Interaction Techniques using Head Mounted Displays and Handheld Devices for Outdoor Augmented Reality

by Rahul Budhiraja

A thesis submitted in partial fulfillment of the
requirements for the Degree of
Master in Human Interface Technology
University of Canterbury
2013

To Charanjit and Madhu Budhiraja, the pillars of support in my life.

Acknowledgements

The author wishes to express sincere appreciation to Professor Billinghurst for his excellent supervision, guidance and support throughout the thesis project. The author also wishes to thank Dr. Lee for his supervision , support in answering some of the technical aspects of this thesis and for his valuable insights. In addition, the author would like to acknowledge the staff and students of the Human Interface Technology Laboratory NZ for their considerable support and contribution in creating a comfortable research atmosphere.

Abstract

Depending upon their nature, Outdoor AR applications can be deployed on head mounted displays (HMD) like Google glass or handheld Displays (HHD) like smartphones. This master's thesis investigates novel gesture-based interaction techniques and applications for a HMD-HHD hybrid system that account for advantages presented by each platform. Prior research in HMD-HHD hybrid systems and gestures used in VR and surface computing were taken into account while designing the applications and interaction techniques. A prototype system combining a HMD and HHD was developed and four applications were created for the system. For evaluating the gestures, an application that compared four of the proposed gestures for selection tasks was developed. The results showed a significant difference between the different gestures and that the choice of gesture for selection tasks using a hybrid system depended upon application requirements like speed and accuracy.

Contents

List of Figures

Chapter 1

Introduction

Augmented Reality (AR) is technology that overlays virtual information on the user's view in real time such that the information appears to be part of the environment. Its ability to provide contextual information in real time opens up possibilities to a wide range of application areas such as gaming, education, navigation, military, tourism, entertainment, and archaeology [9, 10, 11, 14, 18, 23].

Outdoor Augmented Reality is a particular promising area for AR use. In this case, outdoor AR technology can allow users to discover the world around them, enhance their awareness and increase understanding of their environment, and greatly impact their everyday lives.

Outdoor AR applications were initially deployed on wearable-computing systems where users would view virtual content on a head mounted display (HMD) that was connected to a backpack containing a laptop and sensors such as GPS and gyroscopes [14].

HMDs themselves have undergone a rapid change in the past couple of years, with emergence of "Head mounted computers" like Google Glass [30] (see figure 1.1) that combine GPS, processor, memory and a display into a single-wearable unit. The Google Glass unit doesn't need a plug or connection to a smartphone or backpack, reflecting a new form-factor.

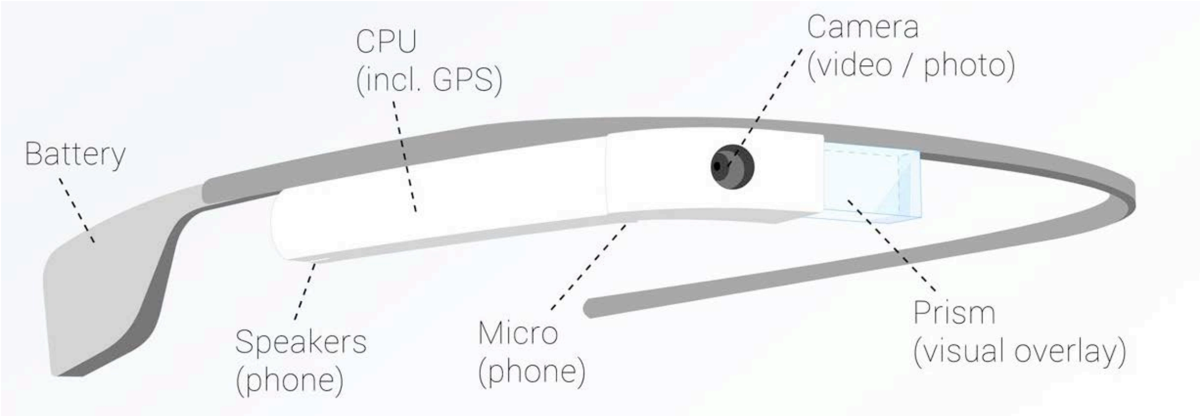


Figure 1.1 : Google Glass

The increase in computation power and portability of handheld computers has made real-time AR applications feasible on smartphones. Thus, every Outdoor AR application has a least two major choices for display – (1) Head Mounted Display (HMD) where virtual information is augmented directly on the user's view or (2) a handheld device (HHD) (e.g. smartphone) where graphics are overlaid on a live video display from the device's camera, so that it appears that you are looking through a transparent window to the world beyond. Each display type has distinctive properties and potential areas of application. However, in our research we are interested in hybrid systems for outdoor AR that combine a HMD and a handheld device.

While plenty of prior research in tracking technologies and interaction methods exists for both head-mounted and handheld AR, less research has been done on systems combining handheld devices with HMD. In earlier research, handheld devices have been used to show a map, display additional information about virtual content on the HMD and to provide input [9]. Handhelds have been used as a gestural input device to translate virtual objects in 3D that are viewed on the HMD [20]. They have also been used to provide play controls to a presentation being viewed on a HMD [10]. This prior research shows a few interaction use-cases of using a handheld device with an HMD – as a source of information, manipulating virtual content, and providing presentation controls. However, there are topics that could be researched further, including the intuitiveness of using such a system, methods to interact with the virtual content, information splitting across the displays, and multi-dimensional views using different displays, among others.

The goal of this thesis is to explore a range of interaction possibilities in a HMD-HHD hybrid system. While the focus is mainly on interaction methods, intuitiveness, information splitting, and multi-display viewing using this system are also explored. With wearable displays getting lighter and the growing popularity of smartphones (one billion as of Oct 2012) [31] the hope is that the techniques discussed in this thesis would prove the usefulness of using a wearable display and a handheld display at the same time.

Thesis Summary

This master's thesis begins with presenting background research in handheld and outdoor AR, with an emphasis for systems that combined smartphones and HMDs. Due to the lack of prior research in outdoor AR, the search domain was extended to include hybrid systems that were deployed indoors. In the next chapter we describe a basic prototype system we have developed combining a HMD and smartphone. Prototyping at an early stage was done to understand the technical difficulties in assembling such a hybrid system which helped us in choosing the right software and hardware platform .The software framework we developed was capable of sending simple messages between the phone and HMD that was useful for understanding the feasibility of real-time performance while using our system.

With the basic framework setup, a set of high-level questions were formulated that governed the research direction of this thesis. The primary goal was to develop a set of gestures that could be performed on the handheld device to interact with virtual content on the HMD. The secondary goals were measuring the intuitiveness of using this system and developing hybrid views combining both displays. With a clear direction, we completed a second phase of background research studying the literature of gestural interaction in virtual reality and surface computing. A set of gestures were chosen and additional gestures utilizing the uniqueness of the hybrid system were also added to this gesture set. Information splitting across multiple displays was also part of our background research.

At the end of the background research phase, the basic framework prototyped earlier was utilized in creating separate systems for each research goal. To evaluate the proposed gesture set, a user study was conducted to compare the gestures for selection tasks. The results of the evaluation showed significant differences in both the measured data and the users opinions of the gestures. This proved that the best choice of gesture for a Handheld-HMD hybrid system depends upon the nature of the application and the desired usage. A gesture suited for fast selection was different from one that required accurate selection. Analysis of the measured data revealed that different gestures would be suited for different purposes, however the user opinions revealed a preference for a specific gesture. These results showed that the choice of best gesture for a HMD-HHD system depends upon a variety of factors like user preferences, nature of application, and may depend on characteristics like hand size.

Chapter 2

Background Research

This thesis explores how a touch screen on a handheld device (HHD) can be used to provide intuitive input into an Augmented Reality system using a head mounted display (HMD). As such, this research extends earlier work in Augmented Reality, handheld and head mounted displays, and touch input. In this chapter we review related work in each of these areas and describe the novel directions of our research.

2.1 Augmented Reality

The origins of Augmented Reality (AR) go back to 1968, when Ivan Sutherland created the first Optical See-Through Head Mounted Display [19]. Due to the limited capabilities of processors at that time, only wireframe drawings could be augmented on the user's view and there were very limited input options. The use of AR since then was more or less confined to military applications until the 1990s where better processing power facilitated an increase in AR research in academic settings. Prototype AR systems we developed for medical [25, 26], industrial [27], education [29] and entertainment applications [28] etc.

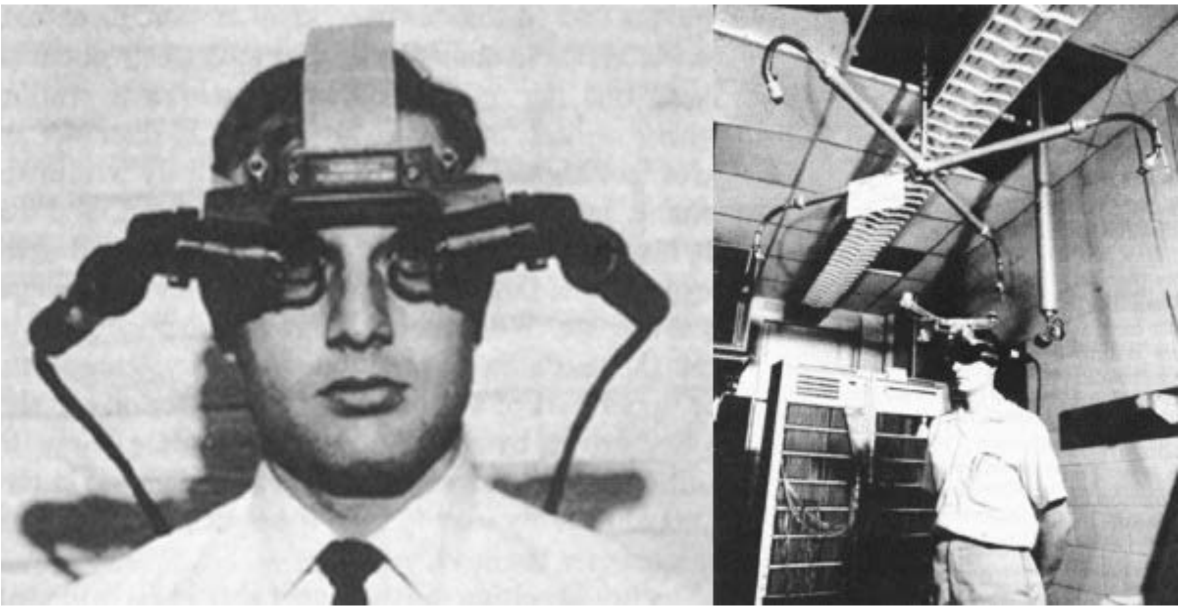


Figure 2.1 The first Optical See-Through HMD

In 1997, Ron Azuma published a comprehensive survey of the existing AR research at that time and provided insights to future research directions [2]. One of the main highlights of the survey was Azuma's definition of Augmented Reality that has been widely accepted and used even today. Azuma defines AR systems as those that have three essential qualities: (1) Systems that combine the real and virtual, (2) are interactive in real time and (3) are registered in 3D.

Around the late 1990s, various sub-research areas within AR have emerged such as Handheld AR (AR applications running on Smartphones and handheld computers), Outdoor AR (AR applied in an outdoor setting) and Spatial AR (information projected on the real world using digital projectors). The following section provides a brief introduction to Outdoor and Handheld AR and covers prior work relevant to this thesis.

2.2 Outdoor and Handheld AR

One of the main application areas for AR is for outdoor use, where virtual content is overlaid on the users surrounding environment. The original outdoor AR applications used large backpack or wearable computers. However, increased computation power, and ubiquity of various location aware sensors in smartphones have led to an increased interest in Handheld AR research. While the focus of this thesis is on interaction between handheld devices and Head mounted displays for Outdoor AR, earlier research in AR systems combining handheld devices and HMDs in an indoor setting is also very relevant.

The sections below provide a brief overview of the wide range of Outdoor AR projects, types of input devices used in Mobile AR (Outdoor and Handheld AR systems) and previous research combining Head Mounted Displays with touch input devices.

2.2.1 Range of Outdoor AR Applications

The first AR system to be developed for outdoors was the **MARS** system in 1996 whose primary function was navigation through an environment and provide information regarding key places [9]. MARS was based on a backpack system and used a handheld tablet for input and HMD for viewing AR content. Using the HMD, users could navigate the Columbia University campus and view information about key buildings. The handheld tablet was used to provide input and view additional details about a particular building.

Figure 2.2 Touring Machine

The MARS system used GPS and compass sensors to locate the user and provide AR overlay on the real world. However, other sensors can be used to improve outdoor tracking. For example, the **ARCHEOGUIDE** (Augmented Reality-based Cultural Heritage On Site GUIDE) project [8] combined computer vision with GPS and compass sensors to more accurately overlay 3D models of buildings that once existed in their current location. ARCHEOGUIDE aimed to develop new interactive methods for accessing cultural heritage information. Users could interact with menu items shown on the HMD through a gamepad.

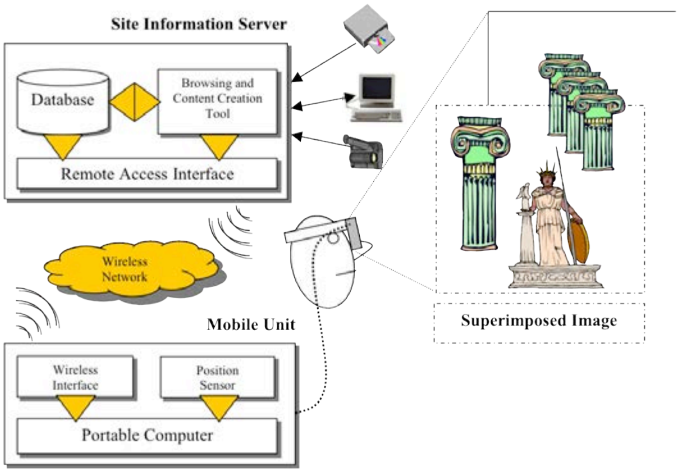


Figure 2.3 The ARCHEOGUIDE System

Outdoor AR has also been explored for military applications. The **BARS** or Battlefield Augmented Reality System [23] used a military navigation and localization system that allowed soldiers to see information such as inside buildings or the position of snipers.

Figure 2.4 BARS System

Visualizing buildings before they are built is another exciting Outdoor AR application. UM-AR-GPSROVER [5] aimed to provide architects the ability to view virtual models of construction in an urban environment. In this case architects view virtual buildings in a HMD before construction so that they can see what they want to build in the future.

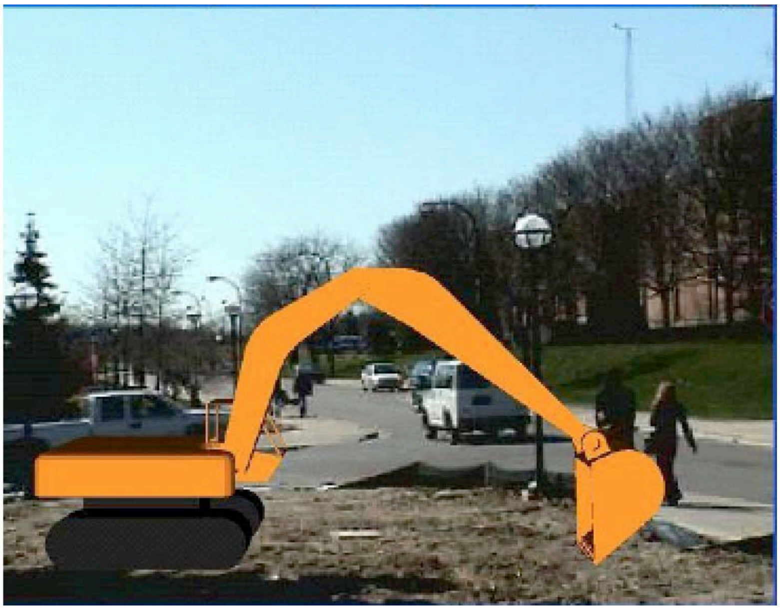
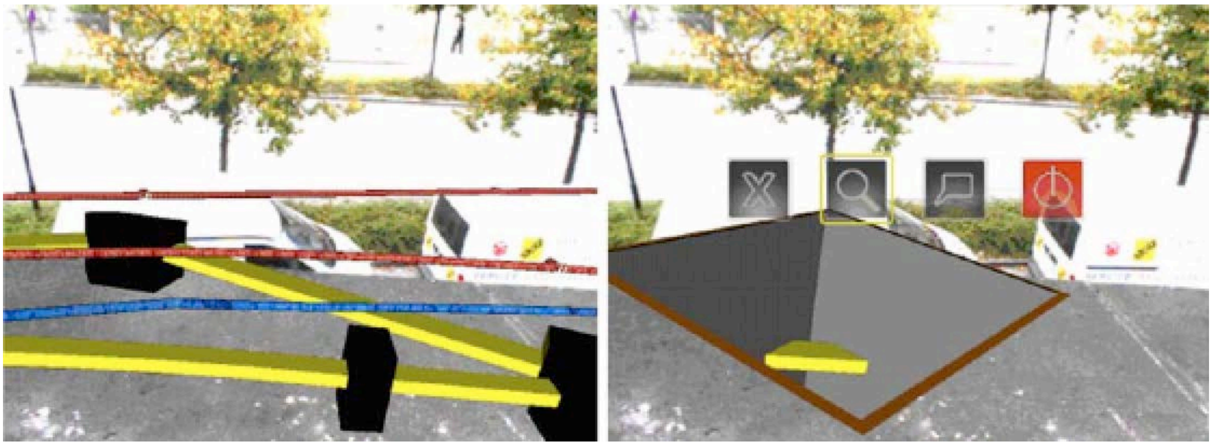


Figure 2.5 The UM-AR-GPSROVER System

Outdoor AR also has applications for gaming and entertainment. ARQuake [14], the first Outdoor AR game, was an adaptation of the popular PC game Quake where players had to shoot virtual enemies that looked like they were part of the real world. ARQuake used a HMD combined with a toy gun input device that allowed players to shoot virtual monsters.

Figure 2.6 ARQuake

Early outdoor AR systems also used handheld systems to show AR content. For example, Schall et al. [18] developed a system that registered a 3D visualization of underground networks on handheld devices in real-time. In this way AR could be used to visualize geo-referenced data that is not directly visible. The handheld device used joystick-like controls using thumbs and index fingers that allowed the user to hold the device for extended periods of time while interacting with the system. Users could use their thumbs to access buttons on the joystick which allowed them to annotate the overlaid 3D model.



a) View on the Vespr screen

b) Input devices for Vespr.

