multi-user lab study)

The purpose of this experiment is to compare paper map and CROSSFLOW system for indoor wayfinding tasks. The whole experimental procedure and timetable were described in the spread sheets.

<u>Setting up the formal experimental environment (Devonshire Building)</u>

- Remove the stuff on the atrium of Devonshire Building
- Set up **2 projectors** and **1 tripod** on the 4th floor. One projector mounted on one tripod projects the grid image with virtual walls to configure the real foam walls. The other projector use wooden frame (check the previous setting photos).

Projectors synchronization: via server-client wireless connection.

Step 1: Run Arrow server program on one wireless connected laptop.

Step 2: Create a shared file.

Step 3: Run Arrow client program from the shared file on another wireless connected laptop.

Note: If the displays are unsynchronized, synchronize projectors manually.

- Connect **2 laptops** as the servers with 2 projectors, at least **1 web camera** to monitor the CROSSFLOW pattern projected on the floor
- Configure the **28 polystyrene foam walls** on the atrium
- Set up **2 video cameras** with **2 tripods** on the 3rd floor; prepare at least **4 blank video tapes**
- Get **1 digital camera** to take experiment photos
- Prepare **6 smartphones,** including the software settings.

7 or 8 destinations? (Depending on the whole time):

7 destinations; 5600ms total time; 800ms per time slot---decided!

8 destinations; 5600ms total time; 700ms per time slot

How long the vibration should be? (Depending on the findings in psychology research)---400ms vibration per time slot ---decided

Phone synchronized with projection---we have not realized the Bluetooth connection yet, now we just synchronize manually.

- Decorate **8 out of the 18** target walls using **cover papers**, **numbered papers**, **envelopes, tapes…**

- Set up **4~5 computers with NASA TLX software installed**

- Set up **2 computer loudspeakers** and **1 laptop** to prepare a public address system

- Prepare the **recording of the public announcement** for secondary task of the experiment

- 2 Extension Cable Reels

The material needed:

- For the subjects:

1) Subjects recruitment ads
2) Training and formal experiment instruction.
3) Bird view map of the training area in Map condition
4) Bird view map of formal experimental area in Map condition, only including target locations in Map condition
5) Subjective mental load questionnaire: **NASA TLX software**
6) **A paper or Word version of Post Experiment Questionnaire** for each participant (assess the user experience and capture the suggestions about improvement).
7) **2 task books** for each participant
8) **6 colorful hats** for marking the participants
9) **Questionnaires testing the announcement memorization after each condition**

- For the observer:

1) Observer's instruction,
2) Bird view map of the training area in Map condition and
3) Bird view map of the formal experimental area, including target locations in both Map and CROSSFLOW conditions and distracted locations.

Setting up the training environment (VR room in Devonshire Building)

The physical configuration should base on the training map.

Stuff need:

1) 1 projector with 1 tripod,
2) 1 laptop with Arrow program installed,
3) 1 smart phone,
4) 4 low tables,
5) Some chairs.

Training:

In the training section, the participant should be individually called to enter a large room and be trained to use a paper map and CROSSFLOW system (a mobile phone combined with projected light pattern) to find a target location in the room. Give the participant the paper instruction once he/she is free during this section. Until all the participants in the same group finished the training section, then guide them to Devonshire Building to attend the formal experiment.

Formal experiment:

In the Devonshire Building, the group has two sets of tasks in two conditions:

1) Find a set of target walls within this maze area through a paper map (Map condition),
2) Find another set of walls through CROSSFLOW system (CROSSFLOW condition).

Each time, all group subjects should start wayfinding tasks at the same time at the same location. Please give the participants the instruction and the material (labelled hats, maps, task notebooks, paper maps or smart phones) and tell them when and where to start their tasks.

Note:

1) **Before the whole experiment start**, make sure the subjects use the wayfinding tool correctly and efficiently. Indicate that they should finish their own tasks as quickly as they can once they start their tasks, but **do not run**! In addition, Emphasize that

they are **not** expected or allowed to **compete and communicate with each other** during the experiment.

2) **Before the whole experiment start**, make sure the video camera always works and records the whole experiment procedure.
3) **Before the whole experiment start**, make sure the subjects know how to do once they find a target wall. This is the instruction for them: "When you find a target wall, please remove the cover paper to see the number on the wall to confirm whether you have found the right wall or not. Once you find the right wall, please rip off the task page and drop it in the envelope attached to that wall to demonstrate that you have finished the task. Then you can start the next task. "
4) **Before each condition start**, play the public announcement recording. When the first subject finishes his/her tasks in the condition, stops the recording immediately. Let the finished subject wait until all the other subjects finish their tasks. Then let all of them answer the questionnaire testing how much they can memorize about the public announcement. (Make the subject expose to the hearing for the same long time, and after the same long time to test their auditory memory. Although in the second condition, the subject will pay more attention to the hearing, but we counterbalance this effect by running reversed condition on the other group of subjects.) After that, ask them to start the next condition.
5) **Before CROSSFLOW condition start**, make sure the projected flows of 2 projectors synchronize with each other, the flows point to the right wall and the smart phones are synchronized with the projected flows.

At last:

Once a subject finished all his/her the wayfinding tasks, guide him/her to a computer to complete the NASA TLX questionnaire and a Post Experiment Questionnaire, then finish all his/her experiment.

Post experiment clear up:

1) Collect NASA TLX questionnaire documents and Post Experiment Questionnaires
2) Collect and label the video tapes and digital photos clearly.

3) Mark or take photos of the spatial positions of all equipment, including projectors, laptops, video cameras, 28 foam walls on the atrium...
4) After the whole experiment finished, restore all the facilities to the original places…

Figure A.1.1 Formal experiment environment (in GtkRadiat, in Grid 64, 1 square is counted as = 1 meter in real world)

A.2 Experiment 2: Experiment introduction for the participants (CROSSFLOW V1 multi-user lab study)

Training:

Good evening, Ladies and Gentlemen. Firstly, thank you for being willing to take part in this experiment. The purpose of this experiment is to compare paper map and CROSSFLOW system for indoor wayfinding tasks. The whole experimental procedure contains two sections. In the first section, you will be called to enter a room and be trained to use a paper map and CROSSFLOW system (a mobile phone combined with projected light pattern) to find a target location in the room. You will be given a detailed instruction about the formal experiment during this section. When you finished, please wait for further guidance of the experimenter.

