



This is to certify that the  
dissertation entitled

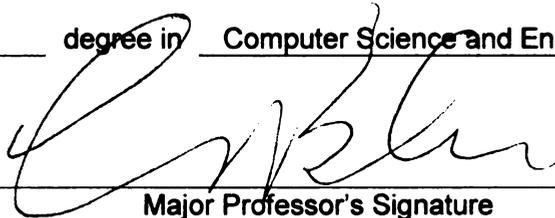
DYNAMIC CONTEXTUALIZATION USING AUGMENTED  
REALITY

presented by

WEI ZHU

has been accepted towards fulfillment  
of the requirements for the

Ph.D. degree in Computer Science and Engineering



Major Professor's Signature

6 - 28 - 2006

Date

*MSU is an Affirmative Action/Equal Opportunity Institution*





**DYNAMIC CONTEXTUALIZATION USING AUGMENTED  
REALITY**

BY

Wei Zhu

A Dissertation

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Computer Science and Engineering

2006

# **ABSTRACT**

## **DYNAMIC CONTEXTUALIZATION USING AUGMENTED REALITY**

**By**

**Wei Zhu**

This thesis investigates the technical possibility and human factors associated with dynamic contextualization using Augmented Reality technologies. Dynamic contextualization, a new term developed during this research, goes beyond traditional context-aware computing. Dynamic contextualization not only discovers and utilizes context information as a factor in the user interface design, but also modifies context using augmentations and diminishments, virtually contextualizing real object with virtual objects that have connections to the real object. Dynamic contextualization is a new method of human computer interaction. Since the context setting can be manipulated by dynamic contextualization, people's perspectives of the real world can be directed and influenced.

Further more, this thesis discusses the application of employing context-aware computing into Augmented Reality systems. Context-aware augmented reality systems can selectively provide more relevant information to users so that the augmentations are more meaningful and useful without disturbing users with excessive information. In addition to providing more relevant information, context-aware computing also provides augmented reality system with more realistic blending if the geometry context and illumination context are taken into

account. For example, the virtual elements will be more realistic if they can be lit and shadowed as if the light source was the same as that of the real world.

This research investigates the context-aware solutions and dynamic contextualization for augmented reality systems based on the PromoPad test-bed system an augmented reality powered shopping assistant. Several user studies conducted using the system demonstrate the effectiveness of dynamic contextualization.

This thesis first presents an overview of context-aware augmented reality technologies. Then the concept of dynamic contextualization is presented in detail. The thesis discusses the design of PromoPad, the test-bed system and user experiments and data analysis that illustrate the effectiveness of the device.

Copyright by  
WEI ZHU  
2006

**To my parents**

## **ACKNOWLEDGEMENTS**

I owe my sincere thanks to Dr. Charles B. Owen, my principle advisor. His continuous support helped me overcome obstacles during my Ph.D. study. I learned a lot from his excellent and patient consultancy and professional editing. His hard work attitude provided a solid ground of this research, and in addition, is a great role model in my future career. It is a great pleasure to work with Dr. Hairong Li. Dr. Li's strong theoretical background of marketing and advertising makes this research and test-bed possible. His generous help also made this research smoothly carried out. I wish to express my sincere thanks to Dr. Mutka and Dr. Esfahanian for spending their valuable time monitoring my work, reading this thesis, and giving thoughtful suggestions.

I am indebted to my parents, Lingfei Zhu and Tengfang Dai, for their unconditional love and emotional support.

Finally, I extend my thanks to my family. My husband, Feng Zhu, has been continuously supporting me by all means. He has been accompanying with me, encouraging me, and supporting me, no matter what happens. A lot of times, he took care of the family so that I could have more time working on my research, while he was also pursuing a Ph.D. degree. Nothing meant more to me than his encouragement and support in the journey of pursuing three advanced degrees. Without him, I would not have been able to reach this point. My children, Anthony and Lillian, are the best gift to me. Thanks for the happiness and laughter they bring to me.

# TABLES OF CONTENTS

|   |             |
|---|-------------|
| <b>LIST OF FIGURES .....</b>  | <b>X</b>    |
| <b>LIST OF TABLES .....</b>   | <b>XIII</b> |
| <b>CHAPTER 1. INTRODUCTION.....</b>   | <b>1</b>    |
| <b>1.1. RESEARCH MOTIVATION.....</b>  | <b>5</b>    |
| <b>1.2. CONTRIBUTIONS .....</b>   | <b>6</b>    |
| <b>1.3. THESIS OUTLINE.....</b>   | <b>8</b>    |
| <b>CHAPTER 2. BACKGROUND AND RELATED WORKS.....</b>                                     | <b>9</b>    |
| <b>2.1. AUGMENTED REALITY .....</b>   | <b>9</b>    |
| 2.1.1. AR DEFINITION.....   | 9           |
| 2.1.2. AR COMPONENTS .....  | 12          |
| 2.1.3. HUMAN FACTORS IN AUGMENTED REALITY .....   | 15          |
| <b>2.2. CONTEXT-AWARE COMPUTING .....</b>   | <b>15</b>   |
| <b>2.3. CONTEXT-AWARE AUGMENTED REALITY SYSTEMS .....</b>                               | <b>17</b>   |
| <b>2.4. AUGMENTED REALITY IN ADVERTISING.....</b>                                       | <b>22</b>   |
| <b>2.5. AR-ORIENTED CONTEXT-AWARE MODELS .....</b>                                      | <b>24</b>   |
| 2.5.1. SPATIAL MODEL .....  | 24          |
| 2.5.2. REGIONAL-BASED MODEL.....  | 25          |
| 2.5.3. RULE-BASED MODEL .....   | 25          |
| 2.5.4. MACHINE LEARNING MODEL.....  | 25          |
| 2.5.5. CURRENT RESEARCH IN CONTEXT-AWARE AR .....                                       | 27          |
| <b>2.6. DISCUSSION AND SUMMARY .....</b>  | <b>28</b>   |
| <b>CHAPTER 3. DYNAMIC CONTEXTUALIZATION: DESIGN AND IMPLEMENTATION OF PROMOPAD.....</b> | <b>29</b>   |