Figure 2.7 The Vespr System a) 3D Visualization of Underground Networks b) Joystick-like Controls allow user to hold the device for extended periods

As can be seen from this section, early outdoor AR research used a variety of input and display devices and a wide range of AR applications were developed. Until the widespread use of GPS and compass equipped smart phones, most mobile AR systems were based around HMDs with custom input devices. In the next section we describe in more detail the types of external input devices that have been used with mobile AR systems.

2.2.2 Mobile AR systems combining HMD and External Input Devices

The Touring Machine [9] was the earliest AR system that made use of external input devices with a HMD. The system used a tablet computer with a stylus for input and a trackpad attached at the back for selecting the menu items. The tablet provided both a second display, and also a surface for pen input.

Reitmayr et al. [16] combined a pen and pad interface with a wearable AR system to build a mobile collaborative AR system. (see figure 2.8) The interface allowed the user to directly interact with virtual objects and collaborate with other users to play games like chess.

Figure 2.8 Mobile Collaborative AR System [16]

Gaze and hand gestures have also been explored as input options for HMD based outdoor AR systems. For example, the Tinmith Project [14] made use of gloves to interact with the AR system (see figure 2.9). The gloves were registered in the virtual world using markers attached to the thumbs and tracked visually. They also had sensors to detect pinching and other gestures. The KIBITZER system [3] tracked the user's eye gaze as natural indicator of attention to identify objects-of-interest. Once objects were selected they could be interacted with using speech input, and non-speech auditory feedback was used to help users to navigate their surroundings. The system used a helmet-mounted smartphone and made use of GPS and compass to track the user's position and orientation respectively. Finally, Bane et al. [4] examined the possibility of giving users virtual x-ray vision (see figure 2.10).i.e. the ability to see through walls by using a multimodal interface to a wearable computing system; combining modalities like vision-based gesture recognition, the Twiddler [1] (a handheld device combining mouse and keyboard) and speech recognition.

Figure 2.9 The Tinmith System

[1] http://www.handykey.com/

Figure 2.10 Virtual X-ray vision to look at building interiors

2.2.3 Research combining Handheld Devices and HMD

While a broad range of devices have been used for both Outdoor and Handheld AR, there has been limited research involving handheld computers or Smartphone based interaction with Head Mounted Displays.

As described above, the Touring Machine made use of a handheld computer to show a map or provide input through stylus and a trackpad (See figure 2.11). Using the Touring machine, the users were able to view information about the places of interest in their vicinity on their HMD and get more information about a particular location (a.k.a. place of interest) by selecting items on the handheld computer using a stylus. Menu items on the HMD could be manipulated using a two-button trackpad mounted on the back of the hand-held computer. Visual feedback is also provided on the HMD as the item that is selected on the handheld is translated down to and off the bottom of the head worn display.

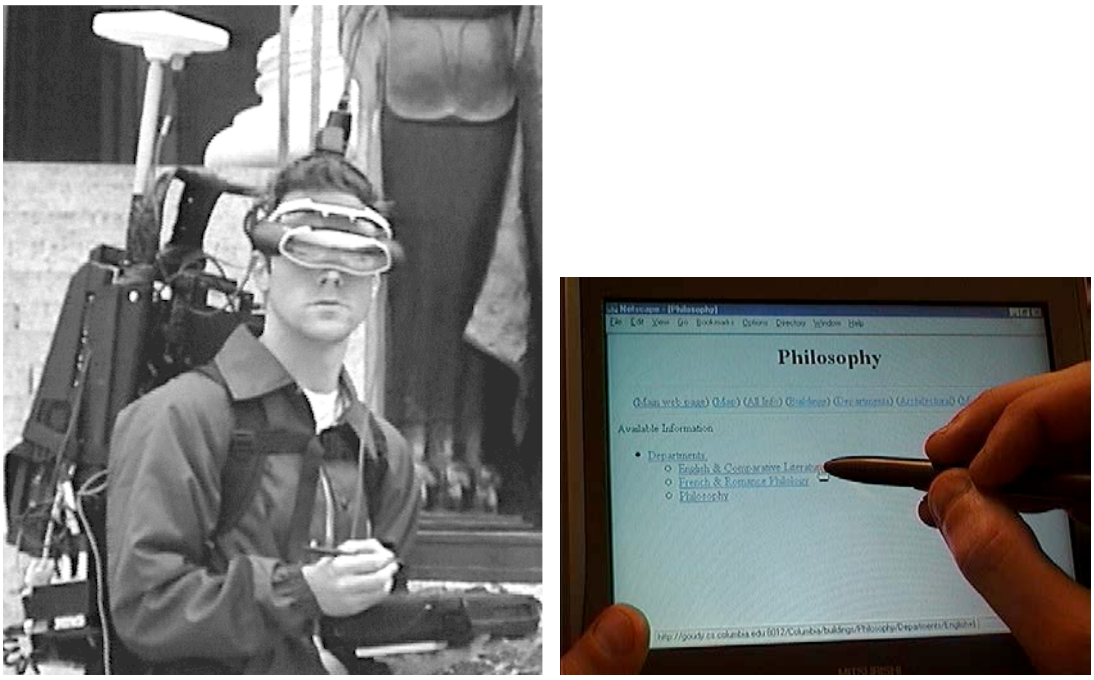


Figure 2.11 Handheld Device for the Touring Machine

Touch input has also been used in mobile AR systems. For example, ARMAR [10] evaluated the benefits of Augmented Reality to support military mechanics conducting routine maintenance tasks inside an armored vehicle. The head mounted display augmented the mechanics view with virtual information that assisted him or her in completing a maintenance task. The mechanic can control the speed of the animated sequence or replay it using a touch-enabled smartphone worn on the wrist that acts as a controller. The user interface of the smartphone had a motion slider that controlled the speed of the presentation, forward and back buttons to navigate between different tasks and start and stop buttons to pause and resume the presentation respectively. The smartphone provided both touch input and also simple visual cues for the user.

Smartphones also allow additional sensors to be used for AR input. For example, AR Wand [20] demonstrated a phone-based 3D manipulation method to translate an object in an AR environment. The system makes use of the phone's orientation (captured through tilt sensors) and the position of the fingertip on the touch screen. This was used to generate a 3D Vector that is then converted to an appropriate 3D motion vector using a transfer function. This converted motion vector is used to translate the target.

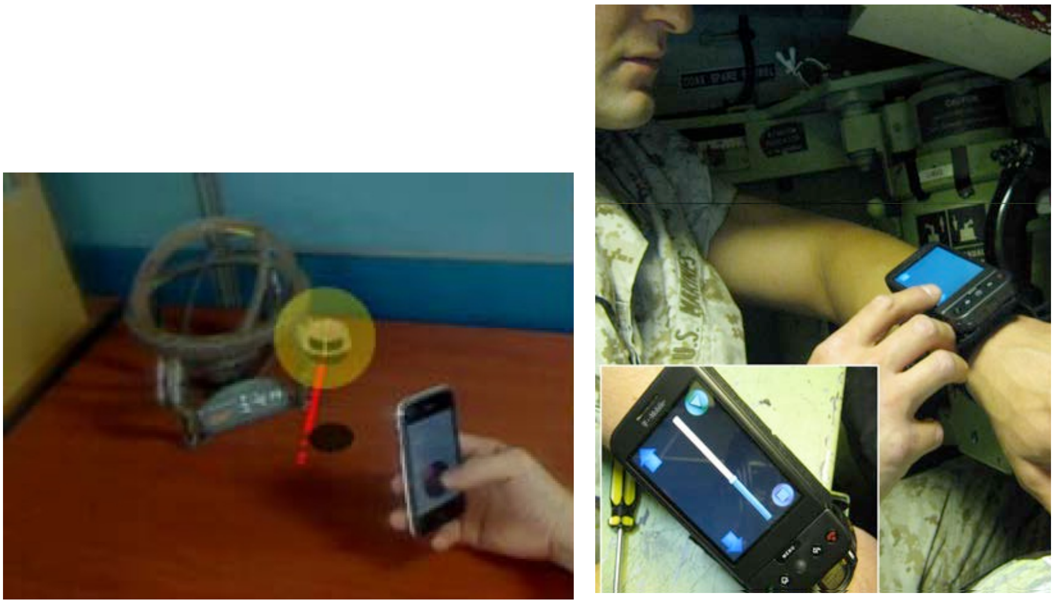


Figure 2.12: Handheld devices used in the ARWand [20] and ARMAR [10] projects

2.3 Splitting Data across Multiple Displays

As shown in the previous section, using a second handheld device with a HMD means that the user may have access to a second display. So, one interesting research question is how to display information across several displays. Various studies have been conducted to determine which information can be best represented on a particular display or combining displays to get the best out of each into a single system.

Wither et al. [21] evaluated a head mounted display, a tablet held as a magic lens and a tablet configured at waist level for AR selection and annotation. Even though time taken by users to complete the tasks was similar in the 3 cases, the user's preferences showed a very different perspective. While the users preferred handheld display to search for real objects, they preferred wearing HMDs while searching for virtual objects.

Figure 2.13 User evaluation of a) Tablet held as Magic Lens, b) HMD and c) Tablet held at waist level

Low et al. [13] combined Head Mounted and Projector based displays to create a hybrid display system combining head-mounted and projector-based displays. The system explored a surgical training application where it is necessary to simultaneously provide both a high-fidelity view of a central close-up task (the surgery) and visual awareness of objects in the surrounding environment. The study examined the effect of two different displays on a user's performance in tasks that simultaneously require the user to concentrate on a central static virtual object while being visually aware of his surroundings. While the studies did prove that hybrid display users are more visually aware of changes in their virtual surroundings, it could not prove that users of the hybrid display can visually and mentally concentrate better on a central static virtual object.

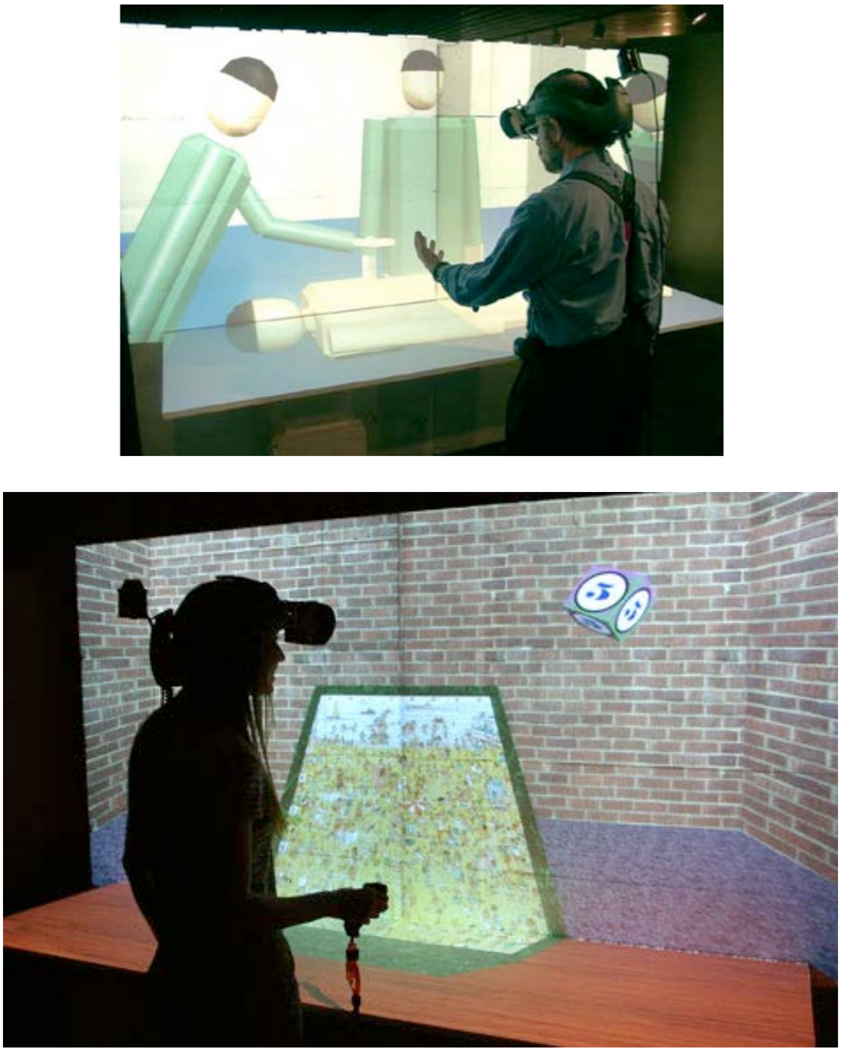


Figure 2.14 a) Hybrid display combining HMD and Projector for Surgery b) User evaluation of the Hybrid System

Ajanki et al. [1] explored contextual information access for AR where the system recognized faces and objects and augmented the scene with relevant information. They conducted a pilot study to compare their system on a HMD with an integrated gaze tracker and a handheld UMPC used in a magic lens configuration. While participants found the system useful, they felt that the HMD was not suitable as its resolution was very low.

2.4 Gestures

This thesis also explores how gestures can be performed on a touch screen to manipulate objects in both 3D and 2D environments. Thus research into how gestures are used in Virtual Reality, AR and Surface Computing are relevant to this research. While the focus of this thesis is mainly on selection techniques, a few other gesture types have also been explored and implemented.

Pierce et al. [15] highlighted four techniques that made use of the 2D image plane concept for selection, manipulation and navigation in virtual environments. Two of these techniques that are of particular significance to this thesis are mentioned below. The *Sticky Finger* technique (see figure 2.15) uses a single outstretched finger to select objects and is useful for picking very large or close objects. The object underneath the user's finger in the 2D image is the object that is selected after the user rests their finger on it for a certain period of time. In the *Head Crusher technique* , (see figure 2.15), the user positions his thumb and forefinger around the desired object in the 2D image to select the object.

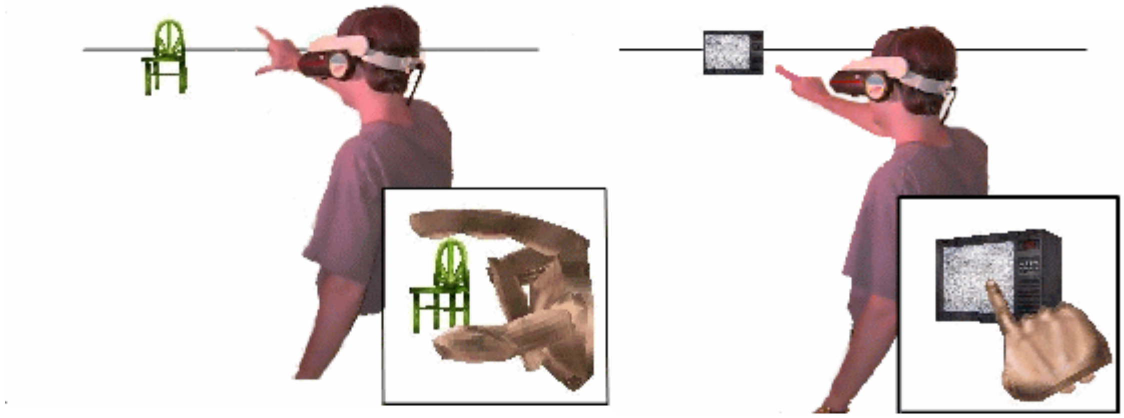


Figure 2.15: Head Crusher and Sticky Finger Gesture [15]

Surface computing gestures are of particular significance to this thesis because gestures used for selecting objects in a 3D environment will be performed on a 2D surface. Wobbrock et al. [22] described a wide range of gestures that can be performed on a 2D surface such as touch screen. These gestures were elicited from non-technical users and were combined with existing knowledge of such gestures. Gestures such as the tap gesture to select an object, drag gesture to move an object and Enlarge/Shrink gestures to scale an object are of particular significance to this thesis.

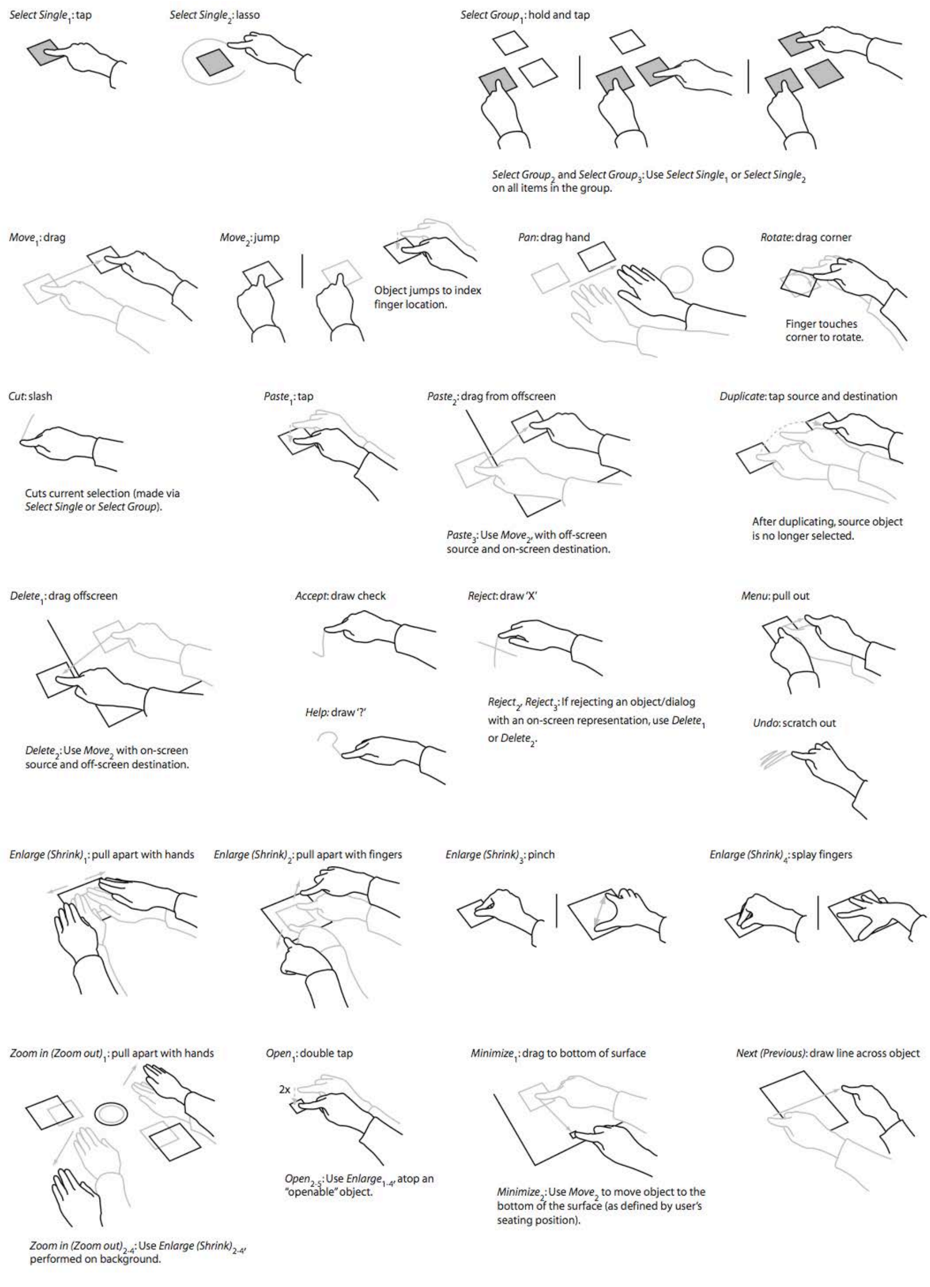


Figure 2.16: Surface Computing Gestures [22]

Combining gestures performed in both 2D and 3D dimensions is also explored in this thesis. In particular, similar gestures that have a different meaning in 2D and 3D contexts are of significant interest. Benko et al. [7] presented a set of basic hand gestures and their interpretations in 2D and 3D environments. Their system used one and two-handed gestures that support the seamless transition of data between co-located 2D and 3D contexts. These gestures were referred to as *Cross Dimensional Gestures*, since the same gesture has a different interpretation in a 2D or 3D environment. While these gestures are not directly applicable to our system, they serve as an example of how a single gesture could be interpreted in multiple environments.