Formal experiment:

In the second section, you together with other participants as a group will be guided to another place. In that place, you will see a maze configured by some polystyrene foam walls located on the atrium of the building. You can imagine this is an Alice in Wonderland game. Your tasks include two parts,

- Find a set of target walls within this maze area using a paper map, and
- Find another set of walls using CROSSFLOW system.

You will be provided with tools (task notebook, paper map or smart phone) before you start. The experimenter will indicate when and where to start your tasks. Once you finish a part, please wait the experimenter for further instruction.

Note:

It is important to use your wayfinding tool correctly and efficiently. Once you start your tasks, you should finish your tasks as quickly as you can using your tools, but do not run! In addition, you are not expected to compete and communicate with other people during the experiment.

When you find a target wall, please remove the cover paper to see the number on the wall to confirm whether you have found the right wall or not. Once you find the right wall, please rip off the task page and drop it in the envelope attached to that wall to demonstrate that you have finished the task. Then you can start the next task.

Once you have finished a set of wayfinding tasks, please wait the experimenter to give you the further instructions.

Please be aware that we will be recording the whole experiment procedure.

At last:

Once you finish all the wayfinding tasks, please complete a computer-based questionnaire and a Post Experiment Questionnaire to provide us with feedback on your experiences. Then you have done! ;-)

If you have any questions, please ask before the formal experiment.

A.3 Experiment 2: Post Experiment Questionnaire (CROSSFLOW V1 multi-user lab study)

===

Thank you for participating in this experiment. Please answer all the questions below.

First Name:

Last Name:

Age:

Gender: M / F (please circle)

Occupation:

Participant code: A / B / C / D / E / G / H / I / J / K (please circle)

1. Have you ever looked for some location or person in a large unfamiliar complex building before, for example, in an airport, a shopping mall or museum? (please tick √)

Yes / No

2. How often do you use these method or tool to find your ways when you enter an unfamiliar complex building?

(1=never, 2=seldom, 3=neutral, 4=often, 5=very often) (please tick √)

Sign:	*1*	*2*	*3*	*4*	*5*
Ask someone in the building:	*1*	*2*	*3*	*4*	*5*
A building layout map:	*1*	*2*	*3*	*4*	*5*
A mobile device based electronic map:	*1*	*2*	*3*	*4*	*5*

Colour coded path: *1* *2* *3* *4* *5*

Other (please specify)________________ *1* *2* *3* *4* *5*

3. If all the wayfinding tools or methods are available, which you prefer to use when you enter a large unfamiliar complex building (e.g. airport, shopping mall, museum etc.) looking for several locations or persons? (please tick √ and answer)

Consult signs in the building
Why?

Ask someone in the building
Why?

Consult a building layout map
Why?

Consult a mobile device based electronic map
Why?

Consult colour coded path
Why?

CROSSFLOW system

Why?

Other (please specify) : ________________

Why?

4. Please rate the following features of CROSSFLOW system:

Compared to other methods or tools, the CROSSFLOW system for wayfinding in an unfamiliar complex indoor environment was

(1=not, 2=less, 3=neutral, 4=more, 5=much more) (please tick √)

Efficient	*1*	*2*	*3*	*4*	*5*
Noticeable	*1*	*2*	*3*	*4*	*5*
Understandable	*1*	*2*	*3*	*4*	*5*
Attractive	*1*	*2*	*3*	*4*	*5*
Enjoyable to use	*1*	*2*	*3*	*4*	*5*
Attention-consuming	*1*	*2*	*3*	*4*	*5*
Spatial memory required	*1*	*2*	*3*	*4*	*5*

5. Compared to the other methods, using CROSSFLOW to determine the best routes during navigation was (please tick √)

1= much easier, 2= easier, 3=neutral, 4=more difficult, 5=much more difficult

Why?

6. Compared to map, has CROSSFLOW helped you avoid obstacles or other people in the experimental environment while you perform the wayfinding tasks? How? Why?

7. What do you **like** most about the CROSSFLOW navigation system in terms of design or features or anything? Why?

8. What do you **dislike** most about the CROSSFLOW navigation system in terms of design or features or anything? Why?

9. Were you able to identify the projected animation, especially the direction flowing to your direction? If not, why?

10. Were you able to adequately understand mobile phone interface of CROSSFLOW and how to interact with the interface? If not, why?

11. Were you able to identify the correlation between projected animation on the floor and vibration from your mobile phone? If not, why?

12. Please give us your suggestions about improvement of CROSSFLOW indoor navigation system:

13. When you are doing your wayfinding task, and listening to the public announcement, do you think in which condition you can pay more attention to the public announcement? (Map condition or CROSSFLOW condition?) Why?

14. What do you do while you are finding your way in public space such as airport, shopping mall, and museum? (please tick √)

On the phone

Talk with your partners

Listen to music from your MP3 player

Others (please specify): ________________, ________________, ________________, ________________, ________________, ________________, ________________, ________________...

For experimenter's use only

Group: 1 2 3 4

Participant code:

A / B / C / D / E /

F / G / H / I / J

Condition:

Notes:

Receipt for your payment of £20.00

Name:
Address:
Tel:
Email:
Date:
Signature:

Appendix B: Experiment 3 and 4 Files

B.1 Experiment 4: Experiment introduction for the participants (CROSSFLOW V2 pilot user study)

Thank you for coming today to help us test the design of the CROSSFLOW V2 system. The test will take about one hour. We are inviting people like you to try out the system so that we can improve its design. Please remember that it is the design we are testing, not you. To find out how easy or difficult CROSSFLOW V2 is to use; we are going to give you some tasks to try to complete. Firstly, you will be trained to use a CROSSFLOW system (a mobile phone combined with projected light pattern) to find a target in the training area. The detailed instruction about the formal experiment will be provided during the training. The training performances will not be measured. Then you will be asked to do a series of tasks in the formal experiment. In the formal experimental environment, you need to find a set of targets in different areas using CROSSFLOW V2. Your performances will be measured.

The experimenter will indicate when and where to start your tasks. Before you start the tasks, please wait the experimenter for the instructions and the tool (i.e. smartphone). After you finish the tasks, please wait the experimenter for further instructions.

