

TANGIBLE USER INTERFACES FOR AUGMENTED REALITY

QIU YAN

(B.Eng.(Hons), Xi'an Jiaotong University)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2003

Acknowledgement

I wish to express my sincere appreciation to Dr. Adrian David Cheok, for all the times that he has guided me, the times that he has pushed me, and the times that he has encouraged me.

I also wish to thank all the research members of Human Interface Technology (HIT) Lab. It was a great experience to work with them. Especially thank Dr. Simon J.D. Prince for his invaluable guidance and K.G. Kumar for his help to develop the hardware.

Above all, I would like to thank God for His everlasting love and tender care through my life.

Contents

Abstract	v
List of Figures	vii
1 Introduction	1
1.1 Augmented Reality	1
1.1.1 Augmented Reality and HCI	1
1.1.2 Augmented Reality Application Domains	4
1.2 Tangible User Interfaces	6
1.3 Wearable Computers	7
1.4 Contribution of this Thesis Work	12
1.5 Organization	13
2 Background and Literature Review	15
2.1 Registration in AR	15
2.1.1 Sensor-based Tracking Methods	15
2.1.2 Vision-based Tracking Methods	16
2.1.3 ARToolKit	19
2.2 Literature Review	21
2.2.1 Sensing	21

2.2.2	Tangible Interfaces in AR	23
3	Non-hardware Tangible Interface	27
3.1	3D Modelling	28
3.2	Implementing Tangible Control	32
4	Interface Enhancement – 3D sound	36
4.1	System Setup	38
4.2	Experimental Goals and Subjects	40
4.3	User Interface	42
4.4	System Assessment Results	48
4.4.1	Performance	48
4.4.2	Questionnaire Responses	49
4.5	Conclusions	55
5	Tilt Pad Based Tangible Interfaces Design	57
5.1	The Design Approach	58
5.2	Tilt Pad–Accelerometer Based Interaction Device	59
5.2.1	Components Description	59
5.2.2	Hardware Description	62
5.3	Tangible AR Map Scrolling System	67
5.3.1	Set Up 3D Virtual Map	67
5.3.2	Visualization in AR Environment	83
5.3.3	Interface with Tilt Pad	86
5.4	Tangible AR Pacman Gaming System	91
6	System Results and Assessment	95
6.1	Results of Map Scrolling System	95

6.2	Results of Pacman Game System	101
6.3	Description of the Results	105
6.4	User Study of Tilt Pad Pacman Game System	106
6.4.1	Aim of this User study	106
6.4.2	Design and Procedures	106
6.4.3	Results of this User Study	109
7	Summary	115
7.1	Our Research Summary	115
7.2	System Limitation and Future Work	117
A	Contents of the CD-ROM	123
	Bibliography	125

Abstract

Augmented Reality (AR) that combines three-dimensional virtual information and real world provides a large space for tangible interface applications. The contribution of this thesis work is twofold: on one hand, it demonstrates a kind of totally novel, tangible and natural interaction method for the HCI research; on the other hand it provides new applications in the areas of Augmented Reality gaming design and military usage. In this project, the main research objectives that have been achieved can be categorized into four parts: A non-hardware tangible interaction AR system: magic book; Some research on 3D sound in AR as an enhancement of the interfaces; A tangible 3D augmented reality map scrolling interaction system with hardware interface tilt pad; And a multimedia augmented reality gaming system interacted with tilt pad. The two hardware based AR interaction systems are distinct from other AR interfaces. Small accelerometer based, low cost hardware tilt pad is adopted as interacting device with PC. It makes the interfaces more natural, physical and independent of keyboards and mice, in the meantime reduces the computational load. This tilt pad uses wireless transmission, therefore, the two systems realize wireless and mobile interaction, which would be potentially able to explore applications for outdoor AR systems or wearable computers. The way using tilt pad to scroll a 3D map system suggests new ideas for military or education purpose as an application of augmented reality. The AR version of the old popular game Pacman is developed with three dimensional characters including Pacman, ghosts, cherries and the maze. Unlike traditional games using keyboard or game stick, here tilt pad is used to control the moving of the Pacman as a tangible and natural interface. And this AR Pacman game is a multimedia entertainment sys-

tem with nice sounds embedded. The demo of this Pacman game has been shown in NetGames 2003 in May held in California.

Key Words:Augmented Reality, Tangible User Interfaces, Tilt Pad, Map Scrolling, Pacman Game, User Study

List of Figures

1.1	Simulated augmented reality medical image	2
1.2	Six degrees of freedom diagram	3
1.3	A wearable computer system	9
2.1	Examples of paper markers.	17
2.2	Input Video(left frame), Thresholded Video(middle frame), Virtual Overlay(right frame)	20
2.3	Image processing in ARToolkit	20
2.4	A Glove application	22
2.5	Augmented groove	23
2.6	Mah-Jongg game implemented in this AR gaming space	25
2.7	A game of AR Chess interfaced with pad and pen	26
3.1	Rough procedure of modeling from images	29
3.2	Different angles of 3D model attached to the marker viewed from head-mounted display(HMD)	32
3.3	Tangible Interaction: Picking and Dropping virtual objects by ma- nipulating physical cards.	33
3.4	A more sophisticated tangible interaction example: virtual Hello Kitty garden.	34

4.1	The physical set-up of our acoustic-inertial tracking AR environment. On the ceiling there are six IS-900 SoniStrips which makes up a three meter by three meter area. A head tracker is mounted on the user’s HMD and a camera is fixed on the head tracker. A wand tracker with four buttons and one joystick was held in the user’s hand as an interaction device.	39
4.2	Looking for the castle through a “magic 3D window”. A virtual plane is augmented on the wand as an avatar for the user. A rectangle window is augmented in front of the wand through which the user can see the virtual land on the ground. When the user is moving around, the 3D landscape is also scrolled in the “magic 3D window”. The user is looking down at the castle as if she is looking down from the sky in the airplane.	42
4.3	Two users are collaboratively fighting the witch hovering around them in the physical world. The user can see the virtual witch which is augmented in three dimensions through HMD. The two users need to reload or transfer their “bullets” by putting their wands together as shown in the first two frames. When the user pressing the button on the wand, the airplane can fire bullets with different color (blue and green in the subsequent frames). The witch is designed to hover in the experimental region. The position and orientation of the wand are used to decide trajectory of “bullets” and whether the witch is shot or not. In the last frame, the witch was shot for ten times and disappeared.	44

4.4	Rescue the princess in the castle with co-player. The first row shows the seamless transition from AR to VR by which the user is “flying” into and “landing” on the VR world. The second row shows one player’s view when they are on the way to rescue the princess in the castle. Here the user uses the joystick to navigate however they can turn the body and head of the virtual character by physically turning their body and head. Meanwhile, the footstep sound comes from the co-player can always be heard. While navigating in the castle, the players are guided by the “Help!” sound of the princess. The last two frames in second row shows the last exciting moment when the player finds the princess and the princess shows her appreciation. In the first frame of the third row, a user can physically look up in the sky to find her co-player who is still in AR mode (before flying into the VR mode). Meanwhile, the last two frames show that the user can still see the co-player and the princess when she is in AR mode by looking down to the land as if she is on the airplane. The second frame shows one player’s view point when both of the players are in VR mode and come to the princess.	46
4.5	The comparison of searching task completion time in the cases of with and without 3D sound aided. The average of completion time is shortened drastically from 133.05 to 88.1 seconds by introducing 3D sound.	49

4.6	(a)The votes of the answers to question 4: Were you aware of your co-players in different places? We can see from the chart that in this AR environment, subjects can often be aware of co-player’s position, however, this result is impossible if there was no 3D sound applied since the subject can only occasionally see co-player due to the narrow angle of view. (b)The votes of the answers to question 5: What has helped your feeling of “human presence and collaboration”? It can be seen from this chart that 3D sound plays the most important role among all the cues. It gets even more votes than vision cues and voice communication cues.	54
5.1	The tilt measurement in x and y axes	63
5.2	Filtered tilt measurement in x and y axes	64
5.3	Schematic diagram of the transmitter part	65
5.4	Schematic diagram of the Receiver part	66
5.5	The tilt pad (10x5x2.5cm)	67
5.6	The red colored segment is determined by the four red colored control points, the green colored segment is determined by the four green colored control points	69
5.7	Process of NURBS rendering in OpenGL	74
5.8	Code flowchart of drawing map surface in OpenGL	75
5.9	Flowchart of texture mapping of the 3D virtual map	77
5.10	Diagram of changing control point values when scrolling	79
5.11	Flowchart of the code to realize map scrolling	80
5.12	Orientation of the tank on the map surface	81
5.13	Flowchart of the code to define the tank	82

5.14	Tracking of multiple fiducials. (a) The four different fiducials are printed on the same paper. (b) A virtual grass is augmented onto the fiducials. (c) Even some of the fiducials are blocked, the correct registration still remains.	85
5.15	Flowchart of the image processing of ARToolKit	87
5.16	(a) Software flowchart on the tilt pad side (b) Software flowchart on the host computer side	89
5.17	Integrated code flowchart of the map scrolling system	90
5.18	Code flowchart of AR Pacman system. This is a loop that executed once every frame.	94
6.1	Different zooming of a 2D map. The user can draw the map close to the camera to view details as shown in the last image.	96
6.2	The top photo is part of the 3D map. The bottom two photos show the tank with different positions and different orientations on the map	97
6.3	Left image is the multiple marker used, right image is the tilt pad fixed on the back side of the marker.	98
6.4	Tilt pad is used for a 2D scrolling map application	99
6.5	Scrolling of the 3D map. The red flag(not part of the map)is added to sign the same place tilted to different positions using tilt pad in four images.	100
6.6	(a) initial game environment with Pacman(the yellow one), ghosts(purple ones) and cherries(in red) (b) user tilts Pacman left (c) user tilts Pacman right (d) user tilts Pacman outward to her (e) user tilts Pacman inward to her	102
6.7	The paper with natural features for tracking	103

6.8	The user is playing the 3D Pacman game. (a) The user is holding the board, on which the colorful picture is pasted. The 3D Pacman maze is now augmented onto the board. (b) The user tilts the board forwards, the Pacman goes forward, to eat the berries and avoid the ghosts. (c) The user tilts the board backwards. (d) The user tilts the board leftwards. (e) The user tilts the board rightwards.	104
6.9	Answers to Q1 – Q8 of the user study. Top left chart is concerning Q1, top right chart is concerning Q2, and continue based on a left-right, up-down sequence	112
6.10	Left: The comparison between the scores of the traditional Pacman game method and the tilt pad Pacman game method. Right: Average score of the two methods	114

Chapter 1

Introduction

1.1 Augmented Reality

1.1.1 Augmented Reality and HCI

Augmented Reality (AR) is a variation of Virtual Reality (VR). The world environment around us provides a wealth of information that is difficult to duplicate in a computer. This is evidenced by the worlds used in virtual environments. Either these worlds are very simplistic such as the environments created for immersive entertainment and games, or the system that can create a more realistic environment has a million dollar price tag such as flight simulators. In contrast to virtual reality, which puts users in a fully computer-generated virtual world, an augmented reality system generates a composite view for the user. It is a combination of the real scene viewed by the user and a virtual scene generated by the computer that augments the scene with additional information. In all applications the augmented reality presented to the user enhances that person's performance in and perception of the world. Figure 1.1 shows a view that the user might see from an augmented reality

system in the medical domain. It depicts the merging and correct registration of data from a pre-operative imaging study onto the patient's head.



Figure 1.1: Simulated augmented reality medical image

Augmented reality technology allows users seamless interaction between the real and virtual world. AR is an interaction paradigm designed to reduce computer interaction using the environment information as an implicit input. Backman [1] describes AR as a combination of the real scene viewed and a virtual scene generated by the computer and a paradigm that augments human perception with supplying information not ordinarily detectable by human senses.

The user's situation is automatically recognized by means of a wide range of methods like time, position, orientation or the object inspection using the computer vision. Specifically, when we talk about orientation and position, the inputs become the six degrees of freedom. The first three directions are the three coordinates needed to determine the position (X, Y and Z). The other three indicate the orientation of the viewer and are known as pitch, roll and yaw as shown in figure 1.2. In a simulation model several interface devices can be used to navigate

and interact with a virtual ambient with six degrees of freedom, depending on the device and the virtual world design [2].

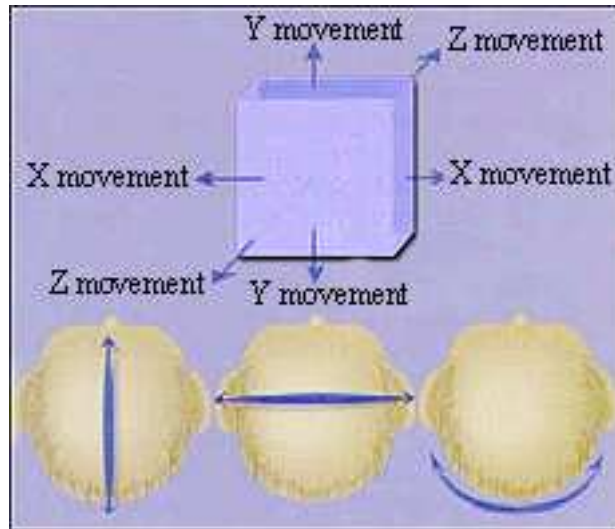


Figure 1.2: Six degrees of freedom diagram

In this way, AR enables the user to remain in contact with his workspace; his attention focus is set on the real world rather than the computer, which allows users to exploit their own visual and spatial abilities. AR provides a new research direction for Human-computer Interaction (HCI) [3]. Human-computer Interaction is affected by the demand of novel interface shaping the nature of future computing. Mobile computing and Ubiquitous communication are required for system developed. Natural and tangible interfaces are expected to make the users more immersive and satisfied. As a research on Human-computer Interaction technology, tangible interfaces based on small hardware are designed. These include a three dimensional map scrolling system and an Augmented Reality Pacman game. By “HCI technology”, I am referring to the computer side of HCI [4].

1.1.2 Augmented Reality Application Domains

Augmented reality systems have been proposed as solutions in many domains. The areas range from medical to engineering design. However, we are not going to cover all of them. Just entertainment and military fields are discussed as they are related to this research work.

Entertainment

A simple form of augmented reality has been in use in the entertainment and news business for quite some time. Whenever you are watching the evening weather report the weather reporter is shown standing in front of changing weather maps. In the studio the reporter is actually standing in front of a blue or green screen. This real image is augmented with computer generated maps using a technique called chroma-keying. It is also possible to create a virtual studio environment so that the actors can appear to be positioned in a studio with computer generated decorating. Examples of using this technique can be found at [5][6].

Movie special effects make use of digital compositing to create illusions [7]. Strictly speaking with current technology this may not be considered augmented reality because it is not generated in real-time. Most special effects are created off-line, frame by frame with a substantial amount of user interaction and computer graphics system rendering. But some work is progressing in computer analysis of the live action images to determine the camera parameters and use this to drive the generation of the virtual graphics objects to be merged [8].

Recently augmented reality gaming is growing up to a key area of entertainment. Due to the nature of AR that interacts with both the real and virtual 3D worlds, players are able to find enhanced game playing experience and get well entertained.

Bruce Thomas built an outdoor AR version of the Quake, a first-person shoot'em up game with the user interface based around a single, first-person perspective screen [9]. In this application [10], both the AR and VR version of the original Quake game were developed. We can see how AR and VR games differ from each other. In the augmented reality gaming mode, the physical world is modelled as a Quake 3D graphical model. The AR information (monsters, weapons, objects of interest) is displayed in spatial context with the physical world. By contrast, in the outdoor virtual reality gaming mode, the physical world is not modelled. The game may be played on any open flat piece of ground.

The tilt pad Pacman game system we developed is also an augmented reality version of the old popular game Pacman. The detail of this system is explained in chapter 5.

Military

The military has been using displays in cockpits that present information to the pilot on the windshield of the cockpit or the visor of their flight helmet. This is a form of augmented reality display. SIMNET, a distributed war games simulation system, is also embracing augmented reality technology. By equipping military personnel with helmet mounted visor displays or a special purpose rangefinder [11] the activities of other units participating in the exercise can be imaged. While looking at the horizon, for example, the display equipped soldier could see a helicopter rising above the tree line [12]. This helicopter could be being flown in simulation by another participant. In wartime, the display of the real battlefield scene could be augmented with annotation information or highlighting to emphasize hidden enemy units.

In this research, A 3D augmented reality map scrolled with tilt pad has its

value in military field. Equipped with this kind of map system, we can imagine soldiers are easy to know the panorama of the particular area. When combined with wearable computer [13], this map system can get more flexible applications due to its natural and mobile interaction manner. We will introduce wearable computer more in detail in section 1.3.

1.2 Tangible User Interfaces

Generally tangible interfaces are systems relating to the use of physical artifacts as representations and controls for digital information. A central characteristic of tangible interfaces is the seamless integration of representation and control, with physical objects being both representation of information and as physical controls for directly manipulating their underlying associations. Input and Output devices fall together.

For more than thirty years, people have relied primarily on screen-based text and graphics to interact with computers. Whether the screen is placed on a desk, held in one's hand, worn on one's head, or embedded in the physical environment, the screen has cultivated a predominantly visual paradigm of human-computer interaction. The last decade has seen a wave of new research into ways to link the physical and digital worlds. This work has led to the identification of several major research themes, including augmented reality, mixed reality, ubiquitous computing, and wearable computing. At the same time, a number of interfaces have begun to explore the relationship between physical representation and digital information, highlighting kinds of interaction that are not readily described by the existing frameworks [14].

Fitzmaurice, Buxton, and Ishii [15] took an important step towards describing

a new conceptual framework with their discussion of “graspable user interfaces”. However, a more common term is “tangible user interfaces” in [16], a growing space of interfaces in which physical objects play a central role as both physical representations and controls for digital information. Let me elaborate this concept further by the prototypical example of abacus. In particular, it is key to note that when viewed from the perspective of human-computer interaction, the abacus is not an input device. The abacus makes no distinction between “input” and “output”. Instead, the abacus beads, rods, and frame serve as manipulable physical representations of numerical values and operations. Simultaneously, these component artifacts also serve as physical controls for directly manipulating their underlying associations.

This seamless integration of representation and control differs markedly from the mainstream graphical user interface (GUI) approaches. Graphical interfaces make a fundamental distinction between “input devices”, such as the keyboard and mouse, as controls; and graphical “output devices” like monitors and head-mounted displays, for the synthesis of visual representations. Tangible interfaces, in the tradition of the abacus, explore the conceptual space opened by the elimination of this distinction.

1.3 Wearable Computers

The locus of computation is now shifting from the desktop in two major directions: i) into the physical environments we inhabit, and ii) onto our skins/bodies. Both are aimed to achieve a heightened legibility and seamlessness of interaction between people and information. Integration of computational augmentations into the physical environment forms the research field of tangible interaction. The tran-

sition to our bodies is represented by recent activities in the new field of “wearable computers” [17]. Commonly augmented reality is implemented on fixed desktop computers. However, it is now an ongoing research problem amongst researchers around the world to implement augmented reality systems onto wearable computers [18]. This will allow augmented reality to follow the user throughout his environment, both indoor and outdoor, and will allow exciting new forms of ubiquitous computing applications [19]. A typical wearable computer is shown in Figure 1.3¹.

¹This computer was developed in the Mixed Reality (MR) lab in National University of Singapore.



Figure 1.3: A wearable computer system

Wearable computers in a loose sense are computers people can wear effortlessly. In a stricter sense [20], wearable computers should run continuously and be operated hands-free [21]. This stricter definition helps distinguish wearable computers from portable ones. Many small laptops are not wearable computers. They do not last long enough to meet the continuous run requirement. Wearable computers can run much longer because of carefully selected power-conservative components and efficient batteries. Wearable computer allows a much closer association between the computer and the user. Wearable computing combines the ideas from mobile computing and ubiquitous computing. In addition it offers an advanced user-interface technologies. People are no longer confined to their desks with the desktop or the laptop. The main characteristics of wearable computers are summarized as following:

- *portable while operational*

The most distinct difference between desktop and laptop computers and wearable computers is that the latter can be used while the user is walking or moving around. The user have uninterrupted access to applications provided by the wearable computer while moving around freely and not constrained to the worktable.

- *hands free use*

The user can use both hands or at least one of his hands for other tasks while using the computer.

- *sensors*

In addition to user inputs, sensors are used for the computer to gather information about its environment. Such sensors include cameras for visual

information and microphones for audio information. The wearable computer should be able to perceive and act upon certain aspect of its physical environment.

- *augmented reality*

Most wearable computers have the augmented reality interface since this technology makes the physical environmental information available in its applications.

- *proactive*

A wearable computer is able to convey information to the user even when it is not actively in use. The computer is always running, working and sensing.

We believe a wearable computer with an augmented reality user interface allows for exciting new applications to be deployed both in an indoor and outdoor environment. The operation of wearable computers in an outdoor setting is hampered by the lack of suitable input devices. Many traditional input devices such as mice and keyboards are not suitable for mobile work, as they require a level flat surface to operate. A second difficulty is the well-known registration problem. The field of virtual reality also suffers from the lack of proper input devices and sub-optimal tracking systems, and as a result, new input devices, interfaces, and trackers are continuing to be developed in an attempt to solve these problems. However, many of these devices require fixed infrastructure and are not useable in mobile environments.

The problem of registering virtual images with the user's view of the physical world is a main focus of AR research. However, there is little previous work in the area of user interfaces for controlling AR systems in an mobile setting. Two

major issues for the development of these user interfaces are as follows: firstly, registration errors will make it difficult for a user to point at or select small details in the augmentation and secondly, pointing and selecting at a distance are known problems in virtual and augmented reality applications (compounded by the fact the user is mobile with less than optimal six degree of freedom tracking of their head and hands).

Therefore, new user interaction techniques are required for wearable computers in augmented reality, and to state the obvious, the input techniques the users are required to use will have a large impact on the usability of the applications. A key element to the new user interactions is that the augmented reality systems have a varying number of coordinate systems (physical world, augmented world, body relative and screen relative) within which the user must work. In an wearable computing application the registration errors of objects at a distance amplify the differences between the physical and augmented world coordinate systems.

1.4 Contribution of this Thesis Work

The contribution of this thesis work is twofold: on one hand, it demonstrates a kind of totally novel, tangible and natural interaction method for the HCI research; on the other hand it provides new applications in the areas of Augmented Reality gaming design and military usage. The novel Human-Computer Interface we introduce here uses physical object such as a pad to control digital information tangibly and eliminates the rely on keyboards and mice of traditional interfaces.

In this project, the main research objectives that have been achieved can be categorized into the following four parts:

1. A non-hardware tangible interaction AR system: magic book.

2. Some research on 3D sound in AR as an enhancement of the interfaces.
3. A tangible 3D augmented reality map scrolling interaction system with hardware interface tilt pad.
4. A multimedia augmented reality gaming system interacted with tilt pad.

The two hardware based AR interaction systems are distinct from other AR interfaces in the following aspects:

1. Small accelerometer based, low cost hardware tilt pad is adopted as interacting device with PC. It makes the interfaces more natural, physical and independent of keyboards and mice, in the meantime reduces the computational load.
2. Tilt pad uses wireless transmission, therefore, the two systems realize wireless and mobile interaction, which would be potentially able to explore applications for outdoor AR systems or wearable computers.
3. The way using tilt pad to scroll a 3D map system suggests new ideas for military or education purpose as an application of augmented reality.
4. The AR Pacman game is a multimedia entertainment system with nice 3D graphic characters and sounds.

1.5 Organization

In chapter 2, some background knowledge related to this thesis work is introduced. This focuses on the registration problem in AR, which is a very important issue. The previous works in the fields of sensing and tangible interfaces in AR are also

reviewed in chapter 2. In chapter 3, a tangible interaction system *magic book* is presented. And an experimental research on 3D sound in AR was carried out in order to enhance our interface design, which is described in chapter 4. Our main research on hardware based tangible interface design including a map scrolling system and an AR Pacman gaming system is introduced in chapter 5. This chapter considers mainly the technical issue. And the results as well as the descriptions of these two systems can be found in chapter 6. Eventually, a summary of this research work is in chapter 7. An appendix is attached to list the contents in the accompanying CD.

Chapter 2

Background and Literature Review

2.1 Registration in AR

One of the most basic problems currently limiting Augmented Reality applications is the registration problem. The objects in the real and virtual worlds must be properly aligned with respect to each other, or the illusion that the two worlds coexist will be compromised. Without accurate registration, Augmented Reality will not be accepted in many applications [22].

There are basically two different approaches to achieve accurate registration and positioning of virtual objects in the real environment: the sensing by long-range hardware trackers, or using various vision-based methods.

2.1.1 Sensor-based Tracking Methods

Sensors, based on magnetic, mechanical, ultrasonic, or optical technologies can be used to track the camera pose as the user moves in the real scene. [23] provides

a comprehensive discussion of the operating principles of the various technologies. Sensor-based tracking systems suffer from some major disadvantages that limit their usefulness for AR applications:

1. These systems are typically very expensive.
2. Extensive infrastructure is required to support the tracking, thus restricting the work area in which an AR system can be deployed.
3. The environment has to be carefully controlled as the sensors are easily affected by perturbations or noise e.g. magnetic sensors are susceptible to electromagnetic interference.

2.1.2 Vision-based Tracking Methods

In recent years, vision-based methods that extract camera pose information from features in the 2D images of the real scene have become increasingly popular for two main reasons:

1. It is convenient and cheap since the 2D images are readily available for a video-based AR system and additional sensors are not required.
2. The camera pose estimates are generally more accurate than those obtained from sensors as the measurement errors are relative to the visually perceived image space units (pixels), not the world space units (meters, inches etc.). Many systems have demonstrated nearly perfect registration, accurate to within a pixel [24] - [28].

The basic principles behind these methods are based on the results and theories developed in computer vision and photogrammetry research. Most previous work

on the AR registration problem using vision-based techniques can be broadly divided into two categories: the Fiducial Marker Tracking Approach, and the Natural Feature Tracking Approach.

Fiducial-based Tracking

In this type of registration algorithms, fiducials such as paper markers are placed in the scene where virtual objects are to be introduced (Figure 2.1).



Figure 2.1: Examples of paper markers.

These paper markers have some nice properties such as known shapes and colors which make them easy to detect and identify in the images. In [27], solid-color circle and triangle stickers were used while the AR system in [29] worked with multi-colored concentric ring markers. The centroids of these markers are the features that are tracked in the video stream and at least three markers have to be detected in the image frame before the camera pose can be computed.

The 3D world coordinates of the marker features are measured *a priori* and given the 2D coordinates of the detected features in the images, a correspondence between 3D and 2D is set up. Pose estimation techniques [30][31] can then be used to estimate the camera pose. These markers are inexpensive to produce and the methods are simple and can be implemented in real-time using normal desktop

computers. However, camera tracking can be easily lost as it is only based on a few features and there is a limited range of camera viewpoints from which the fiducials are visible.

Natural Feature Tracking

Although the fiducial-based tracking methods can achieve up to sub-pixel accuracy while running realtime, these routines all assume that one or more fiducials are visible at all times; without them, the registration can fall apart.

The more challenging job is to perform camera pose tracking in unprepared environments i.e. no modifications to the environment such as placing fiducial markers or sensors are made. Camera pose measurements are obtained based on naturally occurring features such as corner points and edges in the real scene with *a priori* unknown 3D positions. By using natural features, the tracking range and stability are typically greater than fiducial-based tracking systems since there are more features available to track the camera pose from. Natural feature based tracking systems also allow for AR applications e.g. augmenting video archive footage for special effects, which do not permit the placement of fiducials. Furthermore, the user's visualization of the augmented reality is greatly enhanced since virtual objects are introduced into a completely natural setting. The basic procedure to perform camera pose tracking from natural features involves two main steps:

1. Establishing which features correspond to which between different image frames of the incoming video stream.
2. Estimating the change in camera pose between frames based on the change in 2D positions of these features.

One approach to recover the motion field from the tracking of natural features

is reported in [32], where optical flow is used to compute the differential motion estimates between adjacent image frames. However, because it's very computationally heavy to do it this way, it's almost impossible to make it run in realtime on a normal PC. Another approach of real-time tracking algorithm based on homographies was developed to calculate image motion by another research scholar of our lab [33]. In my research work of tangible interactions, this natural feature tracking algorithm was applied in the tilt pad Pacman game in real time, which will be introduced in chapter 6.