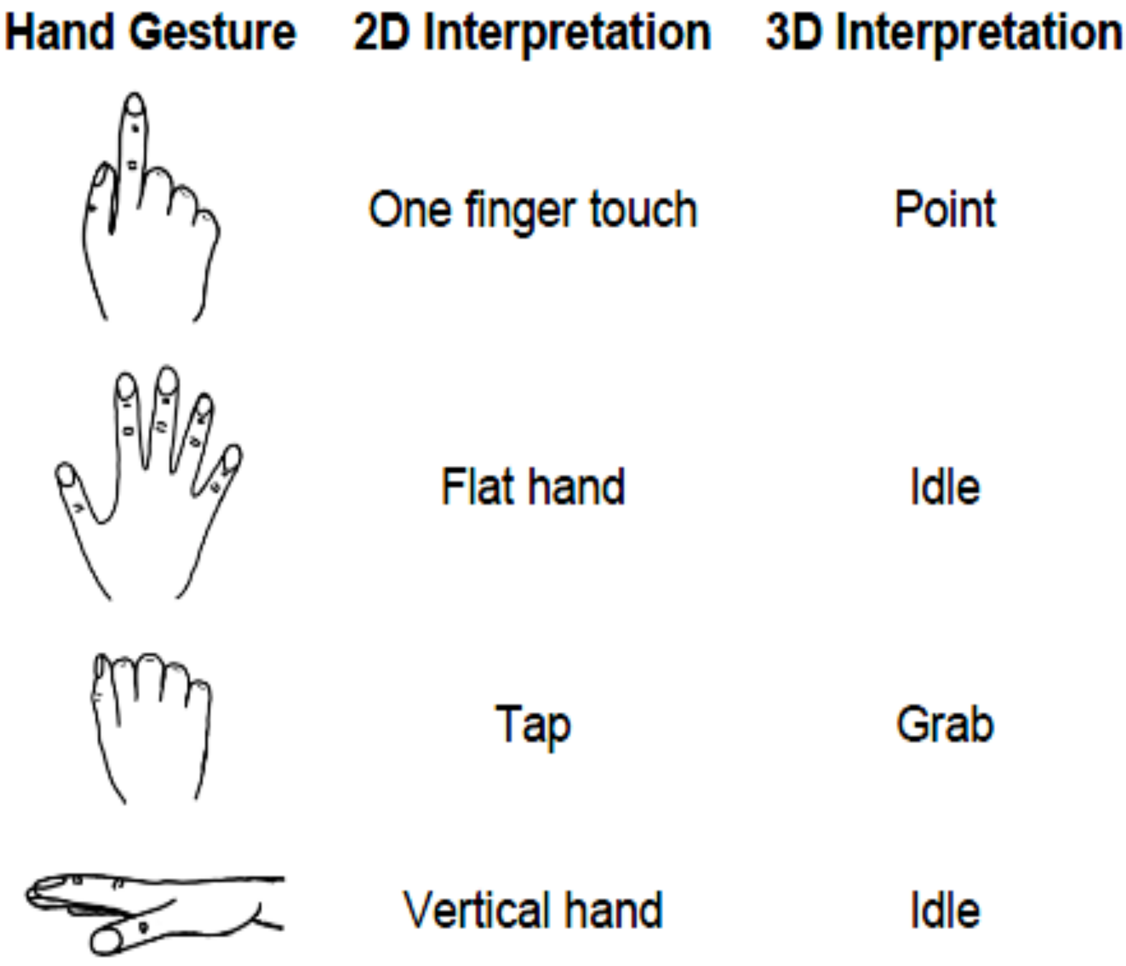


Figure 2.17: Different interpretation of four gestures in 2D and 3D Contexts [7]

2.5 Summary

Previous research in HMD-HHD hybrid systems for Outdoor AR have used the HHD to provide input through virtual and physical buttons, view additional information about the augmented scene, or manipulate an object using gyros of the handheld display. To our knowledge, gestural input using the HHD's touch screen has not been explored earlier. While prior studies do look at comparing multiple displays and evaluating the benefits of using a HMD or HHD for a particular task, there have been no comparative user evaluation studies for HMD-HHD based AR systems. While haptic feedback through physical input devices is explored, visual feedback cues for touch gesture input on both displays has not been explored. Providing an intuitive user experience is a design guideline followed by most AR systems that use either a HHD or HMD, but maintaining intuitiveness specific to a hybrid system has not been looked at. Previous HMD-HHD Outdoor AR research covers different ways of splitting/showing information on a HHD or HMD but simultaneous viewing using both devices has not been explored.

Chapter 3

Research Approach

The overall focus of this master's thesis is exploring different methods of interaction between a Head Mounted Display and a handheld device. Interaction is a broad term that could cover a wide range of possibilities, use cases and application scenarios. The methods proposed in this thesis are different from those covered in prior research and are fairly generic, so they can be applied to different sub-domains and applications. Before finalizing the use-cases, an initial bare-bones version of the prototype was created to help us choose our platform and understand the feasibility of real-time performance using our system. Once this basic prototype was developed and the boundaries of using this system were defined, the use-cases were modified and expanded upon. The final set of fully-formed use-cases led to formulation of research questions that were the key motivators for this thesis. This chapter provides a detailed description of the steps taken in my research approach.

3.1 Differentiation from previous research

To recap, prior research in HMD-HHD hybrid systems typically used the handheld device in one or more of three ways – (1) for presentation control, (2) as an additional information display and (3) for gestural manipulation of virtual content (through tilt sensors and the touch screen). The focus of previous research on gestural manipulation was on translating 3D content using a combination of 2D touch screen and gyroscope input. However, this thesis explores the in-depth use of a handheld touch screen to provide gesture input for a head mounted display. There has been little research before on using touch-screen gestures with HMDs, and none that has used the handheld device to provide rich gesture interaction.

This thesis also explores if using such a hybrid system provides an intuitive user-experience. Using the handheld device as a gestural input surface to interact with virtual content on the HMD must not compromise the overall user experience of using such a system and should feel natural irrespective of the intended application. This particular aspect has not been covered in prior research. Another novelty of our hybrid system is the multi-display viewing of a single scene. While prior research has explored using HMD-projector hybrid systems or HMD and large touch screen systems, multi-display interaction using a handheld display (HHD) and HMD in AR has not been explored. A unique aspect of the HMD-HHD hybrid system is the mobility of a HHD display. The HHD can be placed in front of the user or removed at any time whereas large touch screens or projected screens are static.

3.2 Bare-bones Prototype: Understanding the boundaries

Ideally, the interaction methods discussed in this thesis would be implemented on a head-mounted display (such as the Google Glass) and a handheld device. Due to the unavailability of the intended hardware platform, building an initial prototype was necessary to establish a framework that can be used to test our use-cases. This step was essential as a lot of difficulties were encountered during this phase, which are mentioned in the next chapter. The prototype system was able to send touch positions and simple text messages between the two platforms. This helped us to understand the framework's boundaries and the feasibility of real-time applications on this platform.

3.3 Proposed Interaction Methods

3.3.1 Handheld Device as a Gestural Input device

This is the main focus of my thesis research. The potential of using a handheld device's touch sensitive screen for interacting with virtual content on a HMD is explored. While there is a vast literature of gestural research in both Virtual Reality and Surface Computing, these gestures have not been explored for use on a small touch screen like on a smartphone, and as input for a head mounted display. Some of the gestures (especially those borrowed from VR research) had to be modified for use on a smartphone.

One proposed gesture that is unique to our HMD-hybrid system is the cross–dimensional swipe gesture. The term cross-dimensional gesture has been proposed in earlier research where the same gesture is interpreted differently in different virtual worlds (2D and 3D) [7]. The cross dimensional gesture used in this thesis is unique because the gesture is viewed in the 3D world on the HMD while being performed on a 2D touch screen, but the swipe gesture can be viewed on both displays. Swiping is similar to dragging an object from one display onto another which feels natural on a monocular HMD and smartphone system. For example if the monocular HMD is worn covering the right eye and you are holding your phone in the left hand, the swiping gesture from the right to the left or vice-versa would seem intuitive.

The proposed gesture set that we are going to use is described below.

For all touch based techniques:

In this thesis, we explore three touch interaction methods for object selection: *Sticky Finger, Head Crusher,* and *Tap-again.* All of the methods implemented in this thesis are used in the following way:

- The user touches the HHD touch screen with his or her finger(s).
- The position of touch points on the HHD appears on the HMD as a *circular cursor*.
- Points on the HHD touch screen are absolutely mapped to the HMD view.

However each method has its own unique way of selecting objects in the HMD view, as described below.

Sticky Finger:

To select an object the user uses gesture input to place a cursor on the object and keeps the finger down on the touch screen to keep the cursor on the object. If the user keeps the cursor on the object for a certain period of time, the object gets selected. As the user drags his finger on the HHD, the selected object follows the cursor movement.

Head Crusher:

To select an object, the user touches the HHD with two fingers and places a pair of cursors on top and bottom of the object. As the user drags his or her fingers, the selected object follows the cursor movement.

Tap-again:

With the Tap-again technique, the user uses gestures on a HHD to places a cursor on the object and then lifts the finger up briefly and then taps down again. If the finger is lifted up longer than a certain period, selection doesn’t happen. As the user drags his finger, the selected object follows the cursor movement.

In addition to implementing these techniques for object selection, we also implemented a pinch to scale gesture and cross-dimensional gestures.

Scale :

To scale objects in the HMD view, using the HHD, the user makes gestures on the touch screen and places a pair of cursors on both sides of the object to be scaled. The user pinches his or her fingers which scales the object depending on the how much the user pinches their fingers together. Pinching the cursors together scales down the object whereas pinching them apart scales up the object.

Cross-Dimensional gestures :

Two cross-dimensional gestures are proposed.

a) 3D to 2D

The user selects the object to be selected. A swipe gesture is performed on the touch screen and the cursor appears to move off the edge of the HMD view. Additional Information about the selected object appears on the smartphone screen.

b) 2D to 3D

The user selects a 2D picture of the object on the touch screen. A swipe gesture is performed on the screen and the picture appears to move off the edge of the touch screen. A 3D representation of the object appears in the HMD view.Detailed images illustrating the gestures are shown in the next chapter.

Motion pointing interaction:

In addition to exploring the use of gestures on the touch screen, we were also interested in exploring gestures made with the device itself. Many handheld devices, such as smart phones, incorporate accelerometers or gyroscopes so it is easy to detect device orientation.

For motion pointing interaction, the user holds the phone with one hand in portrait orientation. As the user turns the phone the cursor moves on the HMD screen. The phone works as if it is a laser pointer pointing on a screen except the cursor is displayed on the HMD. The user can tap on the touch screen to do button input to select or drag an object under the cursor. Once selected, the manipulation of the virtual objects with the cursor will work same as the touch-based interaction case.

3.3.2 Changing Visualization based on relative position of HMD and HHD

Another research area explored in this thesis is how to increase the intuitiveness of using a HMD-HHD hybrid system for outdoor AR applications. Combining multiple functionalities into a single hybrid system while preserving the intuitiveness of the system would make a strong case to use this hybrid system.

An example of an outdoor AR application with a very intuitive interface is the Nokia City Lens app for the Nokia Lumia 920 smartphone [33]. This app shows how a handheld system can be made intuitive depending upon how the device is held. It makes use of the handheld's orientation sensor to determine how the device was held, and displays the appropriate user interface view. For example, when the phone is held horizontal it automatically shows a map view (see Table 3.1). In this way the application provides three different viewing functionalities into a single app while preserving the fluidity of the experience, and not requiring the user to explicitly changes viewing modes.

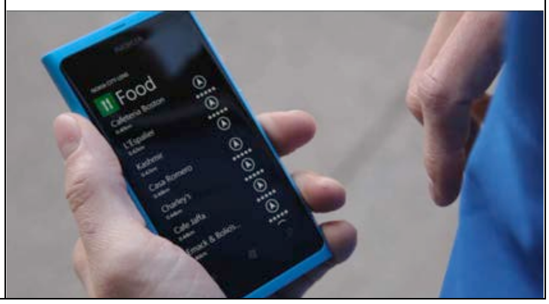
Device Configuration	Device View	Screenshot
Phone held in landscape mode	Map View	
Phone held vertically upwards	AR View	
Phone held at 45^0 angle in portrait mode	List View	

Table 3.1: Nokia Lumia 920: Changing Visualization based on realtive HHD position

This concept of using a handheld device intuitively is extended to the proposed HMD-HHD hybrid system. In our system we track the angle of the HMD and when the user is looking at the HHD, the HMD shows black. This makes the display transparent and it is easier for the user to see the HHD. Thus our system intuitively switches HMD view mode based on the angle of the users head. A table outlining the various details of our extended application is shown below (See table 3.2).

Usage Scenario	Handheld Device	HMD View
User Not looking at Handheld Device	Controller to interact with AR View	AR View
User looking at Handheld Device. Device held in Landscape Mode	Map View	Black/ Transparent
User looking at Handheld Device . Device held in Portrait Mode.	List View	Black/ Transparent
User looking at AR View, Handheld Device is held vertically upwards	Extra Information about Place	AR View

Table 3.2: Changing Visualization based on relative position of HMD and HHD

The proposed use cases are generic and can be applied to a range of outdoor AR domains. Fields such as navigation, visualizations, gaming etc. could benefit from using such a hybrid system as both information and gestures are accessible in the same application.

3.3.3 Multi-display viewing of a single scene

In the proposed hybrid system, the HMD is used to display virtual content on the screen whereas the handheld device is used for gestural input or providing additional information. If the users view on the HMD is augmented with virtual information, it may be possible to use the handheld display to show another view of the AR scene that it is held in front of the face. A research question that this will allow us to get feedback is “Is there a significant use of using the HHD to show another augmentation of the already augmented scene?” A mock-up of this concept is shown below (see figure 3.2).

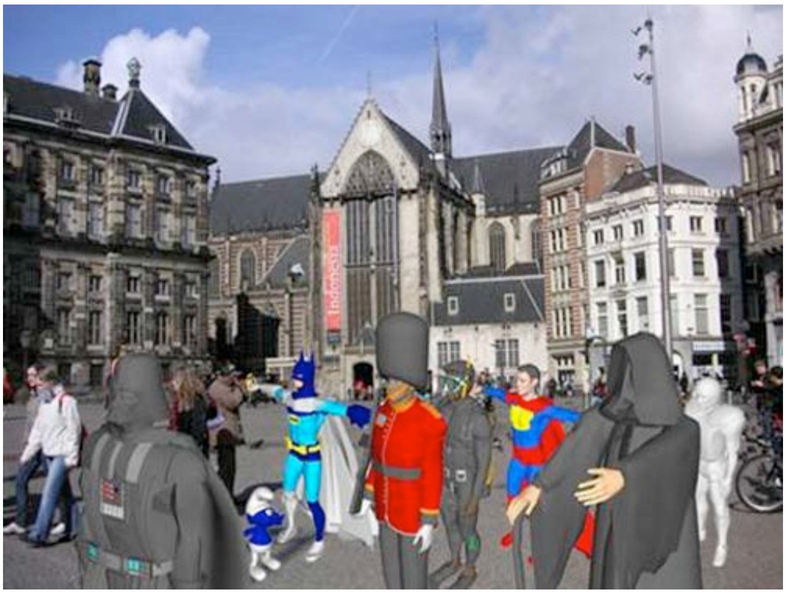


Figure 3.1 AR view as seen from the HMD

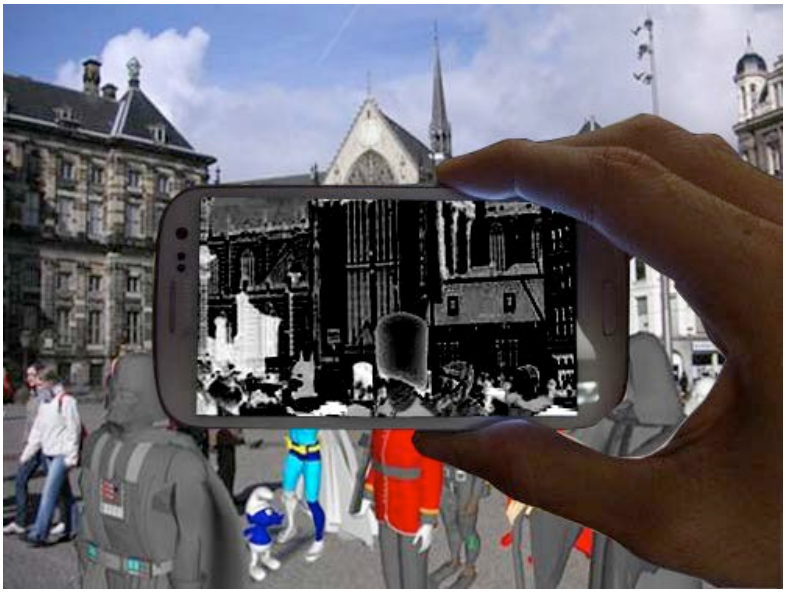


Figure 3.2 Multi-Display view as seen from the HMD

One domain where such multi-display views could be impactful are historical tours. For example, if the user’s view is augmented by a virtual castle, the user could bring the phone in front of his or her view and could look at the interior view of the castle on the HHD, while looking at the exterior view of the castle through the HMD. The handheld is thus like a portable magic lens that when it's brought in front of the users view, and it "augments the Augmented" scene.