Note:

- It is important to use your wayfinding tool correctly and efficiently. Once you start your tasks, you should find the targets as quickly and accurate as you can. However, you should not run, and compete or communicate with other people during the formal experiment.
- When you find a target, please put the **whole corresponding page** of your task book inside the envelope if you find a right target; just put the **right bottom corner** of the corresponding page inside the envelope if you find a wrong target. Then start your next task. Please try your best to find the right targets. If you **CANNOT** find the right targets using CROSSFLOW V2 anyway, please inform the experimenter.
- We will be videotaping the experiment procedure so that we can review and analyse it later. We will not use the tape for any other purpose.

At last, please complete a paper-based Post Experiment Questionnaire. We will not share or distribute the information you provide, and we won't associate your name with your answers.

Do you have any questions before we begin?

B.2 Experiment 3 and 4: Post Experiment Questionnaire (CROSSFLOW V2 single-user lab study)

==

Thank you for participating in this experiment. Please fill in the information below and answer the following questions. We will not share or distribute the information you provide, and we won't associate your name with your answers.

First Name:	
Last Name:	
Age:	
Occupation:	
Gender: (please circle)	M / F
Participant Code: (please circle)	A / B / C / D / E // G / H / I / J / K
Group Number: (please circle)	1 / 2 / 3 / 4

==

Questions

* Have you ever looked for some location or person in a large unfamiliar complex building before, for example, in an airport, a shopping mall or museum?

* What is your first impression of this system (CROSSFLOW2.0)? What do you think of the layout, colors and other visual parts of the display? What do you think of the sound on your mobile device?

* What things did you **like** best about the CROSSFLOW2.0 navigation system? Why?

* What things about the CROSSFLOW2.0 would you most like to change, and how to change them?

* Please rate ease of completing the tasks using CROSSFLOW2.0:

Ease *Difficult*

(please circle a number)

1 *2* *3* *4* *5*

Please write down your explanation for your rate:

* Can you associate the sound on your mobile device with the projected visual displays in the experimental environment **semantically** (e.g. can you associate the sound with wind sound, associate the visual displays with grass flickering in the wind)? Please answer this question and write down your detailed explanation.

* Please rate ease of identifying this **semantic correlation** (意义关联) between sound and the ambient pattern:

Ease				*Difficult*
(please circle a number)				
1	*2*	*3*	*4*	*5*

* When the volume of the sound on your mobile device was increased from zero to maximum, did the increased sound alert you to start to pay attention to the ambient pattern projected in the experimental environment? Did the increased sound long enough to alert you? Please answer these questions and write down your detailed explanation.

* When the volume of the sound on your mobile device was decreased from maximum to zero, could you identify the direction the ambient pattern pointing in the experimental environment? Did the decrease sound long enough to make you identify it within one cycle of the sound? Please answer these questions and write down your detailed explanation.

* Please rate ease of identifying the **time correlation** (时间关联) between sound and the ambient pattern:

Ease *Difficult*

(please circle a number)

1 2 3 4 5

* When the ambient pattern projected in the experimental environment pointed to a direction or a target, was it pointed clearly?

Please rate ease of identifying a direction or a target the ambient pattern pointing:

Ease *Difficult*

(please circle a number)

1 2 3 4 5

Please write down your explanation for your answer and your rate:

* Compared to using handheld Map, using CROSSFLOW2.0 to determine the best routes during navigation is (please *circle* a number)
1= much easier, 2= easier, 3=neutral, 4=more difficult, 5=much more difficult

Please write down your explanation for your rate:

* Compared to Map, can CROSSFLOW help you avoid obstacles or other people in the environment while you perform wayfinding tasks? How? Why?

* What did you think of the CROSSFLOW2.0 navigation system overall?

* Did you enjoy the experience of participating in this experiment? What can we do to improve this experience for future users?

Thank you for helping us to make the technology great!

Appendix C: Experiment Raw Data and Multimedia Files

Online address:

https://docs.google.com/open?id=0B9lL5u8Nnpyrd1NVWkR2YTRCMGM

References

Aarts, E. & Roovers, R., 2003. IC design challenges for ambient intelligence. In *Design Automation and Test in Europe Conference and Exhibition*. IEEE Computer Society, pp. 2–7. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1253578.

Advertiser, G., 2010. Information overload is endangering urban road users. *Advertiser.ie*. Available at: http://www.advertiser.ie/galway/article/32156 [Accessed April 30, 2012].

Allan, L.G. & Kristofferson, A.B., 1974. Successiveness discrimination: Two models. *Attention, Perception, & Psychophysics*, 15(1), pp.37–46.

Allen, G.L., 1999. Spatial abilities, cognitive maps, and wayfinding: bases for individual differences in spatial cognition and behavior. In R. G. Golledge, ed. *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. Johns Hopkins University Press Baltimore, pp. 46–80. Available at: http://books.google.co.uk/books?hl=en&lr=&id=TjzxpAWiamUC&oi=fnd&pg=PA46&dq=Spatial+abilities,+cognitive+maps,+and+wayfinding&ots=T7_Xz2W6L2&sig=cuPguBM5vyfbrFPxTptmfkcnbLU&redir_esc=y#v=onepage&q=Spatial abilities%2C cognitive maps%2C and wayfinding&f=false.

Altinsoy, E., 2003. Perceptual aspects of auditory-tactile asynchrony. In *Proceedings of the Tenth International Congress on Sound and Vibration*. pp. 3831–3838. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.170.4044&rep=rep1&type=pdf.

Altinsoy, E., Blauert, J. & Treier, C., 2002. On the perception of the synchrony for auditory-tactile stimuli [orig. Zur Wahrnehmung von Synchronitat bei auditivtaktil dargebotenen Stimuli]. In *Fortschritte der Akustik - DAGA'02*. Oldenburg, Germany: Deutsche Gesellschaft für Akustik. [in German]., pp. 728–729. Available at: http://www.ias.et.tu-dresden.de/ias/fileadmin/user_upload/akustik/Forschung/altinsoy_alt/41_1_.pdf.

Arthur, P. & Passini, R., 1992. *Wayfinding: People, signs, and architecture*, McGraw-Hill, Incorporated. Available at: http://trid.trb.org/view.aspx?id=367500.