2.1.3 ARToolKit

ARToolKit is a computer vision based augmented reality tracking Toolkit. This Tool Kit is a free software library that enables the development of a wide range of augmented reality applications [34]. It uses computer vision techniques to calculate camera position and orientation relative to physical markers. The markers were directly integrated into the real PC. Because of its robust registration result, ARToolKit is adopted for tracking purpose in this research work, including both the map system and the Pacman game.

ARToolKit uses computer vision techniques to calculate the real camera viewpoint relative to a real world marker. There are several steps as shown in figure 2.2. First the live video image (the left frame) is turned into a binary (black or white) image based on a lighting threshold value (the middle frame). This image is then searched for square regions. ARToolKit finds all the squares in the binary image, many of which are not the tracking markers. For each square, the pattern inside the square is captured and matched against some pre-trained pattern templates. If there is a match, then ARToolKit has found one of the AR tracking markers. AR-

ToolKit then uses the known square size and pattern orientation to calculate the position of the real video camera relative to the physical marker. A 3x4 matrix is filled in with the video camera real world coordinates relative to the card. This matrix is then used to set the position of the virtual camera coordinates. Since the virtual and real camera coordinates are the same, the computer graphics that are drawn precisely overlay the real marker (the right frame). Objects being overlaid could be from OpenGL applications or VRML worlds. The OpenGL API is used for setting the virtual camera coordinates and drawing the virtual images [35].

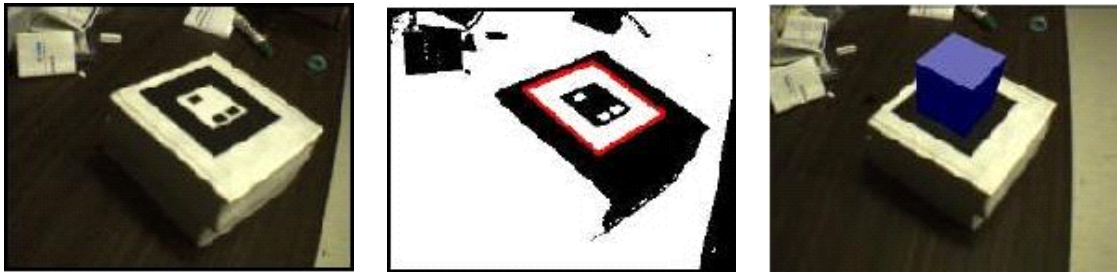


Figure 2.2: Input Video(left frame), Thresholded Video(middle frame), Virtual Overlay(right frame)

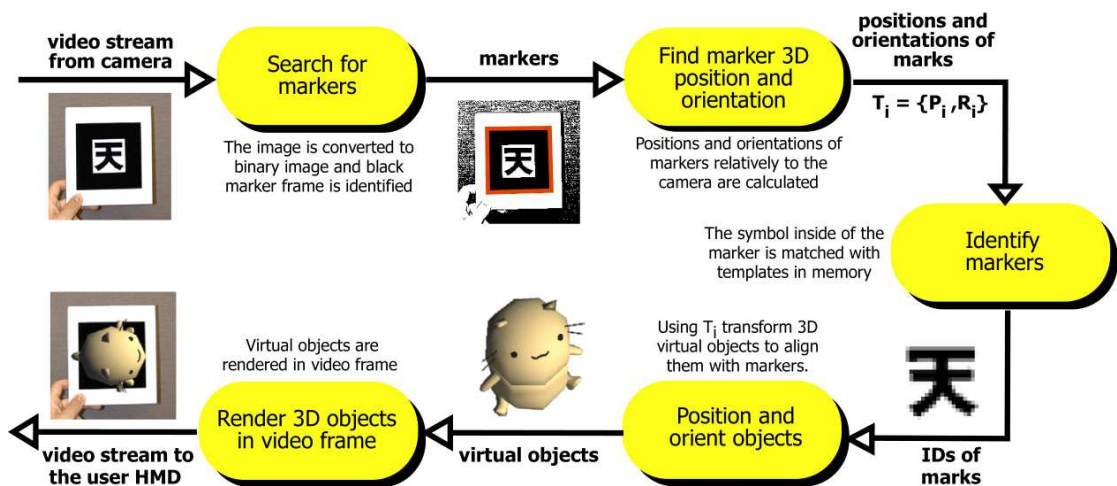


Figure 2.3: Image processing in ARToolkit

The diagram (figure 2.3) above shows the image processing used in ARToolkit

in more detail.

2.2 Literature Review

2.2.1 Sensing

The tilt measurement of the accelerometer is a very important aspect that can be put into effective use in car alarms [36]. Security systems for cars have roped in accelerometers to counter the complexities involved in using magneto-inductive sensors. In the work by Harvey and Lemaire, the ADXL202 is set up to acquire acceleration from 0 to 200 Hz. The accelerometer output is fed into two filters: a low pass filter with a corner frequency at 12.5 Hz used to lower the noise floor sufficiently for accurate tilt sensing, and a band pass filter. The low pass filtered (tilt) output then goes to a differentiator where the determination is made as to whether the accelerometer actually sensed tilt or some other event such as noise or temperature drift. The band pass filter output is integrated and checked for shock/vibration. After the independent analysis of the two signals a decision is taken using a micro-controller, whether or not to set the alarm based on the measured acceleration output from the accelerometer. If the accelerometer tilt has changed by $\pm 5\%$ from the time when the car was parked the alarm is sounded.

Tilt information has also been used to design a mouse pointing device called acceleration sensing glove (ASG) [37]. The Glove has six accelerometers, five on the finger tips and one on the back of the hand. A controller board at the back of the hand formats and transmits the data via a RF transceiver to the host. Before the accelerometer data could be analyzed it had to be calibrated and normalized. The unbuffered data had a noise with standard deviation of 70mg (milli-gravities).

The signal bandwidth of the sensors output has been kept at 220Hz. Since tilt data was enough, simple averaging was carried out to classify the different actions of the fingers. A coordinate transformation was performed to convert the sensor data into polar values (to identify the gestures). Simple gestures like curling of the fingers are used to perform different functions. The work also projects a future work on integrating the communication, processing, power and sensor into a single package which is no larger than 1 mm^3 . The glove is shown in Figure 2.4.



Figure 2.4: A Glove application

The interaction device tilt pad in this thesis work is based on acceleration sensing. Accelerometers are sensors that sense changes in velocity, both static (due to earth's gravitational field) and dynamic (due to motion). A number of micro-machined devices have emerged due to developments in IMEMS (Integrated Micro Electro-Mechanical Systems). The ADXL202 from Analog devices [38] is one of those products. Both the above two projects used ADXL202 as sensors and it is

the one that is used in tilt pad to sense the tilt information.

2.2.2 Tangible Interfaces in AR

Now let's look at some applications developed by other researchers in the area of tangible interfaces in AR and how they lead to our work.



Figure 2.5: Augmented groove

ATR MIC(Media Integrations and Communication) Labs and HIT (Human Interface Technology) Lab collaborated in a joint project called Augmented Groove [17]. The Augmented Groove (figure 2.5) is a musical interface that explores use of augmented reality, three-dimensional (3D) interfaces and physical, tangible interaction for conducting multimedia musical performance. In the Augmented Groove visitors can play music together, with or without use of traditional musical instruments, simply by picking and manipulating physical cards on a table. The physical motions of the cards are mapped to changes in musical elements such as timbre, pitch, rhythm, reverb and others. At the same time, users wearing a lightweight head-mounted display (HMD) can see 3D virtual images attached to the cards

whose shapes, color and dynamics reflect aspects of the music controlled by the visitors. Because in augmented reality environments visitors can see physical world, virtual objects and each other, several visitors can easily join around the mixing table and play together, passing musical sequences to each other in the same manner in which we would pass everyday objects. The music in a sense becomes a physical, tangible object, something we can touch and see as part of our physical environment.

A collaborative gaming augmented reality system has been investigated by other researchers [39]. They built a computer-based gaming environment for home usage which is especially suitable for multi-player board-games. The Mah-Jongg game was implemented as a demonstration of their concepts, to test the setup and interaction techniques. The system architecture consists of a game server maintaining the graphical database of the game and running the simulation based on the input data from magnetic trackers and input devices. Clients for each user render the customized view on the shared environment. Tracker data and information of the buttons on the pens are distributed over the local network using a tracker server. Using a multi-cast group every participant gets synchronized updates of tracker data for rendering. The interaction part of this system allows both social communication and private space as required by a multi-player board game. Figure 2.6 on the next page shows the concept of this gaming space.

The above two projects have introduced the scenery how augmented reality contributes to the build of tangible interactions. The augmented groove is a multimedia musical interface. But strictly speaking, the graphical models here are just a simple representation of the sound. They are adjunctive attribute of the core information. While in our augmented Pacman game, a complete multimedia prototype is provided with nice graphics and sound to give players immersive AR



Figure 2.6: Mah-Jongg game implemented in this AR gaming space

gaming experiences. Augmented groove also is a nice model that illustrates how the digital information such as music is manipulated and controlled physically. This is basically the same idea used in this research work. The latter AR gaming space set foot in the domain of interactive gaming. It extended the conventional multi-person board games to AR environment. Correspondingly, this research makes the AR version of popular PC games.

A wearable augmented reality system with small hardware interface was investigated in [40]. This system supports true stereoscopic 3D graphics. Through a pen and pad interface well known 2D user interfaces can be presented to the user, whereas the tracking of the pen allows to use direct interaction with virtual objects. The Wacom graphics tablet and its pen, The main user interface, are optically tracked by the camera using markers. Moreover the 2D position of

the pen (provided by the Wacom tablet) is also incorporated into the processing to provide more accurate tracking on the pad itself. An experimental game of chess between a mobile and a stationary user was developed as a demonstration of possible applications of this system as shown in figure 2.7.



Figure 2.7: A game of AR Chess interfaced with pad and pen

In the above system, small hardware was used as interaction devices. This is quite an effective way for accurate tracking and natural user interaction. Instead of buying commercial products like them, we use the novel accelerometer-based hardware tilt pad designed by our own lab in this thesis project, which makes it more possible to acquire fully tangible interfaces and to take full advantage of the hardware device.

Chapter 3

Non-hardware Tangible Interface

In this chapter, we present the research we did in non-hardware tangible interface system design for augmented reality.

Augmented Reality interfaces have the following three integral characteristics:

1. They combine the real and the virtual worlds.
2. They are interactive in real time.
3. They are registered in three dimensions.

Although the Augmented Reality technology has come a long way from rendering simple wireframes in the 1960s [22], AR interface design and interaction space development remain largely unexplored. Researchers and developers have made great advances in display and tracking technologies, but interaction with AR environments has been largely limited to passive viewing or simple browsing of virtual information registered to the real world. To overcome these limitations, in this thesis work, we seek to design AR interfaces that provide users with rich interactivity. These kinds of tangible interfaces are supposed to merge the physical space in which we live and the virtual space in which we store and interact with

digital information. In this single augmented space, computer-generated entities would become the directly controlled objects of the physical environment. We would use these entities just as we use physical objects, selecting and manipulating them with our hands instead of with a special-purpose device such as a mouse or joystick. Interaction would then be intuitive and seamless because we would use the same tools to work with digital and real objects.

3.1 3D Modelling

With the previous research on virtual reality, computer graphics has made tremendous progress in visualizing 3D models. Most augmented reality systems nowadays insert 3D virtual objects into real environment rather than simple text annotations in early days. This brings the need of realistic three dimensional models. For user interfaces, graphics appearance is quite an important issue as users perceive and interact with them directly. The problem is that even though the tools that are available for three-dimensional modelling are getting more and more powerful, synthesizing realistic models is difficult and time-consuming.

Generally, there are two ways to generate 3D virtual objects.

1. Create from imagination.
2. Model from planar images.

There are many mature softwares that support original 3D models creations such as Maya, 3D Studio Max, etc. While modelling from images keeps relatively sophisticated as it requires 3D reconstruction based on computer vision knowledge. The rough procedure of modelling from images is illustrated in figure 3.1. Some

commercial softwares realize this algorithm to produce final 3D models like Image Modeller.

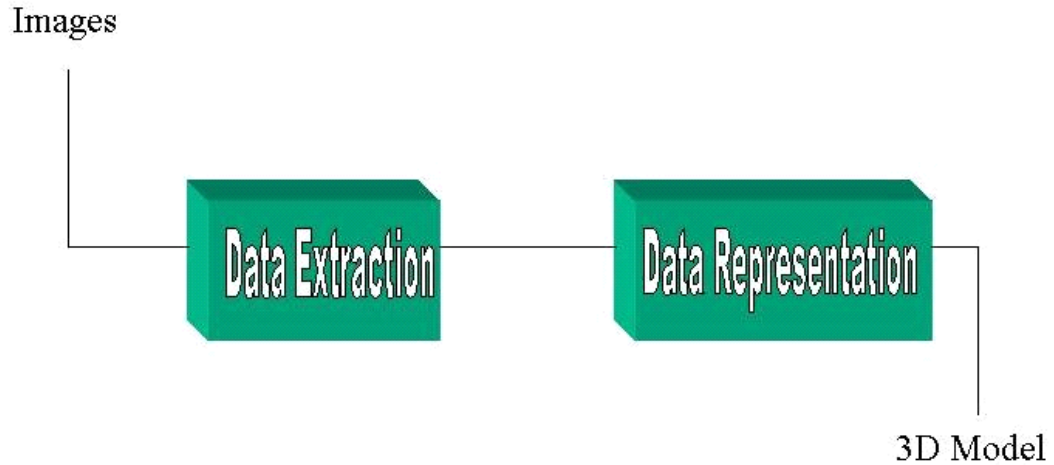


Figure 3.1: Rough procedure of modeling from images

3D models have various formats according to how they are created and the requirement of the application. VRML (Virtual Reality Modeling Language) is a standard language for describing interactive 3D objects and worlds delivered across the Internet. It is a powerful protocol for describing shapes, sensors and scripts and multimedia. VRML is widely used in augmented reality and especially those applications programmed in ARToolKit platform since VRML is the format that ARToolKit supports. It is also the format that we use to create the 3D objects for this *Magic Book* tangible interface system.

In VRML, a 3D scene is described by a hierarchical tree structure of objects (nodes). Every node in the tree represents some functionality of the scene. There are 54 different types of nodes. Some of them are shape nodes (representing real 3D objects), and some of them are grouping nodes used for holding child nodes. Here are some examples:

- Box node – Represents a box in a scene.
- Transform node – Defines position, scale, scale orientation, rotation, translation, and children of its subtree (grouping node).
- Material node – Corresponds to material in a scene.
- Directional Light node – Represents lighting in a scene.
- Fog node – Allows you to modify the environment optical properties.

Each node contains a list of fields that hold values defining parameters for its function. Nodes can be placed in the top level of a tree or as children of other nodes in the tree hierarchy. When you change a value in the field of a certain node, all nodes in its subtree are affected. This feature allows you to define relative positions inside complicated compound objects.

The Hello Kitty model used in *magic book* was made from a real Hello Kitty cartoon toy using its photographs. ImageModeller is the software we used to aid in creating the 3D model. Generally, five steps are involved in creating this model:

1. **Import Images:** We took 9 photos of the toy model in JPG format from different points of view using a typical consumer digital camera to capture all the surface and spatial information of the object. Eight photos were taken from the toy's side faces, with 45 degrees difference between each two photos, which means that photos were taken from front, left-front, left, left-back, back, etc. The last photo was taken from the top side. Since the bottom surface information was not important in this application, no photo was taken from the bottom. One point to note here is that every two sequential photos are supposed to have some common areas to connect with each other and to provide potential spatial relationship.

2. **Calibrate photos:** Given that every two sequential photos have enough overlapping parts, we are able to mark some common features of the imported photographs in this step (about 9 to 20 features for each photograph). After that, the software will help to calibrate the cameras.
3. **Create the model:** Model the object using standard polygon model tools. Each face of the polygon was extruded from points with spatial coordinates.
4. **Extract textures:** Capture textures from original images from face to face.
5. **Export:** Export the 3D Hello Kitty model with full textures to VRML format.

Besides establishing 3D models from images, which outputs models exactly resembling real objects, making models purely relying on the creator's imagination or based on certain characters is also applied to get the other 3D models (frog etc.) in *magic book*. Maya and 3D Studio Max are such kinds of software that allow users to realize their 3D imagination. In this approach, we separate a 3D object into several main parts and use a polygon to describe each part, e.g. head, body, leg, etc. Then we combine these parts together to finalize the whole model by defining the relative spatial coordinates of each part. Finally, a 3D model available for various applications is obtained by exporting the whole scene to the required format.

The following figure 3.2 displays the Hello Kitty model which we have created for *magic book*. This model is made from images.



Figure 3.2: Different angles of 3D model attached to the marker viewed from head-mounted display(HMD)

3.2 Implementing Tangible Control

Based on the tracking methods provided in ARToolKit that we described in chapter 2, applications for Augmented Reality can be developed to create Tangible User Interfaces using our unique 3D models. Real world objects (like cards with fiducial markers printed on top) can be manipulated by the users, with virtual objects superimposed upon them. Manipulation of these real world objects can then be used to interact with the computer in a natural way.

The relative distance between the centers of the two markers, as well as the angle between the two marker surfaces, are calculated in real time. Corresponding actions are then applied to such situation. A simple application of such a tangible interaction in this thesis work is the *Picking* and *Dropping* of the virtual object by controlling the distance and the angle between the markers (Figure 3.3).

- *Picking* is defined when the fiducial marker **A** (the transportation paddle) is been moved very closely to the fiducial marker **B** (the original position for the 3D object), while remaining horizontal comparing to marker **B**, that is, angle between these two marker surfaces is small enough. When *Picking* happens, the system will stop rendering the 3D object on the fiducial marker

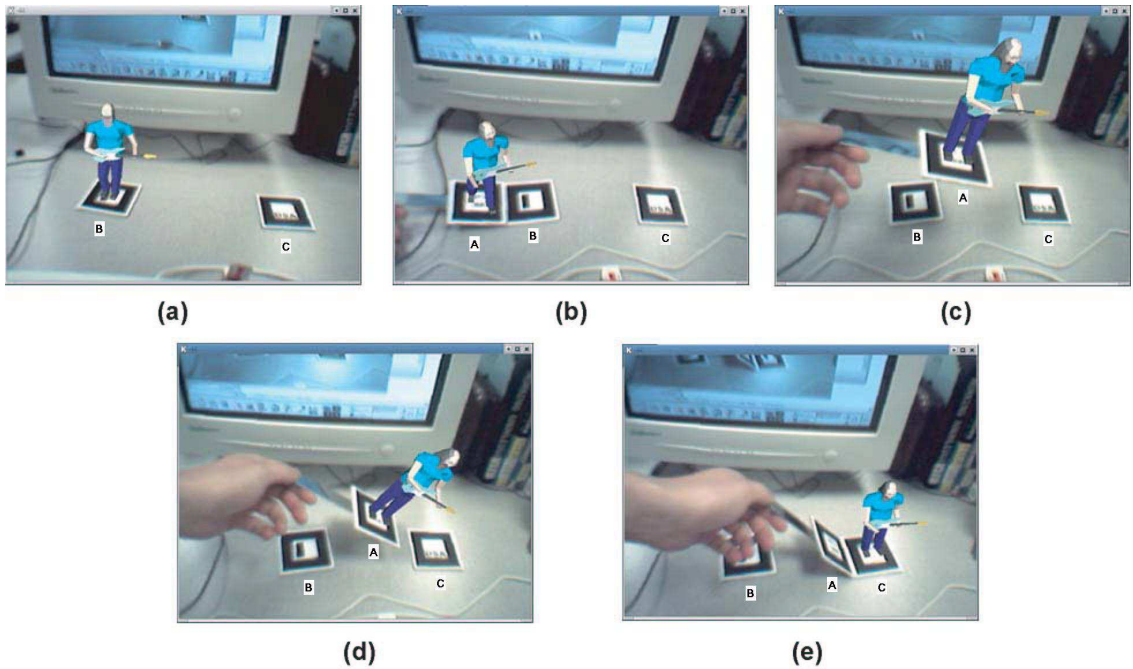


Figure 3.3: Tangible Interaction: Picking and Dropping virtual objects by manipulating physical cards.

B, but on the marker **A** instead (Figure 3.3-b).

- After *picking* up the virtual object, the user can then move it around the environment, or even move it near to his eyes, which is the camera on his head-mounted display (HMD) now, to have a close look of these virtual characters.
- *Releasing* or *Dropping* is defined when the fiducial marker **A** (the transportation paddle) is been moved very closely to the fiducial marker **C** (the destination for the 3D object), and tilt it up to a certain angle (e.g. 50 degrees) relative to the surface defined by marker **C**. When it happens, the system will stop rendering the 3D object on the fiducial marker **A**, but on the marker **C** instead (Figure 3.3-e).

A more sophisticated application of these ideas is presented in a prototype desktop AR system *Magic Book* (Figure 3.4). Here, a catalogue of three dimensional

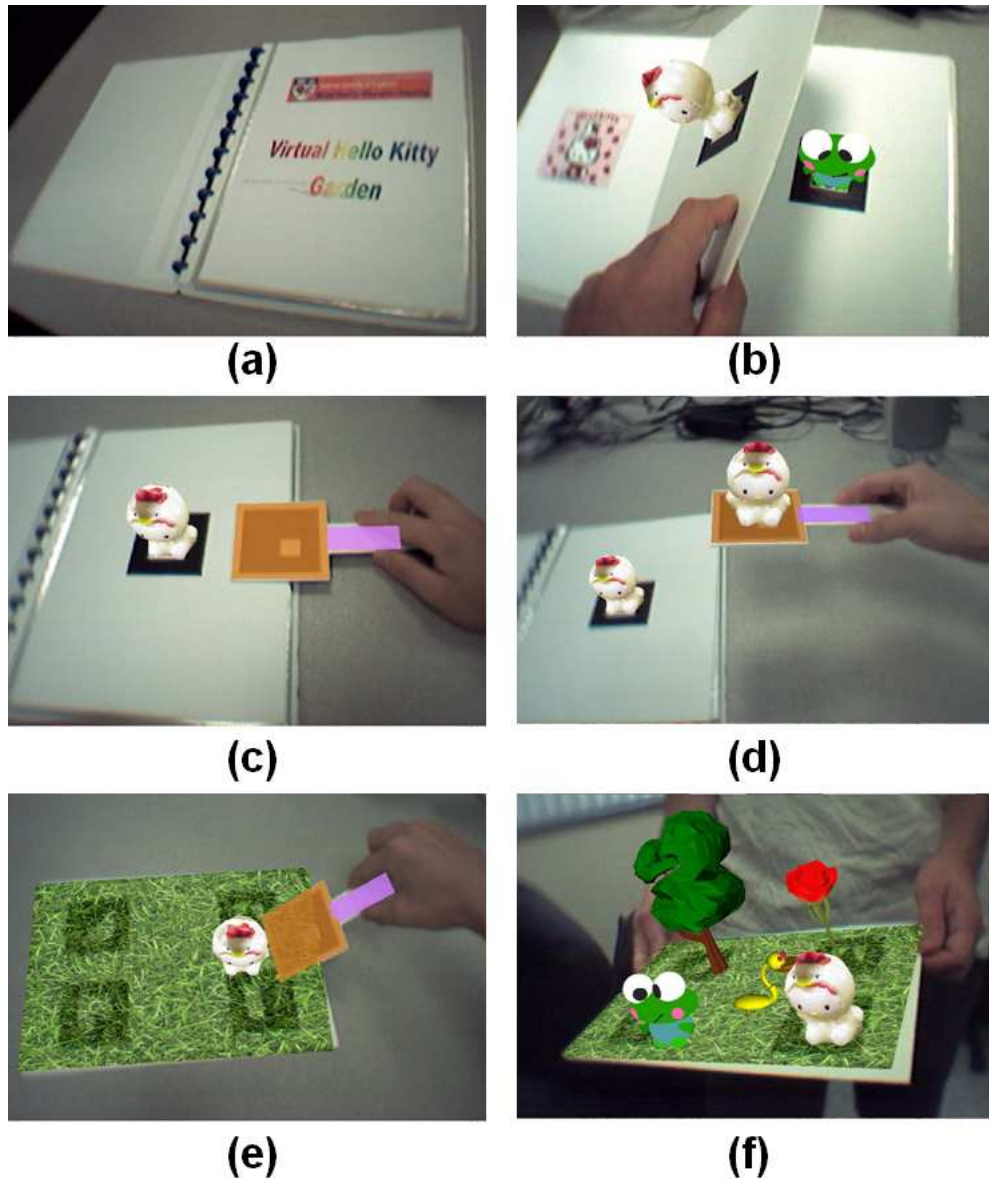


Figure 3.4: A more sophisticated tangible interaction example: virtual Hello Kitty garden.

cartoon models is created by attaching a different fiducial marker to each page of a real book (Figure 3.4-a). Users can select the model that they are interested in by simply turning the pages (Figure 3.4-b). Again, this is a radical departure from scrolling through a list of three dimensional models on a computer screen. The interface is the real-world object which is a book in this case. The user also has a virtual paddle (Figure 3.4-c). When he brings the paddle adjacent to the three dimensional object, the object is copied to the paddle (Figure 3.4-d). It can then be moved through the environment. The object can be released and placed on a surface by tilting the paddle relative to that surface (Figure 3.4-e). In this way, the user can place the virtual objects on a virtual surface to create a model garden (Figure 3.4-f).

We implemented this Magic-Book system on a normal Pentium III PC running Linux, which allows updates at 30 frames per second.

Chapter 4

Interface Enhancement – 3D sound

Although AR is often associated with visualization (starting with the first head-mounted display by Sutherland [41]), augmentation is also possible in the aural domain [42][43]. At this point, a real-time registered three dimensional sound could also be a paradigm of AR applications, and is termed augmented audio reality (AAR) [42]. Such AAR systems may be used for a variety of applications: memory and comprehension improvement in video conferencing [44], interface enhancement, video gaming, etc. Therefore, as a research on exploring tangible interfaces for augmented reality, we feel the need to study the role of 3D sound in AR interfaces. In this chapter, we present an experimental study of the effect of 3D sound and the results would be quite valuable in developing multimedia tangible user interfaces.

3D sound, often termed spatial sound, is sound as we hear it in everyday life. Sounds come at us from all directions and distances, and individual sounds can be distinguished by pitch, tone, loudness, and by their location in space. The spatial location of a sound is what gives the sound its three-dimensional aspect. In order to

gain a clear understanding of spatial sound, it is important to distinguish monaural and stereo sound from 3D sound. A monaural sound recording is a recording of a sound with one microphone. No sense of sound positioning is presented in monaural sound. Stereo sound is recorded with two microphones several feet apart separated by empty space. Most people are familiar with stereo sound; it is heard commonly through stereo headphones and in the movie theater. When a stereo recording is played back, the recording from one microphone goes into the left ear, while the recording from the other microphone is channeled into the right ear. This gives a sense of the sound's position as recorded by the microphones. The recording and the split of the stereo sound can be realized with electronic instruments that provide a good spatialization and a high sound quality. Listeners of stereo sound often perceive the sound sources to be at a position inside the listener's head – this is due to the fact that humans do not normally hear sounds in the manner they are recorded in stereo, separated by empty space. The human head is there acting as a filter to incoming sounds.

The application of digital sound within computers, and in particular the manipulation of sound in virtual space and multimedia systems, is considered to be an effective way of interface enhancement. Spatial sound could help increase the sense of presence in virtual environments by relaying information about the environment and the objects within it. Such environmental awareness could be very beneficial in increasing the user's orientation and immersiveness in virtual environments. Served as a compliment to manipulating visual information in computing via *Tangible User Interfaces*, the utility and control of sound makes the interaction diversified and more natural, for it abundantly activates the sensory organs of human being. We aim to investigate the role of 3D sound in augmented reality, which would support our future work of interface design.

Audio feedback has proven itself to be an effective addition to human-computer interfaces in many studies not related to augmented reality [45]-[49]. Sounds were used to present information that would otherwise be difficult to convey through other means. For example, when the user’s attention needs to be focused away from a visual display, auditory cues can be useful to help participants monitor the status of ongoing processes [45]. Some applications reported that 3D sound increases the immersive feeling in virtual reality environments which provide a more realistic virtual environment [46]. Furthermore, the sound can have impacts on task performance. Mereu [47] suggested that sound could improve depth perception in 3D interfaces. Wolfson [48] argued that different combination of sound and color could have significant effect on task performance. Markino [49] indicates that spatialized sound from virtual acoustic display helps fast and accurate navigation. The arkola simulation by Gaver [45] suggested that auditory icons can play a significant role in multiprocessing and collaborative system. However, very little work has examined the research issues of the advantages of sound in AR systems. It is important to rigorously examine these issues and experimentally verify the assumption about any benefits.