3.4 Research Questions

Based upon the interaction use-cases discussed in the previous section, a number of research questions were formed that helped set the direction of this thesis. These questions are described in this section.

In a HMD-HHD **hybrid system, what are the range of gestures that can be performed on the handheld device to interact with virtual content on the HMD? Which of these gestures is best suited for selection tasks?**

This is the primary research question of my thesis. In the HMD-HHD hybrid system, gestures are performed using the handheld device to manipulate objects viewed on the HMD. Some handheld devices, such as Smartphones, have a high-resolution touch screen and sensors such as accelerometers, gyroscopes etc. So, there can be different types of gestures that can be done using a smartphone. A HMD-HHD hybrid system has two displays – one display can view the 3D world and the other can view high-resolution 2D content. Gestures that make use of this dimensional-divide and can influence both worlds have also been looked at. From the proposed gesture data-set, four gestures have been evaluated to understand the users' preferences of gestures while doing selection tasks on a HMD-HHD hybrid system.

Can a HMD-HHD hybrid system provide an intuitive user experience?

With a wide range of gestures that can be developed for use with HMD-HHD hybrid systems, our system must provide an intuitive user experience to justify combining the two devices together. Failure to do so would result in the handheld being an accessory/additional input device for the HMD rather than the two being part of a complete system.

Is it useful to view a multi-display scene that fuses the real world view and virtual content shown on the HMD and handheld display?

This is a novel concept explored in this thesis. The user wears a HMD and his/her view is augmented with virtual content. Using the above system, is it useful to create a 'tribrid' or an 'augmented view of the augmented scene' where the handheld is placed in front of the HMD and shows different virtual content which seems like a blend of three different worlds in the same scene.'Tribrid' here refers to the view combining the real world seen through his eyes,the virtual world shown on the HMD and the virtual world shown on the smartphone.While there are scenarios where such an approach would be desirable, the prototype developed must effectively convey this concept to have a significant impact.

Chapter 4

Prototype Design and Development

This chapter describes details of the prototype system and the different features of interaction built for the system. The reasons influencing the choice of software and hardware platform are explained first followed by the system architecture and applications built for the system.

4.1 Choice of hardware and Software Platform

4.1.1 Hardware Platform

Ideally, our intended hardware platform is a smartphone and Google Glass i.e. an independent head mounted computer. Since Google Glass is currently unavailable, a considerable amount of time had been spent in assembling a prototype system similar to the intended target platform.

The first attempt involved connecting a HMD to a smartphone where the HMD would display the information shown on the smartphone’s screen. This smartphone was connected via Bluetooth to another phone that was held in the users hand and could be used for providing gestural input. The difficulty in connecting external devices (such as cameras and orientation sensors) to a smartphone prohibited us from using this setup. However, this preliminary setup helped understanding that Bluetooth is not fast enough for transferring touch positions between two mobile devices.

The above-mentioned difficulties led us to look at Ultra Mobile PC's (or UMPCs) as a possible alternative. UMPCs can connect to external devices (like HMDS and cameras) through USB and smartphones through wireless networks. These devices are lightweight, and can be easily worn using small bags (see Figure 4.1). Additionally, UMPCs run the Windows operating system that makes it easier for programs written on desktop PCs to be ported to the UMPC.

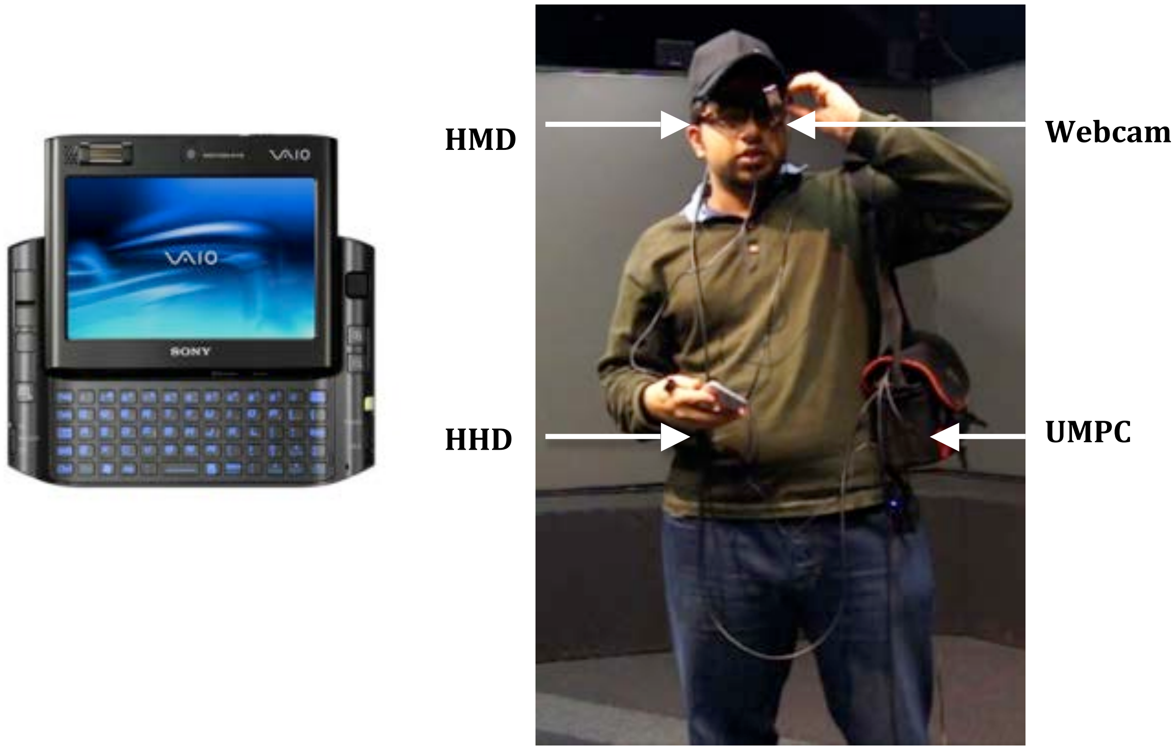


Figure 4.1 : A UMPC and user wearing the system

4.1.2 HMD selection

Our HMD selection criteria required the HMD to be monocular and to preferably have a head orientation tracker. The monocular HMD is preferable because it allows people to more easily see the real world, while still viewing additional information on the HMD. This type of HMD is also useful for mobility. The Brother Air Scouter HMD [2],(see left of Figure 4.2) fulfills one of the above criteria and uses retinal projection display technology similar to Google Glass, which is why it was considered for usage with our system. However, the absence of an integrated Tracker made it difficult to use this HMD with our

[2] http://www.brother.com/en/news/2011/airscouter/

system. The Vuzix iWear 900 [3](see right of Fig. 4.2) has an integrated head tracker and can be easily modified from binocular to a monocular display. Although the display technology was not as good as the Air Scouter, the modified Monocular Vuzix HMD was used for the prototype system since it fulfilled our criteria.

Figure 4.2: The Brother AirScouter and Vuzix HMD

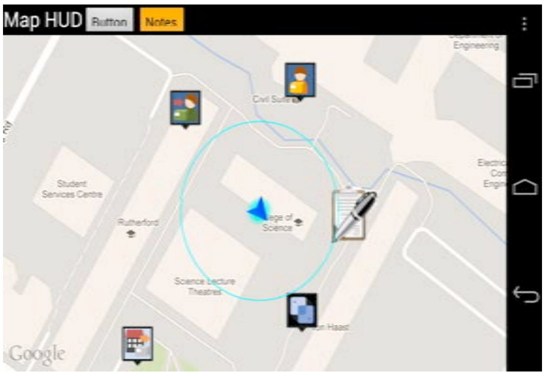
Figure 4.3: Modified Vuzix HMD

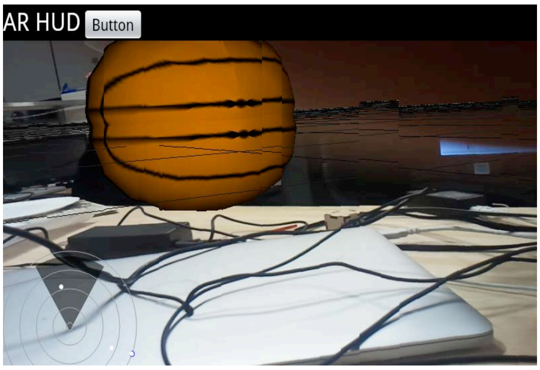
4.1.3 Software platform

The initial prototype system used two Android smartphones that communicated with each other through Bluetooth. One phone had a combination of a *Map View* that showed points of interest on a map and a *List view* that listed information about those points of interest. The other phone, that we planned to feed video signal to a HMD, had a dedicated AR view that tracked the position of the phone using GPS and augmented virtual models on the

[3] http://www.vuzix.com/consumer/products_vr920.html

Camera Feed corresponding to the points of interest. The HITLabNZ Outdoor AR library[4] was used to prototype this system.





Map View **AR View**

Figure 4.4 : Initial platform with Map and ARView on different HHDs

Moving to a smartphone-UMPC system required a complete change in software platform. Choosing a portable software platform was necessary so that we could easily migrate our code. Prototyping different applications quickly without sacrificing robustness was also an important criteria, which prompted us to look at creative coding platforms. Creative coding platforms allow the developer to quickly prototype programs, without any knowledge of low-level libraries. This simplifies programming and facilitates rapid prototyping. For the prototype system, the author quickly prototyped applications without an in-depth knowledge of OpenGL, Model Loading libraries, Computer Vision etc.

Initially, the Processing platform [34] was chosen because of its ease of use and the speed with which one could prototype ideas. Because Processing does not offer real-time performance on a UMPC, the Openframeworks platform was finally chosen. Openframeworks [35] is a creative coding platform that uses C++ and is much faster than Processing which is Java based. Its real time performance on the UMPC and portability to multiple platforms like iOS, Android, Windows, Linux and Mac OS X were suitable for the prototype system.

[4] http://www.hitlabnz.org/mobileAR

4.1.4 Communication

The most important feature required in our HMD-HHD hybrid system was that the positions of the user's fingers must be transmitted in real time and mapped correctly to the users HMD – which is why choosing the right communication protocol was critical for our system. Both speed and accuracy were desired since any delay in transmitting the touch positions would affect the user experience and could lead to incorrect selection of virtual content shown on the HMD.

The User Datagram Protocol (UDP) was chosen because of its speed. By shortening the length of the UDP messages, transmission of finger positions in real time was possible without any problem of lost packets/messages. Since smartphones are capable of hosting mobile hotspots and sending and receiving messages over the hosted network without any attenuation, no external network was required.

4.2 Software Architecture

The software of the prototype system consists of two modules: one on the smartphone and the other on the PC (see Figure 4.5). The software module on the smartphone captures user interaction and sends this information to the module on the PC. The software module on the PC draws the virtual scene while the user interaction information from the smartphone is processed.

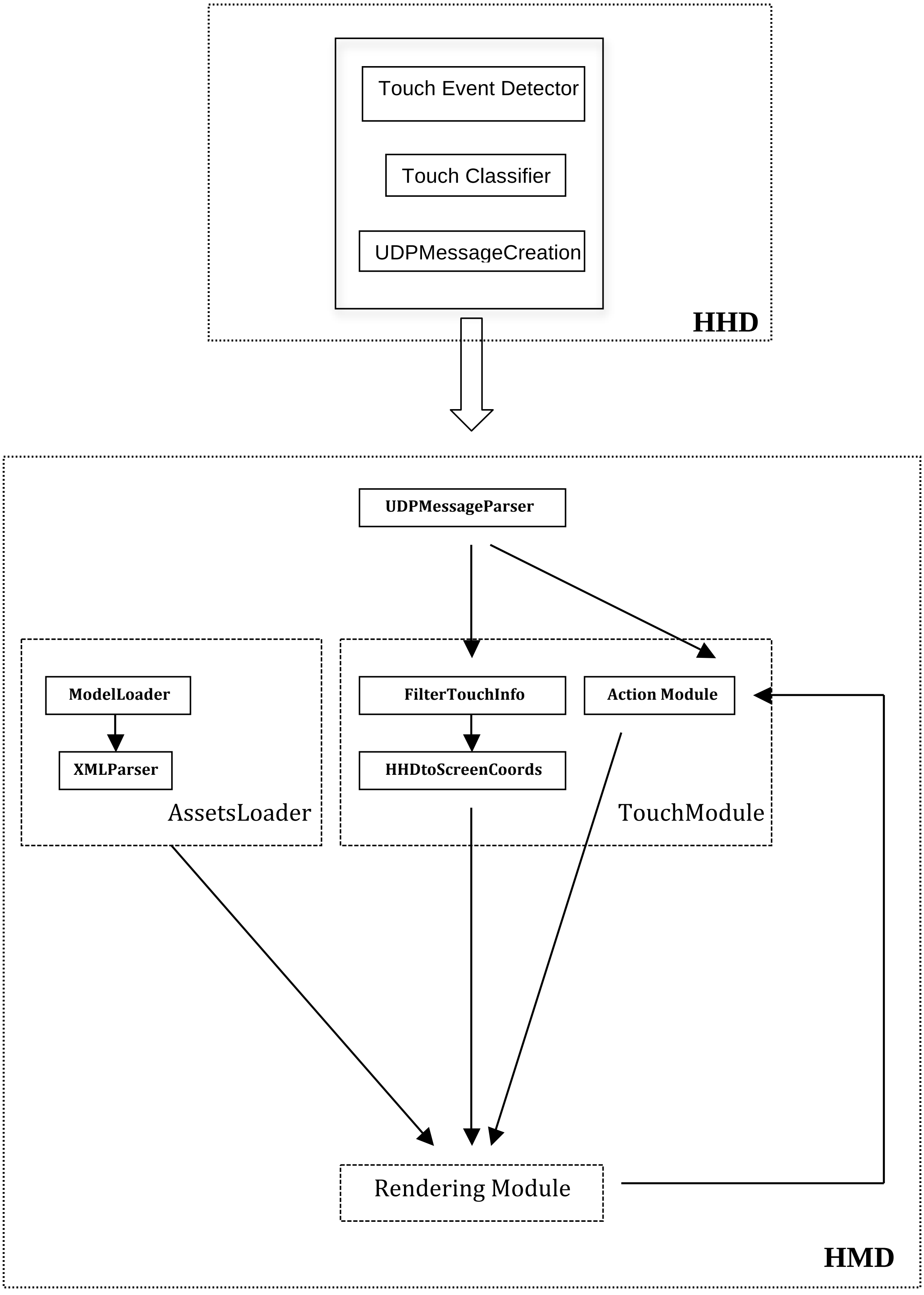


Figure 4.5 : Software architecture of the prototype system

4.2.1 HHD Modules

The Touch Event Detector Module detects finger touches on the smartphone that are sent to the Touch Classifier. The Touch Classifier determines the type of Gesture performed and creates a Message. This message contains details like number of Fingers, type of Gesture Performed, specific details of the Gesture (like duration of the touch while gesture was performed) etc. The touch positions are then packaged into a message, which is sent to the PC by the UDPMessageCreation Module.

4.2.2 PC Modules

The state of the different models present in the virtual scene are stored in an XML file that is parsed using an XMLParser and used by the ModelLoader to load the models and set its properties like scale, position and rotation.

The PC receives the message sent by the smartphone through the UDPModule and the message is parsed and filtered by the FilterTouchInfo module to get information regarding the touch positions. This information is used by the HHDtoScreenCoords module to convert the touch positions from HHD screen coordinates to HMD screen coordinates. The touches are then sent to the Rendering Module that draws them as circles on the user's HMD view.

Information about any gesture performed on the HHD is filtered from the UDPMessageParser by the Action Module and is used with the position of virtual objects provided by the Rendering Module to create an Action Event. This action event could be selection, translation or scaling the virtual object depending upon the nature of the target application and is sent to the Rendering Module.

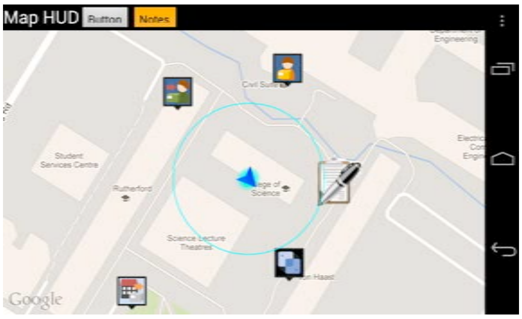
4.3 Interaction Design

4.3.1 Note Taking and Augmenting on the Move

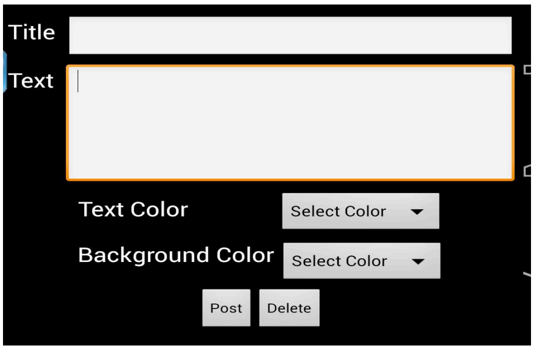
The first interaction is note taking and editing on the move. The interaction for creating notes involves creating notes on the handheld device and viewing it on the HMD. On the smartphone, the user places a cursor on the map where he wished to add a note (see Figure

4.6b) and clicks the add note button. This button opens a view that allows the user to enter details like Note title, description and allow him to customize the colour of the note text and the plane(see Figure 4.6c). The note is stored and appears on the map as an icon. The note is also visualized in the AR view on the HMD as a 3D plane with the note at the GPS position specified by the cursor in the previous step . The added note could also be edited on the fly and the changes would appear immediately in the AR view.

a.



b.

