AUGMENTED REALITY APPLICATION BUILDER: INTEGRATION OF FINGER TRACKING FOR INTERACTION IN AUGMENTED REALITY SYSTEM

Allen Choong Chieng Hoon

Master of Science
2010
AUGMENTED REALITY (AR) APPLICATION BUILDER: INTEGRATION OF FINGER TRACKING FOR INTERACTION IN AR SYSTEM

ALLEN CHOONG CHIENG HOON

A thesis submitted in fulfilment of the requirements for the Degree of Master of Science (Cognitive Science)

Faculty of Cognitive Sciences and Human Development
UNIVERSITI MALAYSIA SARAWAK
2009
ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Ng Giap Weng for his helpful advice, guidance and comments.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS i  
TABLE OF CONTENTS ii  
LIST OF FIGURES vi  
LIST OF TABLES ix  
ABSTRACT x  
ABSTRAK xi  

## CHAPTER 1 INTRODUCTION

1.0 Overview 1  
1.1 Background 2  
1.2 Problem Statement 4  
1.3 Research Objective 5  
1.4 Conceptual Framework 6  
1.5 Scope 7  
1.6 Significance of the Study 7  
1.7 Limitations 8  
1.8 Definition of Terms 8  
1.9 Summary 9  

## CHAPTER 2 LITERATURE REVIEW

2.0 Overview 11  
2.1 Augmented Reality (AR) 12  
2.2 Vision-Based Registration 17  

ii
2.3 Image Processing
  2.3.1 Greyscale Transformation
  2.3.2 Thresholding
  2.3.3 Labelling
  2.3.4 Edge Detection
2.4 Template Matching
2.5 Virtual Objects Generation
  2.5.1 Shading
  2.5.2 Texturing
  2.5.3 Animations
  2.5.4 Graphic API (Application Programming Interface)
2.6 Interaction
  2.6.1 Bare Finger Interaction
  2.6.2 Colour Segmentation
  2.6.3 Finger Tracking
2.7 Summary

CHAPTER 3 METHDOLOGY

3.0 Overview
3.1 Research Design
  3.1.1 System Specifications
  3.1.2 Collection of the Objects
  3.1.3 Construction and Modification of the Objects
  3.1.4 System Construction and Modification
### 3.1.5 Testing

### 3.1.6 Implementation

### 3.2 Evaluation

### 3.3 Summary

#### CHAPTER 4 DESIGN AND DEVELOPMENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0 Overview</td>
<td>62</td>
</tr>
<tr>
<td>4.1 Video Capturing with DirectShow</td>
<td>62</td>
</tr>
<tr>
<td>4.1.1 Sample Grabbing</td>
<td>66</td>
</tr>
<tr>
<td>4.2 Image Processing</td>
<td>69</td>
</tr>
<tr>
<td>4.3 AR Registration</td>
<td>70</td>
</tr>
<tr>
<td>4.4 Virtual Objects Generation with OpenGL</td>
<td>74</td>
</tr>
<tr>
<td>4.4.1 Shading</td>
<td>75</td>
</tr>
<tr>
<td>4.4.2 Texturing</td>
<td>78</td>
</tr>
<tr>
<td>4.4.3 Animations</td>
<td>79</td>
</tr>
<tr>
<td>4.5 Finger Tracking</td>
<td>82</td>
</tr>
<tr>
<td>4.6 Summary</td>
<td>84</td>
</tr>
</tbody>
</table>

#### CHAPTER 5 FINDINGS AND DISCUSSIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0 Overview</td>
<td>86</td>
</tr>
<tr>
<td>5.1 Video Capturing</td>
<td>87</td>
</tr>
<tr>
<td>5.2 Image Processing</td>
<td>89</td>
</tr>
<tr>
<td>5.3 Virtual Object Generation</td>
<td>90</td>
</tr>
<tr>
<td>5.4 AR Registration</td>
<td>93</td>
</tr>
<tr>
<td>5.5 Interaction</td>
<td>98</td>
</tr>
</tbody>
</table>
5.5.1 Finger Tracking