4.1 System Setup

In our research system [50][51][52], we use the InterSense IS-900 system [53] for precise motion tracking. The IS-900 system is a hybrid acoustic-inertial six degree of freedom (6 DOF) position and orientation tracking system. It tracks changes in orientation and position by integrating the outputs of its gyroscopes and accelerometers, and corrects drift by using a room-referenced ultrasonic time-of-flight range measuring system. It can obtain a resolution for position of 15 mm.

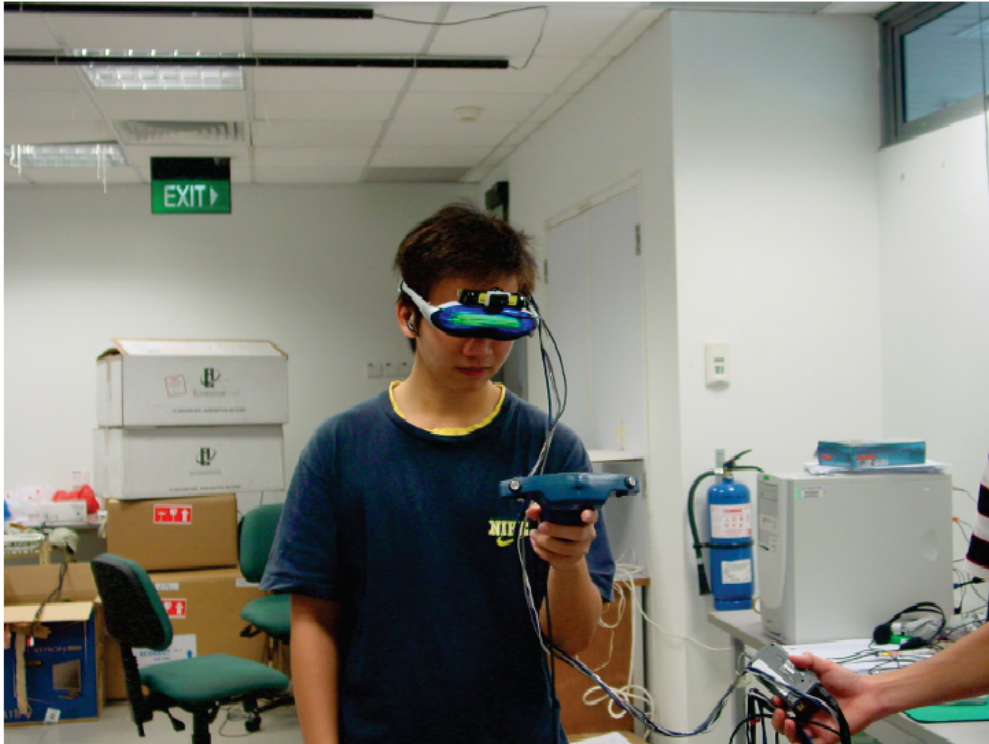


Figure 4.1: The physical set-up of our acoustic-inertial tracking AR environment. On the ceiling there are six IS-900 SoniStrips which makes up a three meter by three meter area. A head tracker is mounted on the user's HMD and a camera is fixed on the head tracker. A wand tracker with four buttons and one joystick was held in the user's hand as an interaction device.

For the experimental system, as shown in Figure 4.1, we mounted six IS-900 SoniStrips on the ceiling of a 3 meter by 3 meter space. This system can support up to four tracking stations. We used four tracking stations to track two users' heads and hands respectively. A head tracker was mounted on the user's HMD and a camera is fixed on the head tracker. A wand tracker with four buttons and one joystick was held in the user's hand as an interaction device. The system superimposes graphics on video sequences captured from the camera and displays the combined images on the HMD. We employed OpenAL [54] for generating 3D sounds.

Normally the space of AR environments is relatively small, and some applications are even restricted to desktop. If the virtual sound sources are limited in this small space, then the attenuation of sound with distance can be too vague to be distinguished. In most 3-D sound systems, the intensity of sound would follow incremental reductions of 6 dB for each successive position of doubled distance [46]. For example, the intensity of sounds could be perceived to be the same when the source is 35 cm or 40 cm in front of you. In this sense, located in the small AR space, the 3D sound provides weaken information about the distance dimension. In order to avoid this, we need to exaggerate the intensity difference of 3D sound according to distance changes of virtual sound sources.

4.2 Experimental Goals and Subjects

In the experiment, we examined the effect of 3D sound in AR environment in the form of a simple game interface. Barfield [55] identified three primary benefits of auditory spatial information displays as: relieving processing demands on the visual modality, enhancing spatial awareness, and directing attention to important spatial events. In our research system, we aim to find out:

1. Whether 3D sound helps in task performance, or more specifically, whether it helps to shorten a searching task completion time and by how much.
2. Whether 3D sound gives more immersive and realistic perception to the users and whether it helps the subject to distinguish different objects in AR environments.
3. Whether 3D sound increases the feeling of “human presence and collaboration”.

Both quantitative and questionnaire studies were conducted in the experiment. Our first experiment aim stated above was expected to find out through quantitative analysis; While the questionnaire was designed to examine the other two assumptions. The study took place over a four-week period in the summer of 2002.

Forty subjects (13 Females and 27 Males) were each paid Singapore \$8 (approximately \$4 USD) per hour to participate in this study. All the participants were undergraduate students at the National University of Singapore. Most of them were freshmen or juniors majoring in computer engineering. The subjects were on average 21.3 years old and had used computer for 8.7 years. They were asked to mark from 1 to 7 about how often (1 means very often while 7 means very seldom) do they use computers and how often do they play games. The averages were 1.7 and 4.5 respectively, which means that they often use computers, but play games only occasionally. All of the subjects reported clear-vision and normal hearing abilities.

In the experiment, participants were given the opportunities to play a game with three stages. After the game, 20 subjects (seven females and thirteen males) were asked to conduct the searching task with the aid of 3D sound while the other 20 subjects (six females and fourteen males) conducted the same task without the 3D sound. Before doing the experiments, all the subjects were given a brief introduction about augmented reality and the tasks to be performed. Each of the subjects was given five minutes to initially try on the HMD and experience the 3D sound effect. After the experiments, all subjects were given 15 minutes to finish a questionnaire based on the experiments they had participated in.

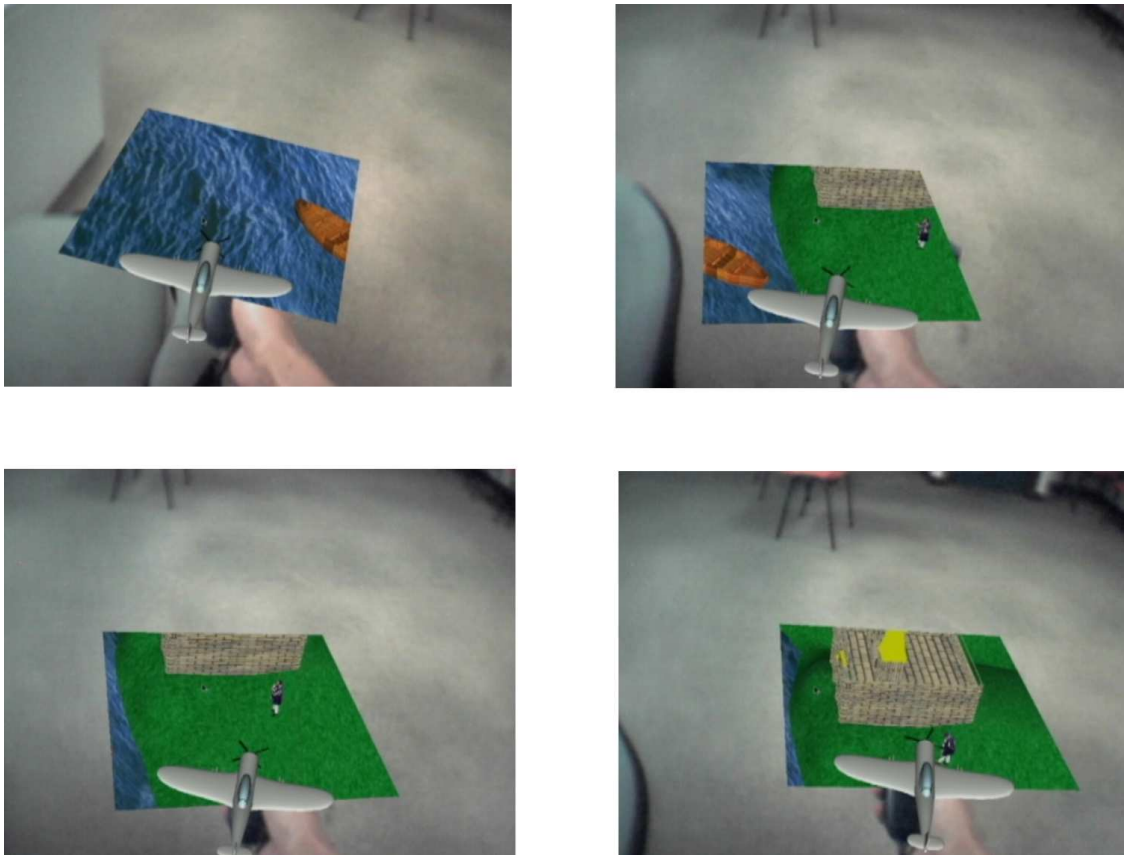


Figure 4.2: Looking for the castle through a “magic 3D window”. A virtual plane is augmented on the wand as an avatar for the user. A rectangle window is augmented in front of the wand through which the user can see the virtual land on the ground. When the user is moving around, the 3D landscape is also scrolled in the “magic 3D window”. The user is looking down at the castle as if she is looking down from the sky in the airplane.

4.3 User Interface

Our experimental system allows two users to play a simple game at the same time. The objective of this game is to rescue the princess who was captured by an evil witch. In this experimental game, three stages were used to investigate the effects of 3D sound: searching castle, beating witch and rescuing princess. In the first stage, users are allowed to walk around in a three by three meter region and to explore a virtual land augmented on the physical ground through a small rectangle

virtual window displayed on the tracked wand. Figure 4.2 shows the scrolling 3D landscape when the player is moving her hand. A small virtual plane is attached to the wand as an avatar for the user, as if the user is exploring the game space by manipulating the plane. The player is looking for the castle through the “magic 3D window”. Two 3D sounds are applied in this stage and can be heard by both players. The first one is a 3D sound “Help!” which was applied to the princess located inside the castle. The second sound is a constant buzz applied to the wand tracker of co-player as if it is generated by the engine of co-player’s airplane. Each player can hypothetically localize his/her co-player by localizing the buzz sound.

In the first stage the player is allowed to become more and more familiar with the searching task and the perception of 3D sound. We purposely design this stage as a training process. When the player walks closer to the sound sources (the castle and the princess therein or the co-player), the sound becomes louder while becomes softer when the player walks away from the castle. An exaggeration of the intensity difference described in *System Setup* was undertaken here as well. Furthermore, it should be noted, when the player rotates her head or body, she can perceive that the sound source is also moving relative to her movements. We updated the position and orientation of the head tracker 25 times per second which allows a quite smooth feeling of sound source movement.

After the castle is found, players move to the second stage. They battle with a virtual witch who has captured the princess. The virtual witch is augmented in the physical world and is hovering in the experimental region. Players can move freely in the physical space to get the best shooting angle and position. They need to reload or transfer their bullets by putting their wands together which encourages the collaboration in the game, as shown in first two frames of Figure 4.3. Apart from the sounds used in stage one, we applied the other two 3D sounds, the cry

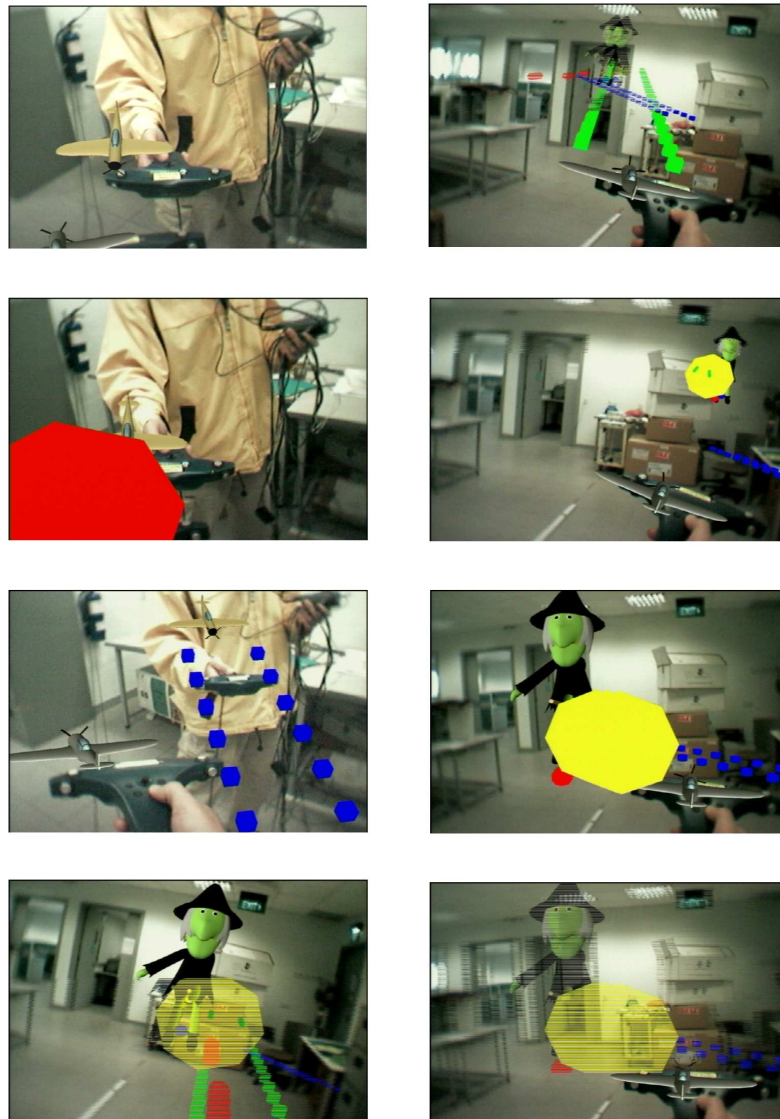


Figure 4.3: Two users are collaboratively fighting the witch hovering around them in the physical world. The user can see the virtual witch which is augmented in three dimensions through HMD. The two users need to reload or transfer their “bullets” by putting their wands together as shown in the first two frames. When the user pressing the button on the wand, the airplane can fire bullets with different color (blue and green in the subsequent frames). The witch is designed to hover in the experimental region. The position and orientation of the wand are used to decide trajectory of “bullets” and whether the witch is shot or not. In the last frame, the witch was shot for ten times and disappeared.

of witch when it is hit and the explosion sound when it is exploded after being hit ten times.

After destroying the witch, the players come to the last stage where they are allowed to fly into the virtual scene and rescue the princess from the castle. This enables a seamless transition from AR environment to VR environment. After the user flies into the virtual scene, she enters a virtual environment in which players can use joystick on the wand to move around, while they can turn over their head or body to change the direction of viewpoint. In this stage we still keep the princess's cry "help" but change the sound of co-player from a buzz of engine into a footstep sound. Both of the players can be either in AR or VR mode and can continuously be aware of each other by vision and sound. Figure 4.4 shows the views in this stage.

To examine the impact of 3D sound on task completion time and subjective feelings in this AR environment, 3D sounds were applied in this experimental game. Their functions are summarized as follows:

1. **Players' sounds:** A constant buzz sound is applied to the co-player's plane when the users are exploring the augmented land. Since the plane is attached to the wand, the 3D positions can be easily obtained from the tracker in the wand. By applying this sound, the user can always hear where is the location of the co-player, even when the co-player is not visually in the scene. She can still roughly tell to what direction is her co-player located and whether the co-player is getting closer or further to her. This buzzing sound is changed into a footstep sound when the co-player enters the virtual reality environment. The different sound gives the hint of whether the co-player is in the AR or VR environment as well.

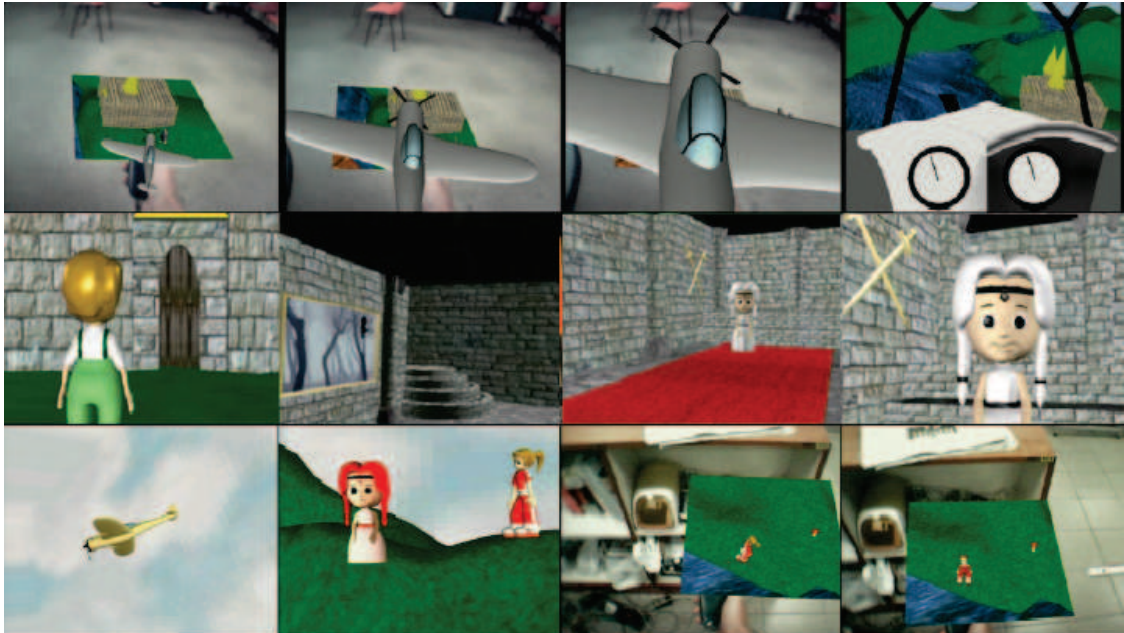


Figure 4.4: Rescue the princess in the castle with co-player. The first row shows the seamless transition from AR to VR by which the user is “flying” into and “landing” on the VR world. The second row shows one player’s view when they are on the way to rescue the princess in the castle. Here the user uses the joystick to navigate however they can turn the body and head of the virtual character by physically turning their body and head. Meanwhile, the footstep sound comes from the co-player can always be heard. While navigating in the castle, the players are guided by the “Help!” sound of the princess. The last two frames in second row shows the last exciting moment when the player finds the princess and the princess shows her appreciation. In the first frame of the third row, a user can physically look up in the sky to find her co-player who is still in AR mode (before flying into the VR mode). Meanwhile, the last two frames show that the user can still see the co-player and the princess when she is in AR mode by looking down to the land as if she is on the airplane. The second frame shows one player’s view point when both of the players are in VR mode and come to the princess.

2. **Human voice:** A constant human voice “Help! Help!” is applied to the princess which is the target of the searching task. This sound helps the player to turn and face to the correct direction and walk to the correct location where she can see the princess through the small window.
3. **Entertainment sounds:** To make the game more entertaining, we also applied other sound effects. A shooting sound is played when user press the shooting button on the wand. A scream sound and an explosion sound were played when the witch is destroyed. A landing sound of engine is played when the user transits from AR environment to VR environment. A female voice saying “congratulations” and some greeting words was played when the user finished her mission – find the princess and rescue her.

All these three experiment stages were designed in order to introduce the participants to an augmented reality world with both visual images and 3D sound information. With such an explorer, users were asked to fill in a questionnaire paper to give their subjective perception of 3D sound. Furthermore, they would be required to perform a task searching mission after these three stage. Instead of searching the castle, they need to search for the princess by walking around and scrolling the window. The princess was taken outside the castle and was put at an open area on the land. Each subject was given a trial to see how the princess looks like in the small window. Then the princess was put at another position which is the same for all subjects. Every subject started from the same position and orientation in the space and began to search for the princess. The completion time of this searching task was recorded in order to make corresponding numerical analysis. Twenty subjects (seven females and thirteen males) were asked to conduct the searching task with the aid of “Help!” 3D sound while the rest twenty

subjects (six females and fourteen males) conducted the same task without any sound. Every subject performed this task only once either with or without sound aided. With the data of task completing time and the questionnaire responses, we made comparison and analysis to examine the experiment aims we brought forward forenamed.

4.4 System Assessment Results

4.4.1 Performance

In this experiment the subjects performed a searching task as described in section 4.3. The chart in Figure 4.5 compares the completion time (in seconds) of the searching task for two groups in the cases with 3D sound and without 3D sound aided. The average task completion time of the group without 3D sound aided was 1.5 times as long (133.05 versus 88.1) as that with 3D sound aided. The effect of 3D sound on shorten the task completion time is significant, $F(1, 38) = 6.7, p < 0.025$. Here using a 0.05 significant level, a one-way ANalysis Of VAriance (ANOVA) [56] of the completing time was used. In an ANOVA, the F-ratio is the statistic used to test the hypothesis that the effects are real. In other words, the means of completing time are significantly different with the effect of 3D sound.

No significant difference is found between the genders in this experiment as well. For the task completion time with 3D sound aided, males have a mean value of 88.4615, and females have a mean value of 87.4286, $F(1, 18) = 0.0018, p > 0.1$. For the task completion time without 3D sound aided, although females' mean value is higher than that of males' (166.33 versus 118.79), no significant difference is found at the significant level of 0.05, $F(1, 18) = 3.03, p > 0.05$.

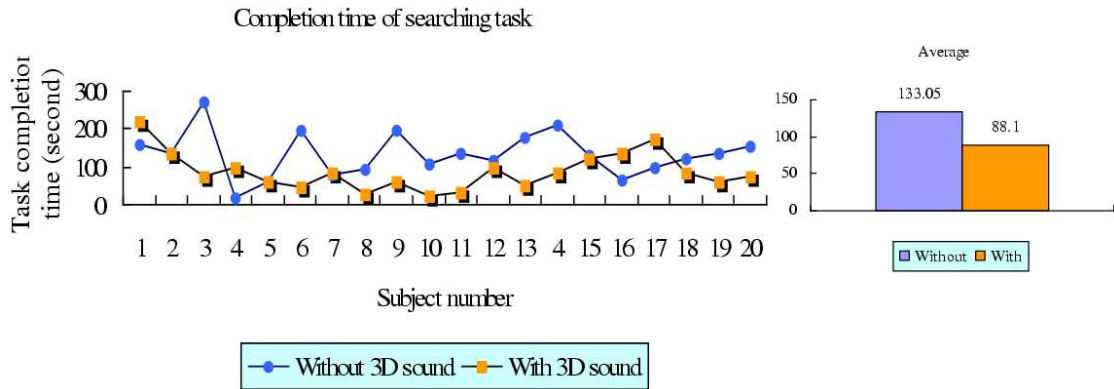


Figure 4.5: The comparison of searching task completion time in the cases of with and without 3D sound aided. The average of completion time is shortened drastically from 133.05 to 88.1 seconds by introducing 3D sound.

From the results given above we can see that, 3D sound helps significantly to shorten (about by half) the searching task completion time in this AR environment.

4.4.2 Questionnaire Responses

The purpose of the questionnaire questions is to examine the impact of 3D sound on the subject's psychological feelings. We wanted to examine whether 3D sound improves the feeling of immersiveness and human presence and collaboration, whether it helps to identify spatial objects, and how much impact it has on these feeling. The main questions in questionnaire about this experiment were as follows:

1. Does 3D sound improve your perception of immersiveness in the augmented reality environment? Do you think it is more realistic with 3D sound than without it? Why?
2. Does the 3D sound improve your feeling of human presence and collaboration as you can hear the sound coming from your collaborator at anytime?
3. Does the 3D sound help you to identify different information signals (spatial

objects)?

4. Were you aware of your co-players in different places?
A. always B. very often C. often D. barely E. never
5. What has helped your feeling of “human presence and collaboration”? (can be multiple choices, please order the factors by the order of decreasing importance)
A. small talk between co-players B. vision cues C. 3D Sound cues D. Others

Question one is designed to examine whether 3D sound functions in improving the immersiveness of an AR environment. Testers’ personal feeling regarding this are requested, as well as the reasons. Apart from immersiveness aspect, we also aim to investigate the feeling of collaboration and human presence with the aid of 3D sound. Question 2, 4 and 5 are designed for this purpose. Question two emphasizes on the effectiveness of 3D sound in improving the feeling of collaboration and human presence. The issue in question four is a kind of situation of human presence. Collaborations would be possible only if users could perceive the presence and proper position of his cooperator. And question five asks testers to rank the function of 3D sound among some other means including visual information and collaborator’s direct communication. The results would help system designers to adopt optimized means to improve such kind of feeling given different conditions and requirements. If 3D sound is tested to be effective in improving users’ feeling of the presence and location of the sound source, we would like to further test the situation given that one sound corresponds to one certain object, which results in object identification. Question 3 is put there to evaluate the function of 3D sound in identifying different objects with different sound.

According to question one, almost all subjects (39 of 40 subjects) claimed that 3D sound did improve the perception immersiveness. A total of 38 subjects claimed

that it is more realistic with 3D sound, and one subject claimed “a little better”. The main reasons they provided to support their viewpoint are as follows:

1. *“It is natural feeling and more realistic with 3D sound.”*
2. *“You can feel the sound from all directions which gives a more 3D feeling.”*
3. *“You can tell where the collaborator is positioned, or to know the surrounding objects in the augmented reality environment.”*
4. *“The feeling is closer to real world and 3D sound helps to navigate.”*
5. *“I can hear and judge distance from the object producing the sound.”*

From these responses, we can find that 3D sound does play an important role complementing the visual AR environment. Although this result seems to be natural since the same conclusion has been drawn in virtual reality environment (VRE), it proves that this conclusion is still tenable in AR environment when the implementation method has been changed and especially when the interfaces preserve our social interactions which are weak in VR.

There were also negative comments. One subject suggested that there should be more background sound like noises to make it more realistic. The subject said, “The sound field is too quiet to be realistic.” In fact, this is consistent with the argument given by Begault [46]: “Even with non-spatial cues, one can tell something about the distance of different sounds and their context, . . . , but the inclusion of 3-D sound techniques and spatial reverberation cues can greatly increase the immersiveness of a simulation, as well as the overall quality and the available nuance in representing different environmental contexts”.

The important result we obtained from the responses to question two is that 3D sound improves the feeling of “human presence and collaboration” when the

subject is performing this experimental task. In this question, 39 of 40 subjects answered “yes”. Some of the comments were:

1. *“At least I can find whoever it is.”*
2. *“The sound constantly reminds us of the presence of the collaborator.”*
3. *“Yes, like someone is actually in the world with you.”*
4. *“It is really interesting that you can turn to face him when the clop-clop is getting closer to you in your back.”*

One interesting comment was that “I can know whether he is still in the airplane or has landed down and walking around” (which meant that the user could know if the other player was in AR or VR modes). This is actually one function of sound which is serving as an auditory icon [57]. A good negative comment suggested that “it can be improved by adding speech communication, because a running motor does not sound very human, so it is like an object somewhat”.

Apart from applying 3D sound to the collaborator, we also applied 3D sound to the princess who is the target of the task. The responses to question three suggest that 3D sound, when applied to different spatial objects (information signals) in AR environment, can help the user to identify different information carriers. In our experiment 36 subjects claimed that they could constantly be aware of both the other collaborator and the princess simultaneously. This result suggests that by introducing more than one 3D virtual source sources in AR environment, it is possible for the users to identify different meanings of them and to distinguish them by the localization of the sound sources. As one subject described, *“yes, very much so. For instance, if the collaborator sounds from the left earpiece and princess*

cries are heard at right earpiece, the directional news determines the positions of the collaborator and princess.”