|   |   |           |
|---|---|-----------|
| <b>3.1.</b>   | <b>INTRODUCTION.....</b>                                      | <b>29</b> |
| <b>3.2.</b>   | <b>THE PROMOPAD SYSTEM.....</b>                               | <b>32</b> |
| <b>3.3.</b>   | <b>AUTOMATED CONTEXT-AWARE ASSISTANCE.....</b>                | <b>34</b> |
| 3.3.1.  | USER'S LOCATION AND ORIENTATION .....                         | 35        |
| 3.3.2.  | USER PROFILE.....   | 36        |
| 3.3.3.  | PRODUCT CONTEXT .....   | 37        |
| <b>3.4.</b>   | <b>TECHNICAL ISSUES.....</b>                                  | <b>38</b> |
| 3.4.1.  | IN-STORE TRACKING .....                                       | 38        |
| 3.4.2.  | VIDEO SEE-THROUGH SYSTEMS.....                                | 39        |
| 3.4.3.  | REALTIME INVERSE LIGHTING.....                                | 47        |
| 3.4.4.  | PERFORMANCE ANALYSIS.....                                     | 51        |
| 3.4.5.  | WORKING SYSTEM.....   | 54        |
| <b>3.5.</b>   | <b>SUMMARY .....</b>  | <b>59</b> |
| <b>CHAPTER 4. DYNAMIC CONTEXTUALIZATION AND MARKETING PERCEPTIONS .....</b> |   | <b>62</b> |
| <b>4.1.</b>   | <b>COMPLEMENTARY PRODUCTS.....</b>                            | <b>63</b> |
| <b>4.2.</b>   | <b>DYNAMIC CONTEXTUALIZATION OVERVIEW .....</b>               | <b>64</b> |
| <b>4.3.</b>   | <b>DYNAMIC CONTEXTUALIZATION WITH AUGMENTED REALITY .....</b> | <b>66</b> |
| 4.3.1.  | AUGMENTING CONTEXT.....                                       | 67        |
| 4.3.2.  | DIMINISHING CONTEXT.....                                      | 69        |
| <b>CHAPTER 5. EMPIRICAL STUDIES .....</b>                                   |   | <b>71</b> |
| <b>5.1.</b>   | <b>INTRODUCTION.....</b>                                      | <b>71</b> |
| <b>5.2.</b>   | <b>USER STUDY 1: PRODUCT CONTEXTUALIZATION .....</b>          | <b>73</b> |
| 5.2.1.  | EXPERIMENT DESCRIPTION.....                                   | 73        |
| 5.2.2.  | METHODOLOGIES .....   | 75        |
| 5.2.3.  | PARTICIPANTS.....   | 77        |
| 5.2.4.  | PROCEDURE.....  | 77        |

|                           |   |            |
|---------------------------|---|------------|
| 5.2.5.                    | DATA ANALYSIS.....                                  | 78         |
| 5.2.6.                    | EXPERIMENT SUMMARY .....                            | 82         |
| <b>5.3.</b>               | <b>USER STUDY 2: DIMINISHING CONTEXT .....</b>      | <b>82</b>  |
| 5.3.1.                    | EXPERIMENT DESCRIPTION .....                        | 83         |
| 5.3.2.                    | METHODOLOGIES.....                                  | 85         |
| 5.3.3.                    | DATA ANALYSIS.....                                  | 86         |
| <b>5.4.</b>               | <b>USER STUDY 3: FUNCTIONAL COMPLIMENTARY .....</b> | <b>91</b>  |
| 5.4.1.                    | METHODOLOGIES.....                                  | 95         |
| 5.4.2.                    | PARTICIPANTS.....                                   | 95         |
| 5.4.3.                    | PROCEDURE.....                                      | 95         |
| 5.4.4.                    | DATA ANALYSIS.....                                  | 96         |
| <b>5.5.</b>               | <b>USER STUDY 4: 3D VIRTUAL CONTEXT .....</b>       | <b>102</b> |
| 5.5.1.                    | EXPERIMENT DESCRIPTION .....                        | 102        |
| 5.5.2.                    | METHODOLOGIES .....                                 | 104        |
| 5.5.3.                    | PARTICIPANTS.....                                   | 104        |
| 5.5.4.                    | PROCEDURE.....                                      | 104        |
| 5.5.5.                    | DATA ANALYSIS.....                                  | 104        |
| <b>5.6.</b>               | <b>USER STUDY 5: USAGE PATTERN ANALYSIS.....</b>    | <b>106</b> |
| 5.6.1.                    | TIME PATTERN .....                                  | 106        |
| 5.6.2.                    | MOVEMENT PATTERN .....                              | 109        |
| <b>5.7.</b>               | <b>USER STUDY 5: FEASIBILITY ANALYSIS .....</b>     | <b>110</b> |
| <b>5.8.</b>               | <b>SUMMARY .....</b>                                | <b>112</b> |
| <b>CHAPTER 6.</b>         | <b>SUMMARY AND FUTURE WORKS.....</b>                | <b>114</b> |
| <b>BIBLIOGRAPHY .....</b> |   | <b>117</b> |
| <b>APPENDIX A.</b>        | <b>SUMMARY OF SURVEY QUESTIONS .....</b>            | <b>125</b> |
| <b>A.1</b>                | <b>PRE-EXPERIMENT SURVEY QUESTIONS.....</b>         | <b>125</b> |

|            |   |            |
|------------|---|------------|
| <b>A.2</b> | <b>POST-EXPERIMENT SURVEY QUESTIONS .....</b> | <b>126</b> |
| <b>A.3</b> | <b>SUMMARY OF DATA ANALYSIS.....</b>          | <b>129</b> |

## LIST OF FIGURES

|  |    |
|--|----|
| Figure 1 The PromoPad system .....   | 4  |
| Figure 2 “Virtuality continuum” by Milgram. [13].....                                | 10 |
| Figure 3 Architecture of AR systems.....   | 12 |
| Figure 4 The KARMA system at Columbia University [33].....                           | 19 |
| Figure 5 The Archeoguide system [42] .....   | 21 |
| Figure 6 Virtual advertising image samples from PVI [55].....                        | 24 |
| Figure 8. Clustering user preference by machine learning technology [41].....        | 27 |
| Figure 9 Using the PromoPad in a store setting.....                                  | 33 |
| Figure 10 the experimental shelf with fiducial images .....                          | 36 |
| Figure 11 Perspective camera projection model .....                                  | 40 |
| Figure 12 Perspective view of frustums .....   | 42 |
| Figure 13 Vertical 2-D view of the perspective frustums .....                        | 44 |
| Figure 14 Tablet PC displays from different viewpoints .....                         | 46 |
| Figure 15 Occlusion model .....  | 47 |
| Figure 16 A frame from the captured video sequence .....                             | 56 |
| Figure 17 Illustration of number of points versus performance .....                  | 57 |
| Figure 18 A frame with a virtual ball and teapot lit with estimated light source ... | 59 |
| Figure 19 Augmenting the box of spaghetti with cooked spaghetti and sauce....        | 68 |
| Figure 20 Augmenting the background .....  | 69 |
| Figure 21 Diminishing context.....   | 70 |
| Figure 22 Spaghetti and sauce can .....  | 74 |
| Figure 23 the view in the PromoPad for two treatment levels .....                    | 75 |
| Figure 24 Histogram of effect on product association .....                           | 78 |
| Figure 25 Box plot of effect on product association .....                            | 79 |

|  |     |
|--|-----|
| Figure 26 Histogram of effect on purchase intent .....   | 81  |
| Figure 27 Box plot of effect on purchase intent.....   | 81  |
| Figure 28 Wines.....   | 83  |
| Figure 29 Two levels of treatment with wines .....   | 84  |
| Figure 30 Histogram of effects on product promotion status.....                                      | 87  |
| Figure 31 box plot of effects on product promotion status.....                                       | 87  |
| Figure 32 Histogram of effects on purchase intent.....   | 89  |
| Figure 33 Box plot of effects on purchase intent .....   | 90  |
| Figure 34 Functional complementary of camera (tripod) and wine (wine glasses)<br>.....               | 92  |
| Figure 35 Original shelf with real focal products (digital camera and wine).....                     | 94  |
| Figure 36 High involvement complementary treatment.....  | 94  |
| Figure 37 Low involvement complementary treatment.....   | 94  |
| Figure 38 Histogram of rating on digital camera with two levels of complementary<br>involvement..... | 97  |
| Figure 39 Box plot of rating on digital camera with two levels of complementary<br>involvement.....  | 97  |
| Figure 40 Participants rating on cameras and tripods in pair .....                                   | 98  |
| Figure 41 Histogram of rating on wine with two levels of complementary<br>involvement.....           | 100 |
| Figure 42 Box plot of rating on wine with two levels of complementary<br>involvement.....            | 100 |
| Figure 43 Participants rating on wine and glasses in pair .....                                      | 102 |
| Figure 44 Virtual context .....  | 103 |
| Figure 45 Scores on likableness .....  | 105 |
| Figure 46 Box plot of likableness .....  | 105 |
| Figure 47 Start tracking time .....  | 107 |