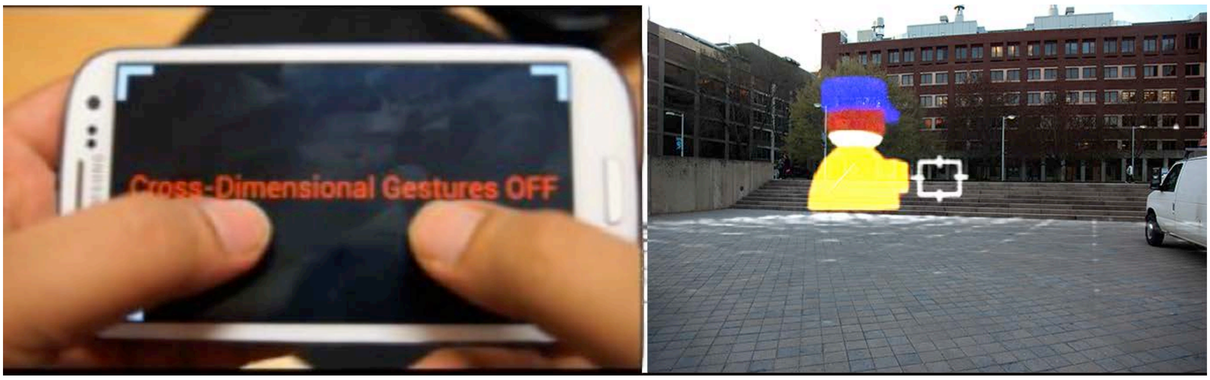


c.

Figure 4.6: Note Taking on the HHD application a) Login Screen b) Adding note at a particular location c) Editing details of the note like Text and color.

4.3.2 Translating and Scaling Models through Gestural Input

Another interaction developed for the prototype system allows the user to manipulate the content viewed on the HMD. The user uses touch gestures on the smartphone to translate or scale virtual objects (see figure 4.7). Three types of gestures are designed by this application, the details of which are mentioned below.



HHD View **HMD View**

Figure 4.7: Manipulating 3D models on the HMD using touch gestures on the smartphone

The left part of the Figure 4.7 shows the smartphone view and the right part shows the view seen through the HMD. For every finger touch on the smartphone a green circular cursor appears on the HMD view.

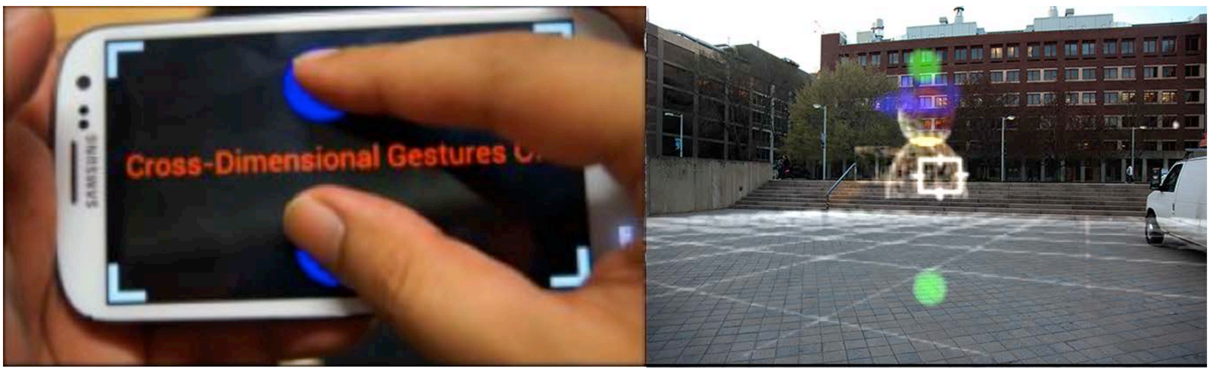
1) Sticky Finger Gesture for Translating Virtual Models

With Sticky Finger gesture, the user selects the virtual object by placing a cursor on the virtual object and holding it still for a second (see the top of Figure 4.8). When the object is selected, it is replaced by a wireframe version of the same model indicating that the object is selected. Once the object is selected, the user translates it by dragging his finger along the smartphone. To stop translating the object, the user lifts his finger off the smartphone (see the bottom of Figure 4.8)

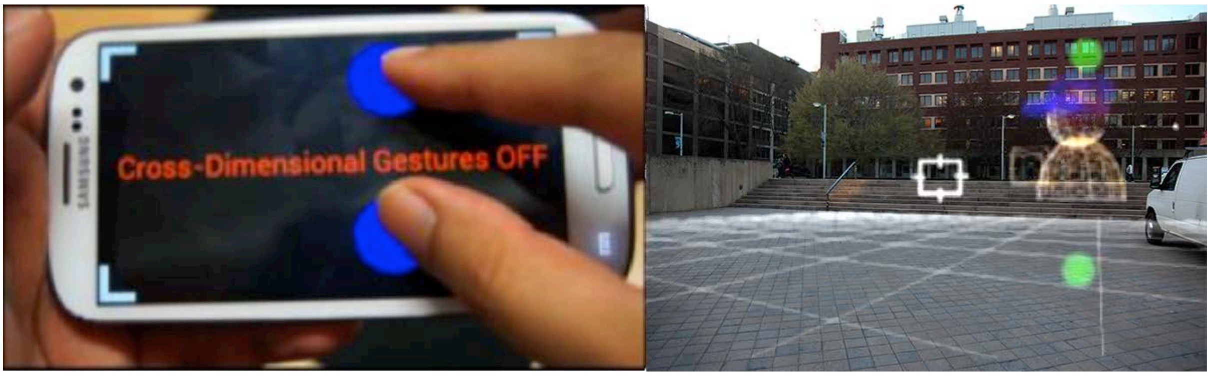
Figure 4.8: Sticky Finger gesture

2) Head Crusher Gesture for Translating Objects

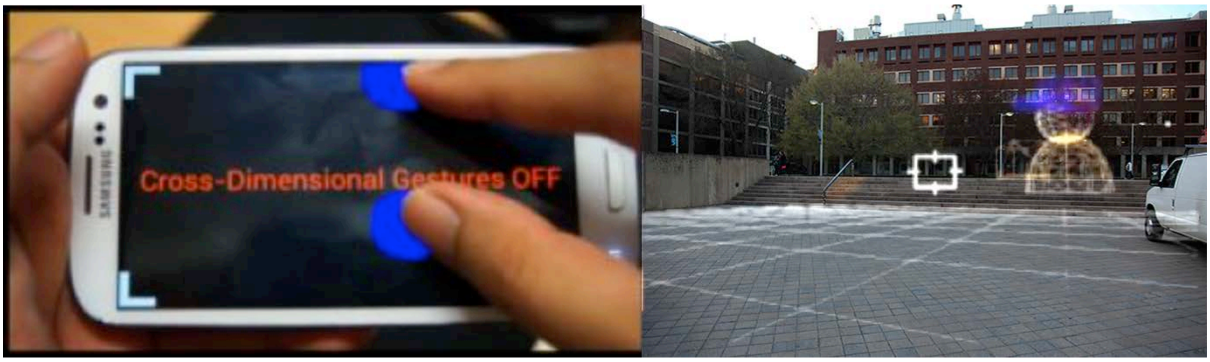
With the head crusher gesture, the user selects the virtual object by placing a pair of cursors on top and down of the virtual object and holds still for a second (see Figure 4.9a) When the object is selected, it is replaced by a wireframe version of the same model indicating that the object is selected. Once the object is selected, the user translates it by dragging two of his fingers along the smartphone (see Figure 4.9b). If the user wishes to use only one finger for translating the object, he simply removes one finger from the smartphone screen while keeping the other finger pressed on the screen. This way he/she has the option to select a model by the Head Crusher gesture and translate the model by using one finger (see Figure 4.9c and Figure 4.9d). To stop translating the object, the user lifts his finger off the smartphone.



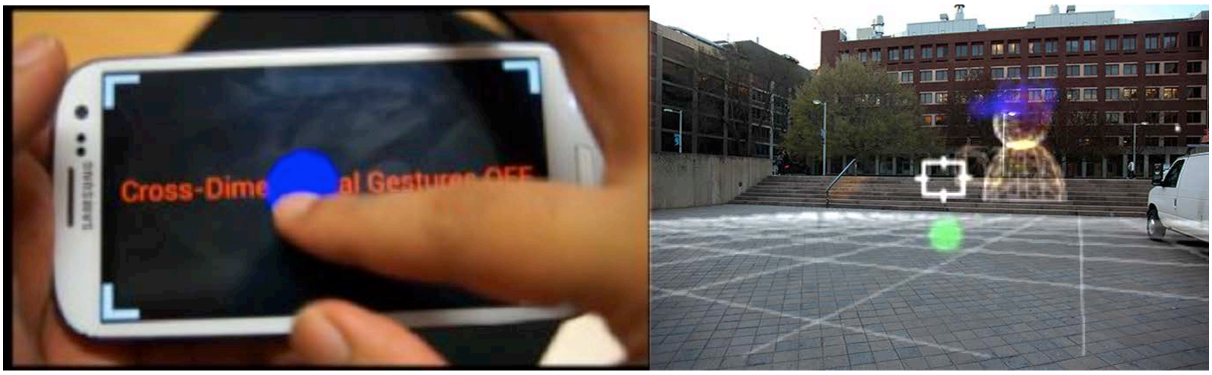
(a)



(b)



(c)



(d)

Figure 4.9 :Head Crusher Gesture

3) Pinch Gesture for scaling up/down objects

For the pinch gesture, the user places a pair of cursors on the side of the virtual object (see top image of Fig 4.10). By pinching his fingers outwards ,the user can scale up the model (see bottom image of Fig 4.10) and by pinching inwards ,the user can scale down the model.

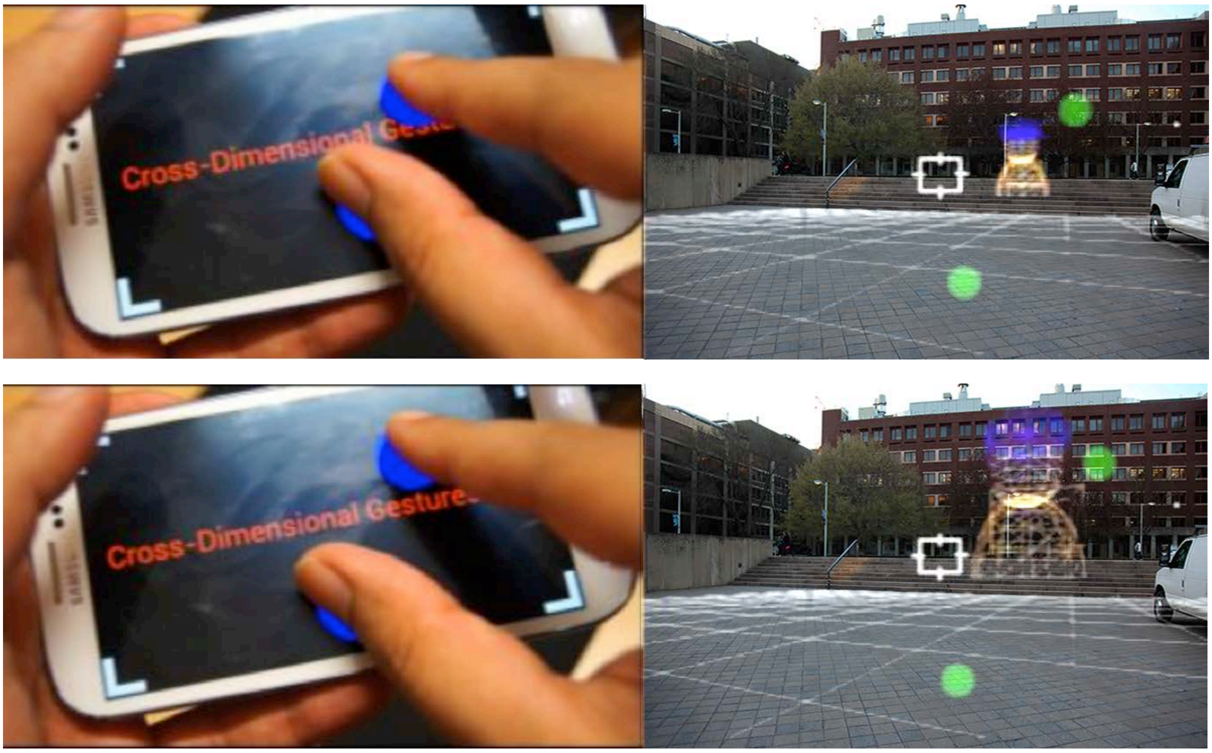
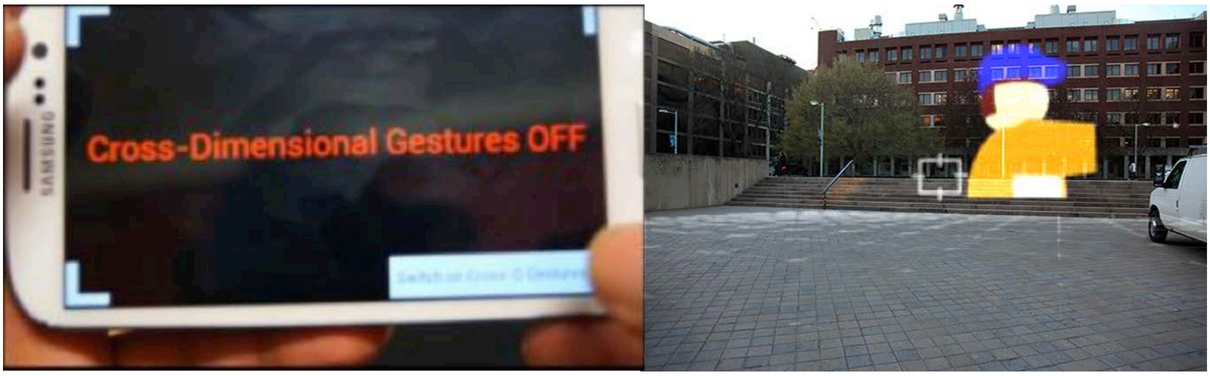


Figure 4.10 : Pinch Gesture

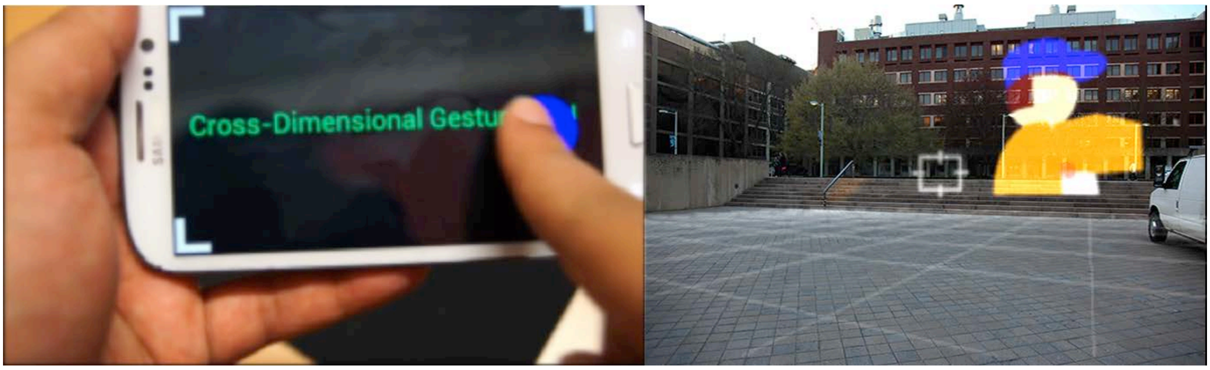
4.3.3 Cross dimensional gestures for Inter-display interaction

The cross-dimensional gestures were designed to allow users to quickly view additional details about an object on the smartphone, without affecting the actual position of the virtual object. The details of the gesture are described below.

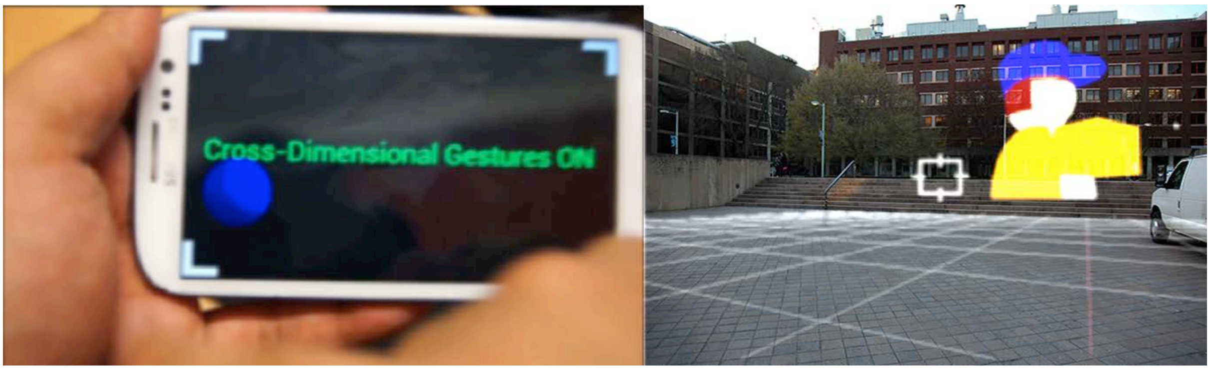
To perform the cross-dimensional gesture, the user switches to the cross-dimensional gesture mode through a menu option (see Figure 4.11a). The user places a cursor on the object for which he wants additional details and performs the swipe gesture while dragging his finger towards the end of the screen (see Figure 4.11b and Figure 4.11c). Additional information about the object appears on the smartphone (see Figure 4.11d).



(a)



(b)



(c)

(d)

Figure 4.11 : Swipe Gesture

4.3.4 Tribrid viewing application combining Real and Virtual Worlds

Smartphone screens have a higher PPI (Pixels per inch) compared to HMDs, making them suitable for viewing high-resolution graphics that are not very clear on the HMD. We developed a tri-brid viewing application that combines information from real time camera feed and two separate virtual worlds – one shown on the HHD and the other on the HMD.

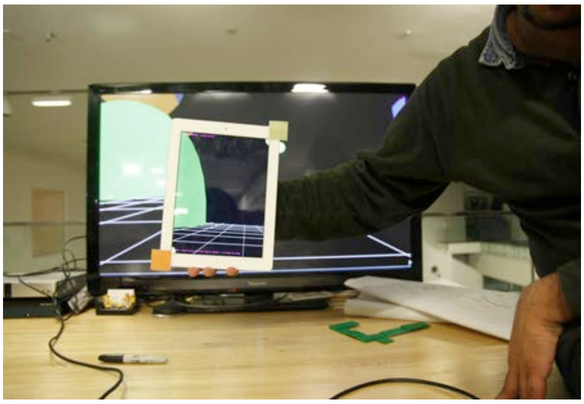
This is useful in situations where you want to look at different versions of the virtual object without pressing any additional button to switch the view or if you want to see a different and higher resolution graphic of a particular object. For example, if you want to look at the interior of a building or an older version of an augmented building.