Bannon, L.J., 1993. CSCW: an initial exploration. *Scandinavian Journal of Information Systems*, 5, pp.3–24. Available at: http://iris.cs.aau.dk/tl_files/volumes/volume05/no1/01_bannon_p3-24.pdf.

Baskaya, A., Wilson, C. & Özcan, Y.Z., 2004. Wayfinding in an Unfamiliar Environment: Different Spatial Settings of Two Polyclinics. *Environment and Behavior*, 36(6), pp.839–867. Available at: http://eab.sagepub.com/cgi/doi/10.1177/0013916504265445.

Baus, J., Cheverst, K. & Kray, C., 2005. A Survey of Map-based Mobile Guides. In L. Meng, T. Reichenbacher, & A. Zipf, eds. *Map-based mobile services-Theories, Methods, and Implementations*. Springer-Verlag, pp. 193–209. Available at: http://www.springerlink.com/content/mu516h0611126438/.

Berger, S. et al., 2005. Using Symbiotic Displays to View Sensitive Information in Public. In *Third IEEE International Conference on Pervasive Computing and Communications*. IEEE, pp. 139–148. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1392750.

Bertelson, P. & Aschersleben, G., 2003. Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. *International Journal of Psychophysiology*, 50(1-2), pp.147–155. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0167876003001302.

Bertelson, P. & Gelder, B. de, 2004. The psychology of multimodal perception. In C Spence & J Driver, eds. *Crossmodal Space and Crossmodal Attention*. Oxford University Press, pp. 141–178. Available at: http://books.google.co.uk/books?hl=en&lr=&id=6UceYiqZGg0C&oi=fnd&pg=PA141&dq=The+psychology+of+multimodal+perception&ots=KeKZcp5AhW&sig=OInqgoP_0vj1f5oQI7sBuBng73w&redir_esc=y#v=onepage&q=The psychology of multimodal perception&f=false.

Bertenthal, B.I., Banton, T. & Bradbury, A., 1993. Directional bias in the perception of translating patterns. *Perception*, 22(2), pp.193–207. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8474844.

Brignull, H., 2005. *Understanding and designing for the voluntary adoption of community displays*. Doctoral dissertation. University of Sussex. Available at:

http://zh.scribd.com/doc/4811340/Understanding-and-Designing-for-the-Voluntary-Adoption-of-Community-Displays.

Brown, L.M., 2007. *Tactons: structured vibrotactile messages for non-visual information display*. Doctoral dissertation. University of Glasgow. Available at: http://www.dcs.glasgow.ac.uk/~stephen/papers/theses/lorna_brown_thesis.pdf.

Burke, J.L. et al., 2006. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In *Proceedings of the 8th international conference on Multimodal interfaces ICMI '06*. Banff, Alberta, Canada: ACM Press, p. 108. Available at: http://portal.acm.org/citation.cfm?doid=1180995.1181017.

Cadiz, J.J. et al., 2002. Designing and deploying an information awareness interface. In *Proceedings of the 2002 ACM conference on computer supported cooperative work CSCW '02*. ACM Press, pp. 314–323. Available at: http://portal.acm.org/citation.cfm?doid=587078.587122.

Cao, H., Olivier, P. & Jackson, D., 2008. Enhancing privacy in public spaces through crossmodal displays. *Social Science Computer Review*, 26(1), pp.87–102. Available at: http://dx.doi.org/10.1177/0894439307307696.

CEA, 2008. Photosensitive Epilepsy [Online]. *Canadian Epilepsy Alliance*. Available at: http://www.epilepsymatters.com/english/faqphotosensitive.html#whylights [Accessed November 16, 2012].

Chang, A. et al., 2001. LumiTouch: an emotional communication device. In *Conference on Human Factors in Computing Systems: CHI'01 extended abstracts on Human factors in computing systems*. pp. 313–314. Available at: http://people.cs.vt.edu/~wangr06/touch review origanization/ChaRKWetal01.pdf.

Ciavarella, C. & Patern, F., 2004. The design of a handheld, location-aware guide for indoor environments. *Personal and Ubiquitous Computing*, 8(2), pp.82–91. Available at: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s00779-004-0265-z.

Dictionary.com, 2012. cue. *Dictionary.com, LLC.* Available at: http://dictionary.reference.com/browse/cue [Accessed November 14, 2012].

DiFranco, D.E., Beauregard, G.L. & Srinivasan, M.A., 1997. The effect of auditory cues on the haptic perception of stiffness in virtual environments. In *Proceedings of the ASME Dynamic Systems and Control Division*. pp. 17–22.

Dix, A. et al., 2004. *Human-computer interaction* 3rd ed. A. Dix, ed., Prentice-Hall. Available at: http://cognac.ai.ru.nl/studie/MMI_summary.pdf.

Dixon, N.F. & Spitz, L., 1980. The detection of auditory visual desynchrony. *Perception*, 9(6), pp.719–721.

Dogu, U. & Erkip, F., 2000. Spatial Factors Affecting Wayfinding and Orientation: A Case Study in a Shopping Mall. *Environment and Behavior*, 32(6), pp.731–755. Available at: http://eab.sagepub.com/cgi/doi/10.1177/00139160021972775.

Driver, John & Spence, Charles, 1998. Crossmodal attention. *Current Opinion in Neurobiology*, 8(2), pp.245–253. Available at: http://eprints.ucl.ac.uk/4553/.

Driver, John & Spence, Charles, 2004. Crossmodal spatial attention: Evidence from human performance. In Charles Spence & John Driver, eds. *Crossmodal space and crossmodal attention*. Oxford University Press, USA, pp. 179–220.

Dyer, C., 1995. *Beginning research in psychology: A practical guide to research methods and statistics*, Blackwell Malden, MA. Available at: http://www.getcited.org/pub/103198805.

Eriksson, E., Hansen, T.R. & Lykke-Olesen, A., 2007. Reclaiming public space: designing for public interaction with private devices. In *Proceedings of the 1st international conference on Tangible and embedded interaction*. New York, NY, USA: ACM, pp. 31–38. Available at: http://doi.acm.org/10.1145/1226969.1226976.