However, one valuable criticism complained that *“the volume differs when you are far but when you are near I could not hear a difference especially in terms of direction”*, the other one complained that *“the sound sometimes seemed to be coming from all directions and I was uncertain whether to turn left/right, only when I was close did the volume increase, but in a jump, not gradually”*. These comments point out the drawbacks of our exaggeration process of 3D sound intensity in this AR environment. We designed the auditory environment so that when the user becomes closer to the virtual sound source in the three by three meters region, the 3D sound intensity is exaggerated. This will be maximum intensity when the user is at the position of the virtual sound sources (distance to the object is zero). In this case, the binaural cues play a very important role for distance and direction perceptions. Because of the closeness to the object when near zero, even a slight movement of the user may cause a totally opposite feeling of the source direction (from left to right). When the movement is very fast, the intensity and sound direction may change rapidly. This problem can be partly solved by designing a better sound-scape. However the situation when the user is overlaid on the sound source is still a difficult problem unless we apply some physical constraints on the user’s movement near to virtual sound sources.

Chart (a) and (b) in Figure 4.6 shows the subjects’ responses to question four and question five respectively. We can see from chart (a) that nine of 40 subjects claimed that they were “always” be aware of co-players while five subjects claimed “very often” and 18 claimed “often”. This leads to the improvement to the feeling of collaboration. A total of 39 subjects agreed that 3D sound did increase their feeling of collaboration from the result of question two.

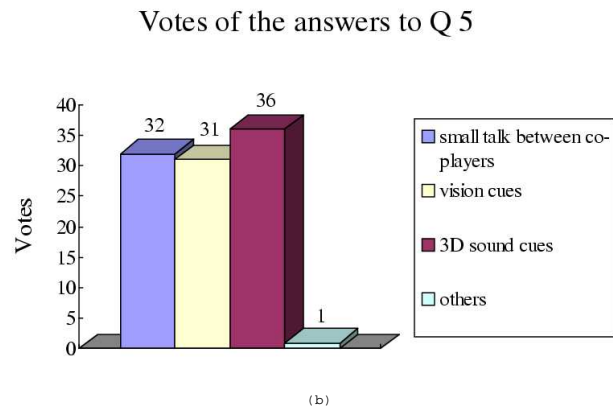
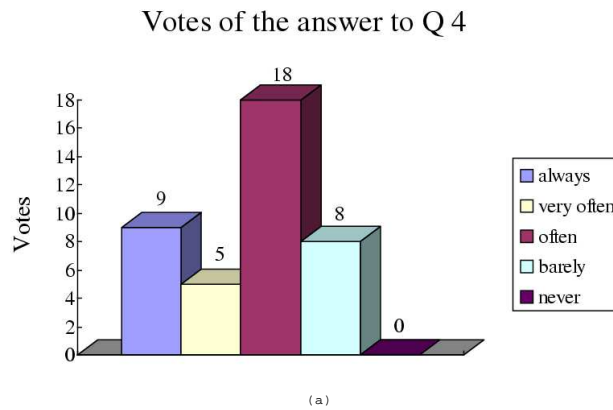


Figure 4.6: (a)The votes of the answers to question 4: Were you aware of your co-players in different places? We can see from the chart that in this AR environment, subjects can often be aware of co-player’s position, however, this result is impossible if there was no 3D sound applied since the subject can only occasionally see co-player due to the narrow angle of view. (b)The votes of the answers to question 5: What has helped your feeling of “human presence and collaboration”? It can be seen from this chart that 3D sound plays the most important role among all the cues. It gets even more votes than vision cues and voice communication cues.

To examine the importance of 3D sound among all cues, subjects were required to choose the cues they thought to have helped them and to order them by the importance. 3D sound cues obtained the highest votes (36 votes) among all the cues. More importantly, a total of 14 subjects regarded 3D sound cues as the most important cue while “small talk” and “vision cues” only obtained five and four votes

respectively. This result suggests that 3D sound does play a very important role in improving the subject's feeling of collaboration. It could even dominate among all possible cues in AR environment, especially when other cues are restrained by hardware settings, for example, angle of view is restrained by HMD in common AR system. In this situation, the visual cues are restrained due to the narrow viewpoint of HMD, and hence their effects on collaboration are limited. On the contrary, the 3D sound can be generated from any points in the AR space. When 3D sound is applied to the user and given different meaning at different situations, it can play the most important role to help the user's feeling of "human presence and collaboration".

4.5 Conclusions

From the above numerical and questionnaire analysis, we can make our conclusions based on the experiment aim preestablished now.

1. 3D sound helps significantly to shorten (about by half in our experiment) the searching task completion time in AR environment.
2. 3D sound does improve the perception immersiveness as it complements the visual AR environment to satisfy another basic demand of human perception. It is suggested that by introducing more than one 3D virtual source sources in AR environment, it is possible for the users to identify different meanings of them and to distinguish them by the localization of the sound sources.
3. 3D sound improves the feeling of human presence and collaboration.

These positive results could be a theoretical foundation of sound utility in AR environment and interfaces. And convinced feedback is shown when sound is ap-

plied in our tilt pad AR Pacman gaming system that is detailed in next chapter.

Chapter 5

Tilt Pad Based Tangible Interfaces Design

According to [58] interaction can roughly be categorized into *object manipulation*, *navigation*, and *system control*. Tangible interfaces stand for the trend of the representation and manipulation of digital information objects through physical means, which makes a radical departure from navigating, displaying and controlling them by mice and keyboards on a computer screen. The utility of this technology blurs the boundary between digital computer world and our everyday physical world, leading to the appearance of intuitive and natural human-computer interactions since the transformation of the objects dominating methods between digital and physical worlds is eliminated now.

In this chapter, we will demonstrate two AR tangible interfaces based on small hardware tilt pad: one 3D map scrolling system and one Placeman gaming system. Both systems are developed in AR environment employing fiducial marker tracking of Acropolis and embedded with hardware device. The technology is detailed and the results of these two systems are displayed in next chapter.

5.1 The Design Approach

Interfaces that combine Reality and Virtuality are not new. However, Ischia summarizes the state of AR research when he says that AR researchers are primarily concerned with “the purely visual augmentation” rather than the form of the physical objects those visual augmentations are attached to [59]. If we are to create more usable AR interfaces then researchers must have a better understanding of design principles based on form as well as function.

In our augmented reality work we advocate designing the form of physical objects in the interface using established Tangible User Interface design methods. Some of the tangible design principles include:

- Object affordable should match the physical constraints of the object to the requirements of the task.
- The ability to support parallel activity where multiple objects or interface elements are involved at once.
- Support for physically based interaction techniques (such as using object proximity or spatial relations).
- The form of objects should encourage and support spatial manipulation.

Physical interface attributes are particularly important in interfaces designed to support face-to-face collaboration. In this case people commonly use the resources of the physical world to establish a socially shared meaning. Physical objects support collaboration both by their appearance, the physical affordance they have, their use as semantic representations, their spatial relationships, and their ability to help focus attention. In an AR interface the physical objects can further be

enhanced in ways not normally possible such as providing dynamic information overlay, private and public data display, context sensitive visual appearance, and physically based interactions. In the next two sections we describe how the Tangible Augmented Reality approach was applied in table-top AR experiences combined with small hardware controlling.

5.2 Tilt Pad–Accelerometer Based Interaction Device

5.2.1 Components Description

The main components required for interface designs in this work are as follows.

- Sensor for providing context information. In our case this will be ADXL202E, which is the accelerometer from Analog Devices.
- A processing and control unit which would serve as an arbiter of data transaction between the sensing device and the host computer system. In our case this will be BS2Sx micro-controller.
- Power regulation. In our case it is ADM666A from Analog devices.
- A wireless transmitter for data transmission between the control unit and the host device. In our case it is TX3 radio transmitter from Radiometrix.
- A wireless receiver for data reception at the host device. In our case it is RX3 radio receiver from Radiometrix.
- A simple communication protocol for establishing the link. In our case it is RS232 protocol.

Let us take a look at the devices used in the designs in some detail.

ADXL202E accelerometer

Here ADXL202E accelerometer is used as the sensor to sense the tilt information. Apart from acceleration measurement, accelerometers provide a simple and efficient means to detect tilt of bodies. When the accelerometer is perpendicular to gravity, its output changes by 17.5 mg per degree of tilt. So, when the accelerometer is oriented such that its axes are parallel to the earth's surface, it can be used as a two axis tilt sensor with a roll and a pitch axis. Once the output from the accelerometer is converted to acceleration that varies between -1 g and +1 g, the output tilt in degrees is calculated as follows, where 'P' and 'R' are the pitch and roll respectively.

$$P = \arcsin\left(\frac{A_x}{1g}\right) \quad (5.1)$$

$$R = \arcsin\left(\frac{A_y}{1g}\right) \quad (5.2)$$

BS2SX Micro-controller

It is one of the recent embedded micro-controllers from Parallax Inc. It has a Scenix processor as the core. It has an interpreter chip, internal program memory (32 bytes of scratchpad RAM and 16K of EEPROM, arranged as 8 pages of 2K each), a 5-volt regulator and 16 general purpose I/O's. This device measures the pulse width of the accelerometer output by polling the I/O pins and transmits the data to the host computer via the wireless transmitter in RS232 format.

AM666A regulator

ADM666A is a linear, precision voltage regulator with a maximum quiescent current of $9\ \mu\text{A}$, which makes it ideal for low power applications. It can operate in 3.3V/5V fixed mode or in adjustable mode with external resistors. The input range is from 2V to 16V with an output current of 100mA. It also has current limiting, thermal shutdown and low battery power detection features. It is used to provide a constant 5V supply to the micro-controller, sensors and the wireless transmitter.

TX3 wireless transmitter

The TX3 is a miniature UHF transmitter module, which allows implementation of data links at speeds up to 50 kb/s and distances up to 30m in buildings and 120m over open ground. It has a crystal locked PLL, and the output is frequency modulated. The operation voltage is from 2.2V to 12V at 7.5mA. The RF output power is 1mW at 3V supply. It comes in two fixed carrier frequencies: 869MHz for 64Kbps data rate and 914MHz for 50Kbps data rate.

RX3 wireless receiver

The RX3 module is a single conversion FM superhet receiver capable of handling data rates up to 50 kb/s. It can operate from a supply of 2.7V to 12V and draws 9.5mA when receiving. It has a SAW front end filter which provides good immunity to interference. It has a RSSI (Received signal strength indicator) output voltage which can be used to monitor the strength of the incoming signal. This signal can be used to detect EMI (Electro Magnetic Interference) from external devices also.

Table 5.1: Tilt pad packet structure

Byte Number	Description
1(Header)	1 (in decimal format)
2	X Axis
3	Y Axis of accelerometer 1
4	Y Axis of accelerometer 2

5.2.2 Hardware Description

The tilt pad is a device, which can provide information about the orientation of an object in a augmented reality environment and can be used as an unobtrusive interface. The tilt pad has two accelerometers mounted in the same plane 8 centimeters apart and their outputs are connected to 4 I/O ports of the BS2SX micro-controller. The separation is provided to provide a smoother tilt pattern in the applications. This is due to the fact that the edges has more angular sweep than the center of the pad. The BS2SX measures the ON time periods and transmits them in 8-bit, no parity, inverted form to the host computer. This tilt information is communicated wirelessly to the host computer using the TX3 and RX3 (914MHz carrier) IC's at 50Kbps. The data transfer from the tilt pad to the computer is done using the data packet structure as shown in Table 5.1. The tilt in x axis is the same for both the accelerometers. In the y axis the tilt is calculated by using the average of the two accelerometers. The optimal threshold for tilt in x and y axis to be detected was taken as ± 0.3 g. A typical plot of the tilt in x and y directions are shown in Figure 5.1. Due to sudden hand movements, unwanted dynamic acceleration is produced and they are filtered by calculating a simple moving average over 10 samples. The filtered results of the above test data is shown in Figure 5.2.

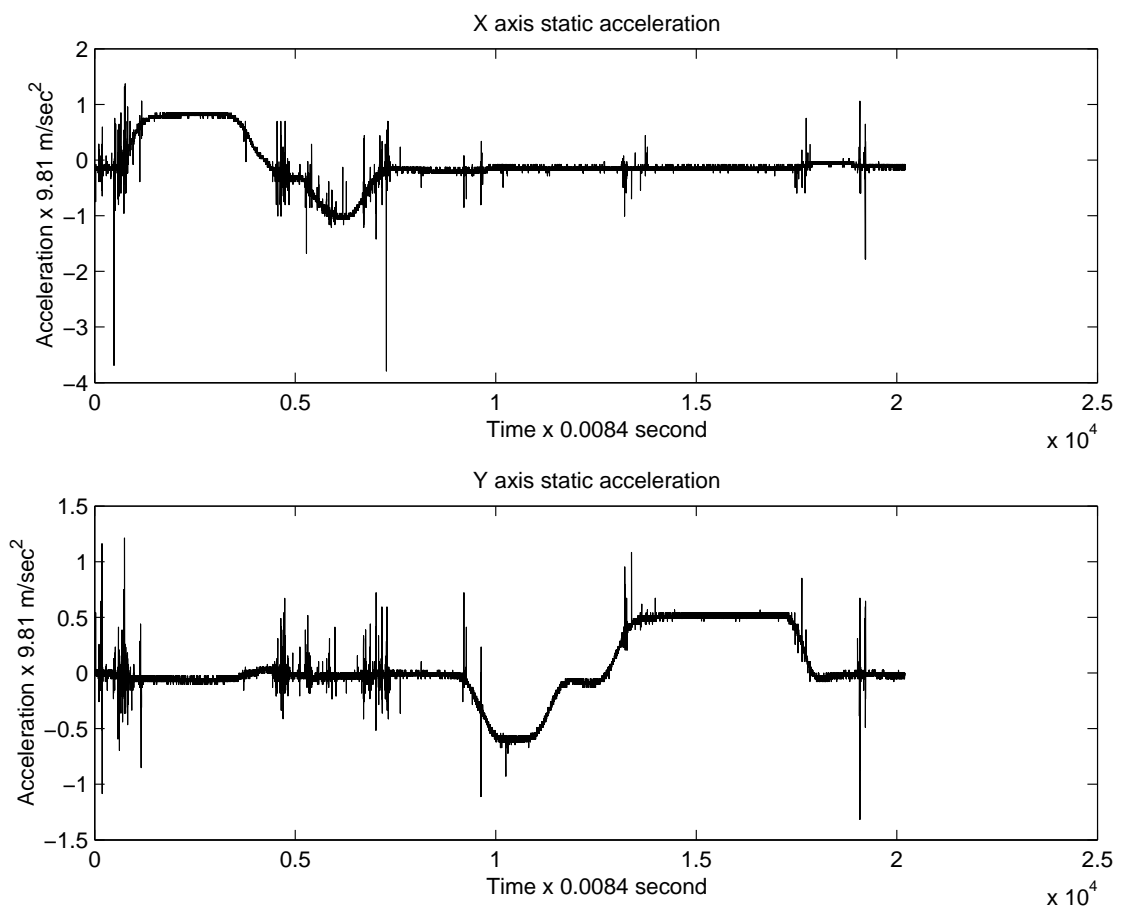


Figure 5.1: The tilt measurement in x and y axes

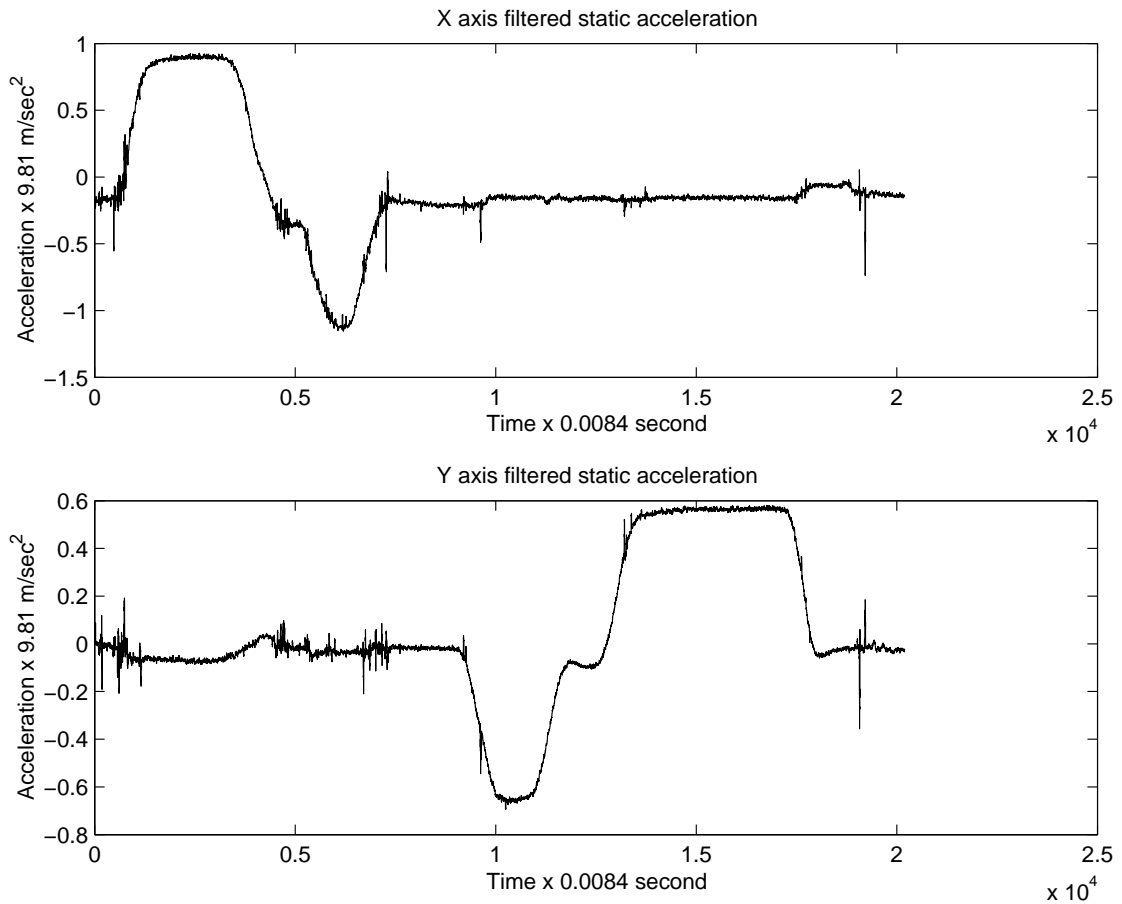


Figure 5.2: Filtered tilt measurement in x and y axes

Since the signal transmission from the transmitter TX3 and the receiver RX3 is wireless, tilt pad could control the digital information away from the computer at a certain distance, which enables its application in mobile computing and wearable computers. The TX3 transmitter on the tilt pad and the RX3 receiver connected to the PC through serial port are shown in schematic diagram in figure 5.3 and figure 5.4. And the tilt pad is shown in figure 5.5.

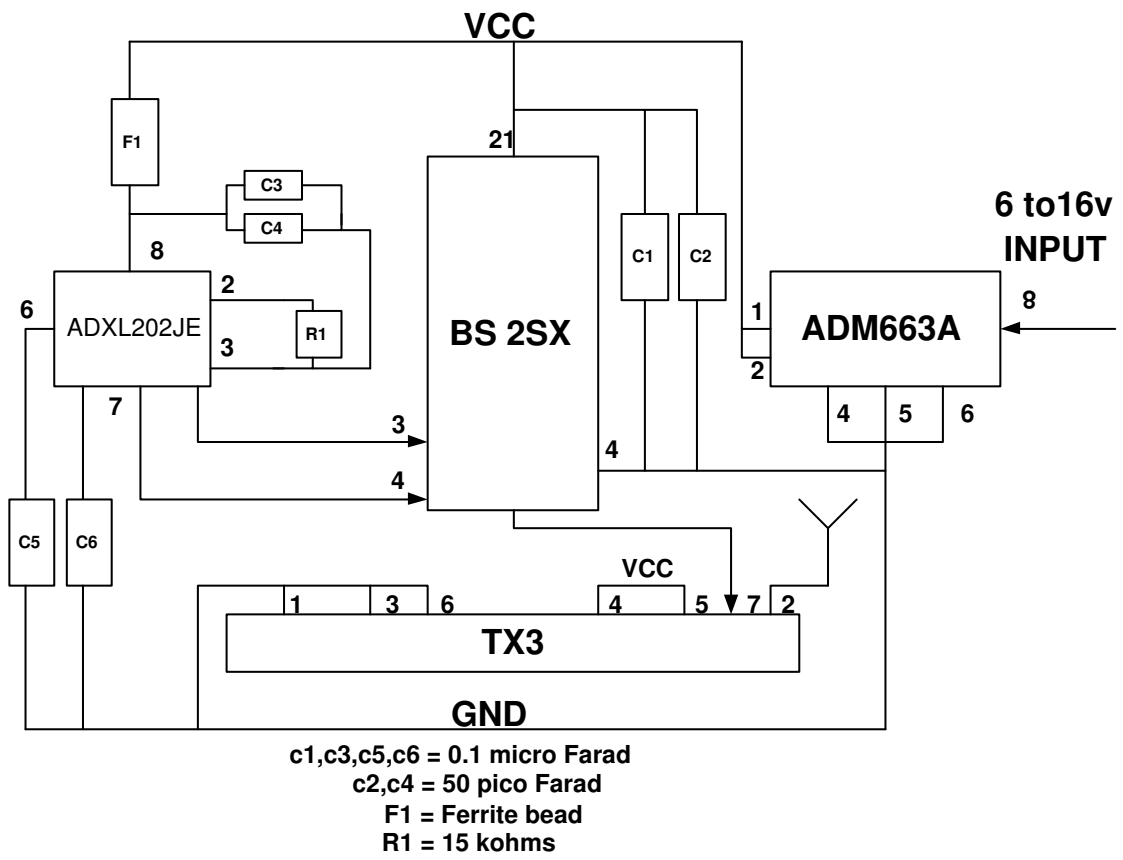


Figure 5.3: Schematic diagram of the transmitter part

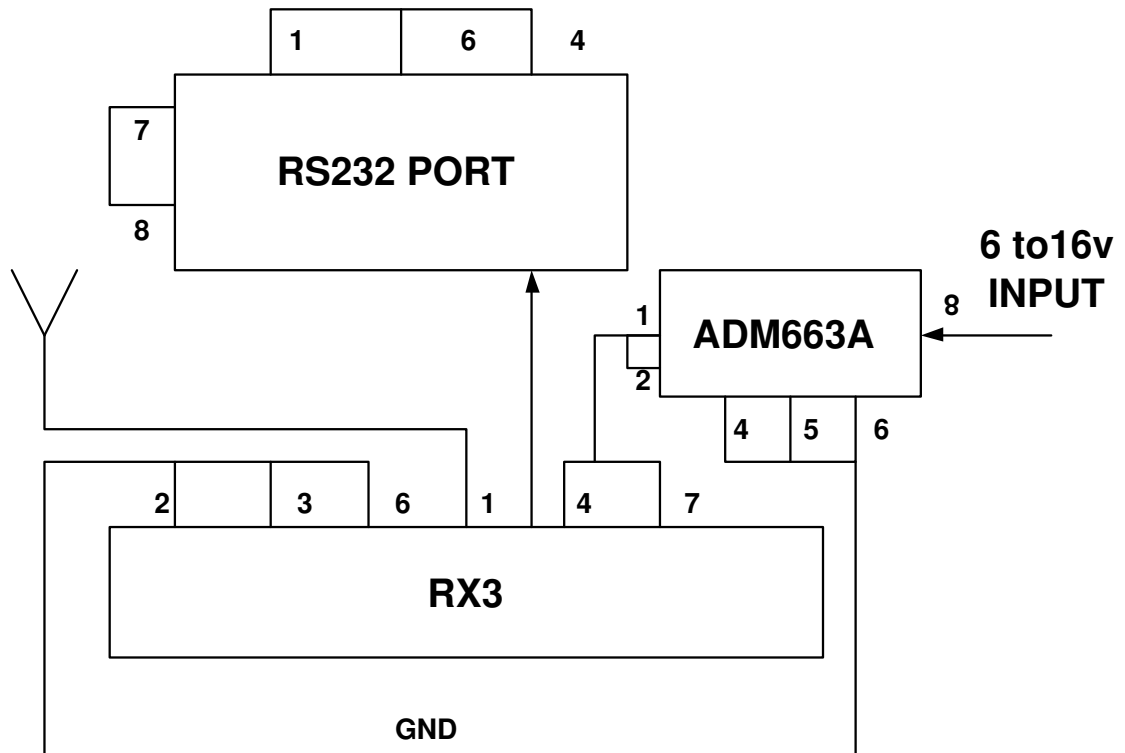


Figure 5.4: Schematic diagram of the Receiver part

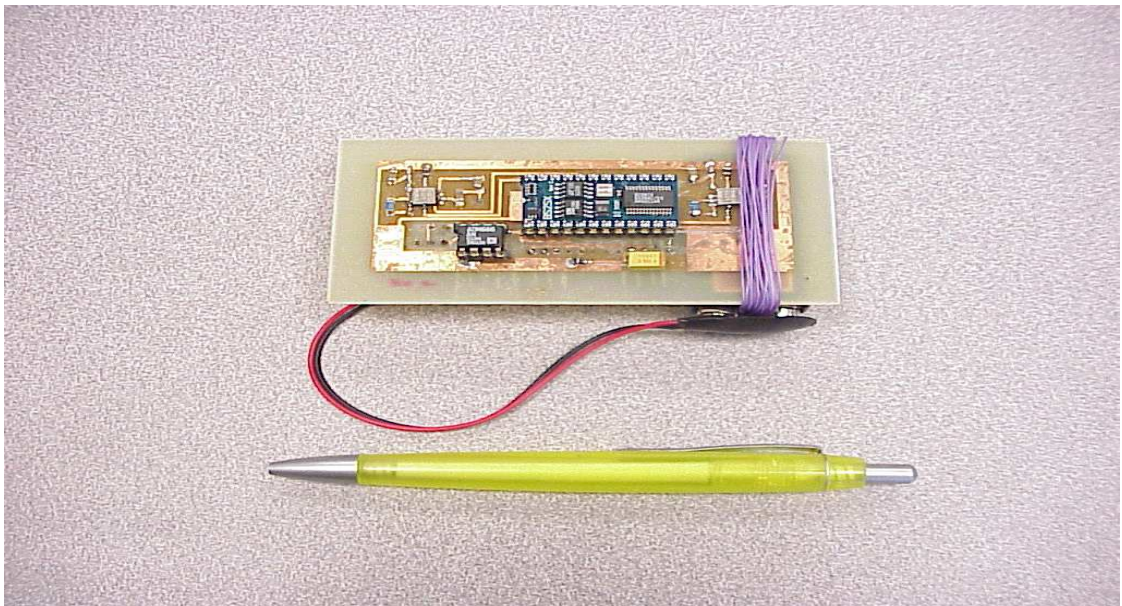


Figure 5.5: The tilt pad (10x5x2.5cm)

5.3 Tangible AR Map Scrolling System

In this section, we depict an AR three-dimensional map system physically manipulated by tilt pad, detailing the 3D map formation, augmented reality visualization, and hardware combination.

5.3.1 Set Up 3D Virtual Map

Based on the knowledge of 3D computer graphics, we define a 3D virtual fluctuant relief map as NURBS surface with full texture, and realize the partly-display of the surface in order to allow scrolling. More technology is detailed here below.

4.3.1.1 Representation of the map surface

Here we use the parametric representation method to get our desired curved surfaces, which are basically “curved quadrilaterals” with control points to specify the exact shape of the surface. Let’s explain this technology of surface modelling from the representation of curves since surfaces are said to be generated from the Cartesian product of two curves.

B-spline curves

A B-spline is a complete piecewise cubic polynomial consisting of any number of curve segments. It is a cubic segment over a certain interval, and going from one interval to the next, the coefficients change. The B-spline formulation is:

$$\mathbf{Q}_i(u) = \mathbf{U}\mathbf{B}_s\mathbf{P} \tag{5.3}$$

$$\mathbf{Q}_i(u) = \begin{bmatrix} u^3 & u^2 & u & 1 \end{bmatrix} \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_{i-3} \\ \mathbf{P}_{i-2} \\ \mathbf{P}_{i-1} \\ \mathbf{P}_i \end{bmatrix} \tag{5.4}$$

where \mathbf{Q}_i is the *i*th B-spline segment and \mathbf{P}_j is a set of four points in a sequence of control points. Alternatively we can write:

$$\mathbf{Q}_i(u) = \sum_{k=0}^3 \mathbf{P}_{i-3+k} \mathbf{B}_{i-3+k}(u) \tag{5.5}$$

where *i* is the segment number and *k* is the local control point index – that is the index for the segment *i*. The value of **u** over a single curve segment is $0 \leq u \leq 1$. Using this notation we can describe *u* as a local parameter – locally varying over

the parametric range of 0 to 1 – to define a single B-spline curve segment.