|   |     |
|---|-----|
| Figure 48 Effective in use time.....                                      | 108 |
| Figure 49 Total time .....  | 108 |
| Figure 50 Camera movement on shelf background (with augmentations) .....  | 110 |
| Figure 51 Camera movement on shelf background (without augmentations) ... | 110 |
| Figure 52 Histogram of feasibility scores .....                           | 111 |
| Figure 53 Box plot of feasibility scores.....                             | 112 |

## LIST OF TABLES

|  |     |
|--|-----|
| Table 1 Convergence time (ms) comparison for group A (random initial guess) and group B (initial guess supplied with our strategies) ..... | 53  |
| Table 2 Number of points versus performance .....  | 57  |
| Table 3 Product complementary examples .....   | 64  |
| Table 4 Examples of augmentations and diminishments .....  | 70  |
| Table 5 ANOVA table for perception of product connection.....  | 80  |
| Table 6 ANOVA table for consumer's purchase intent.....  | 82  |
| Table 7 ANOVA table for perception of wines.....   | 88  |
| Table 8 ANOVA table for purchase intent of wines .....   | 90  |
| Table 9 Experiment scenario .....  | 92  |
| Table 10 Experiment settings for each treatment.....   | 93  |
| Table 11 Summary of time pattern.....  | 108 |
| Table 12 Summary of feasibility analysis .....   | 111 |

# CHAPTER 1. INTRODUCTION

The *context* of an object or event is the surroundings in which it exists. The value of a product on a store shelf is influenced by the products and advertising that surround it. The perceived quality of a tool is influenced by how it is being used and the quality of the materials it is used on. Context has a significant influence on perception, a fact well known by those who package, advertise, and sell. They use context to positively (or sometimes negatively) impact the perception of products.

Heretofore, the modification of context of physical items was also physical, static, and competitive. The context of a physical item is, itself, a physical setting involving placement of real items and media around the item to define that context. The context is set and did not change and the use of moving elements is severely limited by physical constraints. And, all objects within the range of perception become context for all other objects, so emphasizing one product in a store setting often requires deemphasizing another. Context becomes a careful and expensive balancing act.

The advent of augmented reality computer technologies allows for a new concept in contextualization wherein the perceived is continuously and automatically modified. This thesis introduces *dynamic contextualization*, the computer-mediated modification of perceived context using augmented reality. *Augmented reality* (AR) technologies enhance our perception of reality through the employment of computer-generated augmentations [1]. These augmentations can include appearance, sound, touch, and other sensations. In this thesis we will employ augmentations involving 3-D graphic images

that appear to coincide with real-world imagery (though the general concepts could be applied to augmentation of other senses as well). Augmented reality blends computer-generated virtual elements with images captured in the real world. A user perceives both the real world and the augmentations, ideally as if the augmentations were real elements of the world as well.

AR differs from VR (virtual reality) in that there is no attempt to escape or replace the real environment. Instead, AR enhances our perception by contextualizing individual objects we encounter in reality so that these objects can become more meaningful, useful and appealing. Therefore, in an AR environment, users can interact with the real world and move around in the environment.

A goal of *context-aware computing* [2, 3] is to free users from being flooded with excessive information by selecting appropriate information for the current user *context* or modifying the presentation of that information to more effectively suit the current context. Recent advances in mobile technologies such as wireless networks and communications allow new ways of using computing devices. Computing devices are, by no means, restricted to offices and homes any more. People use different computing devices to capture outside information and utilize that information to assist their daily lives. As this quantity of information grows, context-aware computing technologies have emerged to help automatically manage the information and provide the most relevant information to the user. As opposed to information filtering, context-aware computing is concerned with appropriate information *selection*. Admittedly, this is a fine distinction, but this does require a somewhat different approach to the problem.

As key AR technologies, such as tracking and composition, continue to mature, AR will soon be available to a large range of applications, from entertainment to military training. Context-awareness, however, has gotten less attention in AR systems research. In many AR applications, especially mobile AR systems, context-awareness helps improve information presentation to users, while still allowing users to safely move within and interact with the real world.

Dynamic contextualization uses AR technologies to modify the context of a real (physical) object. Additional virtual elements can be placed around the object to give it context. Also, elements of the real world can be removed so as to eliminate distracting or competing context. New virtual elements can be rendered that complement the object under contextualization. *Virtual contextualization*, one specific instance of dynamic contextualization, connects real objects with virtual objects that have a certain *connection* between them. For example, golf balls, hats, and shirts can virtually contextualize golf clubs, even though there are no balls, hats, and shirts physically presented. The relationship in this example is a functional one in that the items are normally used together. With dynamic contextualization, one or more real objects can be highlighted from others by surrounding virtual objects. Thus, dynamic contextualization is able to manipulate users' interests by augmentations and diminishment.

PromoPad is a prototype AR-based shopping assistant. PromoPad serves as the test-bed for studying context-aware computing in augmented reality and the feasibility and effectiveness of dynamic contextualization. It is a Tablet-PC device that presents the image of a rear-mounted camera on the display with computer generated augmentations. These augmentations provide additional information to users in the form of augmented

images. The augmentations are based on location, products under inspection, and user preferences. In addition, this specific dynamic contextualization implementation modifies the context through augmentations consisting of complementary products and diminishments of competitive products to reach the needs of store-wide advertising and shopping assistance by AR technologies. Figure 1 shows a user using PromoPad to observe the products on the shelf.

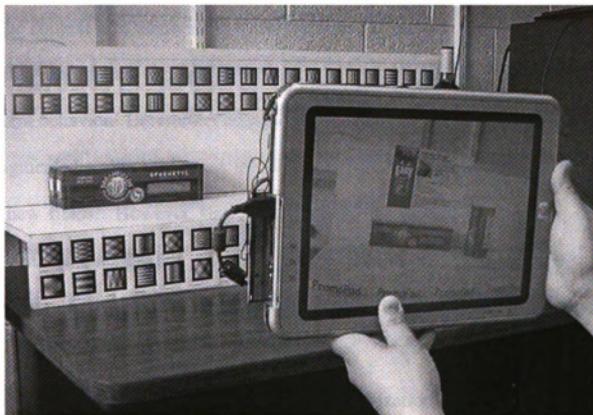


Figure 1 The PromoPad system

This thesis surveys the current advances in the area of augmented reality and context-aware computing, and then discusses the design and implementation of PromoPad. Marketing and advertising issues are addressed as well as the technical issues of implementation. This thesis also analyzes empirical results of user evaluations that explore the effectiveness of dynamic contextualization. Research issues and possible future work in both the area of AR and context-awareness are discussed.