5.6 Evaluation

5.7 Summary

CHAPTER 6 CONCLUSION

6.0 Overview

6.1 Contributions of the Study

6.2 Recommendations for Future Study

6.3 Summary of the Research

REFERENCES

APPENDICES

Appendix 1: Consent Form

Appendix 2: Evaluation Form

Appendix 3: List of Participated Conferences and Exhibitions of this Research Project

Appendix 4: Source Code
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Conceptual Framework of AR Application Builder.</td>
<td>6</td>
</tr>
<tr>
<td>2.1</td>
<td>Examples of 2D square markers used in ARToolKit.</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>Example of markers of ARTag.</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td>Formula of greyscale transformation.</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Example of thresholded image.</td>
<td>25</td>
</tr>
<tr>
<td>2.5</td>
<td>Example of edge detection performed on an image.</td>
<td>28</td>
</tr>
<tr>
<td>2.6</td>
<td>Algorithm of converting RGB to HSV.</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>Combination of reuse model and exploratory model for system development process.</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>System architecture of AR Application Builder.</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>Flowchart of the implementation of AR Application Builder.</td>
<td>57</td>
</tr>
<tr>
<td>4.1</td>
<td>Filters and pins of DirectShow.</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>Snippet of creating filters using DirectShow.</td>
<td>65</td>
</tr>
<tr>
<td>4.3</td>
<td>Snippet of video capture device enumeration in DirectShow.</td>
<td>66</td>
</tr>
<tr>
<td>4.4</td>
<td>Example of Sample Grabber filter in DirectShow.</td>
<td>67</td>
</tr>
<tr>
<td>4.5</td>
<td>Snippet of creating Sample Grabber in DirectShow.</td>
<td>68</td>
</tr>
<tr>
<td>4.6</td>
<td>Snippet of greyscale transformation function.</td>
<td>69</td>
</tr>
<tr>
<td>4.7</td>
<td>Snippet of thresholding function.</td>
<td>70</td>
</tr>
<tr>
<td>4.8</td>
<td>The modification of arDetectMarker() for the compatibility of the system.</td>
<td>71</td>
</tr>
<tr>
<td>4.9</td>
<td>Snippet of marker detection.</td>
<td>72</td>
</tr>
</tbody>
</table>
Figure 4.10  Snippet of converting AR matrix to OpenGL matrix. 73
Figure 4.11  Statements with glNormal*() in OpenGL for the lighting. 75
Figure 4.12  Flat shading algorithm in OpenGL. 76
Figure 4.13  Vertex shading algorithm for OpenGL. 77
Figure 4.14  OpenGL statements of vertex shading. 77
Figure 4.15  Pseudo code of texture mapping in OpenGL. 78
Figure 4.16  Snippet of transformation animation in OpenGL. 80
Figure 4.17  Snippet of vertex animation in OpenGL. 80
Figure 4.18  Formula of interpolation for animation. 81
Figure 4.19  Snippet of interpolation for animation. 82
Figure 4.20  Snippet of colour segmentation based on Kovač, Peer, and Solina (2003). 83
Figure 4.21  Template of fingertip used in finger tracking 84
Figure 5.1  Screenshot of video capturing using DirectShow. 87
Figure 5.2  Screenshot of playing a video file using DirectShow and display in OpenGL as texture mapping. 88
Figure 5.3  Greyscale transformation of the coloured image. 89
Figure 5.4  Result image after thresholding is applied. 90
Figure 5.5  Flat shading of a 3D object in OpenGL. 91
Figure 5.6  Vertex shading of a 3D object in OpenGL. 91
Figure 5.7  Texture mapping on a 3D object in OpenGL. 92
Figure 5.8  The static frames of animation in OpenGL. 93
Figure 5.9  Screenshots of playing video in AR system. 94
Figure 5.10 Screenshots of 3D aeroplane as an AR game. 95

Figure 5.11 Screenshot of KLCC Twin Tower for tourism application. 95

Figure 5.12 Screenshot of animal cell in AR system as an education application for biology. 96