Thus in this notation we see that a B-spline curve is a series of $m - 2$ curve segments that we conventionally label Q_3, Q_4, \dots, Q_m defined or determined by $m + 1$ control points $P_0, P_1, \dots, P_m, m \geq 3$. Each curve segment is defined by four control points and each control point influences four and only four curve segments. This is the local control property of the B-spline curve and its main advantage. The property of sharing four control points for each segment is shown in figure 5.6.

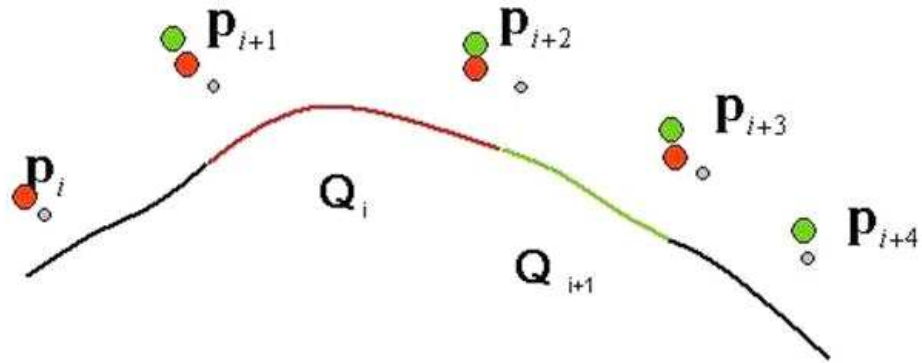


Figure 5.6: The red colored segment is determined by the four red colored control points, the green colored segment is determined by the four green colored control points

We now define the entire set of curve segments as one B-spline curve in u :

$$Q(u) = \sum_{i=0}^m P_i B_i(u) \tag{5.6}$$

In this notation i is now a non-local control point number and u is a global parameter discussed in more detail below.

Non-uniform B-splines

Equation 4.2 shows that each segment in a B-spline curve is defined by four basis functions and four control vertices. Hence there are three more basis functions

and three more control vertices than there are curve segments. The join point on the value of u between segments is called the knot value and a non-uniform B-spline is a curve where the parametric intervals between successive knot values are not necessarily equal. This implies that the basis functions ($B_i(u)$) are no longer translates of each other (which is true in the case of Uniform B-splines) but vary from interval to interval. The most common form of a non-uniform B-spline is where some of the intervals between successive knot values are reduced to zero by inserting multiple knots.

At this stage we can summarize and state that a B-spline curve is made up of $m - 2$ segments defined by the position of $m + 1$ basis functions (also $m + 1$ control points) over $m + 5$ knot values. In practice the knot sequence $[0,0,0,0,1,2,\dots,n-1,n,n,n,n]$ is often used for Non-uniform B-splines, where $n = m - 2$. That is, interpolation is forced at the end points but uniform knots are used elsewhere.

Summary of B-spline curve properties

The properties of B-spline curves are listed, in particular:

- The curve follows the shape of the control point polygon and is constrained to lie in the convex hull of the control points.
- The curve exhibits the variation diminishing property.
- The curve is transformed by applying any affine transformation to its control point representation.
- A B-spline curve exhibits local control - a control point is connected to four segments (in the case of cubic) and moving a control point can influence only these segments.

NURBS curves

A rational curve is a curve defined in four-dimensional space—known as projective space—which is then projected into three-dimensional space. NURBS stands for Non-uniform Rational B-splines and this representation is probably the most popular in CAD. In design applications it admits the following possibilities:

- Interactive placement and movement of control points.
- Interactive placement and movement of knots.
- Interactive control of control point weights.

A NURBS curve is defined as a non-uniform B-spline curve on a knot vector where the interior knot spans are not equal. For example, we may have interior knots with multiplicity greater than 1 (that is knot spans of length zero). Some common curves and surfaces, such as circles and cylinders, require non-uniform knot spacing, and the use of this option generally allows better shape control and the ability to model a larger class of shapes.

Rational B-splines are defined simply by applying the B-spline equation (Equation 4.3) to homogeneous coordinates, rather than normal 3D coordinates. The transformation from These 4D coordinates to 3D coordinates is:

$$(x', y', z', w) \rightarrow \left(\frac{x'}{w}, \frac{y'}{w}, \frac{z'}{w}\right) \quad (5.7)$$

And the inverse transform is:

$$(x, y, z) \rightarrow (xh, yh, zh, h) \quad (5.8)$$

where \mathbf{h} is an additional dimension introduced by homogeneous coordinate to a Cartesian coordinate, usually being unity.

Thus our 3D control point, $\mathbf{P}_i = (x_i, y_i, z_i)$, becomes the homogeneous control point, $\mathbf{C}_i = (x_i h_i, y_i h_i, z_i h_i, h_i)$. A NURBS curve is thus defined as:

$$\mathbf{R}_H(u) = \sum_{i=0}^n \mathbf{B}_{i,k}(u) \mathbf{C}_i \quad (5.9)$$

The perspective map of such a curve in three-dimensional space is defined as:

$$\mathbf{R}(u) = \frac{\sum_{i=0}^n \mathbf{B}_{i,k}(u) \mathbf{P}_i h_i}{\sum_{i=0}^n \mathbf{B}_{i,k}(u) h_i} \quad (5.10)$$

NURBS surfaces

The treatment of parametric cubic curve segments given above is easily generalized to bi-parametric cubic surface patches. A point on the surface patch is given by a bi-parametric function and a set of blending or basis functions is used for each parameter. A bi-cubic B-spline surface patch can be derived by equation 4.3 as:

$$\mathbf{Q}(u, v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{ij} B_{ij}(u, v) \quad (5.11)$$

where \mathbf{P}_{ij} is an array of control points and $B_{i,j}(u, v)$ is a bivariate basis function.

We can generate $B_{i,j}(u, v)$ from:

$$B_{i,j}(u, v) = B_i(u) B_j(v) \quad (5.12)$$

where $B_i(u)$ and $B_j(v)$ are the previously defined univariate cubic B-splines. Thus, we have:

$$\mathbf{Q}(u, v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{ij} B_i(u) B_j(v) \quad (5.13)$$

Accordingly, a NURBS surface of degree (p, q) is defined by:

$$\mathbf{S}(u, v) = \frac{\sum_{i=0}^m \sum_{j=0}^n B_{i,p}(u) B_{j,q}(v) h_{i,j} \mathbf{P}_{i,j}}{\sum_{i=0}^m \sum_{j=0}^n B_{i,p}(u) B_{j,q}(v) h_{i,j}} \quad (5.14)$$

where $B_{i,p}$ and $B_{j,q}$ are the B-spline basis functions, $\mathbf{P}_{i,j}$ are control points, and the weight $h_{i,j}$ of $\mathbf{P}_{i,j}$ is the last coordinate of the homogeneous point $\mathbf{P}_{i,j}^w$.

Drawing of NURBS map surface

From the mathematical principle [60] of NURBS surfaces, we can see that a NURBS surface is determined if we know

- the control point net with $n \times m$ vertices $\mathbf{P}_{i,j}, i = 1, \dots, n, j = 1, \dots, m$.
- the knot vector U and V .
- the weight vector W .

For rendering of this map surface, we use OpenGL Programming interface [61]. This interface consists of about 150 distinct commands that can be used to specify the objects and operations needed to produce interactive three-dimensional applications. We call `gluNurbsSurface()` to generate and render a NURBS surface. And the input parameters include the control points, knot sequence, and order of the polynomial basis function for our NURBS object. The diagram (figure 5.7) shows the process of the rendering in OpenGL.

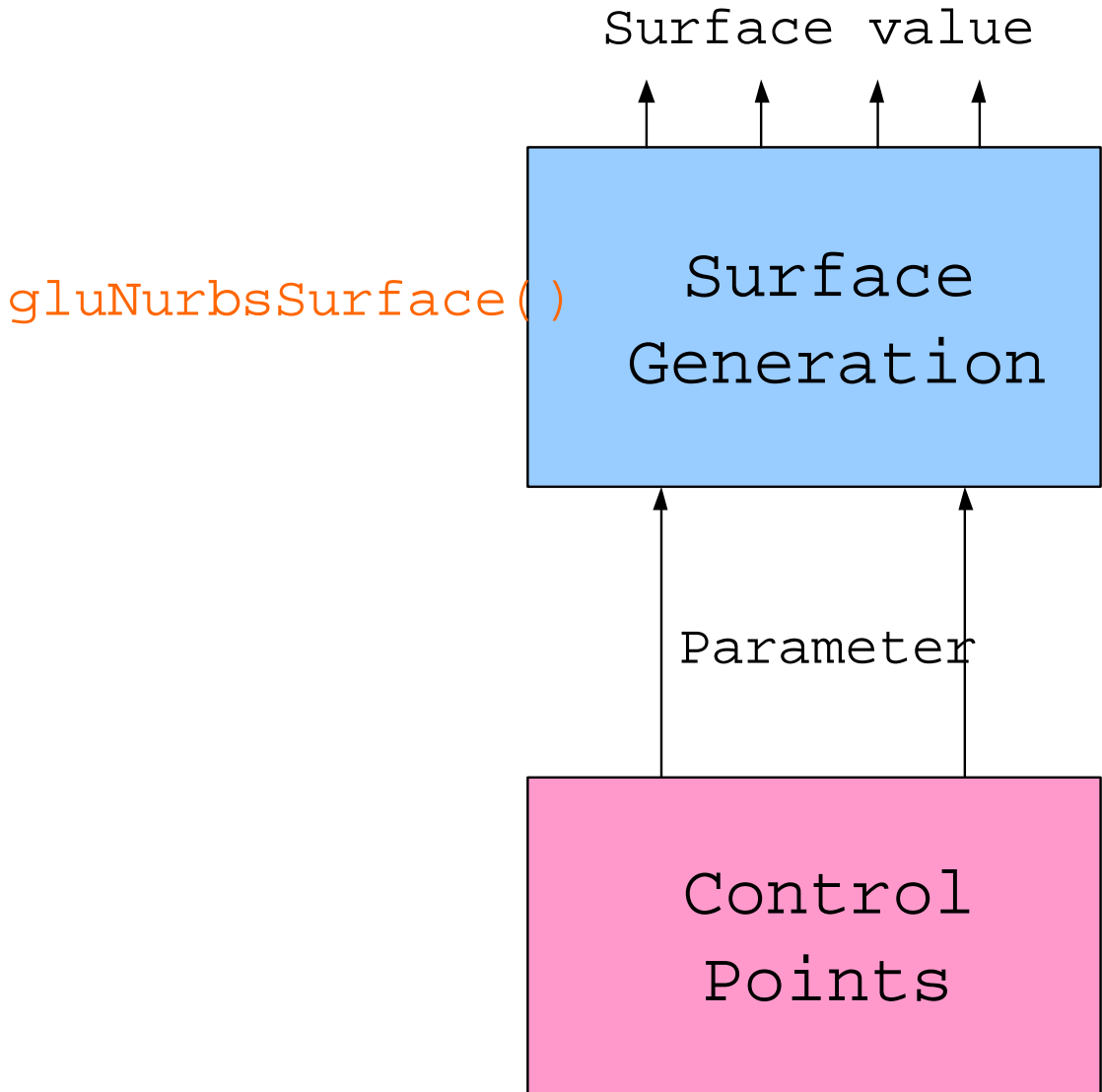


Figure 5.7: Process of NURBS rendering in OpenGL

Concerning our map surface, the orders of the two basis functions are both four. And we use a 10×10 size matrix as the control points set, where each entry of the 14×14 size matrix is the 3D coordinate of one control point. Thus the knot number of each parametric space (u and v) is equal to No of control points + order = $10 + 4 = 14$. And we apply the knot sequence as $[0, 0, 0, 0, 1, 2, 3, 4, 5, 6, 7,$

7, 7, 7] as we described before in mathematical principles. The following flowchart (figure 5.8) displays the programming sequence of drawing a NURBS surface used in our project.

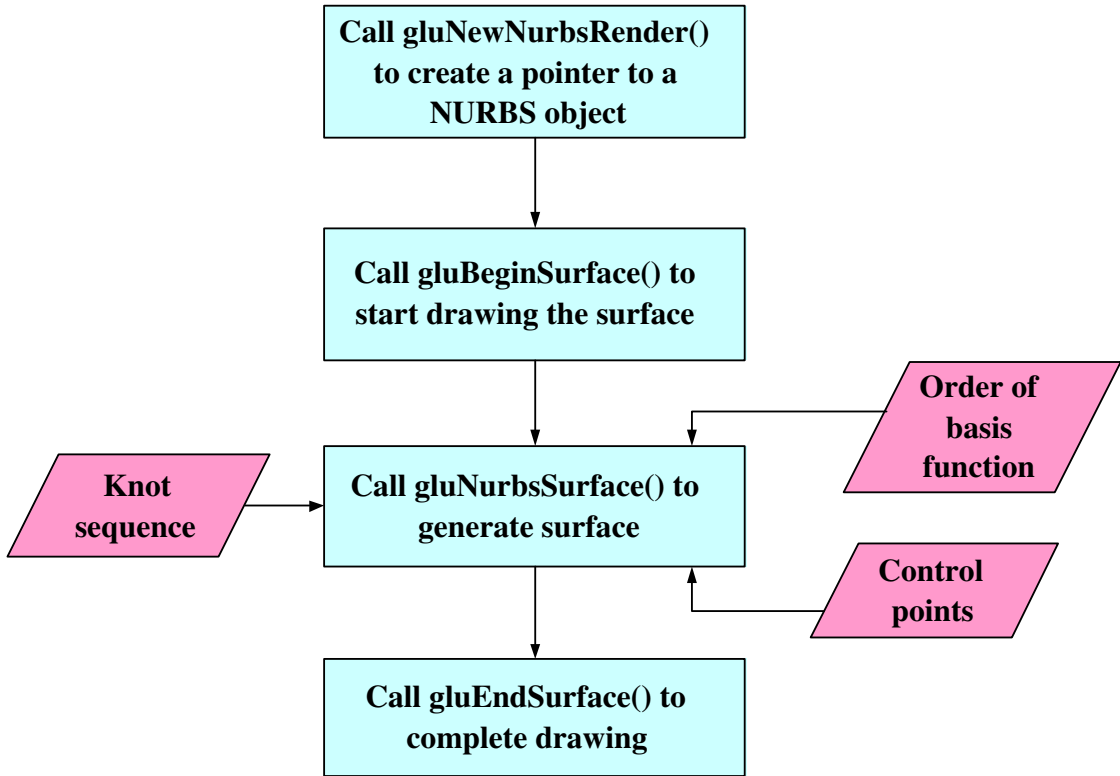


Figure 5.8: Code flowchart of drawing map surface in OpenGL

4.3.1.2 Texture of the map surface

So far, we have drawn the 3D map without texture mapping. Now we are going to deal with texture mapping issue that could make the 3D scenery more visually interesting, realistic or esoteric. Textures are simply rectangular arrays of data - for example, color data, luminance data, or color and alpha data. The individual values in a texture array are often called texels. What makes texture mapping tricky is that a rectangular texture can be mapped to non-rectangular regions, and

this must be done in a reasonable way.

To use texture mapping, we perform these steps.

- Create a texture object and specify a texture for that object.
- Indicate how the texture is to be applied to each pixel. We can choose any of four possible functions for computing the final RGBA value from the fragment color and the texture-image data. the way we adopted is simply to use the texture color as the final color; this is the *decal* mode, in which the texture is painted on top of the fragment.
- Enable texture mapping before drawing the scene. Texturing is enabled or disabled using `glEnable()` or `glDisable()` with the symbolic constant `GL_TEXTURE_1D` or `GL_TEXTURE_2D` for one or two-dimensional texturing, respectively.
- Draw the scene, supplying both texture and geometric coordinates. It is needed to specify both texture coordinates and geometric coordinates as specify the objects in the scene. For a two-dimensional texture map, for example, the texture coordinates range from 0.0 to 1.0 in both directions, but the coordinates of the items being textured can be anything.

For the 3D virtual map texturing, we have several textures consisting of different zooming images and the airscape. And since we are now texturing a NURBS surface, command `gluNurbsSurface()` also suggests a way to calculate and assign texture coordinates by choosing proper parameter. We uses a loop to assign the value in domain $[0.0, 1.0]$ to the texture control points in order to map the whole texture image to the NURBS surface. In this case, our original texture images are in jpeg format, a standard compression format of images. Therefore, we have

to decompress these jpeg files first, which means to rewrite these files to a two-dimensional matrix holding the RGB value of each pixel. This target is achieved by writing a function named *myRead_JPEG_file()* in C programming language. The following flowchart (figure 5.9) shows the texture mapping approach used here. The textured 3D map is displayed in next chapter as the result of this system.

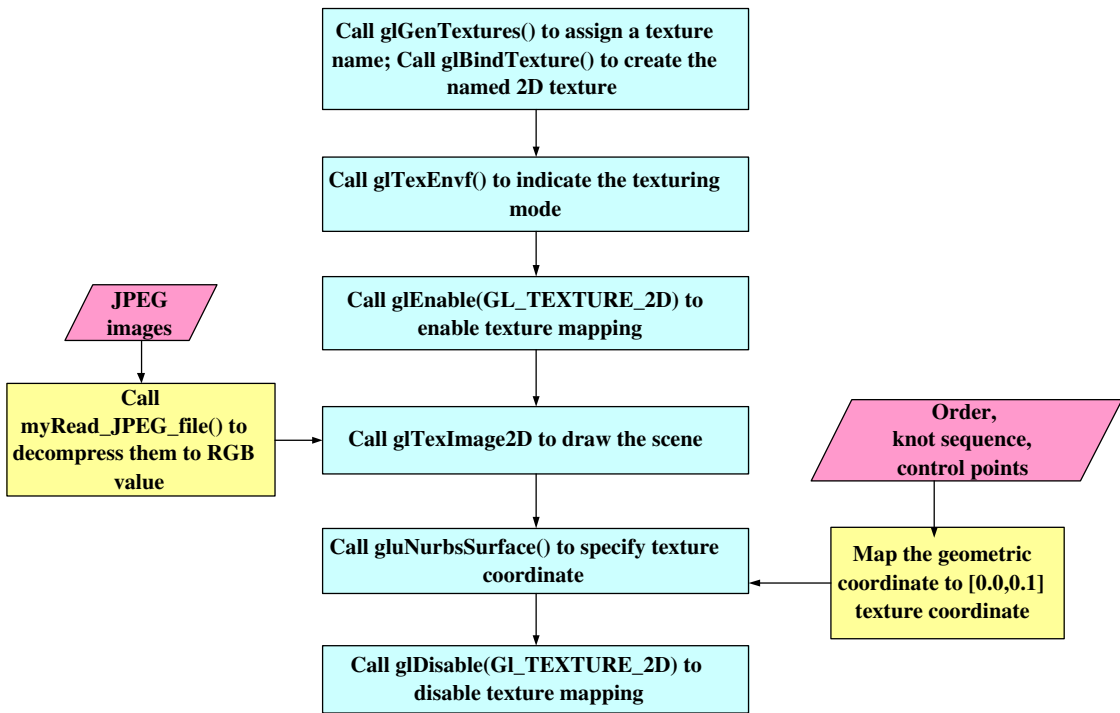


Figure 5.9: Flowchart of texture mapping of the 3D virtual map

4.3.1.3 Scrolling of the map

By far, we have established the 3D virtual map and got it textured. But it is still a static map if without changing its shape while scrolling. Now we are going to discuss the technique about this issue.

Change of the surface's 3D shape

We have explained that the 3D map is defined as a NURBS surface and its shape is determined by the 3D coordinate value of the control points in the previous contents. We also stated that the control points of the map have a dimension of 10×10 , which means that there are 100 control points stored in a 10×10 matrix. Therefore, in order to get different shape of the map surface when scrolling, we need to only change the values of the control points, in another word, to change the value of this 10×10 matrix. That is also the prominent advantage of adopting NURBS as the representation of the map surface due to realizing full control conveniently. And the trick here is that although every entry of this 10×10 matrix has three values of the (x, y, z) coordinates, it is only the third value z needed to be changed since it stands for the depth information. We define a large 35×35 matrix to store all the depth information for the control points and get its 10×10 subset every time for display. The illustration of this scrolling technology is shown in figure 5.10.

M_L is the large matrix with dimension 35×35 . M_S is the 10×10 matrix that holds the control point values. M_S moves in M_L to get distinct values. When the map scrolling to the right, it moves to the right. When the map scrolling to the bottom, it moves down. And the same for left and up. When it moves to the edge of the large matrix, movement is set to stop until scrolling back to the opposite direction is activated, which simulates the situation when scrolling to the edge of the map.

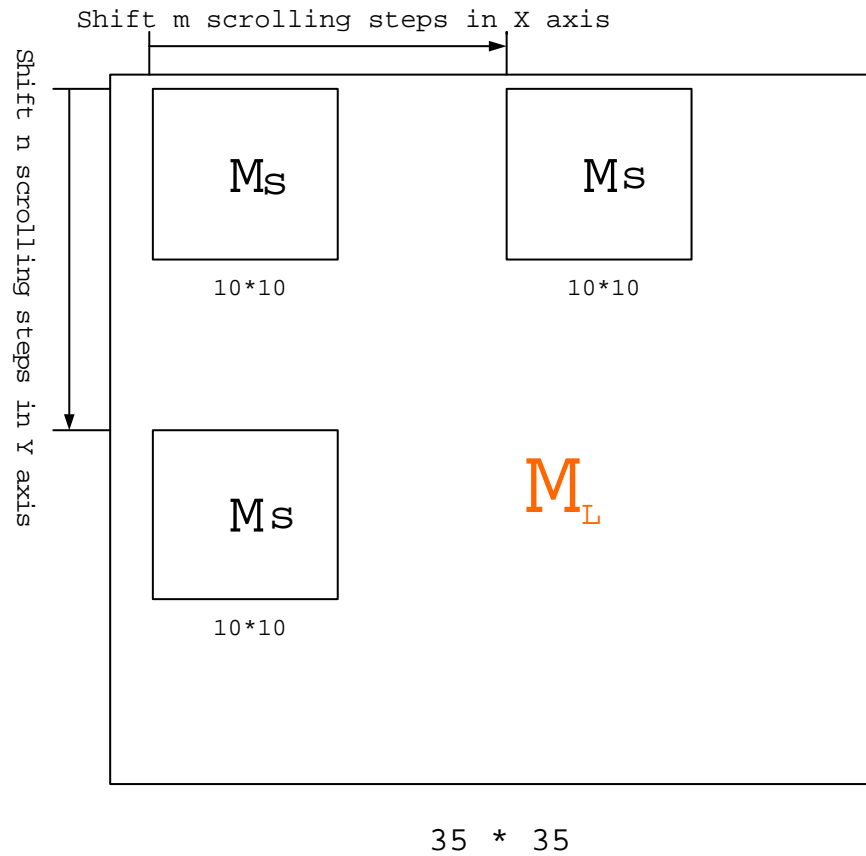


Figure 5.10: Diagram of changing control point values when scrolling

Change of the texture

In addition to changing the 3D shape of the surface, it is necessary to shift the texture of the surface simultaneously during scrolling of the map in order to get acceptable and convincing visual effects. It has been introduced that our original texture image is in JPEG format. And it has 1024×1024 pixels. What we do is to use part of this image as texture once, say 512×512 pixels, and shift this small window when scrolling. Basically this is the same method as 3D shape change. But the focus of the problem is to ensure that the shape of the surface and the texture of the surface shift synchronously, which means that the texture is immobile relative to

the surface when they move together as the situation we scroll a real physical map. After repeated experiments, the synchronization is achieved by shifting 26 pixels for the change of one control point. The flowchart (figure 5.11) of the algorithm to scroll the map is given here. The screen shots of map scrolling are presented in the next result chapter and a video of this is attached in the accompanying CR-ROM of this thesis.

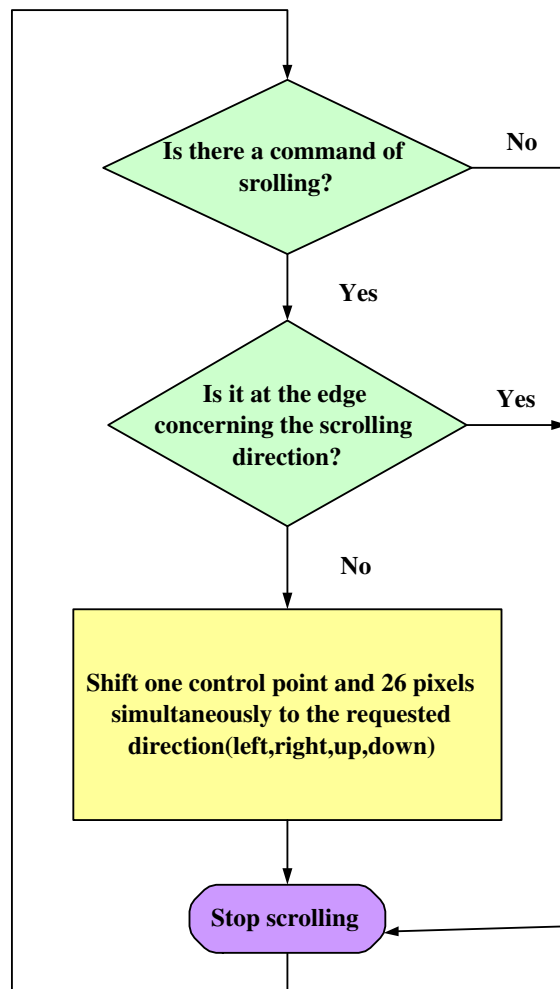


Figure 5.11: Flowchart of the code to realize map scrolling

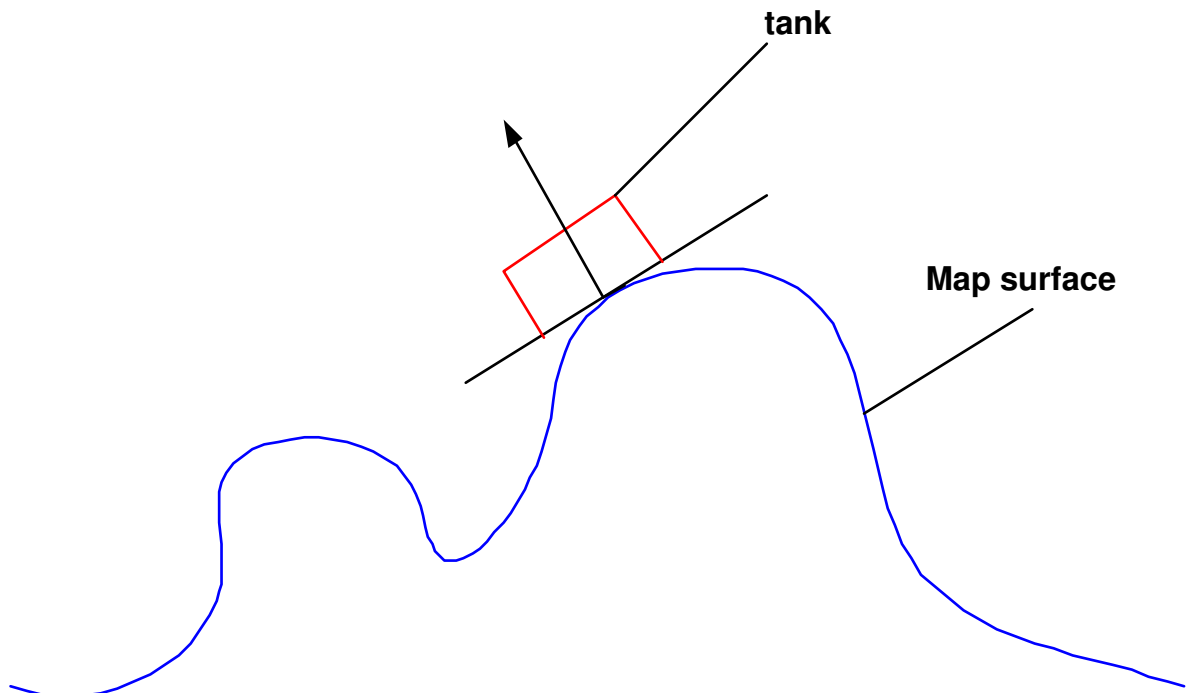


Figure 5.12: Orientation of the tank on the map surface

4.3.1.4 Defining of the tank

By far, we have develop a complete 3D virtual map with full texture and realized the map scrolling. To make the map more interesting and multi-functional, a tank is added to the map surface. Unlike the whole map surface that is generated from OpenGL, this tank is a VRML format virtual object loaded in. It can drive along the 3D landform with proper orientation like a real vehicle does by ensuring that the perpendicular of the map surface and the perpendicular of the tank's body are parallel (figure 5.12). At the same time, the animation of the tank is created. The tank is defined to keep running around a circle route by the following equation:

$$u = r \cos \theta + U_C \quad (5.15)$$

$$v = r \sin \theta + V_C \quad (5.16)$$

where (u, v) are the image coordinate of the tank and (U_C, V_C) is the coordinate of the circle. Accordingly, synchronization is also required for the tank when scrolling the map. This means that the center point of the circle route is fixed to the map and when scrolling the tank will follow moving until out of sight. This could be realized by using the same control variable for map scrolling to change the coordinate of the tank in a proper step. The flowchart of this part of the code is listed in figure 5.13.