## 1.1. RESEARCH MOTIVATION

AR technologies not only enhance a user's perception of the real world [1, 4], but also introduce a new form of human computer interaction [5]. Context-aware computing seeks to select appropriate information for presentation to a user based on the current user state; the *context* of the user. These two technologies are related in that both are highly dependent on the physical context parameters such as location and orientation. In addition, the context-area elements of system design can include statistical elements such as past behavioral habits or population trends, or emotional elements derived from a user's current behavior. Heretofore, context-aware computing has not been studied in the field of augmented reality other than the simple issue of location awareness, as both are relatively new fields. However, there is a potential for more powerful and useful AR systems through the application of context-aware technologies. PromoPad is a test-bed that allows us to experiment with the application of context-aware computing to augmented reality.

AR technologies make dynamic contextualization possible. Dynamic contextualization is a new term proposed during the design of PrompPad [6, 7]. In stead of augmenting the "where" and "when" to augment, as most AR systems do, dynamic contextualization is more interested in "what" to augment, and the effect on the user of this augmentation, specifically the effect on user motivation. Augmented reality is a natural application for context-aware computing because the amount of augmentations that can be practically perceived by a user is limited and a large amount of context is available to applications to drive a context-aware engine. Context-aware computing has typically been limited to applications that are providing information to users. In an AR

application, the computing device is providing not only information to the user, but also a potentially modified context. Hence, the context-aware system can be part of the context delivery. User context is very much dependent on perception and an AR system can modify perception and, consequently, the perceived context. Dynamic contextualization not only utilizes context, as most context-aware systems do, but also modifies the context using AR technologies. It is not only context-aware, but also context-modifying.

As an AR-enabled shopping assistant, PromoPad is a good test-bed to experiment with dynamic contextualization and context-aware computing in AR systems. Virtual experiences and 3-D product visualization have previously been proven to be able to stimulate customer learning of product characteristics and better understand of the product [8-10]. Augmented reality can provide a useful medium for the realization of the types of virtual experiences and 3-D product visualization. The PromoPad project also has practical value since in-store advertising is a major factor driving potential impulse purchases [11] and annual retail grocery shopping in United States alone is a huge volume business [12]. Dynamic contextualization can also be extended and adapted to other application domains such as education, training, and tourism. Hence dynamic contextualization is general purpose concept, not limited to the shopping scenarios.

## **1.2. CONTRIBUTIONS**

The major contribution of this thesis is the development and evaluation of dynamic contextualization as a new method for using augmented reality in advertising applications. The PromoPad system was developed as a test-bed to experiment with context-aware AR and dynamic contextualization. The thesis explores one specific

application of augmented reality to the field of advertising as a test-bed for context-aware computing ideas and dynamic contextualization. A number of technical issues in realizing context-aware in AR systems have also been examined and resolved, including realistic blending of augmentations, in-store tracking, and context sensing. These issues will be discussed in detail in Chapter 3.

Dynamic contextualization, the modification of context to influence users, is studied both theoretically and practically. Based on theories in marketing and advertising, several scenarios for dynamic contextualization were designed and experiments were conducted to evaluate the effectiveness and feasibility of the solutions. Empirical studies were carried out to test the effect of dynamic contextualization in the form of augmenting context, diminishing context and functional complementary. Data analysis and statistical tests show that dynamic contextualization has significant effect on influencing users' perception of the real objects, directing users' interests. Studies in usage pattern of PromoPad with and without augmentation revealed that users tend to spend more time on focal products with augmentations. Feasibility analysis showed that dynamic contextualization is readily deployable. The results of the analysis are listed here; detailed user evaluation and data analysis is presented in Chapter 5.

1. Augmenting context has positive effect on influencing consumers' perception of the focal products.
2. Augmenting context has positive effect on influencing consumers' purchase intent.
3. Diminishing context has positive effect on highlighting the focal products.
4. Virtual functional complementary has positive effect on influencing consumers' attitude towards the focal products.

5. 3D virtual context has positive effect on influencing consumers' perception of the focal products compared to 2D virtual context.
6. Users tend to spend more time on focal products with augmented imagery than those without augmented imagery.

Although the concept of dynamic contextualization was implemented and tested on the test-bed system, an AR powered shopping assistant specifically focused on advertising and product promotion, it can be adapted to other application domains such as education, training, tourism guiding, and so on.

### **1.3. THESIS OUTLINE**

Chapter 2 introduces the concept and key components of augmented reality and context-aware computing. Research projects in context-aware augmented reality systems and AR-oriented context-awareness models are surveyed in this section as well.

The design of the PromoPad is discussed in Chapter 3. The concept of dynamic contextualization is introduced in this chapter and discussed in detail. Possible realizations of dynamic contextualization using augmented reality technologies are addressed. Some technical details involved in the development of PromoPad are also presented in Chapter 3.

Chapter 4 presents the marketing perspective of dynamic contextualization, some related concepts in consumer behavior, psychology and advertising. Chapter 5 discusses the methodologies, procedures, and analysis of user evaluations. The results of data analysis are illustrated. Chapter 6 summarizes this thesis and points our future research directions.

## **CHAPTER 2. BACKGROUND AND RELATED WORKS**

The work of this thesis, and the PromoPad prototype design specifically, combines work in augmented reality, context-aware computing, advertising, and human factors design. This work drew upon a deep well of practical and theoretical support in each of these fields. This chapter presents much of the background material surveyed and utilized during the design of PromoPad and as the basis for the presented experimentation.

### **2.1. AUGMENTED REALITY**

Augmented reality enhances the perception of the real world by augmenting it with computer-generated elements, be they sound, visual, or any of the senses. In the past decade, AR technologies have drawn much attention and an alternative to virtual reality that allows interaction with the real world rather than an alternative reality. Considerable ongoing research is making AR technologies feasible in various application domains. This section briefly introduces the definition, components, and application of AR systems.

#### **2.1.1. AR DEFINITION**

Azuma's definition of AR [1] has been widely adopted in the AR community for applications involving augmented vision. In his 1997 survey [1] and supplemental work on 2001 [4], Azuma defines AR as having the following three characteristics:

1. Combines real world and virtual objects

2. Interacts with the real world in real time
3. Virtual objects are registered with real world in 3-D

The first characteristic indicates the position of AR systems in the “virtuality continuum” proposed by Milgram [13] and illustrated in Figure 2. With real environments and virtual environments at the two opposite ends of the continuum, any systems in between are considered to fall into the domain of AR.