Due to limitations in streaming video from PC to mobile devices, this application is currently not in real time (2 frames/second) and the user has to explicitly trigger a stream event by pressing a button. Details of the application are mentioned below.

Procedure

1) Capture real time camera feed through a camera mounted on the HMD.
2) Hold the HHD in an upright position and track colored blobs around its border in the camera feed. This will give the position of the smartphone in the scene.
3) Replace the tracked HHD portion of the video image with a black rectangle. This is done to occlude the HHD image in the HMD so that the graphics shown on the smartphone will blend well with the real world and HMD view.
4) Stream the tracked portion that is blocked by the blobs on the HMD view to the application running on the HHD. The HHD then renders the streamed image and can make any modification if necessary.

See Figure 4.12 for results.



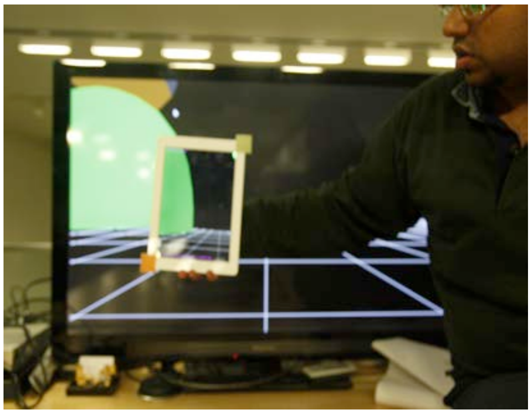


Figure 4.12: A prototype system showing 3D scene on multiple displays registered to each other

4.3.5 Changing Visualization based on relative position of HMD and HHD

This interaction was designed for the prototype system aimed at providing an intuitive user experience while using the HHD–HMD hybrid system. The interaction makes use of the relative orientation of the HHD and HMD. Showing different visualizations depending upon the relative position of HMD and HHD can also provide the combined functionality of several different apps. For instance, Table 4.1 illustrates an Outdoor AR application that integrates a map view, a gesture input mode and an additional information mode while maintaining the overall intuitiveness of the system.

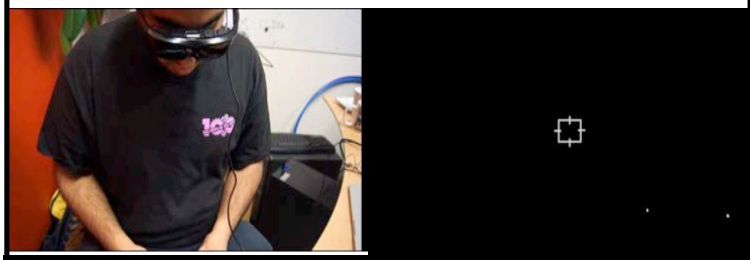
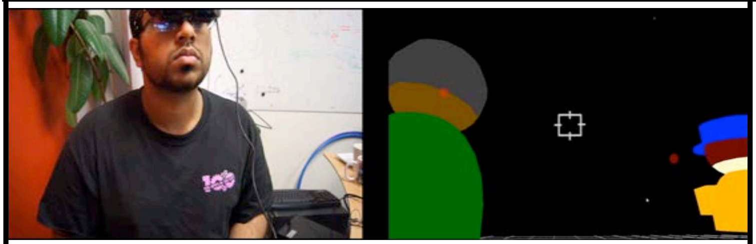
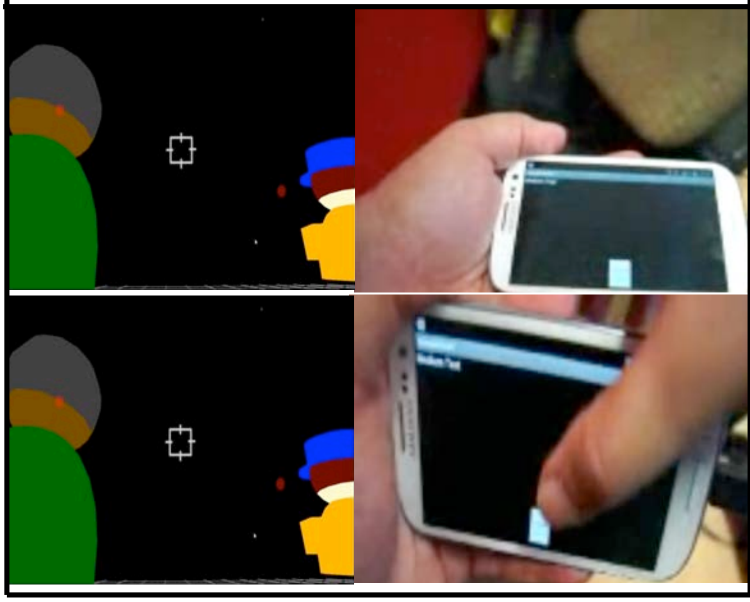
Screenshots	Description
	When the user looks down at the smartphone held in a flat landscape position, the HMD view shows black and the smartphone shows a Map View.
	When the user looks away from the smartphone, the HMD view shows the AR View and the smartphone acts as a gestural input device
	When the user puts the smartphone in the upright position, he can see additional details of the virtual object on the smartphone. Additionally, pressing the freeze button will retain the additional details shown on the screen irrespective of the orientation of the smartphone i.e Map View and Gestural Input will be disabled unless the user presses the unfreeze button.

Table 4.1 : Changing Visualization based on relative position of HMD and smartphone

Chapter 5

Evaluation

This chapter describes the details of an experiment conducted to evaluate the prototype system. The goal and design of the experiment are described first, followed by the results and analysis. All documents and forms that were used in the experiment are found in Appendix A.

5.1 Evaluation goal

The primary goal of the evaluation was to test the proposed HMD-HHD prototype system to determine whether touch based gestures are better than motion based gestures for selection tasks. To achieve this, the author conducted a user experiment that compared touch and motion gestures in terms of accuracy and time taken to complete a particular task. A second goal of the experiment was to collect data to find which touch gesture would be best suited for selection tasks in a HMD-HHD hybrid system. Our criterion for evaluation was the time taken to complete the task, selection accuracy and subjective user feedback.

5.2 Experimental Design

5.2.1 Hypothesis

The overall hypothesis of the user experiment was: **There is no significant difference between using touch screen gesture and device motion for selecting content viewed on a monocular display in a HMD-HHD hybrid system.**

5.2.2 Experimental Procedure

Prior to the experiment, the participants were handed a questionnaire that helped determine their prior experience with mobile devices, 3D interfaces and Outdoor AR applications (See Appendix A). After completing the questionnaire, the participant was then given an introduction to the features of our system.

Once familiar with the technology, each participant then performed a selection task using four different types of gesture. The four gesture conditions used were:

1. Sticky Finger
2. Head Crusher
3. Tap Again
4. Motion Gesture

The order of conditions was counterbalanced using Latin Square so that there was no bias toward a particular gesture type or order effects.

Participants had a trial of performing the task under each condition, and each trial was divided into two parts. The first was a training phase where participants selected 5 targets using a particular gesture. The training phase helped participants to familiarize with the gesture and eliminated any bias that could occur from the participant's earlier knowledge of a gesture. The second part was the actual experimental task itself where participants had to select 20 targets. For the second part of each trial we measured the error in selecting a target and time taken to complete the task. At the end of each condition, we recorded quantitative feedback using questions with Likert-scale ratings and qualitative feedback through questions asking the participants opinion of the technology used (see Appendix A). At the end of the experiment, we recorded the participant's overall preferences, opinions, comments and miscellaneous feedback through a variety of quantitative and qualitative questions.

5.3 Evaluation Prototype

For the purpose of the evaluation, the author developed an application where users had to select 20 targets using four selection methods: Tap again, Sticky Finger, Head Crusher and Motion Gestures. Details of each gesture are described in the previous chapter and the design of the user study mentioned in the next chapter. The flow of the user study application on the smartphone is illustrated below:

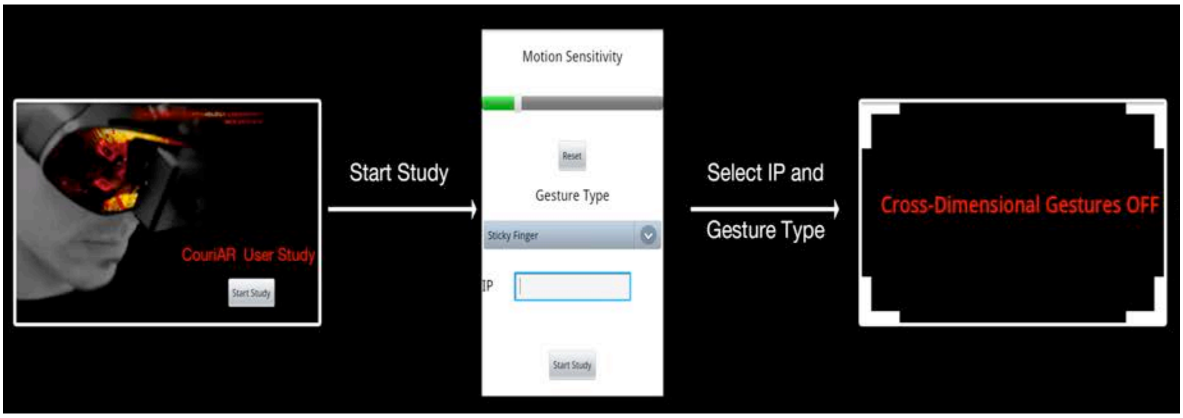


Figure 5.1: The flow of the prototype system built for the user study

The experimenter launches the application on the smartphone and selects the appropriate settings depending upon the type of gesture that will be evaluated. Once these settings are chosen, the experimenter specifies the IP address of the PC and connects with the PC application. The PC application uses information provided by the HHD and maps the gesture to the HMD view. Figure 5.2 is a screen shot of the application.

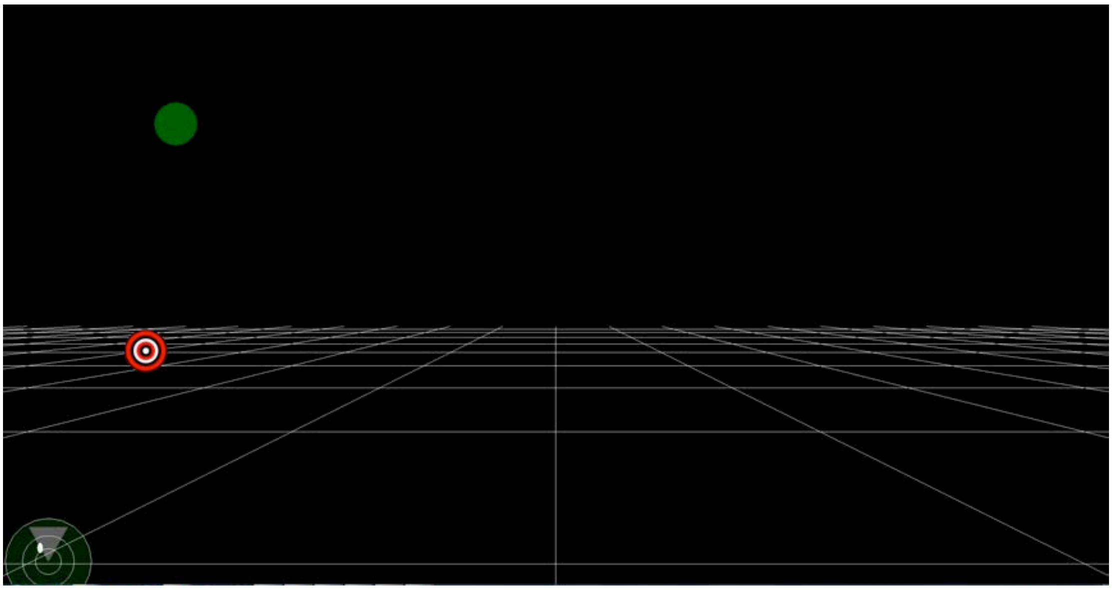


Figure 5.2: A screenshot from the users HMD view

Here, the green circle represents the position of the user's finger on the HHD that has been mapped to the HMD. The Bull's-eye represents the target object that the user must hit. The user would hit 20 of these to complete one iteration of the study.

5.4 Quantitative Results

The study had 12 participants, all of whom were university students between the ages of 19 to 32. The participants' knowledge of outdoor and mobile Augmented Reality varied from none to frequent usage. The participants' had a wide range of 3D motion interface experience varying from using none of the interfaces we had listed to those who had used all of the listed interfaces. All participants had prior experience of using touch screen smartphones and HHD.

The study was a within group study and the order of completing the tasks for each participant was counterbalanced using a Latin square approach. During each trial of the study, we measured time taken to complete the task and error while selecting the target. At the end of each trial, the participant's opinions were recorded using Likert scale questions (ranging from 1 to 9).

5.4.1 Measured Results

The quantitative measures for each phase in the experiment included error (in pixels) and the time taken to complete the task (in milliseconds). Table 5.1 summarizes the average measured error results for all the gestures, and the performance times. The maximum possible error was 50 pixels (the radius of the target).

Gesture Type	Time (in ms)		Error (in pixels)	
	Mean	Std Dev	Mean	Std Dev
Sticky Finger	7324	1084	17	4.23
Head Crusher	6752	1162	25	5.96
Tap Again	5515	1506	23	5.40
Motion	8373	2233	26	4.80

Table 5.1: Table showing mean performance time and error for different gestures

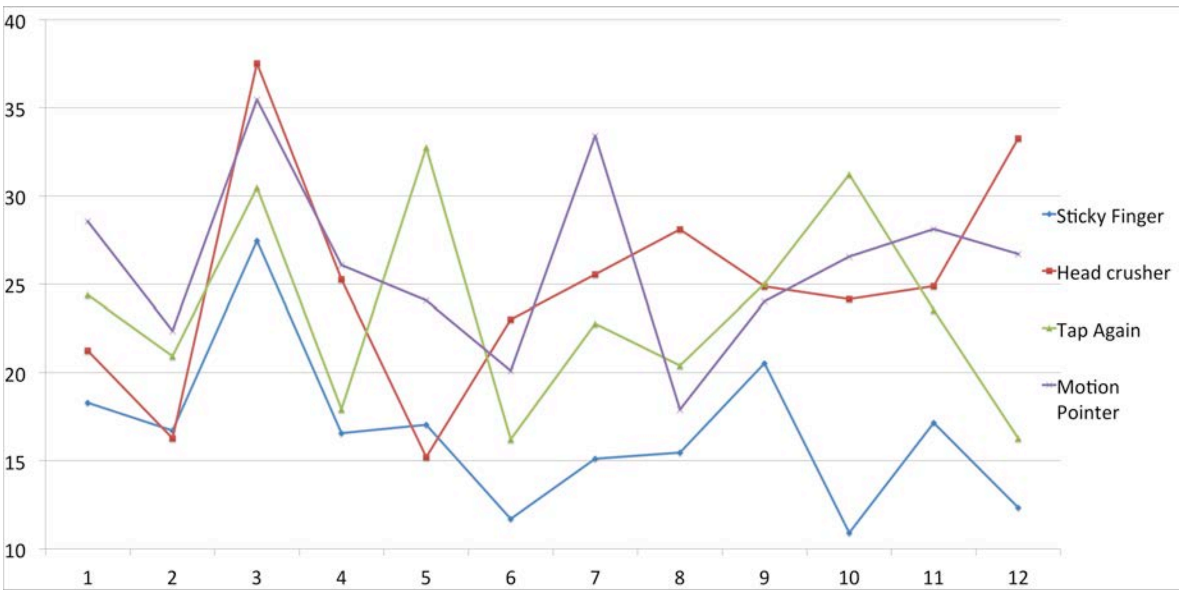


Figure 5.3: Graph showing mean error for the different gestures

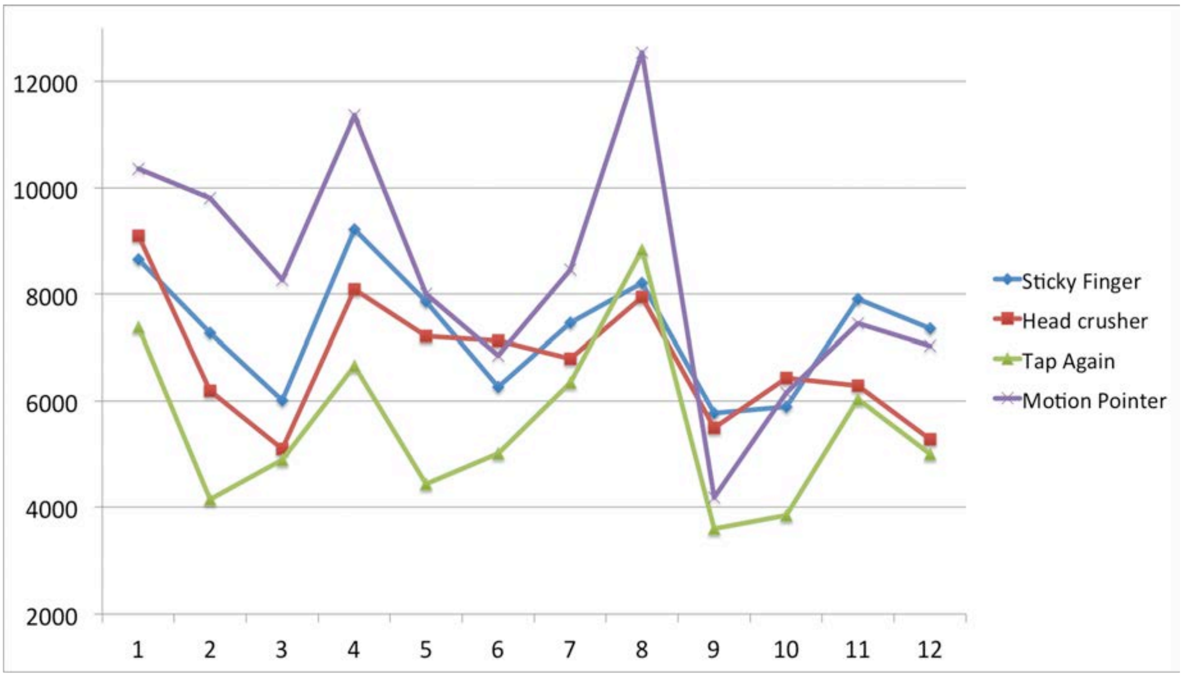


Figure 5.4: Graph showing mean time for the different gestures

5.4.2 Repeated Measures ANOVA Results

Since the errors in selecting the target and time taken to complete the tasks are not discrete but continuous values, a repeated measures ANOVA was conducted to determine whether there were any significant differences. The repeated measures ANOVA test requires that the variances of the differences between all combinations of related groups must be equal i.e. the *Sphericity* assumption. Our data violated the assumption of sphericity, so a Greenhouse-Geisser correction was applied. A post-hoc comparison using the Bonferroni correction was done to see if there were significant differences in the pairwise comparisons of the different gestures. The results of the repeated measures ANOVA are mentioned below:

Error in Target Selection

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that mean errors differed statistically significantly between the different gestures ($F(2.008,22.088) = 10.411$, $p = 0.001$). Post hoc tests using the Bonferroni correction revealed that there is a statistically significant difference between the Head Crusher and Sticky Finger (Z=-2.831, p=0.005), Tap Again and Sticky Finger (Z=-3.071, p=0.002) and Motion and Sticky Finger (Z=-3.063, p=0.002) gestures. Overall Sticky finger was the technique producing the least error. There were no other significant differences between the other conditions.