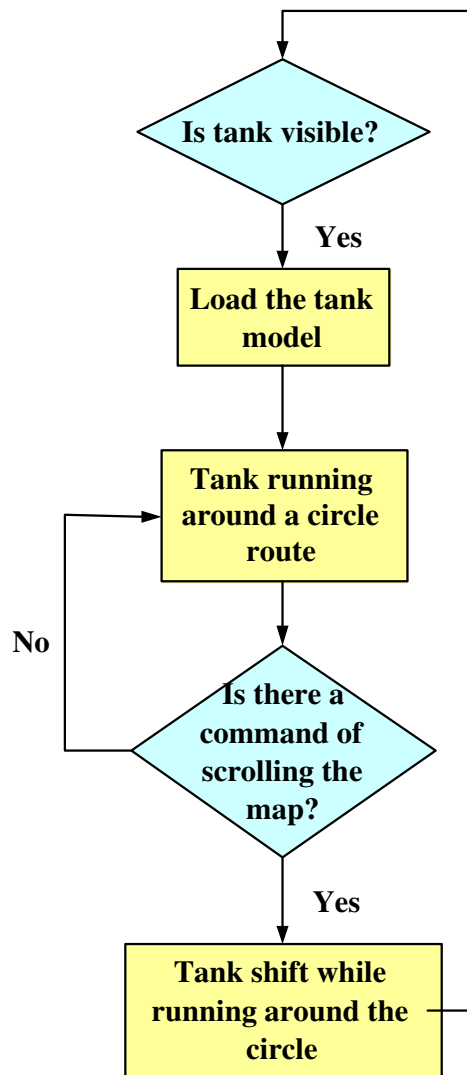


Figure 5.13: Flowchart of the code to define the tank

5.3.2 Visualization in AR Environment

Till now, we have discussed the development of the virtual 3D map issue. In this subsection, how this map is displayed in AR environment is explained. Basically, this consists of the multiple marker tracking and the ARToolKit visualization process.

Multiple marker tracking

It has been described how fiducial marker is used for the purpose of tracking in ARToolKit in chapter 2. However it uses a single relatively large square marker as a fiducial. So if a hand or other object to even partially overlapped the fiducial the tracking was lost. This decreased the robustness of tracking under the conditions where a hand could overlap the fiducials. Also if there is some distance between tracked fiducials and displayed virtual objects, tracking errors strongly influence the registration accuracy. That is, using a single fiducial decreases the accuracy of registration under the conditions where virtual objects need to be displayed around on the table.

To solve this problem, we have developed a new tracking method in which multiple markers are used as fiducials and pose and position are estimated from all of the detected fiducial marks. This means that many of the fiducial, as long as not all of them, can be covered up without losing tracking. This is critically important, as in a typical tangible AR application, it's not avoidable that the user's hand will cover some of the fiducials when he/she is interacting with the virtual objects.

If there is a fiducial marker in the image, it is possible to estimate 3D pose and position using our earlier method in ARToolKit. However if there is more than one fiducial markers are visible, we can achieve more robust tracking if we estimate pose

from all of available features. In order to do this we adopt following procedures:

(step 1) The biggest fiducial marker is selected in the image. 3D pose and position are initially estimated from it. This information is represented as the following transformation function from marker coordinates to camera coordinates:

$$(x_c, y_c, z_c) = trans(x_w, y_w, z_w) \quad (5.17)$$

where $(\mathbf{x}_w, \mathbf{y}_w, \mathbf{z}_w)$ is a position in world coordinates and $(\mathbf{x}_c, \mathbf{y}_c, \mathbf{z}_c)$ is the same position in camera coordinates.

(step 2) The positions of all other fiducial markers are then estimated in screen coordinates by using the above transformation function, a projective function and the 3D positions of these markers in the world coordinates:

$$(x_s, y_s) = perspect(trans(x_w, y_w, z_w)) \quad (5.18)$$

where the function *perspect* is a projective function. This function consists of perspective projection parameters and image distortion parameters.

(step 3) The actual screen coordinates of the detected markers are compared to the estimated positions. Using the positions of all successfully matched fiducial markers, the 3D pose and position are re-estimated. For this calculation, the initial transformation function is used and modified as the amount of errors between the actual feature positions in the image and the estimated positions goes to minimum using a hill-climbing method.

The result of this proposed method is shown in Figure 5.14. As the user moves his hand, and covers some of the fiducials on the desk, the virtual object augmented on the table surface is still correctly registered. However, because only one camera

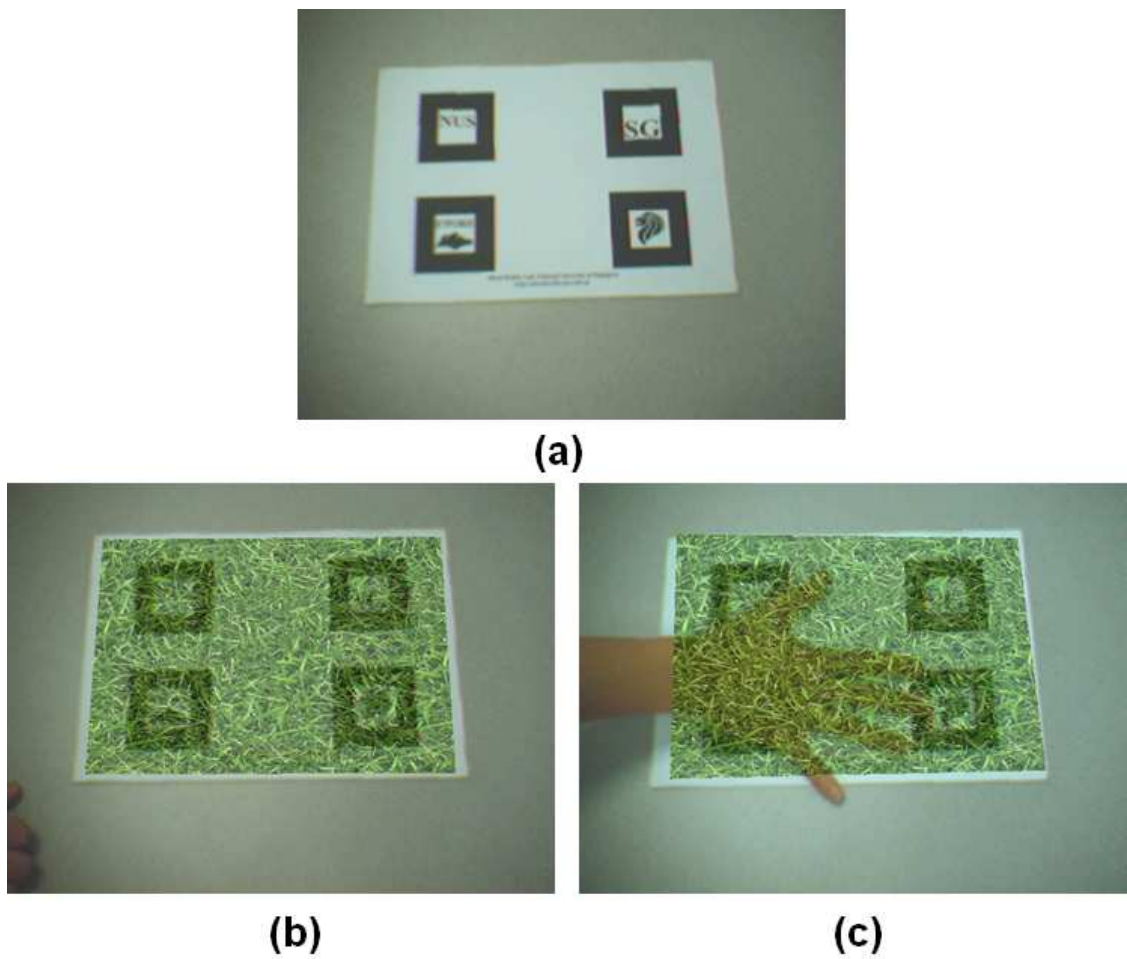


Figure 5.14: Tracking of multiple fiducials. (a) The four different fiducials are printed on the same paper. (b) A virtual grass is augmented onto the fiducials. (c) Even some of the fiducials are blocked, the correct registration still remains.

is used, the depth information is not available, problem with incorrect occlusion results. As shown in Figure 5.14-c, the virtual grass incorrectly appears in front of the hand. It is possible to solve this problem by getting depth information from stereo cameras. However, it will also increase the computational load. Because of the objective of this thesis work is to build a realtime AR system for normal desktop PC, or even laptop computers and small wearable computers, we decided to take the mono-view approach.

Visualization in ARToolKit

After solving the tracking technology, it is ready now to visualize the virtual 3D objects above on the marker. ARToolKit [34] is applied. Figure 5.15 on the next page shows this image processing approach.

Steps 2 through 5 are repeated continuously until the application quits, while steps 1 and 6 are just done on initialization and shutdown of the application respectively. In addition to these steps the application may need to respond to mouse, keyboard or other application specific events. In our case, these events include the tilting of the pad, as well as keyboard.

5.3.3 Interface with Tilt Pad

In this subsection, we are going to introduce the interface technology to involve hardware tilt pad in controlling the digital information. After the software captures the video from the camera and detects the fiducials. The serial port of the host computer is initialized with the necessary settings for interrupt based, 8-bit, no parity reception. When a byte is received, it is checked whether it is the header. If true, then the following packets are stored in an array. Until the next interrupt

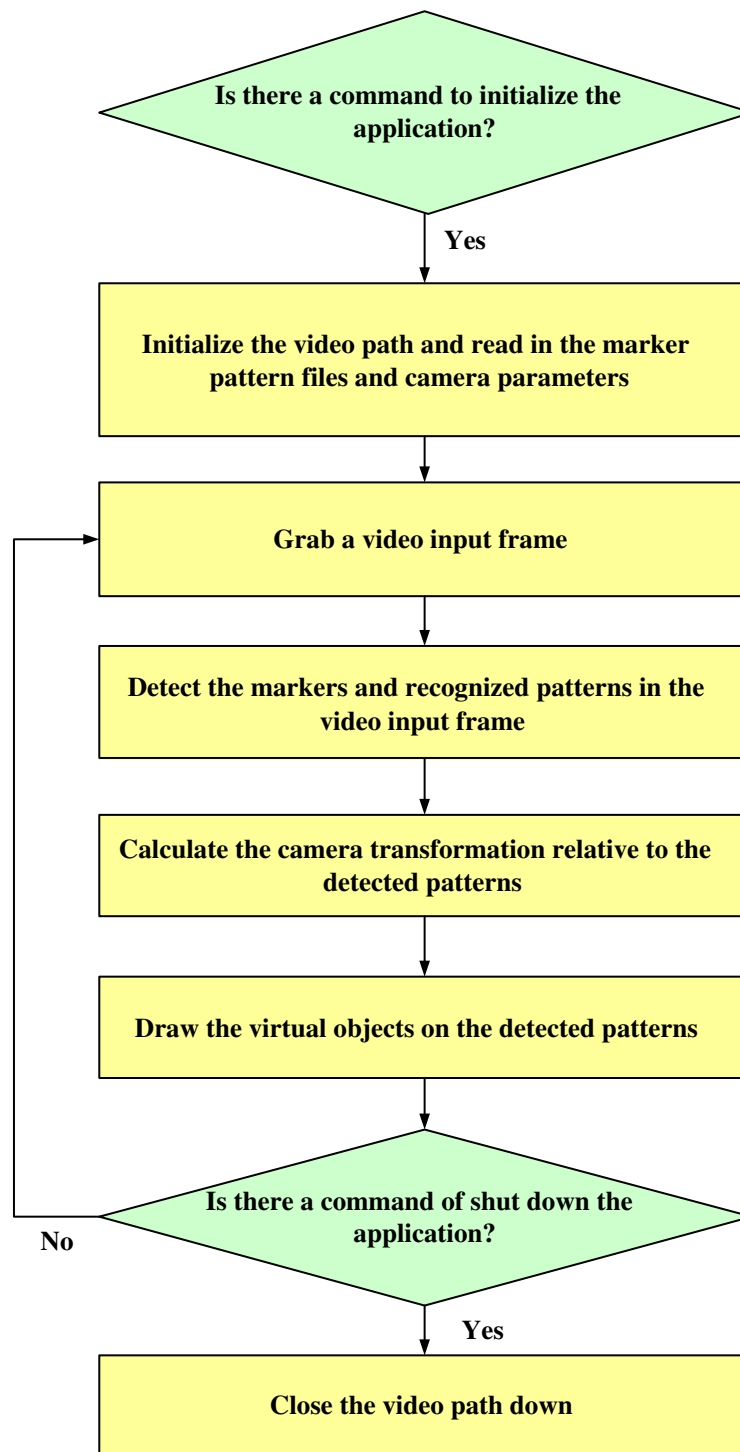


Figure 5.15: Flowchart of the image processing of ARToolKit

arrives, the data is then used to calculate the average tilt in x and y directions. Then the data is moving averaged over previous 9 samples. If the x tilt value is greater than 0.3g then a flag for tilt in +x direction is set, and if it is lesser than -0.3g, the flag is cleared. Similarly, if the y tilt value is greater than 0.3g then a flag for tilt in +y direction is set, and if it is lesser than -0.3g, the flag is cleared. This information is used to overlay augmented images over the fiducial to the corresponding position and if continuous tilting is sensed, the virtual object moves in the virtual space correspondingly. The resulting graphics is sent to the HMD. The software that is implemented on the tilt-pad as well as the receiver computer software to process the tilt data is outlined in Figure 5.16.

To summarize all the procedures of develop this augmented reality 3D map scrolling system interfaced with tilt pad, we list the integrated code flowchart (figure 5.17). The loop part of the code is executed once every frame time.

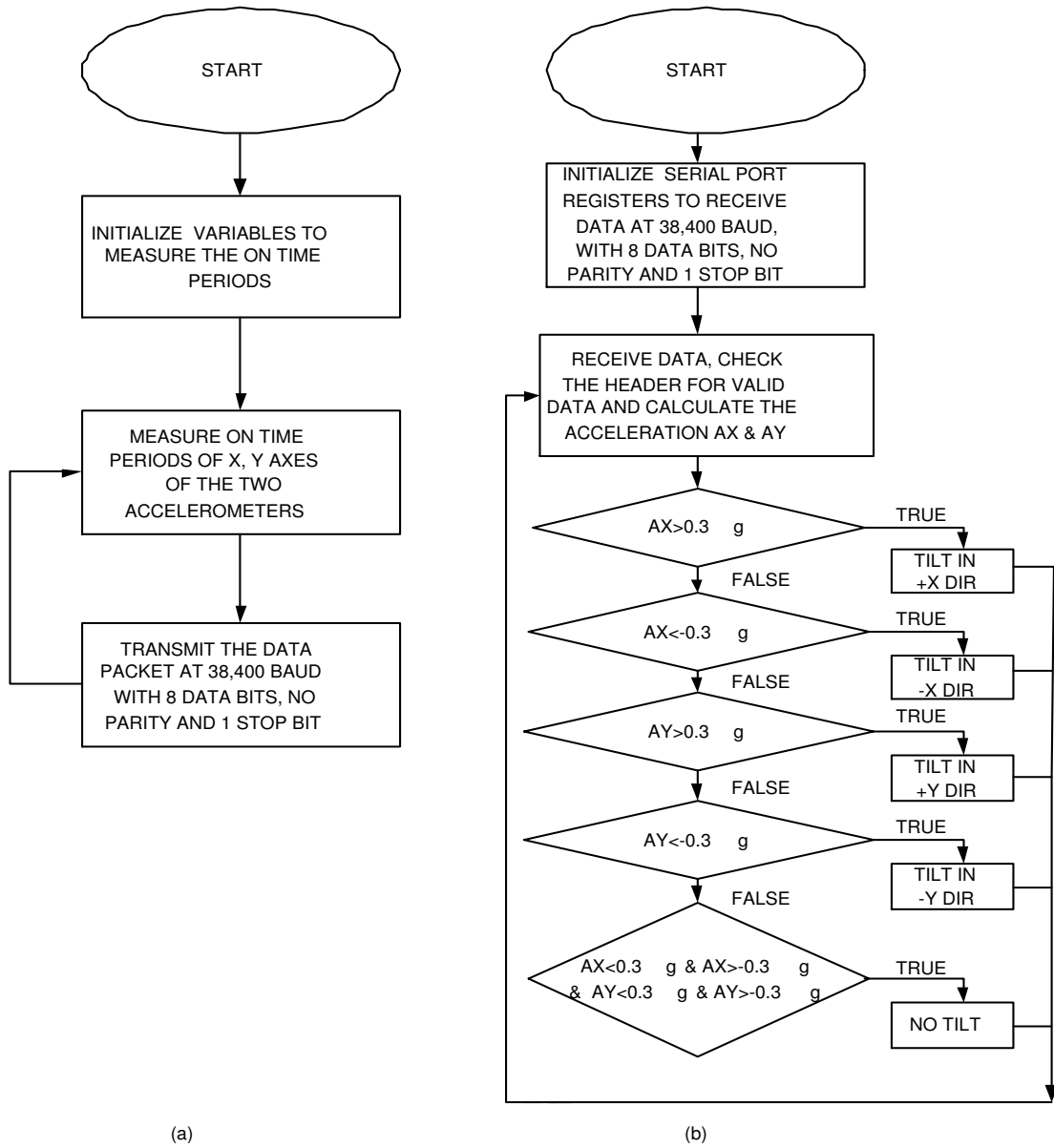


Figure 5.16: (a) Software flowchart on the tilt pad side (b) Software flowchart on the host computer side

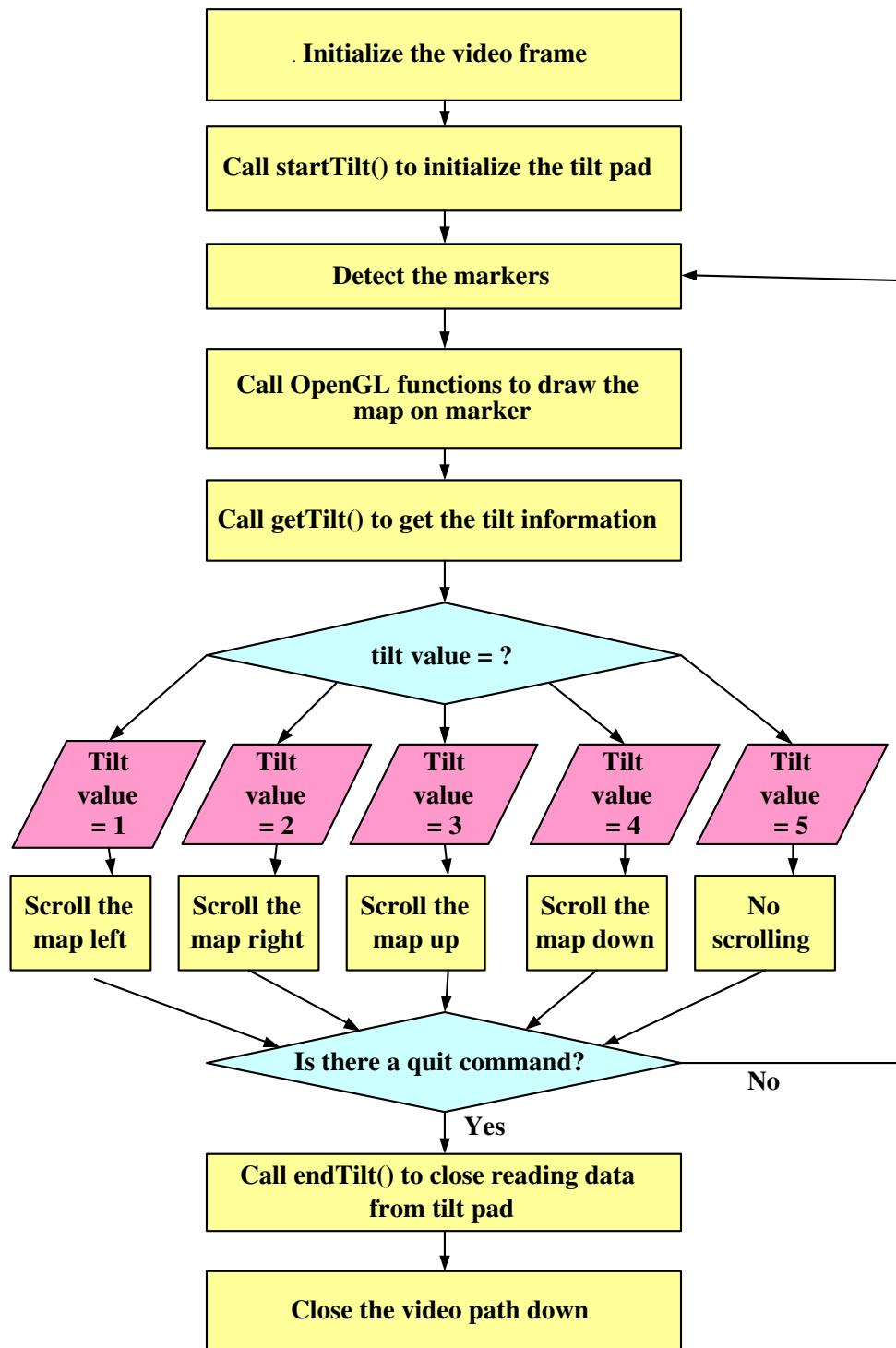


Figure 5.17: Integrated code flowchart of the map scrolling system

5.4 Tangible AR Pacman Gaming System

In addition to the above AR map scrolling system, an AR Pacman game is also developed to demonstrate our concept of tangible interface with tilt pad. Pacman game is an old popular PC game that is designed to let Pacman walk around the maze in order to eat food and also to avoid ghosts. In our design, the game method remains the same, while all these characters including Pacman, maze, food, and ghosts are all 3D virtual characters which makes the graphics more realistic and attracting. The main features of this system is listed:

- It's a tangible interaction game that uses tilt pad to control the moving of Pacman instead of conventional keyboard, mouse and joystick.
- It is an AR game, that see both virtual gaming environment and real physical environment through HMD.
- Marker tracking is used. That is, the gaming environment is attached to a fiducial marker.
- All game characters are made as 3D cartoon models.
- It is a multimedia system that both graphic and sound information is provided.

The most exciting improvement that AR technology brings to small-space PC games is that users can experience the surprising gaming method to have the gaming environment just held on their own hands with the marker. This is quite a unique feature of AR game that can't find in the games using monitor including PC games, even virtual reality games. With tilt pad attached on the back side of the fiducial marker, the user can simply tilt his marker to control the moving of

Pacman in the virtual vision just augmented on top of the physical marker. The effect looks as if the user is tilting the digital graphic information directly, which achieves the tangible interaction.

We have testified that 3D sound is effective on enhancing the interfaces for AR systems in chapter 4. Based on this experimental result, we create sound interface for this AR gaming project. Although it is not 3D sound that is used here, the improvements to game's attraction and users' immersiveness after adding the sounds is quite obvious reported by the subjects that have ever played this AR game. The sounds we introduce to this Pacman game include Pacman death, Pacman eating, and Pacman walking (four different sounds for four walking directions). These sounds are embedded in the main source code by using Allegro library [62]. Allegro is a game programming library for C/C++ developers distributed freely, supporting various platforms including Linux, Windows, etc. It provides many functions for graphics, sounds, player input (keyboard, mouse and joystick) and timers. It also provides fixed and floating point mathematical functions, 3d functions, file management functions, compressed datafile and a GUI.

The game is carried out in a virtual maze with narrow pathway. The Pacman and two ghosts are set to move only along the path; cherries are distributed also along the path. When Pacman meets cherry on the path, he is supposed to eat it. When Pacman meets ghost instead of cherry, he will be eaten. Pacman's death is visualized by loading an animated model that he gets smaller and smaller until invisible. Therefore, the pacman's mission in the maze is to look for cherries and to avoid ghosts. The moving of the Pacman is manipulated with tilt pad. While ghosts patrol in the maze automatically at the speed of one step per frame. If they reach a crotch, the next direction ghosts turn to is decided as a function of the system clock and an equal chance is given to each possible direction.

The other technical parts including marker tracking, visualization with AR-ToolKit and hardware combination are the same as we introduced above for the map system, therefore we will not repeat them mechanically here. The code flowchart (figure 5.18) involves the graphic part mainly. And the screen shots of this gaming system are pasted in next chapter, as well as video provided in the accompanying CD.

In the demo video, HMD point of view of the game, third person point of view of the game and the hardware close-up shot are screened respectively. The first part of the video shows the HMD point of view of the Pacman game, which means that this part of the video recorded what a user exactly saw through his HMD when he tried the system. Therefore, the shots of this part are immersive and concentrated on the gaming environment. If you are watching the video, you can see that our main character Pacman is in yellow color and his move are controlled by the tilt direction of the pad. Please kindly note that due to the video editing effect, the screen presents a discontinuous movement of the Pacman. On the other hand, ghosts in purple are patrolling the maze. The player is supposed to tilt the pad wisely in order to avoid Pacman meeting the ghosts. By successfully evading the ghosts, Pacman can enjoy the deliciousness of cherries in red while he passes over them in his path. The scene of Pacman eating cherries can be clearly seen in this part of video when he moves towards left along the upper most path. The second part of the video is taken from the third person point of view. What a third person observed when the girl was playing the game is shot here. He would see the girl's behavior: holding and tilting the pad with marker on the front and tilt sensing board on the back, as well as the HMD screen which is split to display on the monitor screen also. At last, a close-up shot of the receiver part of the hardware connected with PC is shown, followed by the close-up shot of the sensing

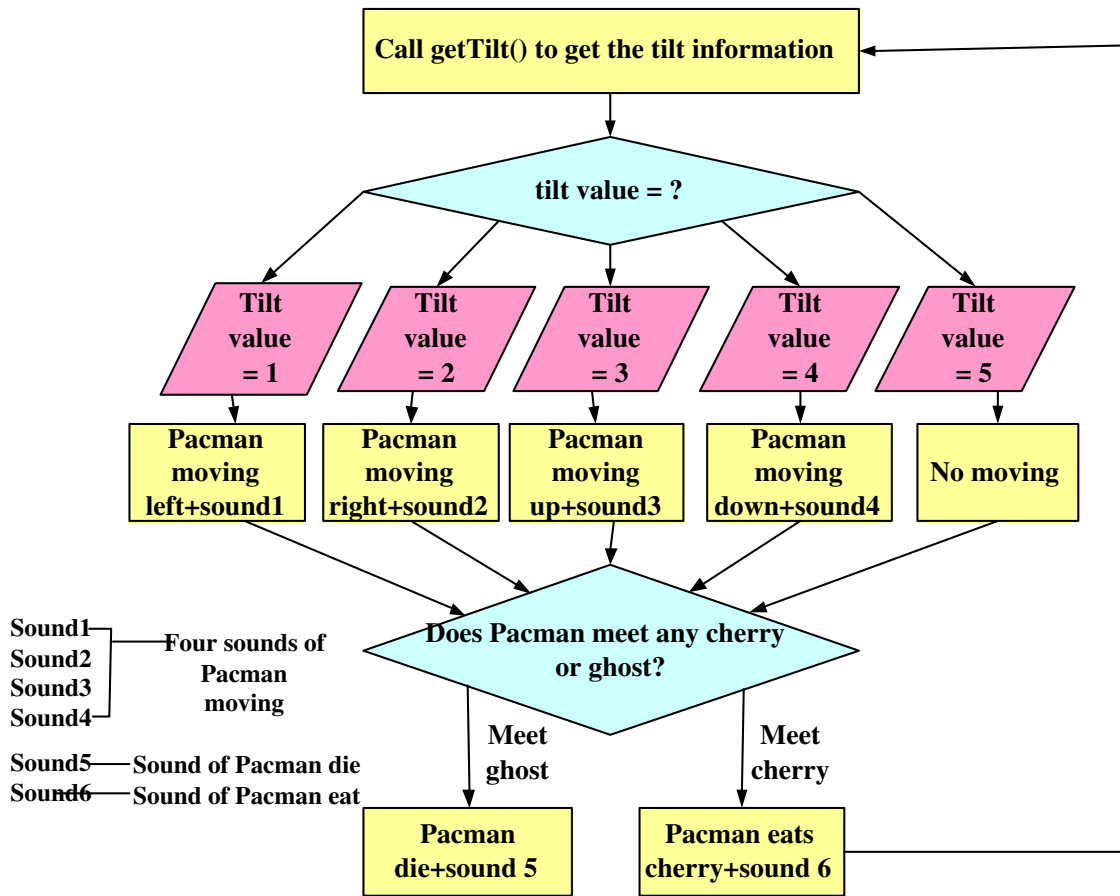


Figure 5.18: Code flowchart of AR Pacman system. This is a loop that executed once every frame.

part of the hardware.