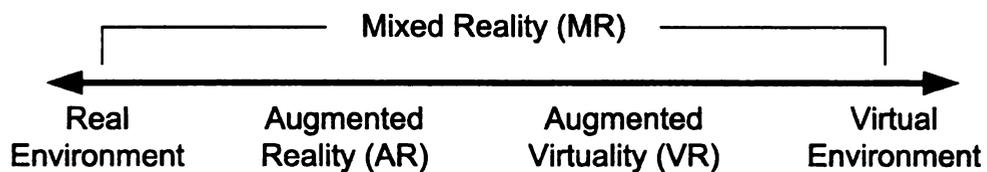


Figure 2 “Virtuality continuum” by Milgram. [13]

The second characteristic addresses the real time requirement of AR systems. This is the key point that distinguishes AR systems from the augmented imagery utilized in films that blend virtual objects with real scenes such as “Who Framed Roger Rabbit?” and “Jurassic Park”. In AR systems, virtual objects are superimposed on real world objects in real time; unlike in the films, in which the blending is done offline. This definition is somewhat limiting in that others consider augmented reality to include off-line composition technologies, particularly those requiring 3-D registration. However, for the purposes of this proposal, real-time will be considered a requirement in that context-aware systems are not practical unless on-line and able to process context in real time.

The third characteristic, 3-D registration, ensures that AR systems present both the virtual and real information in a seamless form such that both paradigms properly align to each other. This requires the systems to be aware of the 3-D position and orientation of the user relative to the environment. This characteristic also distinguishes AR

applications from applications that overlay 2-D virtual image over live video such that the overlay is not registered in 3-D with the reality.

Azuma's definition of AR systems doesn't limit AR systems to the use of Head-Mounted Displays (HMDs). Any displaying technologies can be used as long as the three essential characteristics are present. This definition also allows the augmentation of senses other than sight. AR technologies could be extended to 3-D sound, haptics, or even other senses in the future.

Mackay [5] depicted AR as a new paradigm of human and computer interaction. In her vision, AR is a revolution in computer interface design that changes the way we think about and use computers. She described three approaches to augment reality: 1) augment the user; 2) augment the physical object; and 3) augment the environment surrounding the user and object. Augmenting the user is accomplished using a variety of devices that users wear to see both the real world and virtual elements. Typically, these devices are head mounted displays that allow the user to see-through and interact with the real world with virtual elements superimposed on it, though the augmentations could be through other devices such as haptic gloves that present virtual haptic feedback to the user. These devices make AR possible by providing a means that the virtual elements can be seen or felt or otherwise perceived by the users. Augmenting a physical object refers to small electronic devices such as sensors, logical devices, etc. that are attached to the objects of interest directly. Those devices provide cues of the position and orientation of the user related to the objects of interests and thus allows an AR system to register virtual elements with these real world objects. Augmenting the environment refers to mechanisms that use independent devices to collect and provide information about the

surrounding environment. Examples of these devices are cameras, scanners, projectors, etc.

### **2.1.2. AR COMPONENTS**

A typical AR system takes original video as the input, generates virtual elements based on the original video image (modeling), accurately blends the virtual elements with of the real world as if the virtual elements are parts of the reality (registration), and merges the virtual elements and real world together (composition). Owen et al. give a detailed explanation on the components and architecture of augmented imagery [14] as illustrated in Figure 3.

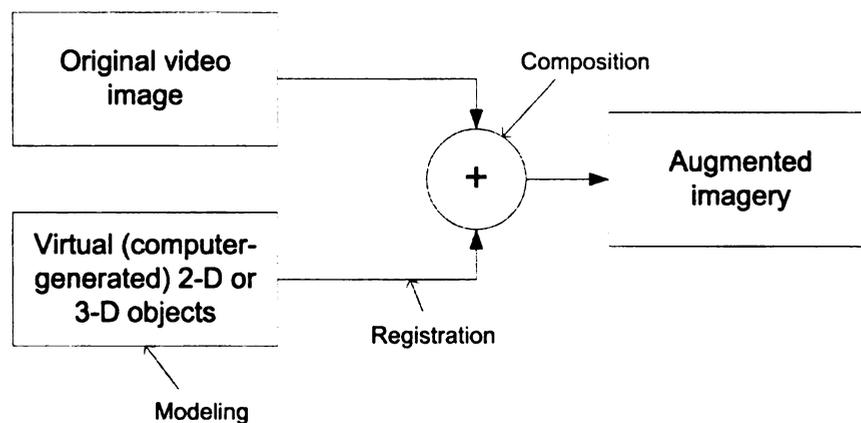


Figure 3 Architecture of AR systems

#### **2.1.2.1 MODELING**

Modeling is the description of virtual and real elements in the environment using either data structures or mathematical concepts. Most AR applications require objects to be rendered from different point of view since the users are potentially in motion. Thus 3-D modeling techniques from computer graphics [15], such as polygon meshes and scene graphs, are needed to represent models. Involving both virtual elements and real elements in modeling allows the system to know the geometric relationship between the

real environment and virtual environment. Virtual elements are modeled in such a way as to support rendering to an image. Real elements are modeled so as to support placement of the virtual objects relative to the real objects, occlusion of virtual objects by real objects, and tracking of real objects. An ideal AR system integrates virtual elements and the real world seamlessly as if the virtual elements are a part of the real world. Graphics systems such as OpenGL and Direct3D provide APIs to render realistic virtual objects. In addition to geometric descriptions, which are relatively easy to model in an AR environment, other descriptions such as lighting, material, etc. should be consistent with the real environment.

### **2.1.2.2 REGISTRATION**

To seamlessly integrate virtual and real elements in an AR environment, virtual elements should be properly aligned to their real counterparts. This alignment is referred to as *registration*. Registration determines the relationship between the real world and virtual elements so that the real and virtual parts are properly aligned as if they were in the same frame of reference. Many AR systems, such as medical or military training systems, require accurate registration. A large number of research projects are examining this area in an effort to simplify registration environmental requirements (such as markings or physical instrumentations) and to reduce registration errors [16].

Registration is the determination of the relationship between the real world and the virtual elements. The alignment is done by transformation from one model to the other. This transformation maps the point in virtual model to the corresponding point in the real model. Transformations in AR applications can be 2-D to 2-D or 3-D to 2-D. The former refers to simply replacing a planar region in the image. The latter assumes a 3-D

model that will be aligned to the 3-D real world and then projected to a 2-D display surface.

To accurately register the real world and its virtual augmentations, the AR systems need to be aware of the location of the user or the objects of interests relative to the entire scene. This requires a robust and accurate tracking system. Azuma addressed several basic tracking requirements for AR systems [17] in the Communications of ACM special issue of Augmented Reality in 1993. As Azuma summarized, AR requires the tracking to be 1) accurate in orientation and position; 2) having very small combined latency of the tracker and the graphics engine; and 3) able to working at long ranges. Registration is complicated by the extraordinary sensitivity of the human visual system to registration errors. Over a decade has passed since this seminal work and significant improvement in tracking systems has been achieved due to the effort of a considerable number of researchers. Considerable ongoing research is working on the use of ultrasonic, RFID, and infrared technologies to achieve location-awareness (as reviewed in [18, 19]). Vision-based tracking systems use vision cues in the scene to compute the user or object position and orientation. Tracking systems are in development for both prepared and unprepared environments. In *prepared environments*, placed fiducial marks such as circle-based black and white images [20] , multi-ring color images [21] , or ARToolkit-based fiducial images [22, 23] provide vision cues to the tracking system; while in *unprepared environments*, the tracking system uses natural features to extract the position and orientation of user and object. Vision-based tracking systems use cameras to acquire the vision cues.