Farmer, E. & Adam Brownson, Q., 2003. Review of Workload Measurement, Analysis and Interpretation Methods. *Methods*, 47(25), p.33. Available at: http://www.eurocontrol.int/integra/gallery/content/public/documents/expt_def_wp2_final.pdf.

Faubert, J., 2002. Visual perception and aging. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expÃ© rimentale*, 56(3), p.164. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12271747.

Fisher, D.L. & Kichhanagari, R., 2006. Reducing the Time to Scan Information-Dense Airport Terminal Signs. In *Transportation Research Board 85th Annual Meeting*.

Washington DC, United States, pp. 06–2452. Available at: http://www.ltrc.lsu.edu/TRB_82/TRB2003-002454.pdf.

Ganz, L., 1975. studies averaged response to sound leading and lagging visual stimuli absolute values. *Journal of Experimental Psychology: Animal Behavior Processes*, 33. Available at: http://psycnet.apa.org/journals/xan/33/1/.

Geldard, F.A., 1960. Some neglected possibilities of communication. *Science*, 131(3413), pp.1583–1588. Available at: http://adsabs.harvard.edu/abs/1960Sci...131.1583G.

Green, K.P. & Kuhl, P.K., 1989. The role of visual information in the processing of place and manner features in speech perception. *Attention, Perception And Psychophysics*, 45(1), pp.34–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2913568.

Greenberg, S. & Rounding, M., 2001. The Notification Collage: Posting Information to Public and Personal Displays. In J. Jacko, ed. *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM Press, pp. 514–521. Available at: http://portal.acm.org/citation.cfm?doid=365024.365339.

Gunther, E. & O'Modhrain, S., 2003. Cutaneous Grooves: Composing for the Sense of Touch. *Journal of New Music Research*, 32(4), pp.369–381. Available at: http://www.informaworld.com/openurl?genre=article&doi=10.1076/jnmr.32.4.369.18856&magic=crossref.

Hairston, W.D. et al., 2003. Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience*, 15(1), pp.20–29. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12590840.

Hall, R., 1998. Within-Subjects Designs. *Psychology World*. Available at: http://web.mst.edu/~psyworld/within_subjects.htm [Accessed November 16, 2012].

Hart, S.G. & Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, 1, pp.139–183. Available at: http://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf.

Heiner, J.M., Hudson, S.E. & Tanaka, K., 1999. The information percolator: ambient information display in a decorative object. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*. ACM Press, pp. 141–148. Available at: http://portal.acm.org/citation.cfm?id=320719.322595.

Henderson, J.M., 1997. Transsaccadic memory and integration during real-world object perception. *Psychological Science*, 8(1), pp.51–55. Available at: http://pss.sagepub.com/content/8/1/51.

Heylighen, F., 2004a. Complexity and Information Overload in Society: why increasing efficiency leads to decreasing control. *Technological Forecasting and Social Change*, pp.1–19. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.6233&rep=rep1&type=pdf.

Heylighen, F., 2004b. Tackling Complexity and Information Overload: from intelligence amplification and attention economy to the global brain. *Technological Forecasting and Social Change*, pp.20–44. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.6233&rep=rep1&type=pdf.

Hollands, J.G. & Wickens, C.D., 1999. *Engineering Psychology and Human Performance* 3rd ed. N. Roberts, ed., New Jersey: Prentice Hall. Available at: http://webfiles.ita.chalmers.se/~mys/HumanAspects/WickensHollands/0_Wickens_Index_Preface.pdf.

Holmquist, L.E. & Skog, T., 2003. Informative art: information visualization in everyday environments. In *Proceedings of the 1st international conference on computer graphics and interactive techniques in Australasia and South East Asia*. ACM, pp. 229–235. Available at: http://portal.acm.org/citation.cfm?id=604471.604516&coll=GUIDE&dl=ACM&CFID=81993695&CFTOKEN=77468377.

Huang, E.M. & Mynatt, E.D., 2003. Semi-public displays for small, co-located groups. In G. Cockton & P. Korhonen, eds. *Proceedings of the ACM conference on Human factors in computing systems CHI 03*. ACM Press, pp. 49–56. Available at: http://portal.acm.org/citation.cfm?doid=642611.642622.

Hulusic, V., Aranha, M. & Chalmers, A., 2008. The influence of cross-modal interaction on perceived rendering quality thresholds. In V. Skala, ed. *Proceedings of the 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*. Vaclav Skala - Union Agency, pp. 41–48. Available at: http://wrap.warwick.ac.uk/28245/.

Intille, S.S., 2002. Change Blind Information Display for Ubiquitous Computing Environments G. Borriello & L E Holmquist, eds. *UBICOMP 2002: UBIQUITOUS COMPUTING Lecture Notes in Computer Science*, LNCS 2498, pp.193–222. Available at: http://www.springerlink.com/index/12VJNM652JTARW8T.pdf.

Ishii, H. et al., 1998. ambientROOM: integrating ambient media with architectural space. In *CHI 98 conference summary on Human factors in computing systems*. ACM, pp. 173–174. Available at: http://portal.acm.org/citation.cfm?id=286652.

James, V. et al., 1999. Change-blindness as a result of “mudsplashes”. *Nature*, 398(6722), p.34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10078528.

Jansen, M., 2009. *Unobtrusive interfaces: preventing information overload in Ambient Interfaces*. University of Twente. Available at: http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Jansen-Michel.pdf.

Jaskowski, P., 1996. Simple reaction time and perception of temporal order: dissociations and hypotheses. *Perceptual and Motor Skills*, 82(3 Pt 1), pp.707–730. Available at: http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0030161189&partnerID=40&rel=R6.5.0.

Jones, Q., Ravid, G. & Rafaeli, S., 2004. Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration. In *IEEE Engineering Management Review*. INFORMS: Institute for Operations Research, pp. 194–210. Available at: http://isr.journal.informs.org/cgi/doi/10.1287/isre.1040.0023.

Katsumi, W. & Shimojo, S., 2001. Postcoincidence trajectory duration affects motion event perception. *Attention, Perception & Psychophysics*, 63(1), pp.16–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11304012.

Kim, Y. & Rieh, S.Y., 2005. Dual-Task Performance as a Measure of Mental Effort in Searching a Library System and the Web. *Proceedings of the American Society for Information Science and Technology*, 42(1). Available at: http://www3.interscience.wiley.com/journal/112785720/abstract.