Time taken to complete the task

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that task completion time differed statistically significantly between the different gestures ($F(2.041,22.449) = 18.419$, p <0.0005). There was a statistically significant difference in perceived performance between the Motion and Tap Again gestures (Z=-3.059, p=.002) , Sticky and Tap Again (Z=-2.981, p=.003), and Head Crusher and Tap Again (Z=-2.667, p=.008). Overall Tap Again was the fastest selection technique. There are no other significant differences between the other conditions.

5.4.3 Questionnaire responses

The quantitative portion of the questionnaire consisted of 8 questions answered on a Likert scale ranging from 1 to 9. These questions recorded the participants' opinion of the gesture type and their opinions of using this gesture to interact with virtual content on the HMD. The mean values of the participant responses are shown in table 5.2. A chart showing the bar responses is also plotted below. Details about the individual questions can be found below or in Appendix A.

Gesture Type	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Sticky Finger	6.25	6.00	6.08	7.08	6.42	5.92	4.00	3.25
Head Crusher	6.00	5.25	5.58	6.50	5.33	4.58	4.42	3.25
Tap Again	7.25	7.25	7.25	7.67	7.17	6.25	3.50	2.75
Motion	5.00	3.92	3.75	5.50	4.25	3.92	5.25	4.50

Table 5.2: Mean Likert-scale responses (between 1 to 9) for the questionnaire

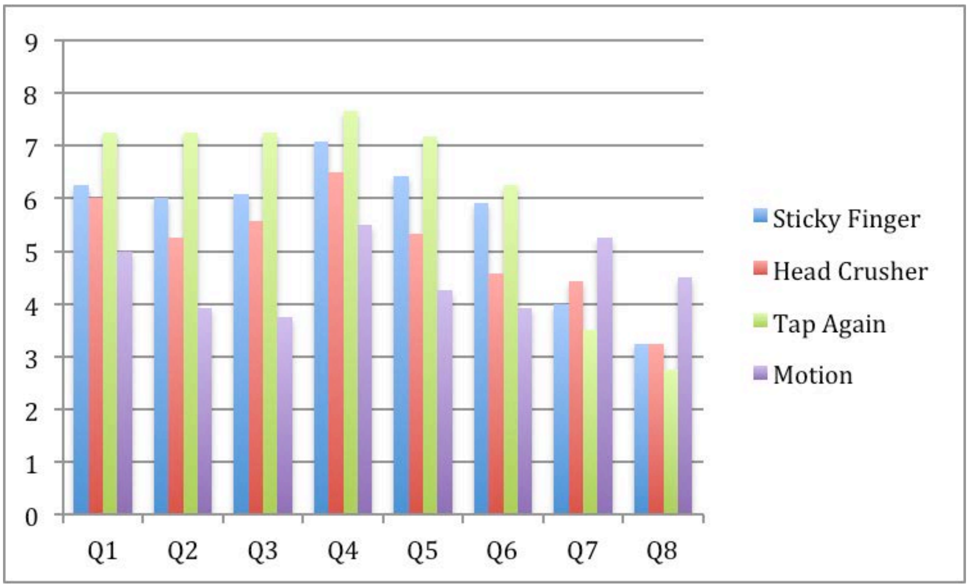


Figure 5.5: Graph showing mean likert-scale responses for the questionnaire

To determine whether there was a significant difference in the users' responses to questions regarding the different gestures, a oneway ANOVA (Friedman test) was conducted. Post Hoc analysis using Wilcoxon signed-rank tests was conducted to determine whether there was a significant difference between any 2 gestures. Since there are 4 gesture types, a Bonferroni correction was applied. The adjustment value used was :

(.05/(number of pairwise comparisons) =(.05/6) =0.0083 ,

So, p<0.0083 was used for the post hoc tests.

The results of the ANOVA test are mentioned below:

Q1: Did you feel you were performing well ?

There **was** a statistically significant difference in the user perceived performance between the 4 gestures. χ^2 (3)=11.882, p=.008. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at p < .0083 . There was a statistically significant difference in perceived performance between the Motion and Tap Again gestures (Z=-2.931, p=.003). Users felt that the Tap Again gesture enabled them to perform better than the Motion gesture. There were no statistically significant differences between the rest of the conditions.

Q2: Did you feel the gesture was useful for completing the task?

There **was** a statistically significant difference in the user perceived usefulness of the gesture in completing the tasks. χ^2 (3)=15.109, *p*=.002. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. There was a statistically significant difference in perceived performance between the Motion and Tap Again gestures (*Z*=-2.966, *p*=.003). Users felt that the Tap Again gesture was more useful for completing the selection task than the Motion gesture. There were no statistically significant differences between the rest of the pair of trials.

Q3: Was the gesture easy to use ?

There **was** a statistically significant difference in the user perceived ease of use between the 4 gestures. χ^2 (3)=15.108, *p*=.002. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. There was a statistically significant difference in perceived ease of use between the Motion and Tap Again gestures (*Z*=-2.834, *p*=.005). Users felt that the Tap Again gesture was easier to use than the Motion gesture. There were no statistically significant differences between the rest of the conditions.

Q4: Was the gesture Easy to Learn ?

There **was** a statistically significant difference in the user perceived ease of learning between the 4 gestures. χ^2 (3)=10.486,p=.015. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.0083$. There were no statistically significant differences between either pair of trials.

Q5: Was the gesture Intuitive ?

There **was** a statistically significant difference in the user perceived intuitiveness between the 4 gestures. χ^2 (3)=14.632, *p*=.002. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. There **was** a statistically significant difference between the perceived intuitiveness between the Motion and Tap Again gestures (Z=-2.756, *p*=.006). Users felt that the Tap Again gesture was much more intuitive than the Motion gesture. There were no statistically significant differences between the other conditions.

Q6: Was the gesture natural ?

There **was** a statistically significant difference in the user perceived natural feel between the 4 gestures. χ^2 (3)=8.297, *p*=.04. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. There were no statistically significant differences between either pair of trials.

Q7: Was the gesture mentally stressful ?

There **was** a statistically significant difference in the users perceived mental stress between the 4 gestures. χ^2 (3)=10.882, *p*=.012. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$ There were no statistically significant differences between either pair of trials.

Q8: Did you feel the gesture was physically stressful?

There **was** a statistically significant difference in the user perceived physical stress between the 4 gestures. χ^2 (3)= 13.373, *p*=.004. Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. There were no statistically significant differences between conditions.

5.5 Quantitative Analysis

5.5.1 Statistical Analysis of Measured Data

Statistically, the Sticky Finger gesture is the least error-prone among the four gesture types and there is a significant difference in the pairwise comparison of the sticky finger gesture with other gestures. With the Sticky Finger technique, the user holds their finger still for a specified period of time (1 second in our case) and has some time to adjust the selection cursor to get closer to the target. Holding the finger still and having this extra time to get closer to the center of the target might why this gesture is the most accurate. The HMD has a low resolution (640x480) and the target has a radius of only 50 pixels, which is why we did not expect that the results would be so conclusive and that a single gesture would be the best choice for error-free selection. Thus, the sticky finger seems to be a suitable gesture for precise selection of objects in a HMD-HHD hybrid system.

The Tap Again gesture was the fastest gesture among the four gestures. This was an expected outcome of the experiment. Although the tap again gesture is a faster way to select an object, it is more error-prone than the sticky finger gesture. This explains why the mean difference in error between the Tap Again and Sticky Finger gestures is so significant.

The Sticky Finger was the most accurate gesture but is quite slow and the Tap Again gesture was fast but is error prone. This prompted us to calculate the correlation coefficient to see if there was any correlation between the error while selecting a target and the time taken to complete the experiment.

The correlation coefficient between the time to complete gesture versus error is shown below:

Correlation Coefficient	Gesture Type			
	Sticky Finger	Head Crusher	Tap Again	Motion
χ	-0.101	-0.451	-0.357	-0.155

Although there is an overall negative correlation between time taken to complete the gesture and the error, the correlation is weak. However, in case of the Head Crusher gesture, there is a moderately negative linear correlation.

While comparing the time taken to complete the task for the different gestures, the standard deviation from mean for the Motion gesture is quite high (2331 ms) compared to its mean (8372 ms). There could be two explanations for such a high deviation:

1) Motion sensitivity: At any given value of the motion sensitivity, users will have varying levels of task completion. This is because every user prefers a different value of sensitivity while using this gesture. The Motion gesture is not reliable for a consistent performance across a wide-range of population.

2) Head Motion: Users mentioned that they had used a combination of head motion with the Motion gesture while selecting the target Those users who used their head motion more than the Motion gesture would have completed the task faster compared than those who used the Motion gesture more which would explain such a large deviation from the mean.

Statistically, the Motion gestures are the most error prone, have the highest task completion time and a large deviation from the mean task completion time. The above reasons make this gesture unsuitable for selection tasks in a HMD-HHD hybrid system.

5.5.2 Questionnaire Response Analysis

The quantitative portion of the questionnaire consisted of 8 questions on a Likert scale. There was a significant difference between the four gestures for **all** of the questions. This greatly strengthens our overall hypothesis.

Pairwise Comparisons: All the significant differences in the pairwise comparisons of the different gestures were among the Motion and Tap again gestures. There was a significant difference in the participants' perception of performing well while they were using the Motion gesture and Tap Again gesture. There was also a significant difference in the participant's perception of usefulness to complete the task while they were using the Motion and Tap Again gestures. Finally, there was a significant difference in the ease of use and intuitiveness between the Motion and Tap Again gestures. From these results, we can conclude that participants strongly favor the Tap Again gesture compared to the Motion gesture.

Among the touch gestures, 8 participants (66%) felt that Tap Again was the best method and two (16%) felt that Sticky Finger was better. One participant felt that both methods were equally good. Based on our results, Tap Again seems to be the most preferred gesture by the users.

Four participants (33%) explicitly mentioned that Head Crusher was not intuitive while 6 (50%) others cited difficulties with this gesture. While one user chose Head Crusher as his most preferred gesture, it seems that Head Crusher is not a suitable gesture for usage across a mainstream audience.

5.6 Qualitative Analysis

5.6.1 Gesture-wise comments

After completing the study, users were able to write their own comments about the gesture conditions. For the Tap Again gesture, people liked that the gesture was fast and there was a minimal delay in selecting the target compared to the Sticky Finger.

While the feedback for the Sticky Finger gesture has generally been positive, 60% of the participants felt the time they had to hold their finger still was too long. For our experiment, we chose one second as the duration, but if the user lifted their finger slightly or moved around a bit while it was down they had to hold still for another second that would potentially cause more delay.

The lesson learnt is that setting the right duration for the gesture is tricky. The duration of a long hold varies from user to user and choosing a universal duration is not feasible. Users should be allowed to choose their own duration.

For the Motion gesture case, one participant liked that they could see the cursor all the time compared to the other gestures, however 75% of the users had issues with the cursor. Some said it was slow or it was too hard to control. Similar to the Sticky Finger case, the rate at which the cursor moves with respect to the orientation of the phone is subjective. Setting the right sensitivity for the Motion gesture is tricky.

Some participants also commented on the up/down direction being swapped for the Motion gesture. We think that this is due to the way the Motion gesture is implemented: They had to tilt left or right to move the cursor which feels like the cursor is meant to be sliding down the surface on the smartphone device, which makes them think pointing down (tilting the upper side of the phone down) is to make the cursor go up. Investigating different types of motion gestures could be looked at in future research.

5.6.2 Overall Comments and Analysis

Three participants (25%) mentioned that head movement was a major factor in selecting the targets i.e. in some conditions, users would rotate their heads, trying to get as close to the target as possible and then use touch/motion gestures to get the cursor closer to the target. Some users have even mentioned that for the Motion Gestures, they preferred to move their head and used that for selection rather than using the smartphone. This is a case that we had *not* anticipated or taken into account. We expected that users would rotate their head and would stop when they saw the target in their view and then use the touch or motion gestures to select the target. Because of the increased rotations of your head, the chance for motion sickness increases.

One user remarked that Head Crusher was best for selection, saying it was ".. very relaxed and felt more precise". The reason given was that placing a finger on the screen was relaxing and sliding it across the mobile device felt comfortable. This comment is further strengthened because Head Crusher was the last gesture in his experiment.

This suggests that the ratio of the size of hand to the size of the device might be a factor that could determine the feasibility of using two-finger gestures on a handheld device. If the size of the hand was smaller compared to the screen, the two-finger gestures might feel comfortable. Using the Head Crusher gesture on a tablet or maybe larger devices might be something interesting to look at further.

For the head crusher technique, it is hard to select a target if it was in the corner of the screen. We could infer that the small screen of a smartphone is not suitable to perform gestures involving two fingers (or would depend upon the ratio of size of hand to device screen as mentioned above). Two finger gestures like pinch-to-zoom could work as they are intuitive but using more than one finger for selection might not be a good option.

One user mentioned that he felt dizzy during the experiment. Two users mentioned that sometimes the background (the wall in this case) was a hindrance while focusing their attention on the image plane. The background plays a role when focusing onto the screen and can be difficult if there is a stark contrast between the two views.

One user mentioned for Head crusher that he had to focus more because of the two fingers shown and the center point in between, while Sticky Finger required less focusing. This means that users would like to focus less while using such a system and less content (i.e. using one finger) should be shown on the screen to reduce the strain/effort while focusing.

5.6.3 Study Specific Comments

Three (25 %) of the users felt that the cable connecting the HMD to the PC actually got in their way while performing the tasks. They felt that that the cable restricted their head movement. These issues will not exist while using wireless HMDs like Google Glass.
An interesting comment by one user was that while Tap Again and Sticky Finger were intuitive, gaze selection (crosshair in the center of the screen can select object) could have been also included. This user remarked that he would have preferred gaze selection compared to Head Crusher and Motion gestures.

Some users had issues with placement of the radar. While some would have preferred to have it in the centre, some felt that the radar should have the objects rotating around (like in computer games). Two users had issues with focusing on the radar.

5.7 Threats to Validity

5.7.1 Possible Internal Threats

A few participants (4) mentioned that the wires connecting the HMD to the PC got in their way while turning their heads during the experiment. Since all participants had the same setup, there wasn't any bias towards a specific participant.

All users had prior experience with touch-based smartphones, which could lead to a possible bias towards touch gestures compared to motion gestures. However, as 80% (10) of the users had prior experience with motion pointing type devices (like the WiiMote), this bias can be eliminated.

One goal of this study was to find gestures that would be suitable for a HMD-HHD hybrid system for Outdoor AR applications. Since the study was conducted indoors, this may not mimic an actual outdoor real world setting. There is a possibility of having different results if the study was conducted outdoors. We anticipate difficulties like the participants not being able to see virtual elements clearly on the HMD, the phones screen not being very visible outside, difficulties in focusing etc. All these factors would affect the feasibility of using such a hybrid system outside, but would not affect our study regarding the differences among the various gestures.

5.7.2 External Threats

All participants in the study were males. It may be possible that gender could influence the results of this study.

The prototype system has a monocular HMD where users could view information on their right eye. Over 80% (10 out of 12) of the participants were right eye dominant which may bias the study since right eye dominant people might perform better than left eye dominant people with our HMD prototype. Additional studies can be conducted to confirm that selection tasks in a monocular HMD-HHD hybrid system are independent of the users dominant eye.

5.8 Conclusions

The primary goal of our evaluation was to determine which gesture performed on a smartphone would be best suited for selection tasks in a HMD-HHD hybrid system. To achieve this goal we developed a prototype that consisted of a monocular HMD and smartphone that supported four types of gestural input; three touch based and one motion based gesture.

The results showed significant differences in both the measured data and the users opinions of the gestures. In particular there were significant differences between the Motion and Tap Again gestures, proving that users prefer touch gestures to Motion Gestures. The study also showed that the choice of gesture depended upon the type of application. For applications requiring fast selection, the Tap Again gesture is most suitable whereas the Sticky Finger gesture is more suited for applications requiring accurate selection. Analysis of users' preferences showed that almost all users preferred the Tap Again gesture for selection, making it a suitable gesture for widespread usage over a large population.

The results are encouraging because they also show that accurate selection is possible using a particular gesture in spite of the low resolution of the display. User feedback such as number of fingers, displaying minimal content to avoid difficulties in focusing are factors that could affect the choice of the gesture and content shown on the HMD.

This study also revealed potential areas for further research. One such area was the possibility of how head movement and gestural input influence selection in a HMD-HHD hybrid system. While the average feedback for the Head Crusher gesture was not very positive, the strong positive feedback given by one participant showed that while Head

Crusher may not be suitable for widespread adoption on smartphones, the size of the device could affect the usability of this gesture. Exploring the relation between size of the smartphone/handheld and users hand size for this gesture might be a possible research area. Another possible research area is setting the right sensitivity for a gesture. While setting the correct sensitivity is difficult for widespread usage of a particular gesture, it does affect user’s performance during selection tasks - especially for gestures like Sticky Finger and Motion gestures. Allowing users to set the sensitivity for gestures like Sticky Finger and Motion gesture, and comparing their performance with our data would strengthen our conclusions.

5.9 Discussion

Based on the results of our user study, the following are a few guidelines to be kept in mind while choosing gestures for a hybrid HMD-HHD system.

1) Every additional virtual element shown on the HMD makes it more difficult for the user to focus on the HMD content. This includes virtual models, radars and finger cursors mapped from gestures performed on the HHD. While designing HMD-HHD systems, it is advisable to have gestures that use fewer fingers so that the user has to focus less and there is a reduced chance of dizziness.

2) While accurate selection using the Sticky Finger gesture is possible in a HMD-HHD system, user preferences indicate that speed is more preferred which is why tap again is the most preferred gesture by the users .Our user studies show that for general selection tasks using a HMD-HHD system, a gesture that is easy to learn and fast to execute is more preferred by users than a gesture which is more accurate.