### **2.1.2.3 COMPOSITION**

Composition is the blending of the virtual and real elements into final output image. The simplest composition is simply overlaying the virtual elements over the real image. In most applications, however, more complicated composition such as alpha mapping or segmentation are necessary to create realistic effects that appear seamless.

### **2.1.3. HUMAN FACTORS IN AUGMENTED REALITY**

As a new paradigm of human and computer interaction [5], AR is capable of providing assistant information in the form of computer generated imagery. Baird [12] and Tang [24] have evaluated the effectiveness of assistance provide by AR systems in assembly task instruction scenario. AR is also capable of directing users' attention by computer generated virtual imagery. Tönnis, et al evaluated the effectiveness of AR visualization for directing a car driver's attention [25]. Bonanni, Lee, and Selker proposed an attention-based design of AR interfaces [26] to improve usability. Biocca, et al built an AR interface that interactively guides a user's attention to any object, person, or places in space and evaluated the interface [27, 28].

## **2.2. CONTEXT-AWARE COMPUTING**

Advances in wireless communications and portable computing devices allow people to move around and access computerized information and network resources "anytime, anywhere". The use of *context* is important in such mobile and interactive applications. The concept of context-aware computing arose as a mobile computing paradigm that collects and utilizes contextual information automatically.

Schilit and Teheimer first introduced the term 'context-aware' in their 1994 work [29] which referred to context as the location of the user, nearby people and objects, and changes to those people and objects. Schilit et al. further stated in their 1994 review of context-aware computing applications [30] that context is the constantly changing execution environment including *computing environment*, *user environment*, and *physical environment*. Pascoe [31] defines context to be the subset of physical and conceptual states of interest to a particular entity. A more general definition of context is given by Dey and Abowd [32] as "*any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*". Chen and Kotz [2] add *time context* to the categorization in the 1994 review by Schilit et al. [30], and define context as "*the set of environmental states and settings that either determines an application's behavior or in which an application event occurs and is interesting to the user*".

In summary, there are four commonly accepted categories of context:

1. Computing context, such as network connections, bandwidth, nearby computing service and resource, for example, printers, servers, etc.
2. User context, such as user's preference, profile, social environment, other people nearby.
3. Physical context, such as temperature, noise, lighting, location
4. Time context, such as season of a year, time of a day, day of a week or month.

Different contexts play different roles in various application domains. Thus, Context is not limited to location or specific objects; context is application-dependent and is the

situation that is relevant to the application domain and users. It changes from situation to situation.

The major goal of context-aware computing is to maximize the relevance of the information that is provided to users with minimal user actions and inputs. Applied in augmented reality systems, the benefit of context-awareness is to maximize the relevance and minimize the confusion of the virtual elements that are presented to the user.

### **2.3. CONTEXT-AWARE AUGMENTED REALITY SYSTEMS**

Since AR systems register virtual and real counterparts in 3-D, they are primitively considered to be an application of context-aware computing [2, 32] . It is obvious that AR systems take into account the location context of the user and the objects that are of interest. Many AR systems, especially mobile AR systems, have sensors that track user's positions and automatically provide relevant information when needed. If the systems flood the user with excessively large amount of information, the information leads to overloaded and may induce confusion and impede the user's ability to interact with the real world, or, even worse, induce safety hazards on users. This is particularly the case when users are wearing head-mounted displays. The amount of virtual information users can receive before being distracted from the real world is limited. Some research activities in AR seek to improve the usability of AR system by providing the most relevant virtual information automatically with context-aware capabilities.

In most context-aware applications, the key points are the *discovery* of context and the *application* of context. System designers and developers of context-aware AR systems are interested in what context is of interests and how to sense the context. The context of

interest is generally application dependent. For example, an AR tourist guide system will be interested in the tourist's current location and the tourist's interests, e.g. historical or natural attractions. This context can be detected by a stable tracking system and the user's interaction with the system or some user profiling algorithm. A useful context in an AR assembly training system will be the trainee's current step of working, which can be acquired by a sensing system or a memory space that records the trainee's activities.

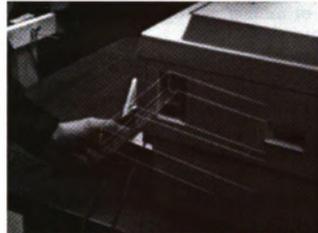
Most AR applications take into account the context of the user, physical objects, or the environment depending on the application domain. Context-awareness narrows down the content of augmentations and therefore provides a neater information representation to the user. This section discusses some representative context-aware AR systems in different application domains. Limited by the length of this report, only a few sample projects are mentioned in this chapter, many more projects are working on augmented reality-oriented context retrieval and utilization.

AR technologies are extensively used in assembly and maintenance applications. If the augmentations occur in a context-sensitive manner, the view of the real world of a skilled worker, technician, or engineer is augmented with textual or graphical information that is related to the individual and his/her current situation.

The KARMA, Knowledge-based Augmented Reality for Maintenance Assistance, is a test-bed system at Columbia University Computer Graphics and User Interfaces Lab [33]. It is one of the representatives of context-aware AR systems that are used to provide assembly instructions. It provides simple laser printer maintenance assistance through a see-through head-mounted display. The KARMA uses rule-based approach to select relevant information to assist a user performing a maintenance task, as shown in Figure 4.



*a. The KARMA system provides simple laser printer maintenance assistance through a see-through head-mounted display.*



*b. The solid line highlights the paper tray as it moves, an arrow indicates the action and direction of pulling the tray, and the dotted line shows the tray's desired destination state.*

Figure 4 The KARMA system at Columbia University [33]

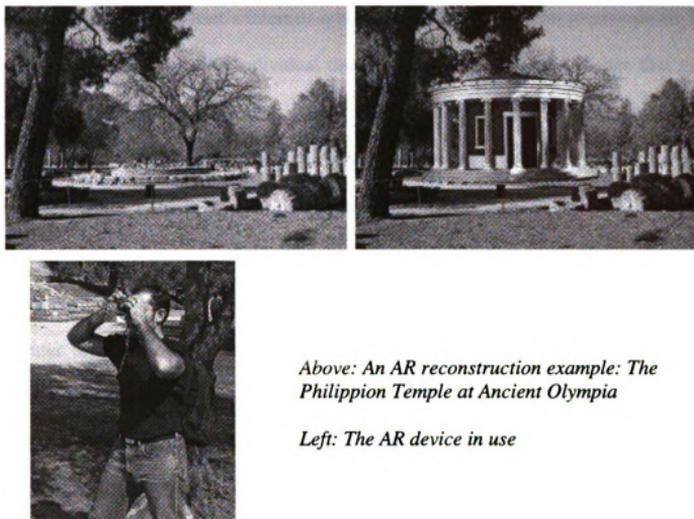
The contexts that are referenced in this system are 1) user position and orientation; 2) inter-object occlusion relationships; and 3) role of object in a specific task. The sensing user position and orientation is the tracking problem that is the part of the registration of AR systems. The KARMA system uses ultrasonic based tracking technology. The transmitter is made up by a triangle with three ultrasonic sources near the corners and receivers are made from triangles with three microphones near the corners. The transmitter and the receivers work together to compute the user's position and orientation. The inter-object occlusion relationships are discovered by 3D geometric processing while the object of interests, for example the toner cartridge, is within the view volume. The third context mentioned above is acquired based on IBIS (Intent-Based Illustration System) [34], which is a rule-based system that designed illustrations. 'Illustration' is a term referred to pictures that are designed to satisfy an input communicative intent. The communicative intent is a list of prioritized communicative goals, which specifies something to accomplish, for example, to show a property of an object or a change in a property. The illustrations generated by IBIS are dynamic, which

means the IBIS is an adaptive system that continuously redesign the picture to best maintain the goals. A more detailed description of how IBIS works will be given in next section.