Klippel, A. et al., 2004. Direction concepts in wayfinding assistance systems. In *Workshop on Artificial Intelligence in Mobile Systems*. pp. 1–8. Available at: http://w5.cs.uni-sb.de/~baus/aims04/cameraready/aims04-proceedings.pdf#page=6.

Koch, C. & Ullman, S., 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), pp.219–227. Available at: http://papers.klab.caltech.edu/104/1/200.pdf.

Kray, C., Kortuem, G. & Krüger, A., 2005. Adaptive Navigation Support with Public Displays. In *Proceedings of the 10th international conference on Intelligent user interfaces IUI 05*. San Diego, California, USA: ACM, pp. 326–328. Available at: http://doi.acm.org/10.1145/1040830.1040916.

Laakso, M., 2004. Ambient Displays and Changing Information. In *Seminar on User Interfaces and Usability*. Helsinki University of Technology, Finland.

Leeper, J.D., 2000. What statistical analysis should I use? *UCLA: Statistical Consulting Group*. Available at: http://www.ats.ucla.edu/stat/spss/whatstat/default.htm [Accessed November 16, 2012].

Levitin, D.J. et al., 2000. The perception of cross-modal simultaneity (or "the Greenwich Observatory Problem" revisited). *AIP (American Institute of Physics) Conference Proceedings*, 517, pp.323–329. Available at: http://link.aip.org/link/?APC/517/323/1&Agg=doi.

Liu, H. et al., 2007. Survey of Wireless Indoor Positioning Techniques and Systems H Darabi, ed. *IEEE Transactions on Systems Man and Cybernetics, Part C: Applications and Reviews*, 37(6), pp.1067–1080. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4343996.

Ljungblad, S., 2003. *EXPLORING THE USE OF AMBIENT INFORMATION DISPLAYS: Designing and Evaluating Informative Art*. Master student essay. Göteborg University. Available at: https://gupea.ub.gu.se/handle/2077/1195.

MacIntyre, B. et al., 2001. Support for multitasking and background awareness using interactive peripheral displays. In J. Marks, ed. *Proceedings of the 14th annual ACM symposium on User Interface Software and Technology UIST '01*. ACM Press, pp. 41–50. Available at: http://portal.acm.org/citation.cfm?doid=502348.502355.

Maglio, P.P. & Campbell, C.S., 2000. Tradeoffs in displaying peripheral information. *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 00*, 2(1), pp.241–248. Available at: http://portal.acm.org/citation.cfm?doid=332040.332438.

Mankoff, J. et al., 2003. Heuristic evaluation of ambient displays. In G. Cockton & P. Korhonen, eds. *Proceedings of the conference on Human factors in computing systems CHI 03*. Florida, USA: ACM Press, pp. 169–176. Available at: http://portal.acm.org/citation.cfm?doid=642611.642642.

Mastoropoulou, G. et al., 2005. The influence of sound effects on the perceived smoothness of rendered animations. *Proceedings of the 2nd symposium on Appied perception in graphics and visualization APGV 05*, p.9. Available at: http://dx.doi.org/10.1145/1080402.1080404.

Matthews, T. et al., 2003. *A Peripheral Display Toolkit*, Intel Research. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.133.1533.

Matthews, T. et al., 2004. A toolkit for managing user attention in peripheral displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. Santa Fe, NM, USA: ACM Press, pp. 247–256. Available at: http://portal.acm.org/citation.cfm?id=1029632.1029676.

Mayhew, D.J., 1999. *The Usability Engineering Lifecycle: A Practioner's Handbook for User Interface Design* S. Card et al., eds., Morgan Kaufmann. Available at: http://books.google.co.uk/books/about/The_Usability_Engineering_Lifecycle.html?id=q_9tlou__DgC.

McCarthy, J.F., Costa, T.J. & Liongosari, E.S., 2001. UniCast, OutCast & GroupCast: Three Steps Toward Ubiquitous, Peripheral Displays. In *Proceedings of the 3rd international conference on Ubiquitous Computing*. Atlanta, Georgia, USA: Springer Berlin/Heidelberg, pp. 332–345. Available at: http://interrelativity.com/joe/publications/UbiquitousPeripheralDisplays-UbiComp2001.pdf.

McCrickard, D.S. & Zhao, Q.A., 2000. Supporting Information Awareness Using Animated Widgets. In *Proceedings of the 2000 USENIX Converence on TclTk*. USENIX Association, pp. 117–127.

McDonald, J.J., Teder-Sälejärvi, W.A. & Hillyard, S.A., 2000. Involuntary orienting to sound improves visual perception. *Nature*, 407(6806), pp.906–908. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11057669.

McGrath, M. & Summerfield, Q., 1985. Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *Journal of the Acoustical Society of*

America, 77(2), pp.678–685. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3973239.

McGurk, H. & MacDonald, J., 1976. Hearing lips and seeing voices. *Nature*, 264(5588), pp.746–748. Available at: http://www.nature.com/nature/journal/v264/n5588/abs/264746a0.html.

McQuire, S., 2006. The politics of public space in the media city. *First Monday Special Issue #4: Urban Screens: Discovering the Potential of Outdoor Screens for Urban Society*. Available at: http://en.scientificcommons.org/59069597 [Accessed November 8, 2012].

Meissner, H. et al., 2011. Best Practices for Mixed Methods Research in the Health Sciences. In E. J. H. Sharon Andrew, ed. *Best Practices for Mixed Methods Research*. National Institutes of Health, pp. 1–39. Available at: http://obssr.od.nih.gov/scientific_areas/methodology/mixed_methods_research/pdf/Best_Practices_for_Mixed_Methods_Research.pdf.

Meredith, M.A., Nemitz, J.W. & Stein., Barry E., 1987. Determinants of Multisensory Integration Neurons . I . Temporal Factors in Superior Colliculus. *The Journal of Neuroscience*, 7(10), pp.3215–3229. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3668625.

Miller, T. & Stasko, J., 2001. The InfoCanvas: information conveyance through personalized, expressive art. In *CHI'01 extended abstracts on Human Factors in Computing Systems*. ACM, pp. 305–306. Available at: http://portal.acm.org/citation.cfm?id=634067.634248.