Figure 5.13 Screenshot of plant cell in educational application in biology. 97

Figure 5.14 Screenshots of interaction in AR system using mouse as an input device. 98

Figure 5.15 Screenshots of finger tracking as an input device for interaction in AR system. 99

Figure 5.16 Screenshots of AR video player and using finger as an input device to operate the virtual player. 101
LIST OF TABLES

Table 5.1 The result of the evaluation on the system functionality of the finger tracking for interaction with the virtual object in AR system. 102
ABSTRACT

AUGMENTED REALITY (AR) APPLICATION BUILDER: INTEGRATION OF FINGER TRACKING FOR INTERACTION IN AR SYSTEM

Augmented Reality (AR) is a technology that generates virtual information to the real world to the users perception and interaction to perform the tasks. This research studied and developed a computer library, AR Application Builder, which was integrated with the finger tracking feature for the interaction with the virtual objects in the AR system. The AR Application Builder allows users to build AR applications. Hence, the users are able to use their bare finger to interact with the virtual objects. The development of the AR Application Builder involved video capturing, image processing, vision-based AR registration, virtual objects generation, and finger tracking. Video capturing was implemented by using DirectShow. An image library was built to perform image processing functions: greyscale transformation and thresholding. ARToolKit libraries were integrated to implement vision-based registration. OpenGL and GLUT were used to draw the 3D computer graphics as the augmented objects. Finger tracking involved skin colour segmentation and the template matching. The template matching function was implemented by using OpenCV. Furthermore, an evaluation was conducted to test the system functionality of finger tracking for interaction with the virtual object in AR system. Thus, an AR video player was built for the evaluation. The result of the evaluation showed that the participants were able to use bare finger to interact with the virtual objects of the AR application. This allows the participants to interact with the virtual objects naturally without any other device.
ABSTRAK

PEMBINA APLIKASI AUGMENTED REALITY (AR):
INTEGRASI PENJEJAKAN JARI UNTUK INTERAKSI DALAM SISTEM AR

CHAPTER 1
INTRODUCTION

1.0 Overview

Augmented Reality (AR) Application Builder is a computer library integrated with the finger tracking for interaction with virtual objects in the AR system. The users are able to develop AR applications by using the library. The library allows users to render the AR environment with 2D and 3D virtual objects in the AR environment. The library requires a video capture device so that the users can capture real-time environment and render the AR environment. When they are rendering their AR environment, they can use a mouse device or a finger to interact with the virtual objects.

Besides developing the computer library, this research also studied on computer
vision techniques for AR registration. AR registration is a method to enable the virtual objects align properly in the real environment so that users can perceive the virtual objects in the correct position and orientation. Then the system will generate 3D computer graphics in the real environment.

AR also emphasises the interaction between the users and virtual objects to perform real-world tasks. The bare finger as an input device will produce a natural interaction for the users. This research implemented the finger tracking for the interaction with virtual objects in the AR system. This provides an AR system that allows the user to interact with the virtual objects with bare finger.

1.1 Background

AR is a deviation of Virtual Reality (VR) (Azuma, 1997). AR is a technology that allows users to see, hear, feel, and smell the virtual objects, which are integrated in the real world (Bonsor, 2001). AR technology generates 2D and 3D computer graphics which are accurately integrated into the real environment (Malik, 2002). AR technology is implemented on wearable computers (Starner, Mann, Rhodes, Levine, Healey, Kirsch, Picard, & Pentland, 1997). Wearable and mobile computers, such as wristwatch, laptop, palmtop, tablet PC, head-mounted display (HMD), and head-up display (HUD) can be worn by the users. Currently, AR is used in medicine, repair and maintenance, labelling, robotics, military, entertainment, education, games, and other fields (Azuma, 1997; Billinghurst,
2002; Thomas, 2003). Besides that, AR will be one of the technologies which emerges with the phone technology in future (Crago, 2008).