There are lot of interesting AR projects designed to improve the productivity and performance in assembly or maintenance depending on the context information. ARVIKA at Siemens [35, 36], Boeing's augmented reality instructional system [37] , the STARMATE project of a consortium of European organizations [38, 39], the SEAR project at Siemens Corporate Research [40], and a lot of others, are making big contribution to the AR community.

Featuring augmented reality and context-awareness, a tourist guide system can have the advantages of both guided and unguided tours, and even goes beyond them. It provides personalized guide to individual tourist, having the flexibility of unguided tours and information retrieval for guided tours. In addition, the personalized guide is not possible for both guided and unguided tours in traditional ways.

Archeoguide (Augmented Reality based Cultural Heritage On-site GUIDE) is an AR-based tourist guide system for personalized tours in cultural heritage sites [41, 42] by a consortium of European organizations. The goal of the Archeoguide is to enhance the tourist's overall experience to a cultural heritage site by reconstructing the site's ruined monuments using augmented reality technologies. This system provides personalized guide by taking into account the tourist's personal preference. The guide is adaptive to the user's location and interaction to the system. Figure 5 shows a sample augmented image of the Archeoguide system and a touring user wearing the system at a cultural heritage site.



*Above: An AR reconstruction example: The Philippion Temple at Ancient Olympia*

*Left: The AR device in use*

Figure 5 The Archeoguide system [42]

The contexts that will most facilitate the tourists are the user's location and the user's preference profile and visiting history. The tracking system used in Archeoguide is a hybrid system. A GPS system gives a very rough positioning of the user at first, and the combination of vision-based and inertial tracking provides the user's exact position. Since the nature of the site (archeological site) makes it impossible to use an arbitrary number of artificial fiducials, a combination of both artificial and natural landmarks are used in for vision-based tracking. Combined with inertial and vision-based technologies, the tracking system is reported to be able to get a better estimation of initial poses and balance errors.

To sense the user's preference and optimize the need of user action and input when using the system, the designers use machine learning techniques in different levels to

dynamically adapt to the user's interests and current situation in order to provide the best possible presentation to each individual user. They define a feature space that maps user's attributes as categorized items. The system predicts the user's interest and chooses to render the objects that are of the most interest to the user. User's cluster of points in the feature space is updated accordingly whenever he/she makes any action and thus affects the future prediction of the system. This is a recursive learning process. The more the user utilizes this system, the more the system knows about the user, and hence the better the system serves the user. A more detailed explanation of the machine learning techniques used in the Archeoguide system will be discussed in Section 2.4.2.

Other representative context-aware AR tourist guide systems include but not limited to the MARS project at Columbia University [43, 44], an audio augmented reality tour guide system [45] proposed by Benjamin B. Bederson at Bell Communications Research. a handheld AR museum guide [46] prototyped by Dieter and Daniel.

AR researches, combined with context-aware computing, are also actively carried out in other application domains, such as medical [47-49], education and training [50, 51], industrial maintenance[39, 40], and others as reviewed in [1, 4]

## **2.4. AUGMENTED REALITY IN ADVERTISING**

Augmented reality in advertising is a young area in comparison to other applications. Augmented reality technologies can be used to realize a virtual experience, a term in advertising research that refers to presentations that stimulate customer learning of the product and leads to better understand of the product [8, 9, 52]. Virtual experience also impacts consumers' behavior as survey by Host [53] in the German furniture market.

Wierzbicki and Margolf [54] pointed out that AR technologies are becoming more and more powerful for commercial presentation and marketing of products, labels and companies themselves.

The PromoPad project at the Media and Entertainment Technologies Laboratory (METLAB) of Michigan State University [6, 7] is an experimental in-store shopping assistant that provides personalized advertising. The concept of *dynamic contextualization* is introduced in this project. Dynamic contextualization, going beyond traditional context-aware computing, not only discovers and utilizes the context information of the customer and object under inspection, but also *modifies* the context using augmented reality technologies. Virtual contextualization, as an element of dynamic contextualization, contextualizes real objects with complementary virtual objects. By augmenting the context of the product, the product is contextualized with its complementary products, which are virtual computer graphics models, to emphasize the focal product. By diminishing the background or competitive products of the product, attention can be drawn to the focal product.

Prince Video Image is a commercial organization that is working on virtual advertising and other computer graphics product [55]. Figure 6 shows some sample images of their virtual advertising videos. Strictly speaking, these are not true augmented reality system since the blending of the virtual and real elements may be done offline. It is still worth mentioning here since the company is moving to live video virtual advertising.



*Virtual Ford Truck in Time Square  
for ABC Rose Bowl*



*Coca-Cola Animation coming out of  
Center Field*

Figure 6 Virtual advertising image samples from PVI [55]

Dynamic Digital Advertising [56] use virtual reality and augmented reality technologies to show a 360-degree view of a product, tour different rooms of a building, or showcase a panoramic view.

## **2.5. AR-ORIENTED CONTEXT-AWARE MODELS**

The role of context-awareness in augmented reality system lies mainly in automatically managing the amount and content of information that are delivered to the user. This is especially useful in mobile AR system, where users need to keep a clear view of the real world so as to ensure safety and allow normal real-world interactions. Therefore, care is being taken to ensure that the display is not cluttered with excessive amounts of information. Filtering crowded information to prevent clutter and improve information presentation is also a major goal of context-aware computing.

### **2.5.1. SPATIAL MODEL**

Benford and Fahlen proposed a spatial model of interaction [57] that supports group interaction in large-scale virtual worlds. This model provides a generic technique to

manage awareness and interaction and fits to almost any system where a spatial metric can be identified, including AR systems. Two sub-spaces work together to determine the awareness. One is a sub-space within which an object can see, and the other is a sub-space within which an object can be seen. Therefore the awareness is not necessarily mutually symmetrical due to the effect of two sub-spaces surrounding each object.

### ***2.5.2. REGIONAL-BASED MODEL***

Julier et al. [58] proposed a regional-based information filtering algorithm based on a spatial model. The regional-based model adds tasks to spatial model. Hence, besides the spatial information, this algorithm assumes that each user is assigned a series of tasks. Two objects can be aware of each other only when their “see” and “been seen” sub-spaces collide, and they have common tasks.

### ***2.5.3. RULE-BASED MODEL***

The KARMA system, as discussed in the previous section, uses a rule-based approach based on the IBIS (Intent-based Illustration System) [34, 59, 60] to select relevant information to assist a user performing a maintenance and repair task. Rule-based model defines a set of rules to determine whether two objects can communicate.