Mitchell, W.J., 2000. *E-Topia: "Urban Life, Jim--But Not As We Know It"*, Cambridge: MIT press. Available at: http://books.google.co.uk/books/about/e_topia.html?id=E2On0MJBItYC.

Morris, M.R., Morris, D. & Winograd, T., 2004. Individual Audio Channels with Single Display Groupware : Effects on Communication and Task Strategy. In J. D. Herbsleb & G. M. Olson, eds. *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. Chicago, Illinois, USA: ACM, pp. 242–251. Available at: http://portal.acm.org/citation.cfm?id=1031607.1031646&coll=DL&dl=ACM&CFID=18420995&CFTOKEN=13435579.

Mynatt, E.D. et al., 1998. Designing Audio Aura. In *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 98*. ACM Press, pp. 566–573. Available at: http://portal.acm.org/citation.cfm?doid=274644.274720.

Mynatt, E.D. et al., 2001. Digital Family Portraits : Supporting Peace of Mind for Extended Family Members. In J Jacko & A. Sears, eds. *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM Press, pp. 333–340. Available at: http://portal.acm.org/citation.cfm?id=365024.365126.

Müller, H. et al., 2012. AmbiX: Designing Ambient Light Information Displays. In *Designing Interactive Lighting workshop at the ACM conference on Designing Interactive Systems 2012*. Newcastle upon Tyne, UK. Available at: http://interactivelight.id.tue.nl/papers/Muller_AmbiX_Designing_Ambient_Light_Information_Displays.pdf.

NASA, 1988. NASA-TLX homepage. *NASA*. Available at: http://humansystems.arc.nasa.gov/groups/TLX/ [Accessed November 14, 2012].

Nowell, L., Schulman, R. & Hix, D., 2002. Graphical encoding for information visualization: an empirical study. In *IEEE Symposium on Information Visualization 2002 INFOVIS 2002*. IEEE Comput. Soc, pp. 43–50. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1173146.

Ollson, R., 2007. *Public Space for Ambient Intelligence*. Unpublished master student essay. University of Waterloo. Available at: http://www.architecture.uwaterloo.ca/faculty_projects/terri/384-comp.html#2007.

O'Neill, M.J., 1991. Effects of Signage and Floor Plan Configuration on Wayfinding Accuracy. *Environment and Behavior*, 23(5), pp.553–574. Available at: http://eab.sagepub.com/cgi/content/abstract/23/5/553.

Park, H.M., 2008. *Univariate analysis and normality test using SAS, STATA, and SPSS*. Working Paper. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University. Available at: http://indiana.edu/~statmath/stat/all/normality/normality.pdf.

Paulos, E. & Goodman, E., 2004. The familiar stranger: anxiety, comfort, and play in public places. In E. D. Erickson & M. Tscheligi, eds. *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, pp. 223–230. Available at: http://portal.acm.org/citation.cfm?id=985721.

Pedersen, E.R. & Sokoler, T., 1997. AROMA: abstract representation of presence supporting mutual awareness. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM Press, pp. 51–58. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22028053.

Prante, T. et al., 2004. Ambient agoras: InfoRiver, SIAM, Hello.Wall. In E. Dykstra-Erickson & M. Tscheligi, eds. *Proceedings of the Extended Abstracts on Human Factors in Computing Systems CHI04*. ACM, pp. 763–764. Available at: http://portal.acm.org/citation.cfm?id=985924.

Prante, T. et al., 2003. Hello.Wall – Beyond Ambient Displays. In Peter Ljungstrand & Jason Brotherton, eds. *Video Track and Adjunct Proceedings of the 5th International Conference on Ubiquitous Computing Ubicomp 03*. Kluwer, pp. 277–278. Available at: http://www.ipsi.fraunhofer.de/ambiente/paper/2003/prante-hello.wall_ubicomp03-withCopyright-letter.pdf.

Proctor, R.W. et al., 2005. Implications of compatibility and cuing effects for multimodal interfaces. In *Foundations of augmented cognition (Proceedings of the 11th Annual Conference on Human-Computer Interaction International 2005)*. pp. 3–12. Available at: https://engineering.purdue.edu/~hongtan/pubs/PDFfiles/C61_Proctor_etal_HCII2005.pdf.

Proctor, R.W. & Vu, K.-P.L., 2005. *Handbook of Human Factors in Web Design* K.-P. L Vu & W. R. Proctor, eds., Lawrence Erlbaum Associates. Available at: http://books.google.com/books?hl=en&lr=&id=d0N6mumoLbAC&oi=fnd&pg=PR11&dq=HANDBOOK+OF+HUMAN+FACTORS+IN+WEB+DESIGN&ots=AGoty49lL8&sig=mbtOIcMNduk8b2E9qrUqXuEEklw.

Puckette, M., 1996. Pure Data: another integrated computer music environment. In *Proceedings of the Second Intercollege Computer Music Concerts*. pp. 37–41. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.3903.

Raghunath, M., Narayanaswami, C. & Pinhanez, C., 2003. Fostering a symbiotic handheld environment. *Computer*, 36(9), pp.56–65. Available at: http://dl.acm.org/citation.cfm?id=942616.

Raubal, M., 2001. Human wayfinding in unfamiliar buildings: a simulation with a cognizing agent. *Cognitive Processing*, 2(3), pp.363–388. Available at: ftp://ftp.geoinfo.tuwien.ac.at/raubal/3442_CognProc01.pdf.

Rekimoto, J., 1997. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*. ACM, pp. 31–39. Available at: http://portal.acm.org/citation.cfm?id=263505.

Reynolds, L., Birnholtz, J., Luxenberg, E., Gutwin, C., & Mustafa, M., 2010. Comparing awareness and distraction between desktop and peripheral-vision displays. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, pp. 3571–3576. Available at: http://dl.acm.org/citation.cfm?id=1754020.

Robson, C., 1994. *Experiment, Design and Statistics in Psychology* 3rd ed., Penguin Group (USA).

Sarter, N., 2006. Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, 36(5), pp.439–445. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0169814106000217 [Accessed March 1, 2011].

Sekuler, R., Sekuler, A.B. & Lau, R., 1997. Sound alters visual motion perception. *Nature*, 385(6614), p.308. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9002513.