Chapter 6

System Results and Assessment

After detailed the technology applied in the hardware based tangible interaction design in last chapter, we list the results obtained from the map scrolling system and Pacman game system, both controlled with tilt pad. How these developed systems accord with the definition of tangible user interfaces is also explained. A user study of the tilt pad Pacman system is conducted to assess the system performance.

6.1 Results of Map Scrolling System

In this augmented reality map system, we achieved following functions:

- Map zooming.
- 2D map scrolling using tilt pad.
- 3D map scrolling using tilt pad.
- Tank patrols on the 3D map.



Figure 6.1: Different zooming of a 2D map. The user can draw the map close to the camera to view details as shown in the last image.

Figure 6.1 shows the zooming effect of this map. These different zooming 2D maps are augmented on top of the multiple marker that can be handled on user's hands. And user can take the virtual map quite close to his eyes (which is camera now) to see the detailed content of the map.

We can see the tank on the marker from Figure 6.2. The screen shot on the top is part of the 3D AR map in the vision of a user wearing HMD. And the tank superimposed on the map is shown in the bottom two screen shots. Actually the tank is animated to move around a circle on the map, which can be seen from the video we appendix in the CD-ROM. These two screen shots give the images of the different positions when tank moves and the different orientations when tank drives

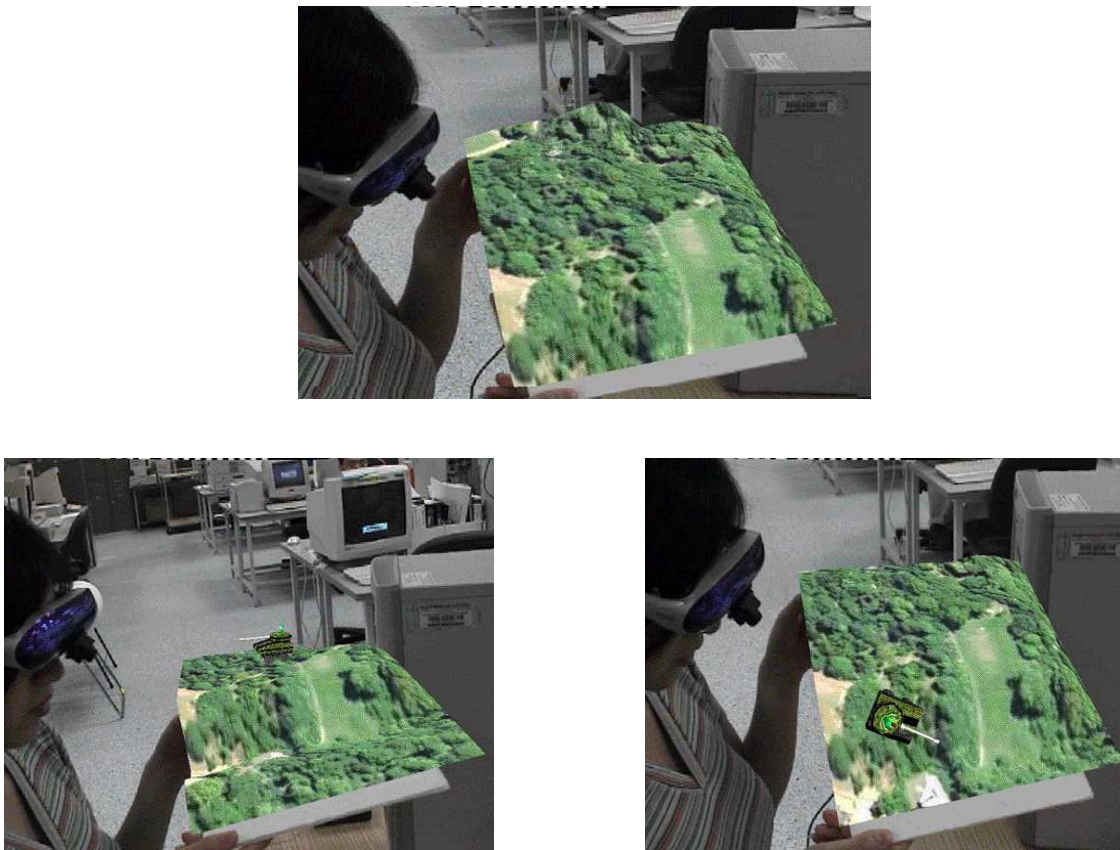


Figure 6.2: The top photo is part of the 3D map. The bottom two photos show the tank with different positions and different orientations on the map

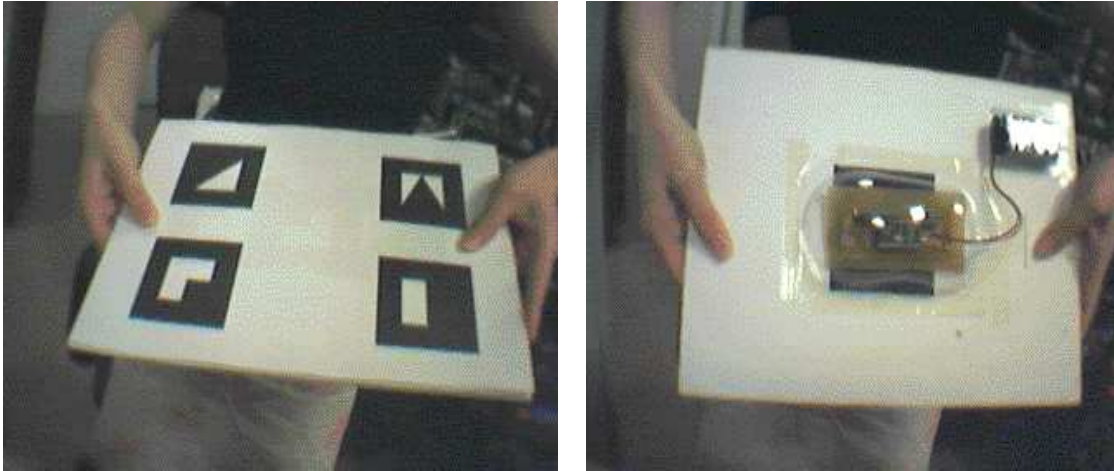


Figure 6.3: Left image is the multiple marker used, right image is the tilt pad fixed on the back side of the marker.

on distinct slopes of the surface.

In order to scroll the map with tilt pad, we attach tilt pad on back side of the fiducial marker. Figure 6.3 shows the multiple marker we use and the tilt pad fixed on the backside. Figure 6.4 are eight screen shots when a user is scrolling a 2D map to different positions using tilt pad. Each screen shot is taken from diverse part of the map texture. And figure 6.5 displays the scrolling condition in a 3D case. The peaks signed with the flag in the four images are the same place of the map, but scrolled to different positions by tilting. The peak is tilted up in the top left image; The peak is tilted down in the top right image; And it is tilted left in the bottom left image; tilted right in the bottom right image. In order to explain the relative positions during scrolling clearly, that is to include the same surface point in all the four tilting positions, just part of the map is displayed in the screen shots. The scrolling of the whole map can be viewed from our video in the CD-ROM.

This map scrolling system is implemented on a Pentium 4 PC under Linux OS, which allows a speed of 22 frames per second. The system captures the camera's video stream at 640×480 pixel resolution.



Figure 6.4: Tilt pad is used for a 2D scrolling map application

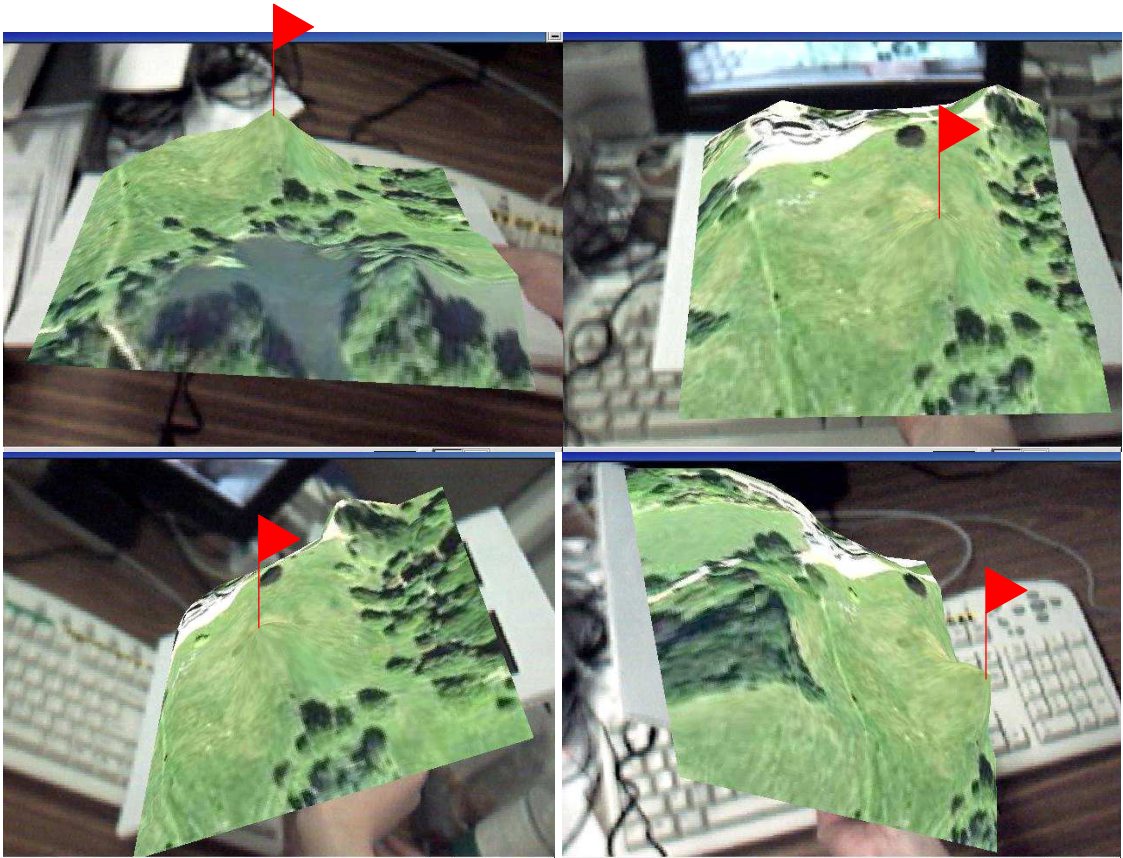


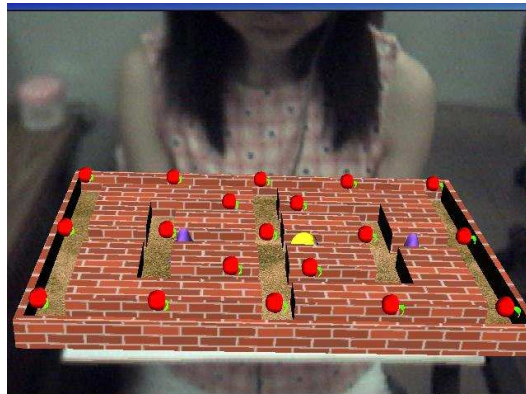
Figure 6.5: Scrolling of the 3D map. The red flag(not part of the map)is added to sign the same place tilted to different positions using tilt pad in four images.

6.2 Results of Pacman Game System

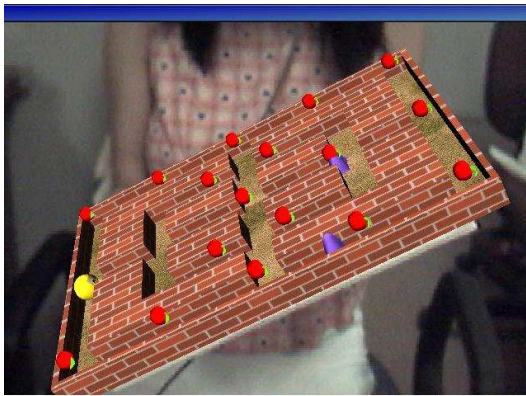
In the AR Pacman game system, the gaming environment is designed as a 3D maze superimposed on the fiducial marker. Pacman and ghosts are set to move along the path of the maze, as well as cherries spread on the path. The move of Pacman is controlled by tilt pad as shown in figure 6.6. Image (a) shows the initial game environment with Pacman, two ghosts and all the cherries. In image (b), the user is moving the Pacman to the left side. Tilt pad is fixed on the back side of the marker as shown in figure 6.3. And she moves Pacman to the right side in image (c). In image (d) and (e), Pacman is tilted outward and inward to the user, respectively. It can also be obviously observed that the cherries are getting less and less when Pacman runs around the maze.

This gaming system is set up on a normal Pentium 4 desktop computer running Linux with a speed of 25 frames per second and 640×480 pixel resolution. We also install this system on a Dell Pentium 4 laptop. To achieve this, some minus change is made to the hardware. That is to increase the amplitude of the waveform transmitted between the tilt pad and the receiver connected to the laptop during wireless transmission because the laptop requires a higher voltage of the transmitting signals than normal PC. However, the speed is improved on the laptop up to 30 frames per second. It also increases the mobility of this system by running on a laptop. Further research can even consider combining this system with wearable computers as one of the entertainment facility of wearable computers.

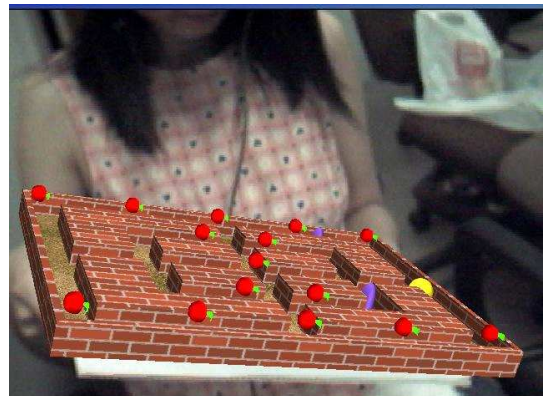
In addition to applying fiducial marker tracking, we also use natural feature tracking to register the virtual contents. The tracking algorithm is developed by another research scholar of our lab [33]. Figure 6.7 is the paper with natural features that are tracked. The game superimposed on the natural feature paper is



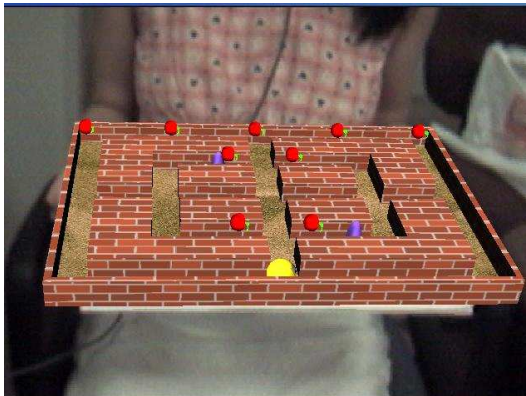
(a)



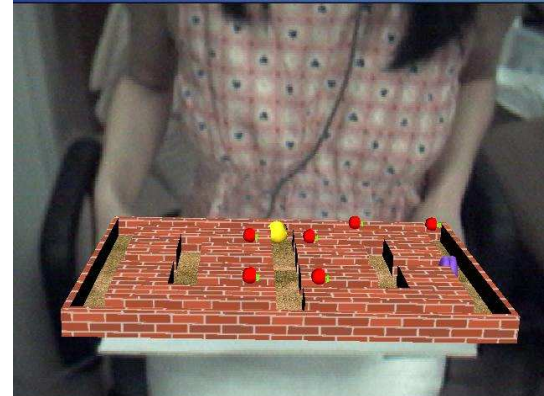
(b)



(c)



(d)



(e)

Figure 6.6: (a) initial game environment with Pacman(the yellow one), ghosts(purple ones) and cherries(in red) (b) user tilts Pacman left (c) user tilts Pacman right (d) user tilts Pacman outward to her (e) user tilts Pacman inward to her



Figure 6.7: The paper with natural features for tracking

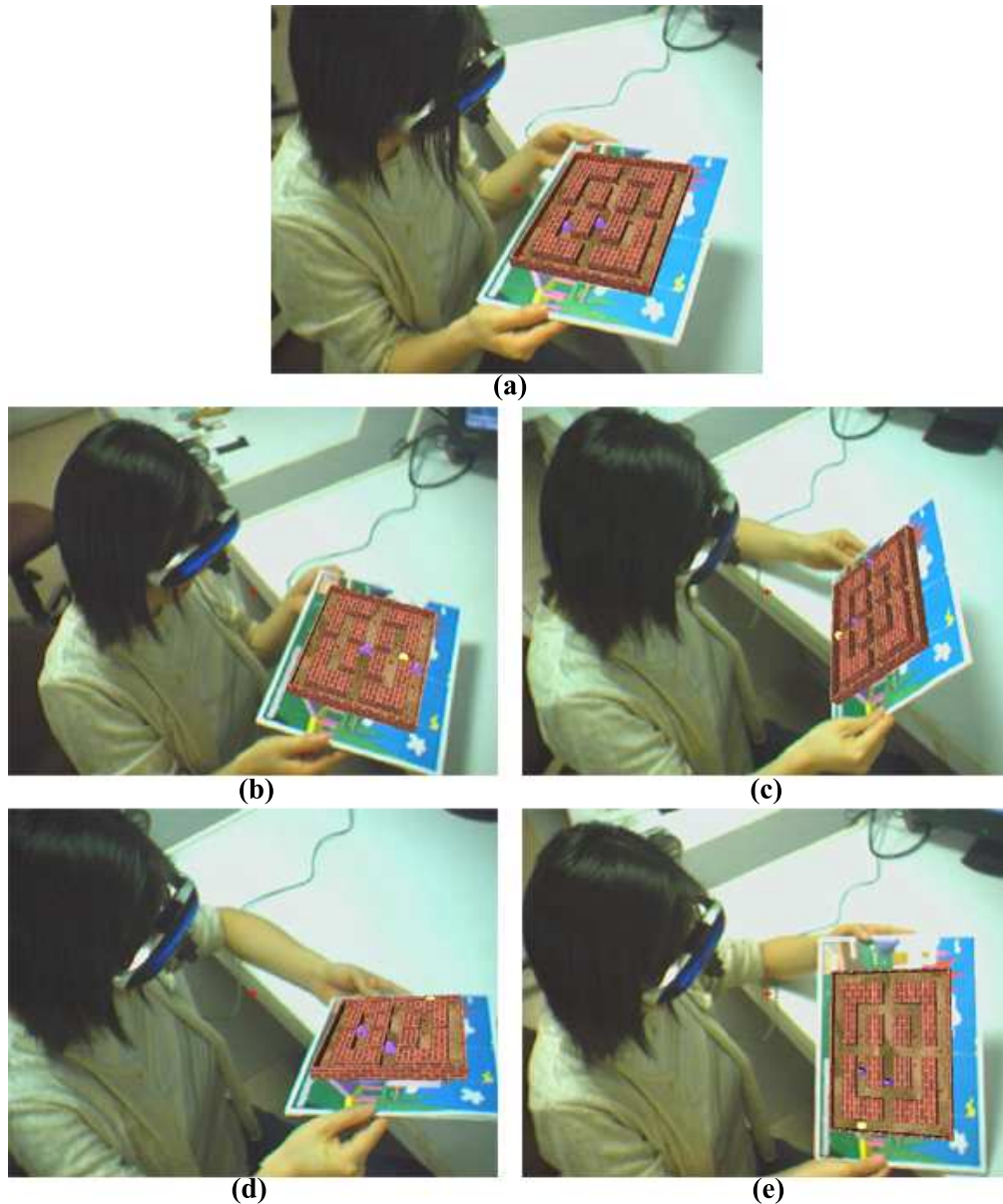


Figure 6.8: The user is playing the 3D Pacman game. (a) The user is holding the board, on which the colorful picture is pasted. The 3D Pacman maze is now augmented onto the board. (b) The user tilts the board forwards, the Pacman goes forward, to eat the berries and avoid the ghosts. (c) The user tilts the board backwards. (d) The user tilts the board leftwards. (e) The user tilts the board rightwards.

shown in figure 6.8.

6.3 Description of the Results

As we have pointed out before, tangible user interfaces are featured with representing and controlling digital information physically. In our two tilt pad based AR interaction systems, the interface is the real-world object which is a board with tilt pad fixed on its back side. Holding this board, 3D virtual objects including map or game environment are attached just over the user's hands. Users can move these objects to view the details from different angles and close distances by simply moving the board around. And the unique interaction method is introduced by the usage of tilt pad. With this hardware, searching, selection, and manipulation of the digital information is achieved by tilting the physical board when a user scrolls the map and plays the game. For example, map scrolling is a kind of information searching; Particular of the map can be selected to have a clear look; and the direction of the map scrolling and the move of the Pacman are both manipulated by tilting the pad. In the case of natural feature tracking Pacman game, since we can use the paper drawn with sketches of the computer 3D characters for tracking, these sketches can be regarded as the physical, emulational representation of the computer world objects and we can handle this representation to control the digital world directly.

From the analysis above, we can see that our uniquely designed interfaces are accord with the features of tangible user interfaces. They have brought a completely new idea to this research field by introducing small and cheap hardware to augmented reality, which also provides the potential applications if combined with wearable computers in the future due to their small-size and mobility.

6.4 User Study of Tilt Pad Pacman Game System

6.4.1 Aim of this User study

We conduct this user study of our developed AR tilt pad controlling Pacman game in order to get feedback from users and to assess the performance of this system to see whether its tangible user interface is valuable in providing novel and natural interaction system to the users. The improvements that may be continuously made in the future work are also expected to obtain from this user study.

6.4.2 Design and Procedures

Twenty subjects (8 Females and 12 Males) were found to participate in this study. Most of the participants were final year undergraduate or postgraduate students major in Electrical and Computer Engineering at the National University of Singapore with an average age of 24.3 years old. Two of them were staff members of National University of Singapore. All of the subjects reported clear-vision and normal hearing abilities.

Before doing the experiments, all the subjects were given a brief introduction about augmented reality and the tasks to be performed. During the experiments, the subjects were asked to play the normal PC Pacman game controlled by keyboards and mice for several minutes first in order to get them familiar with this game and to make it easy and fair to compare our interaction method later with the traditional method. After that they were led to play our AR Pacman game with the instruction in how to use the board. They would be experiencing the Augmented Reality game space wearing HMD and all the game functions such as

Pacman running, Pacman eating and Pacman dying with different sounds. After they tried both traditional Pacman game and our system, they were supposed to fill in a questionnaire paper with 9 questions as follows:

1. Compared to conventional PC games, which interface do you prefer for a Pacman game? Why?
A. keyboards and mice B. our tilt pad C. others(please state):
2. Do you feel this gaming interaction method is or is not a natural and physical like we control our real world objects, e.g. roll a ball on a surface?
A. yes B. no, no difference with traditional methods
3. Do you find that the augmented reality interface characters are more realistic and interesting than common 3D/2D computer game graphics on a screen?
A. yes B. no C. doesn't matter
4. What kind of display method do you prefer for a Pacman game?
A. display on screen B. display through HMD on a board that is graspable
5. What kind of game space do you like more? Why?
A. Virtual Reality (VR) game space with totally computer-generated scenes
B. Augmented Reality(AR) game space with both virtual objects and real world environment
6. Our tilt pad is a wireless device, do you think that using wireless devices can make you enjoy the game more?
A. yes B. no C. doesn't matter
7. Do you think that the sounds added in the game make the game more interesting and make you more entertained?

A. yes,quite nice B. a bit C. doesn't matter D. worse

8. How will you rank the entertaining function of this game? Please kindly give your comments on this game.

A. very interesting and unique B. the interaction method is attractive, but the game complexity need to be improved C. not good, I like conventional PC game

9. What are the improvements to the tilt pad game you can suggest?

Question 1,question 2 and question 6 are designed to evaluate the tangible user interface we introduced by tilt pad. Before the participants were allowed to try our system, they were asked to play the normal mice and keyboards controlled Pacman game for about five minutes. Naturally, their comparisons between these two kinds of gaming interaction methods as well as the reasons are requested in question 1. While question 2's emphasis is on the natural and physical character of our interface. The physical movement such as rolling a ball on a surface is listed to stimulate the readers' association of ideas. Through these two questions, we expect to test the acceptance of this new tangible interaction gaming style. Additionally, the wireless aspect of our tilt pad is worth noticing. So question 6 is added to find out whether a wireless device is necessary in those small-space desktop PC games.

Since this is a research on new interfaces design for Augmented Reality, it is essential to investigate the AR gaming potential through this user study. Questions 3, 4 and 5 are brought forward for this purpose. Question 3 and question 4 focus on the graphics part of AR. The 3D graphics in Augmented Reality are notably different from conventional 2D or even 3D computer graphics on normal monitor screens. Apart form that they are 3D representation, they are also attached to a physical marker that is graspable. Therefore, users can see the details of these 3D characters from any angle that they prefer. If users like this unique feature

of AR graphics reflected from question 3 and 4, we can say that AR gaming has its advantage at least at this unique display method. Historically, Virtual Reality gaming got its beginning earlier than Augmented Reality gaming. So people are more familiar with VR gaming than AR gaming. Question 5 allows participants to give their opinions about these two game spaces that have certain common points.

Question 7 is designed to evaluate the sound effects embedded in this Pacman gaming system. As far as the entertaining function of computer games are concerned, delicately-designed graphics as well as compatible sound effects contribute a lot to the users' sensory perceptions. Through question 7, we expect to see whether our added sounds achieve to improve the gaming effects to some extent. The last two questions of this questionnaire, question 8 and 9, attempt to get the testers' overall assessments of this gaming system and any valuable suggestions they can give, in order to make further improvements in the future.

6.4.3 Results of this User Study

Concerning question 1, question 2 and question 6, we expect to assess the performance of tilt pad in this game. 14 subjects out of 20 said that they preferred our tilt pad to keyboards and mice for a Pacman game. And the reasons why they prefer our pad were like:

- *“It’s a newer interface and novel.”*
- *“I make more movements to play with tilt pad than keyboards and mice.”*
- *“It’s an exciting way of playing game.”*
- *“The interaction with the objects of the game is more interesting. You identify yourself with the Pacman on a high level, more than on the screen.”*

- *“It’s challenging and exciting to play with your pad.”*

Those subjects who liked keyboards and mice almost stated the same reasons as “easy to control” and “save energy”. According to question 2, 16 users (account for 80 percent) claimed that they felt like handling physical objects when using tilt pad to play Pacman game. And regarding the wireless control of tilt pad, almost all subjects (19 out of 20) thought that they enjoyed the game more with wireless control device. Only one thought it doesn’t matter about this. From the answers to these three questions, it is suggested that the majority of the users preferred our tilt pad controlled Pacman game to conventional keyboard and mouse controlled one as they felt that our interface is novel and natural to them. Moreover, the mobile and wireless control method of the tilt pad also contributes to the convenience and smoothness of the interface. From their comments, we can see that our natural and physical interface requires users to make some related physical movements such as hand, arm and even slight body movements when they play the game. This unique method allows them to get more involved in the game and identify more with Pacman as the player’s movement results in Pacman’s corresponding move at the same time. As a kind of completely new gaming manner, some people (still minority) didn’t consider it better than the traditional one is also acceptable due to their custom and personal preferences.