The biggest problem for AR is the registration problem (Azuma, 1997; Zhou, Wang, Yan, & Xu, 2000). Computer vision and image processing techniques are used in AR to solve the registration problem (Azuma, 1997; Klinker, 1999). Computer vision is one of the Artificial Intelligence (AI) fields. The goal of computer vision is to analyse and interpret images (Schiewagen, 2001). Besides that, computer vision also simulates human being’s visual perception (Meer, Stewart, & Tyler, 2000). Perception is the conscious experience when our senses receive stimuli from the environment (Goldstein, 2005). Moreover, perception is the gateway to other cognitions such as memory, attention, reasoning, problem solving, and decision making (Goldstein, 2005). Therefore, a camera is needed to simulate our eyes to receive the information from the real environment.

Moreover, using techniques of computer vision and image processing, the finger can be used as an input device for the interaction in the AR environment (Hardenberg, 2001). For the interaction in the AR environment, wired gloves are one of the devices mostly used (Azuma, 1997). Furthermore, using the finger as an input device can create natural interaction. Natural interaction is the interaction that does not use any devices (Hardenberg, 2001).

ARToolKit is an open-source software library. It is frequently used in developing
Augmented Reality and Mixed Reality applications (Haller, Hartmann, Luckeneder, & Zauner, 2002). ARToolKit uses computer vision techniques to detect a fiduciary marker. Camera orientation and position are computed based on the detected fiduciary marker. ARToolKit provide interaction with the virtual object by using another fiduciary marker on a handheld paddle (Gordon, Billinghurst, Bell, Woodfill, Kowalik et al., 2002). However, ARToolKit does not provide natural interaction such as the bare finger.

1.2 Problem Statement

AR enhances the users' senses by augmenting a virtual layer on the real world especially visual senses (Aaltonen & Lehikoinen, 2006). Interaction is also an important aspect in AR (Azuma, 1997). The virtual objects will not only enhance the visual perception of the users, but also assist the users to accomplish the real-world tasks.

Interaction with the pure virtual information in AR is difficult (Azuma, Baillot, Behringer, Feiner, Jullier, & MacIntyre, 2001). The AR prototypes in various previous research normally used the keyboard or mouse for the interaction. Moreover, some of the AR systems used gesture recognition or tracking 6DOF (degrees of freedom) pointers. The open source ARToolKit library allows the users to use paddle to interact with the virtual object. A marker is needed on the paddle so that the paddle can be tracked (Irawati, Green, Billinghurst, Duenser, &
Ko, 2006). However, bare finger interaction with the virtual object is difficult to be implemented due to the intangibility of the virtual objects. Hence, this research has implemented several computer vision algorithms to make the bare finger interact with the virtual object.

Integration of finger tracking in AR system requires finger detection and recognition. However, the conditions of the environment will affect the accuracy of the finger detection, especially background colour and the lighting of the environment. These factors will affect the finger detection for the interaction.

Furthermore, AR system performs registration, virtual objects generation, and interaction in real-time. If any of these methods is computationally complicated and requires a lot of time for the execution, system latency will be occurred. Hence, integration of the finger tracking in AR system not only requires an optimal algorithm for registration, but also an algorithm for the finger tracking which will not affect the system latency.

1.3 Research Objective

The general objective of this research is to design and develop Augmented Reality (AR) Application Builder, which allows users to build their own AR application and the integration of bare finger interaction with the virtual objects.

The specific objects are as follows:
1. To develop AR Application Builder which allows the users to create the AR application.

2. To implement the computer vision algorithms: colour segmentation and template matching, to detect the bare finger.

3. To integrate the finger tracking with the AR applications, so that the finger can be used as an input device to interact with the virtual objects in the real environment.

1.4 Conceptual Framework

Figure 1.1: Conceptual Framework of AR Application Builder.