Shams, L., Kamitani, Y. & Shimojo, S., 2002. Visual illusion induced by sound. *Cognitive Brain Research*, 14(1), pp.147–152. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12063138.

Sharp, H., Rogers, Y. & Preece, J., 2007. *Interaction design: beyond human-computer interaction* 2nd ed., John Wiley. Available at: http://www.id-book.com/secondedition/.

Shen, C. et al., 2002. Sharing and building digital group histories. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work CSCW 02*. New Orleans, Louisiana, USA: ACM Press, pp. 324–333. Available at: http://portal.acm.org/citation.cfm?doid=587078.587124.

Sheth, B.R. & Shimojo, S., 2004. Sound-aided recovery from and persistence against visual filling-in. *Vision Research*, 44(16), pp.1907–1917. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15145684.

Shneiderman, B. & Plaisant, C., 2005. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* 4th ed., Pearson Addison Wesley, USA. Available at: http://www.lavoisier.fr/notice/frDWOA2L2AARW32O.html.

Shoemaker, G.B.D. & Inkpen, K.M., 2001. Single display privacyware: augmenting public displays with private information. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, pp. 522–529. Available at: http://portal.acm.org/citation.cfm?doid=365024.365349.

Shuttleworth, M., 2009. Within Subject Design. *Explorable.com*. Available at: http://explorable.com/within-subject-design.html [Accessed November 16, 2012].

Spence, Charles & Driver, John, 2004. *Crossmodal space and crossmodal attention* Charles Spence & John Driver, eds., Oxford University Press, USA. Available at: http://discovery.ucl.ac.uk/100265/.

Srinivasan, M.A., Beauregard, G.L. & Brock, D.L., 1996. The impact of visual information on the haptic perception of stiffness in virtual environments. In *ASME Winter Annual Meeting*. pp. 555–559. Available at: http://touchlab.mit.edu/publications/1996_006.pdf.

Stein, Barry E et al., 1996. Enhancement of Perceived Visual Intensity by Auditory Stimuli: A Psychophysical Analysis. *Journal of Cognitive Neuroscience*, 8(6), pp.497–506. Available at: http://www.mitpressjournals.org/doi/abs/10.1162/jocn.1996.8.6.497.

Steinschneider, T., 1990. Search task evaluation and ways to decrease time of visual search. In D. Brogan, ed. *Visual Search: proceedings of the First International Conference on Visual Search*. Taylor & Francis, London, pp. 391–398. Available at: http://books.google.co.in/books/about/Visual_search.html?id=VVRRAAAAMAAJ.

Stevens, J.C. & Marks, L.E., 1965. Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences*, 54(2), pp.407–411. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC219679/.

Stewart, J., Bederson, B.B. & Druin, A., 1999. Single display groupware: a model for co-present collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit CHI '99*. ACM, pp. 286–293. Available at: http://dl.acm.org/citation.cfm?id=302979.303064.

Storms, R.L., 1998. *Auditory-visual cross-modal perception phenomena*. Doctoral dissertation. Naval Postgraduate School, Monterey, California, USA. Available at: http://gamepipe.usc.edu/~zyda/Theses/Russell.Storms.PhD.pdf.

Thibos, L.N. & Bradley, A., 1991. The limits of performance in central and peripheral vision. *Digest of Technical Papers Society for Information Display*, 22, pp.301–303. Available at: http://research.opt.indiana.edu/Library/performLimits/performLimits.pdf.

Tidwell, J., 1999. High-density Information Display [Online]. Available at: http://www.mit.edu/~jtidwell/language/high-density_information_display.html [Accessed November 18, 2012].

Vogel, D. & Balakrishnan, R., 2004. Interactive Public Ambient Displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In S. Feiner & J. A. Landay, eds. *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, pp. 137–146. Available at: http://portal.acm.org/citation.cfm?id=1029632.1029656.

Vroomen, J. & Gelder, B. de, 2000. Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), pp.1583–1590. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11039486.

Wan Bejuri, W.M.Y., Mohamad, M.M. & Sapri, M., 2011. Ubiquitous Positioning: A Taxonomy for Location Determination on Mobile Navigation System. *Signal & Image Processing: An International Journal*, 2(1), pp.24–34. Available at: http://arxiv.org/abs/1103.5035 [Accessed October 18, 2012].

Watanabe, K., 2001. *Crossmodal interaction in humans*. Doctoral dissertation. California Institute of Technology. Available at: http://www.cs.ubc.ca/labs/spin/publications/related/watanabe-thesis.pdf.

Weiser, M., 1991. The computer for the 21st century. *Scientific American*, 265(3), pp.94–104. Available at: http://dl.acm.org/citation.cfm?id=329126.

Weiser, M. & Brown, J.S., 1996. Designing Calm Technology P. J. Denning & R. M. Metcalfe, eds. *PowerGrid Journal*, 1(July), pp.1–5. Available at: http://nano.xerox.com/weiser/calmtech/calmtech.htm.

Werner, S. & Long, P., 2003. Cognition Meets Le Corbusier — Cognitive Principles of Architectural Design. In Christian Freksa et al., eds. *Spatial Cognition III*. Springer Berlin / Heidelberg, p. 1034. Available at: http://dx.doi.org/10.1007/3-540-45004-1_7.

Winkler, S. & Faller, C., 2005. Audiovisual quality evaluation of low-bitrate video. *Proceedings of the SPIE/IS&T Human Vision and Electronic Imaging*, 5666, pp.139–148. Available at: http://link.aip.org/link/?PSI/5666/139/1&Agg=doi.

Wisneski, C. et al., 1998. Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information N. Streitz, S. Konomi, & H.-J. Burkhardt, eds. *Computer*, 1370, pp.22–32. Available at: http://www.springerlink.com/index/744438x187j43417.pdf.

Wright, P., Hull, A.J. & Lickorish, A., 1993. Navigating in A Hospital Outpatients Department: the Merits of Maps and Wall Signs. *Journal of Architectural Planning Research*, 10(1), pp.76–89.

Wu, M. & Balakrishnan, R., 2003. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology UIST 03*. Vancouver, Canada: ACM Press, pp. 193–202. Available at: http://portal.acm.org/citation.cfm?doid=964696.964718.