Question 3, question 4 and question 5 are designed to assess the augmented reality game space. Concerning question 3, all subjects (20) declared that the augmented reality interface characters are more realistic and interesting than common 3D and 2D computer game graphics on a screen. And 15 subjects chose to display game environment through HMD on a board that is graspable rather than on a screen. As a result, 18 subjects liked AR game space with both virtual objects and

real world information more than VR game space with totally computer-generated scenes for question 5. They explained reasons such as: *“I can feel the spatial reference and my own position and interaction with the space.”* and *“You can actually have real physical control while playing the game, hence making the game play more realistic.”* Therefore, we can get the basic standpoint that AR gaming has a large and promising space left to be explored, in respect that AR systems possess attractive and unique display method, nice graphics and the combination of real and virtual worlds.

As a multimedia system, we also value the entertaining and interface enhancing functions of the sounds we add in the game through question 7. As a result, 19 subjects thought that the sounds made the game more interesting and made them more entertained to a great extent and 1 subject felt a bit improved with sounds. This is quite encouraging to apply multimedia in further interface development.

Regarding question 8 and question 9, 11 subjects ranked this game very interesting and unique; 9 subjects commented that the interaction method is attractive, but the game complexity need to be improved. We consider this outcome quite reasonable as compared to professional commercial games, our game complexity is obviously infant. But our emphasis is the interaction method, and it has been accepted and appreciated by the test users even with a simple game programming. The improvements users suggested such as: *“Add wider paths, more monsters, more traps to gain excitement and challenge”* and *“Hardware sensitivity of the tilt pad can be refined.”*

The pie charts of the votes of question 1 to question 8 can be viewed from figure 6.9.

Apart from analyzing the users' questionnaire responses in detail above, we also did some statistical observation on the questionnaire data in order to testify the

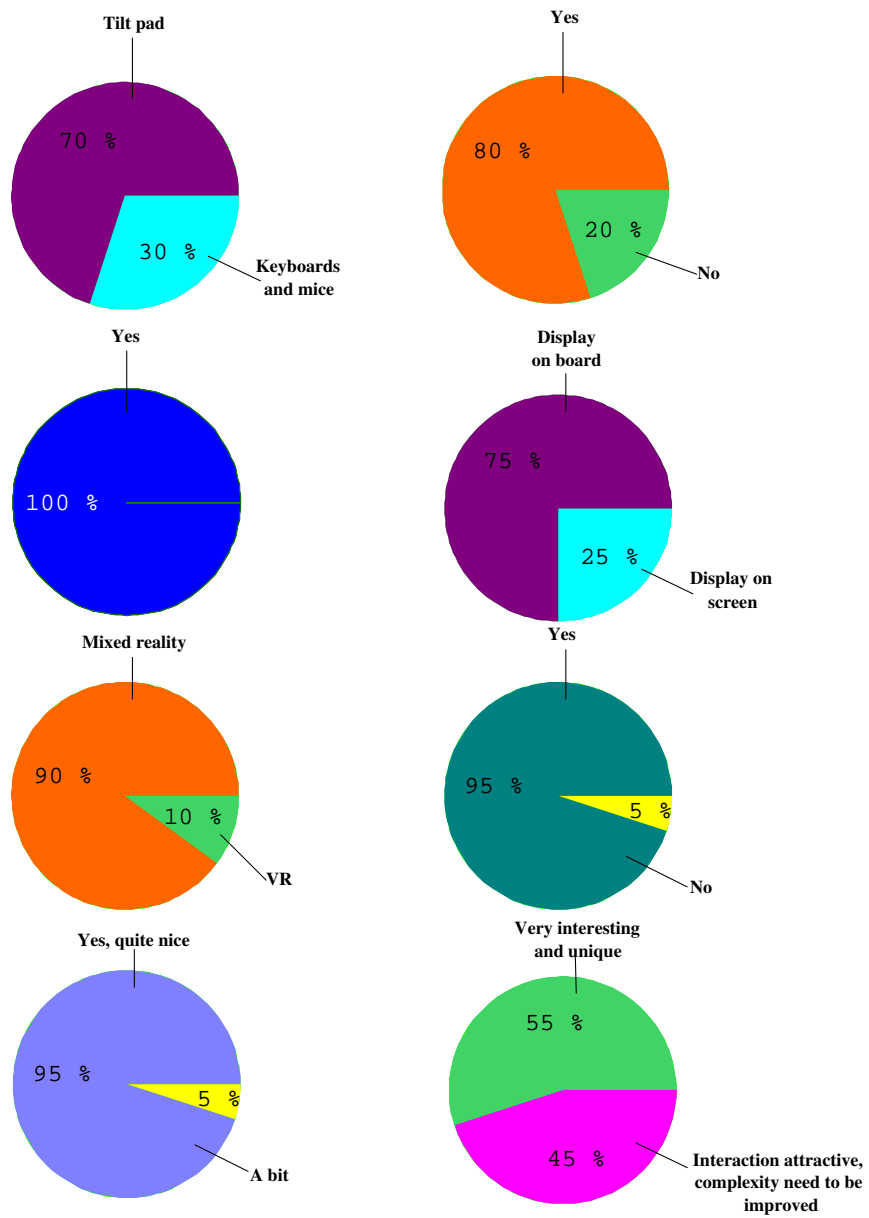


Figure 6.9: Answers to Q1 – Q8 of the user study. Top left chart is concerning Q1, top right chart is concerning Q2, and continue based on a left-right, up-down sequence

effectiveness of our new tangible interface more strongly. Here, we established two sets of numerical data to see whether it makes differences between the traditional keyboard controlled Pacman game and our tilt pad controlled Pacman game regarding the interaction method. We interpreted the questionnaire responses in such a way below. Question 1, 2, 3, 4, 6 and 8 were used here as they required users to make a comparison between these two methods, especially in those features one of them exclusively has. Each question was given 10 points. If the user preferred the traditional Pacman game, traditional method would get all the 10 points and tilt pad Pacman got 0 points. On the contrary, if the user preferred the tilt pad controlled Pacman game, the 10 points would be assigned to tilt pad method instead of traditional one. If the user held a neutral position, the two interaction methods would get 5 points respectively. For example, for question 8, tilt pad controlled Pacman game would gain 10 points if the user chose A; Traditional Pacman game would have 10 points if the user chose C; And each of the two games would get 5 points if the user chose B. After such a process, every user would evaluate these two kinds of gaming interfaces regarding the natural and unique interacting features related to Augmented reality by providing two scores. Altogether we got 20 sets of data from 20 participants and each set of data consisted one score of the traditional method and another score of the tilt pad method. These 20 sets of data are described in figure 6.10.

We would like to point out that the scores here are just relatively representing values and our emphasis is on examining whether the features that tangible AR interface introduces have made significant differences to the users' perception. A one-way ANalysis Of VAriance (ANOVA) [56] was used to test the data statistically. Having set a 0.05 significant level, significant difference was found between the traditional Pacman game and the tilt pad Pacman game, with $F(1, 38) = 190.3, p <$

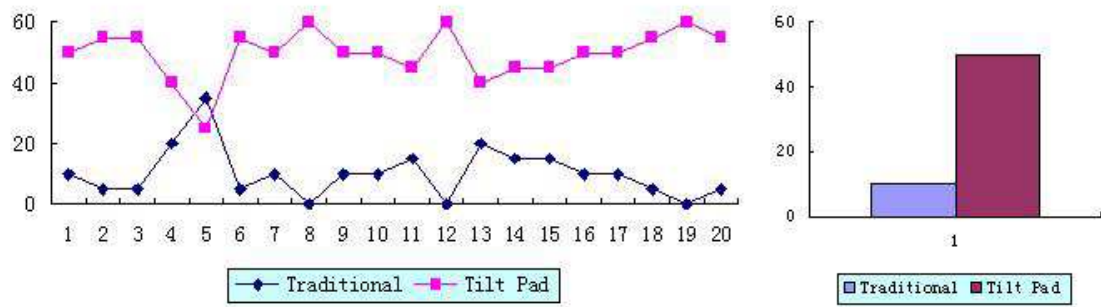


Figure 6.10: Left: The comparison between the scores of the traditional Pacman game method and the tilt pad Pacman game method. Right: Average score of the two methods

0.001. It is clear that as far as the unique features that our tilt pad Pacman system has such as physical and tangible interaction method, wireless controlling and graspable graphics are concerned, users felt the differences and preferred them to a large extent.

Overall, we can conclude now that our tilt pad controlled Pacman game system is testified to have produced tangible, natural and seamless gaming interface to the users through this study. Additionally AR gaming should deserve attention due to its unique double information space and graspable graphics. Anyway, further improvements to this system may be made in increasing the gaming complexity and hardware refinement.

Chapter 7

Summary

7.1 Our Research Summary

Although a lot of research work has been conducted in display and registration problems for augmented reality systems, AR interface design and interaction space development remain largely unexplored. Interaction with AR environments has been largely limited to passive viewing or simple browsing of virtual information registered to the real world. Tangible User Interfaces are defined as those computer interaction technologies that allow representing and controlling digital world through physical objects. Due to their unique features, tangible user interfaces can lead to natural and seamless Computer-human Interactions since the transformation of interaction means between digital world and real world is eliminated. Therefore, the aim of this research work is considered to design tangible user interfaces for augmented reality, which lacks of enough attention previously. Additionally, we takes advantage of the mobility and convenience of small-size as well as cheap hardware to achieve mobile and flexible interaction even with low computational load.

The main research on tangible user interfaces design has been completed involving:

- A desk-top AR system *Magic Book* that provides the physical representation of 3D virtual models as the contents of the book and physical manipulation of these models by picking and dropping using a paddle. Here, a catalogue of three dimensional cartoon models is created by attaching a different fiducial marker to each page of a real book. Users can select the model that they are interested in by simply turning the pages. This is a radical departure from scrolling through a list of three dimensional models on a computer screen. The interface is the real-world object which is a book in this case.
- An experimental research on the role of 3D sound in AR interfaces. A large space AR game embedded with several kinds of sounds is applied as the user interface. Players are requested to finish some certain missions in this game space aided with or without sound. Then, we take both numerical and psychological studies of the subjects. From the experiment results, we can find that 3D sound does play an important role as the complement of the visual AR environment, especially when interfaces preserve social interactions. 3D sound also improves the feeling of human presence and collaboration.
- An AR *map scrolling system* interacted with tilt pad. In this augmented reality map system, we achieved several functions including map zooming, 2D map scrolling using tilt pad, 3D map scrolling using tilt pad and augmentation of tank on the map surface. In order to scroll the map with tilt pad, we attach tilt pad on back side of the fiducial marker. By doing so, the board with both fiducial marker and tilt pad becomes the physical interaction tool with the digital map. User can scroll the map to his desired directions to search the

whole map simply by tilting this board to corresponding directions. Physical, tangible and mobile control of the AR objects is realized.

- An AR *Pacman game system*. Originated from the old popular Pacman game that is designed to let Pacman walk around the maze in order to eat food and also to avoid ghosts, our AR version is featured with 3D virtual characters, hardware interface, marker tracking and adding of entertainment sounds. Here, the same board with tilt pad on the back side as used in the map scrolling system is also employed as the interaction device. The moving of Pacman is controlled by tilt pad instead of conventional keyboards, mice or joysticks. The employment of sound makes the interaction more attractive and increases the immersiveness of users. This AR Pacman game is also implemented using natural feature tracking, which uses the sketches of the gaming characters to register the virtual 3D contents.

From the research outcome that we have listed, it can be concluded that our aim of designing tangible user interfaces for augmented reality is achieved to a quite large extent. And the unique features of the tilt pad based tangible interfaces are summarized in the following table 7.1.

7.2 System Limitation and Future Work

Having conducted this research on AR user interfaces design, we are strongly aware of the importance and promising future of developing natural, tangible and seamless Human Computer Interfaces. More efforts deserve to be put into this area with further wideness and depth. As far as our current small hardware based interaction systems is concerned, due to the limitation of time and other related sources, the

Table 7.1: Features of tilt pad based tangible interfaces

common features	tangible
	natural
	physical
	wireless
	mobile
	small-size
	low cost
	independent of keyboards and mice
	computational load reduced
3D map scrolling system	military application
	educational application
AR Pacman game	multimedia system with sounds
	nice 3D graphic characters
	entertaining

sensitivity of the hardware is not at a high level. User is required to tilt the pad to a predefined fixed angle with the ground and the tile has to be purely in the direction of left, right, front or back. This translates mathematically to mean that the tilting axis is strictly the midline of the tilt pad: the longer midline, or the shorter midline. Other tilting directions or too gentle a tilt do not generate enough acceleration so that the system can not produce the corresponding data to process. Apart from the hardware part, the Pacman gaming complexity and excitement degree is also not at its advanced stage, which reserves further working space to fulfill the improvements. And the user study of the Pacman gaming system is to some extent restricted in sample size, which could be enlarged. As a result, we recommend some directions of future research as follows:

- Enhance the hardware sensitivity so that more than the four directions can be sensed. This improvement could be useful not only in the current Pacman game system but also some other applications such as ball rolling, etc.

- Increase the wireless distance of the tilt pad so that some other applications that allow large-space movement could be implemented.
- Add gaming complexity and stages to the current AR Pacman game system. For example, more complicated path could be designed, more ghosts could be added, and different food with different scores could be distributed along the path.
- Interactive gaming allowing more than one person to play at the same time could be imagined. For example, one pad controls the move of Pacman, while another pad controls the move of the ghost.
- Carry out a more large-scale user study of the system. A larger sample size could be used, more detailed questionnaire could be designed, and a sample with a lower average age could be found in order to stimulate more imagination and innovation in AR.

As a whole, until recently, most AR prototypes concentrated on displaying information that was registered with the world and didn't significantly concern themselves with how potential users would interact with these systems. Prototypes that supported interaction often based their interfaces on desktop metaphors (for example, they presented on-screen menus or required users to type on keyboards) or adapted designs from virtual environments research. We need a better understanding of how to display data to a user and how the user should interact with the data. To date, few have studied such issues. However, we expect significant growth in this area because research AR systems with sufficient capabilities are now more commonly available.

We hope that through this research on tangible AR interface development, more focuses would be attracted to design tangible and mobile interfaces for computer. More and more, the interface design for computer inclines towards human centered design. In the coming generation, people would require computer and internet access at anytime anywhere. So we can say that mobile and ubiquitous computing is one of the main future direction of HCI. Augmented Reality is a powerful UI for this type of computing due to its double information space. As Bill Buxton said: “Rather than turning inward to an artificial world, ubiquitous media encourages us to look outward. It expands our perception and interaction in the physical world.” We can notice that computing tools get smaller, faster, more intimate from its blossom history. Embedded with a tangible and mobile interface, computers would ulteriorly be invisible as the transition between human and computer is designed to be friendly and seamless. While for an tangible AR interface, how can we smoothly transit between spaces still remains a question unanswered. And also when we are talking about mobile AR or the current wearable computers, challenges are encountered on the issue of power supply, weight and tracking. Particularly for an outdoor AR system, extra complexity of the problem appears in respect of wide range of operating conditions and little control over environment. With further research efforts, we expect the appearance of novel tangible AR interfaces combining the advantages of both AR user interfaces and tangible user interfaces, which leads to enhanced display space merged with task space and improved interaction in virtual world.

Collaborative tangible AR is another direction that will draw a lot of attention in the future. For example, AR conferencing is an interesting topic. It moves conferencing from the desktop to the workspace. One of the prominent advantage of AR conferencing is that user would feel the remote collaborator like a real person

facing him due to the real-time augmented graphics and interactions. In another word, AR conferencing increases collaborator presence and communication. Ideally in the future, features are awaited to appear such as shaking hands with your virtual conference partner, mobile video conferencing allowing you to have a face-to-face talk in a car to a person on the other side of the ocean, or opening your virtual report physically by hand to show it to all the conference attendees. Imaginations are always inspiring and invaluable in evoking evolution in the research area of Augmented Reality.

Computer gaming and entertainment, as a multi-billion people involved huge market, offers a unique test-bed and market for advanced concepts in computer graphics, human computer interaction and computer-supported collaborative work (CSCW). In addition, games provide an engaging way to discover and explore research issues in the fields of wearable computing and augmented reality. With the unique display space and user interfaces, Augmented Reality would create fantasy and excitement when meeting computer gaming. While the development of optimized three-dimensional graphics engines has dominated industry interest in the past, one of the major costs of creating current generation computer games is the development of content. Developing an entire 3D world for the player to explore is time consuming and labor intensive. In fact, the complexity of world building required to create a compelling experience is directly dependant on the dimensions of movement allowed to a player's character. With augmented reality, the virtual world is composited with the physical world, resulting in obvious restrictions to the simulation, but not without significant benefits. First, the game designer already has a pre-made environment for his game, that of the user's physical surroundings. Thus, the designer can concentrate on creating just the virtual artifacts that deviate from the physical world in the game. In addition, the system may associate graphics

with the player's physical body, allowing him to shoot fireballs from his hands, envelop himself in a force field, or leave a trail of glowing footprints. In the case of augmented reality gaming, wearable computers composite appropriate artifacts into the user's visual and auditory environments while he is interacting with the physical world. Apart from the gaming contents creation, the issue of gaming interaction method research would also be a strong engine to impel the innovation of Augmented Reality.

We yearn for more research efforts on the area of tangible, mobile and collaborative interfaces design for Augmented Reality in the coming future. And the major obstacles limiting the wider use of AR could be cleared so that the state of the art in AR could mature beyond lab-based prototypes and come into commercial applications.

Appendix A

Contents of the CD-ROM

A CD-ROM is attached at the back of this thesis paper, which provides a soft copy version of this thesis, as well as the source code and result videos for the applications mentioned in the previous chapters. It has the following three main folders:

Folder One: Thesis

- A PDF version of the thesis — ThesisV2.pdf

Folder Two: Source Code and Tutorials

- Map scrolling system using tilt pad
- Pacman game system using tilt pad
- Tutorials on these applications — Tutorial.txt

Folder Three: Videos

- 2D map scrolling — 2Dmap.wmv and 2Dmapuser.wmv

- 3D map scrolling — 3Dmap.avi
- AR Pacman game — Pacman.avi

Bibliography

- [1] A. Backman, “Augmented Reality,” 2000. Available online, <http://www.cs.umu.se/kurser/TDBD12/HT00/lectures/ar.pdf>.
- [2] T. L. Greenbaum, *The Computer User as Toolsmith: the Use, Reuse, and Organization of Computer-based Tools*. Cambridge University Press, 1993.
- [3] Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, Strong, and Verplank, “Human-computer Interaction,” *ACM SIGCHI Curricula for Human-Computer Interaction*.
- [4] B. A. Myers, “A Brief History of Human Computer Interaction Technology,” *ACM interactions*, vol. 5, pp. 44–54, March 1998.
- [5] R. Schmidt, “Blue Screen Tv Studios,” 1996. <http://www.heathcom.no/robert/blue.htm>.
- [6] R. Schmidt, “Weather Studio 96,” 1996. <http://www.heathcom.no/robert/blue/vejr96.htm>.
- [7] G. G. Pyros and B. N. Goren, “Desktop Motion Image Processing for Episodic Television,” *Advance Imaging*, vol. 10, pp. 30–34, July 1995.
- [8] G. Zorpette, “An eye-popping summer,” *IEEE Spectrum*, vol. 31, pp. 19–22, October 1994.

- [9] I. Software, "Quake," 2001. PC Software.
- [10] B. H. Thomas, "Challenges of Making Outdoor Augmented Reality Games Playable," presented at 2nd CREST Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing , 2003.
- [11] E. C. Urban, "The Information Warrior," *IEEE Spectrum*, vol. 32, pp. 66–70, November 1995.
- [12] P. J. Metzger, "Adding Reality to the Virtual," *Proceedings of the IEEE 1993 Virtual Reality Annual International Symposium*, pp. 7–13, 1993.
- [13] S. Mann, "Smart clothing: The wearable computer and wearcam," *Personal technologies*, vol. 1, pp. 21–27, March 1997.
- [14] H. Ishii and B. Ullmer, "Emerging Frameworks for Tangible User Interfaces," *Human-Computer Interaction in the New Millenium*, pp. 579–601, August 2001.
- [15] G. Fitzmaurice, H. Ishii, and W. Buxton, "Bricks: Laying the Foundations for Graspable User Interfaces," *Proc. of CHI*, pp. 442–449, 1995.
- [16] H. Ishii and B. Ullmer, "Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms," *Proc. of CHI*, pp. 234–241, 1997.
- [17] S. Mann, "'smart clothing': Wearable multimedia computing and 'personal imaging' to restore the technological balance between people and their environments," *Proc. of ACM MULTIMEDIA*, pp. 163–174, 1996.
- [18] W. Piekarski and B. H. Thomas, "Tinmith-hand: unified user interface technology for mobile outdoor augmented reality and indoor virtual reality," in *IEEE Virtual Reality*, (Orlando, FL, USA), pp. 287 – 288, IEEE, March 2002.

- [19] M. Tuters, “Variation on a videogame or spatial graffiti; the socio-spatial and futurological implications of augmented reality and location awareness,” in *IEEE Virtual Systems and Multimedia*, (Berkeley, CA, USA), pp. 517–526, October 2001.
- [20] A. Pentland, “Wearable Intelligence,” *Scientific American*, vol. 276, pp. 90–95, November 1998.
- [21] P. Huang, “Promoting Wearable Computing: A survey and future agenda,” *Technical Report TIK-Nr.95, Computer Engineering and Networks Laboratory, Swiss Federal Institute of Technology*, September 2000. <http://citeseer.ist.psu.edu/article/huang00promoting.html>.
- [22] R. T. Azuma, “A Survey of Augmented Reality,” *Teleoperators and Virtual Environments*, vol. 6, pp. 355–385, August 1997.
- [23] W. Barfield and T. Caudell, *Fundamentals of wearable computers and augmented reality*. Lawrence Erlbaum Associates, 2001.
- [24] M. Bajura and U. Neumann, “Dynamic Registration Correction in Video-Based Augmented Reality Systems,” *IEEE Computer Graphics and Applications*, pp. 52–60, September 1995.
- [25] J. P. Mellor, “Enhanced Reality Visualization in a Surgical Environment,” Master’s thesis, Department of Electrical Engineering, MIT, January 1995.
- [26] J. P. Mellor, “Realtime Camera Calibration for Enhanced Reality Visualization,” *Proceedings of Computer Vision, Virtual Reality and Robotics in Medicine*, pp. 471–475, April 1995.

- [27] U. Neumann and Y. Cho, “A Self-Tracking Augmented Reality System,” *Proc. VRST*, pp. 109–115, July 1996.
- [28] A. State, G. Hirota, D. T. Chen, B. Garrett, and M. Livingston, “Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking,” *Proc. SIGGRAPH*, pp. 429–438, August 1996.
- [29] Y. Cho and U. Neumann, “Multi-ring color fiducial systems and a detection method for scalable fiducial-tracking Augmented Reality,” *Proceedings of IEEE 1998 Virtual Reality Annual International Symposium (VRAIS '98)*, p. 212, 1998.
- [30] D. Dementhon and L. Davis, “Model based object pose in 25 lines of code,” *IJCV*, vol. 15, pp. 123–141, 1995.
- [31] M. A. Fischler and R. C. Bolles, “Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [32] U. Neumann and S. You, “Natural Feature Tracking for Augmented Reality,” *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 53–64, 1999.
- [33] K. Xu, A. D. Cheok, K. W. Chia, and S. J. D. Prince, “Visual Registration for Geographical Labeling in Wearable Computing,” *Proc. ISWC*, pp. 109–116, October 2002.
- [34] M. Billinghurst, H. Kato, and I. Poupyrev, “Artoolkit: A Computer Vision Based Augmented Reality Toolkit,” *IEEE VR2000, New Jersey*, 2000. see also <http://www.hitl.washington.edu/share/frames.html>.

- [35] H. Kato, M. Billinghurst, R. Blanding, and R. May, *ARToolKit PC Version 2.11*, December 1999.
- [36] H. Weinberg and C. Lemaire, "Using the adxl accelerometer as a multi-function sensor (tilt,vibration and shock) in car alarms," *iMEMS Technologies/Applications, Analog Devices*, 1995. Technical Note.
- [37] J. K. Perng, B. Fischer, S. Hollar, and K. S. Pister, "Acceleration sensing glove," in *Proc. Third International Symposium on Wearable computers*, (San Fransisco,CA,USA), pp. 178–180, 1999.
- [38] Analog Devices Inc., Santa Clara,CA, ADXL202 Datasheet, 2001.
- [39] M. G. Z. Szalavri, E. Eckstein, "Collaborative Gaming in Augmented Reality," *Proceedings of VRST'98*, pp. 195–204, November 1998.
- [40] G. Reitmayr and D. Schmalstieg, "A Wearable 3d Augmented Reality Workspace," *Proceedings of 5th International Symposium on Wearable Computers*, p. 165, October 2001. Short paper.
- [41] I. E. Sutherland, "A Head-mounted Three Dimensional Display," *Proc. of the AFIPS Fall Joint Computer Conference*, vol. 33, pp. 7.57–64, 1968.
- [42] M. Cohen, S. Aoki, and N. Koizumi, "Augmented Audio Reality: Telepresence/VR Hybrid Acoustic Environments," *Proc. of Workshop on Robot and Human Communication*, pp. 3.61–64, 1993.
- [43] E.D. Mynatt, M. Back, R. Want, and R. Frederick, "Audio Aura: Lightweight Audio Augmented Reality," *Proc. of UIST '97 User Interface Software and Technology Symposium*, pp. 21.1–2, 1997.

- [44] J. J. Baldis, “Effects of Spatial Audio on Memory, Comprehension, and Preference during Desktop Conferences,” *Proc. ACM CHI*, pp. 166–173, 2001.
- [45] W. W. Gaver, R. B. Smith, and T. O’Sheak, “Effective sounds in complex systems: The arkola simulation,” *Proc. of CHI ’91*, pp. 85–90, 1991.
- [46] D. Begault, *3-D Sound for Virtual Reality and Multimedia*. Boston: AP Professional, 1994.
- [47] S. W. Mereu, “Improving depth perception in 3d interfaces with sound,” Tech. Rep. CS95-35, Waterloo, Canada, 1995.
- [48] S. Wolfson and G. Case, “The effects of sound and colour on responses to a computer game,” *Interacting with Computers*, vol. 13, no. 2, pp. 1.83–92, 2000.
- [49] H. Makino, I. Ishii, and M. Nakashizuka, “Development of navigation system for the blind using gps and mobile phone combination,” *Proc. of 18th Conf. IEEE, EMBS*, vol. 18, pp. 1–5, 1996.
- [50] Weihua Wang, Xubo Yang, Adrian David Cheok, M. Billinghurst, and H. Kato, “Touch space: An embodied computing mixed reality game space,” *International Workshop on Entertainment Computing (IWEC2002)*, pp. 4.55–62, 2002.
- [51] Adrian David Cheok, Xubo Yang, ZhiYing Zhou, M. Billinghurst, and H. Kato, “Touch-space: Mixed reality game space based on ubiquitous, tangible, and social computing,” *ACM Journal on Ubiquitous Computing*, to be published.

- [52] A.D. Cheok, W. Wang, X. Yang, S.J.D. Prince, S.W. Fong, M. Billinghurst, and H. Kato, "Interactive theatre experience in embodied and wearable mixed reality space," *International Symposium on Mixed and Augmented Reality, ISMAR'2002*, pp. 59–68, 2002.
- [53] InterSense, "Is-900." online. <http://www.intersense.com/>.
- [54] Loki Entertainment Software, "Openal SDK." [Online]. Available: <http://www.openal.org>.
- [55] W. Barfield, M. Cohen, and C. Rosenberg, "Visual, auditory, and combined visual-auditory displays for enhanced situational awareness," *The International Journal of Aviation Psychology*, vol. 7, no. 2, pp. 1.23–38, 1997.
- [56] D. W. Stockburger, "Introductory statistics: Concepts, models and applications." online, 1998. <http://www.psychstat.smsu.edu/introbook/sbk00.htm>.
- [57] W. Gaver, "Auditory icons: Using sound in computer interfaces," *Human Computer Interaction*, vol. 2, no. 1, pp. 1.66–77, 1986.
- [58] C. Hand, "A Survey of 3d Interaction Techniques," *Computer Graphics Forum*, pp. 269–281, December 1997.
- [59] H. Ihsii and B. Ullmer, "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms," *Proceedings of CHI 97*, pp. 234–241, 1997.
- [60] A. Watt, *3D Computer Graphics*. England: Addison-wesley, 2000.
- [61] M. Woo, J. Neider, T. Davis, and D. Shreiner, *OpenGL Programming Guide*. Addison-Wesley, 1999.
- [62] <http://www.talula.demon.co.uk/allegro/>.