3) Setting the sensitivity of a particular gesture is difficult and depends upon the users perception. This is why some users complained about the sensitivity of the Sticky Finger gesture. In general, gestures requiring some time to register are tricky as different people have different perception of the “right amount of sensitivity” for a particular gesture.

4) The choice of gesture for a particular task affects the user's head movement. Gestures that are difficult to perform on the HHD cause the user to use their head to get closer to the target which in turn cause dizziness. This should be kept in mind while designing gestures for a HMD-HHD system.

Chapter 6

Conclusion

6.1 Summary

Augmented Reality provides contextual information in real time that is often related to environment semantics. This gives AR a vast potential to be applied outdoors as AR can be used to discover new things about the world around us, provide interactive navigational experiences, educate us about our surroundings, visualize structures/things that we cannot normally see etc. AR Tracking is improving continuously as Computer Vision algorithms and GPS continue to get better and Head Mounted Displays are no longer bulky wearable devices but are now available in spectacle-sized form factors. Outdoor AR applications are also available on smartphones, but smartphones and HMDs present different advantages. While HMDs are good for hands-free viewing of a scene and augmenting the live view in front of your eye, smartphones are good for user input and viewing information that cannot be clearly shown on the HMD. This is why creating a hybrid system that utilizes the benefits of both a HMD and smartphone could be useful for AR applications.

The aim of this research was to develop a working prototype of this hybrid system and build several applications using this system that illustrated its usefulness. Prior research in hybrid platforms and gestures used in both smartphones and HMD revealed few applications and we used them as references. A prototype was assembled using off-the shelf components that resembled the target platform. A set of applications were built using this system such as using smartphones for gestural input, tribrid viewing of a scene and relative layouts for intuitive interaction within the system.

For the evaluation participants wore our prototype system and performed four different gestures on the smartphones to select virtual objects viewed on the HMD. The lessons learnt from our evaluation are mentioned in brief in Section 6.3.

6.2 Contributions

- Answers to the general research questions presented in section 3.
- A prototype system that combines a HMD and smartphone that receives input from the smartphone to interact with content viewed on a HMD.
- A set of 5 applications built using the above system.
- An evaluation of the prototype and a formal user study, which supports the research conclusions.
- A modular software architecture that could be reused with other projects
- A gesture set that can be referenced while designing applications for a smartphone-HMD hybrid system.
- A thesis detailing all aspects of the research.

6.3 Challenges and Lessons Learned

There were many lessons learnt during each phase of this research.

During the prototype implementation phase, we learnt that connecting a HMD directly to a smartphone is not feasible since external orientation sensors, cameras and displays are not well supported in smartphones. But apart from connecting external components, smartphones have fast processors, good built-in cameras and are capable of hosting Wi-Fi hotspots with speeds fast enough for external devices to connect and exchange messages with the smartphone. Smartphones are independent devices that have high computational power for a small form factor whose components keep getting better.

During the development of our tribrid viewing application, we realized that streaming images to a smartphone is not fast enough for real time viewing but virtual worlds can be easily shown on smartphones in real time. It is this reason why we feel that instead of streaming the camera image from the PC to smartphone, we could send the coordinates of the virtual world blocked by the blobs on the smartphone to the smartphone which would replicate the virtual world shown on the pc and display only that portion which was blocked by the camera. This method would probably make tribrid viewing in real time.

While the evaluation phase yielded satisfactory results and gave us pointers to future work, there were many lessons learnt about gestures used between a HMD-HHD hybrid system. Surprisingly, accurate selection of content is possible using gestures on the smartphone in spite of the low resolution of the Head mounted display. For gestures like sticky finger and motion gestures, setting the appropriate sensitivity is a hard task. The choice of a gesture for an application depends upon the purpose. For applications requiring a fast method of selection, the sticky finger gesture is best suited whereas the Tap again gesture is more suited for faster selection. However, user preferences suggest that users prefer the tap again gesture more than the sticky finger gesture, which means that tap again should be generally used for selecting content in applications unless the requirements demand higher selection accuracy.

6.4 Future work

This research has looked at the usefulness of a HMD-HHD hybrid system for AR applications. The research is substantiated by an evaluation of four gestures for selection. Additional gestures have also been explored that are shown through different applications. While the emphasis is on presenting different ways of interacting with a hybrid system, the applications that can be built using these interaction methods are manifold. The gestures proposed in this research take into account previous research on both gestures and hybrid systems and some new ones have also been proposed and demonstrated. Further work could be done at expanding this gesture set.

While the results from our user study are conclusive, additional user studies can be conducted to eliminate any potential threats to the conclusions. A study that examines the effect of users dominant eye in selection tasks using a monocular HMD-HHD system would clear any bias associated with eye dominance for selection.

One potential area for further research is the relation between Head Motion v/s type of gesture for selection tasks. Many participants reported that for some selection tasks, they preferred to use their head motion for getting closer to the target and selecting it, rather than the gesture itself. This could potentially give insights on whether the choice of a particular gesture influences head movement. Head motion must be reduced while using this hybrid system since increased head movement could lead to fatigue and motion sickness.

A study that examines user preferences for a gesture's sensitivity during selection tasks could be very useful. A modified version of the sticky finger gesture where users can set the sensitivity compared with the tap again gesture could be worth exploring since these two gestures were most preferred by the participants.

While the head crusher was not a preferred gesture as indicated by many participants, the strong positive preference by one participant evokes some thought. While this participant may be an outlier to the study, there could be a possible relation between hand size and smartphone screen specific to the head crusher gesture. Handheld devices come in varying sizes ranging from smartphones (3.5"-4.9") to tablets (5" to 10"). A study evaluating the head crusher gesture across handheld devices of different sizes might reveal certain devices or specific operations where the head crusher gesture might be useful.

While this research has looked at gestures for object selection, manipulation (translation, rotation and scaling) and cross-dimensional interaction, further studies can be conducted to determine which gestures are best suited for object manipulation using the HMD-HHD system. While combining multiple gestures had been briefly looked at during this research (eg: the Head Crusher for selection and Sticky finger for translation), combining different gestures for selection and object manipulation could be an area worth looking at. This could lead to a minimal set of gestures that could be used for a wide variety of tasks.

During the user study, two users complained about the background affecting their task and a few users mentioned that they preferred their head motion during selection tasks using the motion and head crusher gestures. A user's comment regarding the head crusher gesture suggested that showing two finger-cursors on the HMD made the user to concentrate more and felt harder to focus on the virtual content. Backgrounds that are difficult for the user to focus on the virtual content, gestures that decrease head movement and in turn reduce dizziness, automated methods of adjusting virtual content depending upon the background are some research aspects that could be looked at to enhance the overall user experience of using a HMD-HHD system. Finally, exploring the back of the HHD as a gestural surface could also be worth looking at. Since the user holds the HHD with one/two hands and at least one hand is always behind the HHD, gestures utilizing both the front and back of the HHD could offer interesting possibilities. For example, an alternate version of the cross dimensional gesture would involve selecting the object using a gesture on the front surface and a sliding gesture performed by a finger using the other hand at the back of the HHD.

This research provides an insight into coupling today's devices like smartphones with the future of head mounted computers like Google Glass. While finding a universal gesture for selecting objects in a hybrid system may not be possible, applications can make use of the findings of this research to see which gesture might be best suited for a particular application. Previous AR research has looked at using a HMD or smartphone for an intended AR application, combining the two devices yields significant advantages as shown in this research .The applications developed and results of this research could prove useful for designing and developing AR applications for HMD-HHD systems.

Bibliography

1. Ajanki, A., M. Billinghurst, H. Gamper, T. Järvenpää, M. Kandemir, S. Kaski, M. Koskela, M. Kurimo, J. Laaksonen, K. Puolamäki, et al.,2011: An augmented reality interface to contextual information. *Virtual reality*, 15(2), 161–173.

2. Azuma, R. T., 1997: A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385.

3. Baldauf, M., P. Fröhlich, and S. Hutter, 2010: Kibitzer: a wearable system for eye-gaze-based mobile urban exploration. In *Proceedings of the 1st Augmented Human International Conference*, AH ’10, ACM, New York, NY, USA, 9:1–9:5.

4. Bane, R. and T. Hollerer, 2004: Interactive tools for virtual x-ray vision in mobile augmented reality. In *Proceedings of the 3rd IEEE/ACM Inter- national Symposium on Mixed and Augmented Reality*, ISMAR ’04, IEEE Computer Society, Washington, DC, USA, 231–239.

5. Behzadan, A. H. and V. R. Kamat, 2005: Visualization of contruction graphics in outdoor augmented reality. In *Proceedings of the 37th con ference on Winter simulation*, WSC ’05, Winter Simulation Conference,1914–1920.

6. Benko, H., E. W. Ishak, and S. Feiner Collaborative mixed reality visualization of an archaeological excavation. In *Mixed and Augmented Reality,2004. ISMAR 2004. Third IEEE and ACM International Symposium on*,132-140.

7. Benko, H., E. W. Ishak, and S. Feiner Cross-dimensional gestural interaction techniques for hybrid immersive environments. In *Virtual Reality, 2005. Proceedings. VR 2005. IEEE*, 209-216.

8. Dahne, P. and J. Karigiannis, 2002: Archeoguide: system architecture of a mobile outdoor augmented reality system. In *Mixed and Augmented Reality, 2002. ISMAR 2002. Proceedings. International Symposium on*, 263 - 264.

9. Feiner, S., B. MacIntyre, T. Hollerer, and A. Webster A touring machine: prototyping 3d mobile augmented reality systems for exploring the ur- ban environment. In *Wearable Computers, 1997. Digest of Papers., First International Symposium on*, 74-81.

10. Henderson, S. J. and S. Feiner, 2009: Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality*, ISMAR ’09, IEEE Computer Society, Washington, DC, USA, 135–144.

11. Hollerer, T., S. Feiner, and J. Pavlik Situated documentaries: embedding multimedia presentations in the real world. In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*, 79-86.

12. Kawai, T., J. Hakkinen, T. Yamazoe, H. Saito, S. Kishi, and H. Morikawa,2010:Ergonomic evaluation of ubiquitous computing with monocular head-mounted display.

13. Kok-Lim, L., A. Ilie, G. Welch, and A. Lastra Combining head-mounted and projector-based displays for surgical training. In *Virtual Reality, 2003. Proceedings. IEEE*, 110-117.

14. Piekarski, W. and B. Thomas, 2002: Arquake: the outdoor augmented reality gaming system. *Commun. ACM*, **45**(1), 36–38.

15. Pierce, J. S., A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik, and M. R. Mine, 1997:Image plane interaction techniques in 3d immersive environments. 39-ff.

16. Reitmayr, G. and D. Schmalstieg Mobile collaborative augmented reality.In *Augmented Reality, 2001. Proceedings. IEEE and ACM International Symposium on*, 114-123.

17. Rekimoto, J., Y. Ayatsuka, and K. Hayashi Augment-able reality: situ- ated communication through physical and digital spaces. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, pgs 68-75.

18. Schall, G., D. Schmalstieg, and S. Junghanns Vidente-3d visualization of underground infrastructure using handheld augmented reality.

19. Sutherland, I. E., 1968: A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, AFIPS ’68 (Fall, part I), ACM, New York, NY, USA, 757–764.

20. Taejin, H. and W. Woontack Arwand: Phone-based 3d object manipulation in augmented reality environment. In *Ubiquitous Virtual Reality (ISUVR),2011 International Symposium on*, 44-47 Takada, D., T. Ogawa, K. Kiyokawa, and H. Takemura, 2009: A context- aware ar navigation system using wearable sensors. 793-801.

21. Wither, J., S. DiVerdi, and T. Hollerer Evaluating display types for AR selection and annotation. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, 95-98.

22. Wobbrock, J. O., M. R. Morris, and A. D. Wilson, 2009: User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’09, ACM, New York, NY, USA, 1083–1092.

23. Yohan, S. J., S. Julier, Y. Baillot, M. Lanzagorta, D. Brown, and L. Rosenblum, 2000: Bars: Battlefield augmented reality system. In *In NATO Symposium on Information Processing Techniques for Military Systems*, 9–11.

24. Strength and significance of Correlation Coefficients [www.bized.co.uk/timeweb/crunching /crunch_relate_expl.htm]

25. Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. 1992. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. *SIGGRAPH Computer Graph.* 26, 2 (July 1992), 203-210. DOI=10.1145/142920.134061 http://doi.acm.org/10.1145/142920.134061

26. Andrei State, Mark A. Livingston, William F. Garrett, Gentaro Hirota, Mary C. Whitton, Etta D. Pisano (MD) and Henry Fuchs. “Technologies for Augmented-Reality Systems: Realizing Ultrasound- Guided

Needle Biopsies." Proc. SIGGRAPH 96 (New Orleans, LA, August 4-9, 1996). In *Computer Graphics* Proceedings, Annual Conference Series, 1996, ACM SIGGRAPH, pp. 439-446.

27. Tuceryan, Mihran, Douglas S. Greer, Ross T. Whitaker, David Breen, Chris Crampton, Eric Rose, and Klaus H. Ahlers. Calibration Requirements and Procedures for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics 1*, 3 (September 1995), 255-273.

28. Maes, Pattie. Artificial Life Meets Entertainment: Lifelike Autonomous Agents. *CACM 38*, 11 (November 1995), 108-114.

29. Feiner, Steven, Blair MacIntyre, Marcus Haupt, and Eliot Solomon. Windows on the World: 2D Windows for 3D Augmented Reality. *Proceedings of UIST '93* (Atlanta, GA, 3-5 November 1993), 145-155.

30. Google Glass (http://www.google.com/glass/start/)

31. Strategy Analytics Report (http://blogs.strategyanalytics.com/WDS/post/2012/10/17/Worldwide-Smartphone-Population-Tops-1-Billion-in-Q3-2012.aspx)

32. ARFlashMob : (http://www.sndrv.nl/ARflashmob/)

33. Nokia City Lens (http://betalabs.nokia.com/trials/nokia-city-lens-for-windows-phone)

34. Processing (http://www.processing.org/)

35. Openframeworks (http://www.openframeworks.cc/)

Appendix A

The following questionnaire was given to the participants during the study.

Participant #___ Trial Condition Seq: _______ > _______ > ________ > _________

1. Age:

2. Gender: Male / Female

3. Which hand do you usually use to write?

☐ Right handed ☐ Left handed ☐ Comfortable with both

4. Have you used a **mobile device with touch screen** interface before?

☐ Not at all
☐ Rarely (couple of times a year)
☐ Seldom (couple of times a month)
☐ Often (couple of times a week)
☐ Everyday

5. Please check ALL of the types of **3D motion interface** you've used.

☐ Microsoft Kinect
☐ Nintendo Wii
☐ Sony PlayStation MOVE
☐ Virtual Reality Systems with 3D motion tracking

6. How often have you used **3D interfaces or 3D software**?

☐ Not at all
☐ Rarely (couple of times a year)
☐ Seldom (couple of times a month)
☐ Often (couple of times a week)
☐ Everyday

7. Have you used **Outdoor Augmented Reality Interfaces or apps** before?

☐ Not at all
☐ Rarely (couple of times a year)
☐ Seldom (couple of times a month)
☐ Often (couple of times a week)
☐ Everyday

Thank you!

Participant #____ Trial Condition: Sticky Finger

Per condition questionnaire

	Strongly Disagree				Neutral				Strongly Agree
I was performing well	1	2	3	4	5	6	7	8	9
The given interface was …									
Useful to complete the task	1	2	3	4	5	6	7	8	9
Easy to Use	1	2	3	4	5	6	7	8	9
Easy to Learn	1	2	3	4	5	6	7	8	9
Intuitive	1	2	3	4	5	6	7	8	9
Natural	1	2	3	4	5	6	7	8	9
Mentally stressful	1	2	3	4	5	6	7	8	9
Physically stressful	1	2	3	4	5	6	7	8	9

Please briefly explain what made you feel not easy, not intuitive, unnatural, or stressful, if you felt in that way.

Participant #_____ Trial Condition: Head crusher

Per condition questionnaire

	Strongly Disagree				Neutral				Strongly Agree
I was performing well	1	2	3	4	5	6	7	8	9
The given interface was …									
Useful to complete the task	1	2	3	4	5	6	7	8	9
Easy to Use	1	2	3	4	5	6	7	8	9
Easy to Learn	1	2	3	4	5	6	7	8	9
Intuitive	1	2	3	4	5	6	7	8	9
Natural	1	2	3	4	5	6	7	8	9
Mentally stressful	1	2	3	4	5	6	7	8	9
Physically stressful	1	2	3	4	5	6	7	8	9

Please briefly explain what made you feel not easy, not intuitive, unnatural, or stressful, if you felt in that way.

Participant #_____ Trial Condition: Tap again

Per condition questionnaire

	Strongly Disagree				Neutral				Strongly Agree
I was performing well	1	2	3	4	5	6	7	8	9
The given interface was …									
Useful to complete the task	1	2	3	4	5	6	7	8	9
Easy to Use	1	2	3	4	5	6	7	8	9
Easy to Learn	1	2	3	4	5	6	7	8	9
Intuitive	1	2	3	4	5	6	7	8	9
Natural	1	2	3	4	5	6	7	8	9
Mentally stressful	1	2	3	4	5	6	7	8	9
Physically stressful	1	2	3	4	5	6	7	8	9

Please briefly explain what made you feel not easy, not intuitive, unnatural, or stressful, if you felt in that way.

Participant #____ Trial Condition: Motion pointer

Per condition questionnaire

	Strongly Disagree				Neutral				Strongly Agree
I was performing well	1	2	3	4	5	6	7	8	9
The given interface was …									
Useful to complete the task	1	2	3	4	5	6	7	8	9
Easy to Use	1	2	3	4	5	6	7	8	9
Easy to Learn	1	2	3	4	5	6	7	8	9
Intuitive	1	2	3	4	5	6	7	8	9
Natural	1	2	3	4	5	6	7	8	9
Mentally stressful	1	2	3	4	5	6	7	8	9
Physically stressful	1	2	3	4	5	6	7	8	9

Please briefly explain what made you feel not easy, not intuitive, unnatural, or stressful, if you felt in that way.

Participant #____

Post experiment questionnaire

1. What are the strength and weakness of each interaction method?

	Strength	Weakness
Sticky finger		
Head crusher		
Tap again		
Motion pointer		

2. Which one do you prefer to use if you will have to do a similar task again?

- ☐ Sticky finger
- ☐ Head crusher
- ☐ Tap again
- ☐ Motion pointer

3. Please briefly explain why you chose the interaction method above in question #2.

4. Did you have any problem during the experiment?

5. Any other comments on the interface or the experiment?

Thank you very much!