Figure 1.1 shows the conceptual framework of this research. AR combines the real world with the virtual objects to enhance the users' senses especially visual senses (Aaltonen & Lehikoinen, 2006). However, human visual perception is very sensitive to the error of the registration (Simon & Berger, 2000). Computer vision
algorithms which simulates human visual perception is used to understand the real
world information. Consequently, the registration of the real world and virtual
objects can be performed to generate correct scene to the users. The virtual objects
becomes stimuli for the visual perception of the users. The users can use bare
finger to interact with the augmented objects. In order to understand the
interaction with bare finger, computer vision algorithms are used for finger
tracking to allows the natural interaction happened in AR system.

1.5 Scope

This study focuses on the development of a computer library, AR Application
Builder, which allows users to interact the virtual objects with the bare finger in
an AR environment. The computer vision algorithms were implemented for the
finger tracking. Furthermore, the interaction of the virtual objects were also
implemented.

1.6 Significance of the Study

The key contribution of this research is the integration of the finger tracking in the
AR system for interaction. There were several studies on the interactions of the
AR system. Finger tracking as an input device in AR was done in previous
research (Dorfmüller-Ulhaas & Schmalstieg, 2002). However, this research
implemented the bare finger as an input device for the interaction in AR.
1.7 Limitations

There are several limitations in this research. The bare finger tracking did not provide any depth information. Consequently, the finger cannot interact with the virtual objects with depth information.

1.8 Definition of Terms

Several terms are used in this research. The literal definition of terms used in this research are defined as follows:

- Augmented Reality (AR) – AR is a technology that enhances the users’ senses by augmenting a virtual layer on the real world (Aaltonen & Lehikoinen, 2006).
- Registration – The real and virtual worlds coexist and the objects are aligned correctly (Azuma, 1997).
- Virtual object – The computer generated objects which are superimposed on top of the video of the real world in AR system (Buchmann, Violich, Billinghurst, & Cockburn, 2004).
- Finger tracking – The finger of the user can be tracked by the system and interaction can be performed (Crowley, Berard, & Coutaz, 1995).
- Barehanded interaction – An interaction using bare hand without using any device and wires (Hardenberg & Bérard, 2001).

For the research purpose, the operational definition of terms are defined as
follows:

- **Augmented Reality (AR)** – Generation of virtual objects in the real environment to enhance the meaning of the real world to the users.

- **Registration** – A process to align the virtual object and the real environment properly in order to generate a correct scene in AR.

- **Virtual object** – The computer generated graphics especially 3D computer graphics which uses three-dimensional representation of geometric information in AR.

- **Finger tracking** – Detection of the finger so that the finger can be used as an input device in the AR system.

- **Barehanded interaction** – Finger as a part of the hand to interact with the virtual objects in the AR system without using any other device.

### 1.9 Summary

Augmented Reality (AR) Application Builder is a computer library integrated with finger tracking for the interaction with virtual objects in the AR system. The users are able to develop AR applications by using the library. The users can interact with the virtual objects by using finger as an input device. AR is a technology which augments a virtual layer on the real world especially visual senses for the users. The virtual objects are used to assist users to accomplish real-world tasks. Thus, interaction is needed between the users and the virtual objects. Interaction with the pure information in AR is difficult. This causes bare finger
interaction with the virtual object difficult to be implemented. Therefore, the objective of this research is to design and develop AR Application Builder which is integrated with the bare finger interaction with the virtual objects, and the users are allowed to use the library to build their own AR applications. The scope of this research focuses on the development of the computer library. Computer vision algorithms for finger tracking is implemented. The key contribution of this research is the integration of the finger tracking in the AR system for interaction. However, the limitations of this research is the bare finger tracking does not contain any depth information.
CHAPTER 2  
LITERATURE REVIEW 

2.0 Overview

This chapter discusses Augmented Reality (AR). In this research, vision-based registration for AR was used. Therefore, image processing techniques were used for vision-based registration. The image processing techniques were greyscale transformation, thresholding, labelling, and edge detection. Furthermore, template matching was used to identify the fiduciary marker. The generation of the virtual objects is also discussed. Interaction is implemented in this research. The finger is used as a natural input device in the AR system. The implementation of the finger tracking is discussed in this chapter.