### ***2.5.4. MACHINE LEARNING MODEL***

The Archeoguide system uses recursive machine learning technologies to adapt to a guided visitor’s preference and predict the user’s next action [41]. Since it is expensive to render all the complex 3-D models (i.e. reconstructed ruined buildings) related to the scene in the view, the system needs to selectively render the models that are of the most interest to the user. The designers employ a machine learning model to organized user’s

preferences and predict user's interest. This machine learning model works as follows: User preferences are defined as a feature space with each axis as one attribute of the users. The system maintains a space for each item in the field. Each point in the space indicates a positive or negative value of the item by every user. A positive value means the item is requested or accepted by a user; likewise, a negative value means the item is rejected by a user. The system continuously monitors the points for each user in the feature space. The system dynamic updates the points in the feature space as the user interacts with the system by requesting, accepting, or rejecting items. If the user requests or accepts an item, one positive value will be added to the appropriate position. Hence, this is a recursive learning process. The more the user utilizes the system, the better the system knows him/her, and hence the better the system serves him/her. The next action provided by the system depends on the positive and negative points on the corresponding area of the user in the feature space. The decision could be taken by taking the majority of the points or other decision making algorithms. Figure 7 illustrates a simple 2-D space for one item. The boxed area is the space for a single user. Since there is some extend of uncertainty in the learning process, the system is only able to provide a "best" possible prediction, not the absolutely "right" prediction.

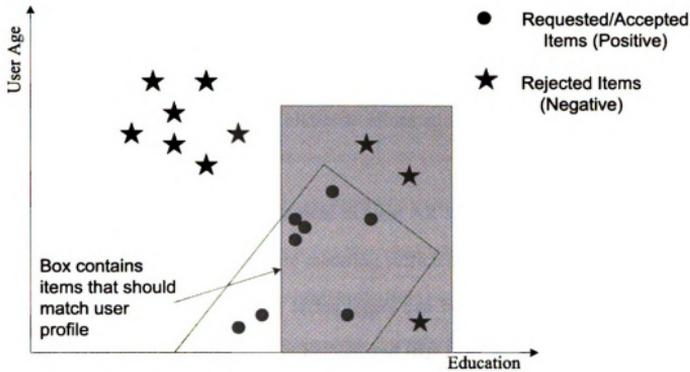


Figure 7 Clustering user preference by machine learning technology [41]

### **2.5.5. CURRENT RESEARCH IN CONTEXT-AWARE AR**

Recent advances in context-aware computing and augmented reality seek to enhance people's view of the real world with augmented graphical information that is most related to the user's current situation.

Previous sections discussed four categories of context-awareness in augmented reality: 1) spatial models; 2) regional-based models; 3) rule-based models, as used in the KARMA system; and 4) machine learning models, as used in the Archeoguide system. Probability and statistical models such as Logistic curves, Markov chains, Bayes rules, and social filters are used extensively in retrieving user context and predicting user's preference [61-63]. Utilizing probability and statistical models helps deal with uncertainty, which is an inherent property in context-aware computing, and maximally reduce the error rate in predicting the user's preference.

## **2.6. DISCUSSION AND SUMMARY**

This chapter has discussed several representative AR systems and key techniques for using context-awareness to prevent cluttered information presentation and information overload.

Location context is an essential element in most AR systems. Tracking technologies, i.e. the sensing of location context, are maturing sufficiently to achieve accurate and low latency tracking. Different tracking technologies can be used in different applications according to the accuracy requirement, budget, and spatial requirement of the application. As mentioned before, tracking technologies have been explored extensively, this proposal is more concerned about the discovery and utilization of other contexts that manage the data density in an AR system.

This chapter surveys the roles of context-aware computing in augmented reality and the approaches of retrieval and utilizing context information in augmented reality systems to provide the most relevant and appealing information to each individual user. Context-aware computing is promising for managing and controlling information presentation in augmented reality systems as the hardware and graphics technologies become mature.

There are certainly some limitations in context-aware augmented reality system. For example, there is a trade-off between automation and user flexibility [64], the more automated the system is, the less flexibility that can be provided to the user. Balancing this trade-off requires defining an interactivity level at the design stage according to the requirement of the application or dynamically adjusting the levels of interactivity based the user's knowledge of using the system.

## **CHAPTER 3. DYNAMIC CONTEXTUALIZATION: DESIGN AND IMPLEMENTATION OF PROMOPAD**

### **3.1. INTRODUCTION**

This chapter presents the design of the *PromoPad*, an Augmented Reality shopping assistant that provides a new method of human computer interaction. Augmented reality technologies enhance people's perception of and interaction with the real world using computer generated virtual objects, changing the way that people interact with both computers, and the real world. Considerable work has been done in the area of augmented reality and human computer interaction in various application domains [5, 65, 66]. The shopping environment, however, poses unique challenges and is, as yet, not well explored. First, a friendly user interface and negligible user interference are essential characteristics for such a system. Second, the amount of information that can be delivered to the user is vast, so effectively providing only the most relevant information to the user without cluttering his/her view becomes a major concern. Display clutter can significantly degrade the quality and performance of the tasks that the user is performing [67]. Third, the users of the system come from diverse backgrounds and possess a wide variety of skill levels. Hence, robustness and stability are key points in the design of a final system. These challenges are deliberated throughout the design and implementation of the system and are addressed in detail.

PromoPad is a prototype hand-held device that provides context sensitive shopping assistance. It has been designed as a test-bed for context-aware computing technologies in an augmented reality environment. Powered with context-aware computing technologies, PromoPad provides relevant information to users as context modifications. These modifications are in the form of augmented imagery using augmented reality (AR) technologies and the content of the assistance is built upon the concept of *dynamic contextualization*.

Augmented reality enhances the perception of reality in this application by contextualizing individual objects that are encountered in the real world with virtual complements so that these real objects become more meaningful and appealing. The PromoPad is a tablet PC with a camera on the back. The display on the tablet provides a modified version of the camera image. This image is modified using augmented reality technologies that can add new imagery placed relative to a focal product or remove elements of the image that may distract from the focal product. Thus augmented reality technologies offer the technical capabilities necessary for realizing dynamic contextualization. In traditional context-aware computing, context plays only a passive role as the situation of the user. With dynamic contextualization, the context can be modified to be more meaningful for the focal objects and more interesting to the users.

Researchers in the field of context-aware computing and e-commerce systems have sought ways to provide handy and natural electronic assistance for shoppers. Kourouthanassis and Roussos developed MyGrocer [68], a pervasive retail system, that can manage shopping lists, monitor the total cost of the cart content, popup promotion information, and help consumers navigate within the store. Project Voyager examines

the use of context-aware computing as a shopping assistant [69]. PSA is another experimental system that provides personalized shopping assistance [70]. These projects focused on discovering the store context so as to provide an electronic and automated shopping aid that can ease or assist the shopping process. The discovered context in these projects, however, plays a passive role as the situation, such as the location, of the user was not specifically known; only the products being selected and placed in the cart. In addition to context-sensitive content delivery, the PromoPad simulates a virtual experience based on the concept of dynamic contextualization that results in more product knowledge, better brand attitude, and elevated purchase intent [9].

A good e-commerce system does not just provide passive information. The Point-of-Purchase Advertising Institute's research shows that 70% of buying decisions are made in the store [11, 71]. Hence a good e-commerce system should also be able to trigger impulse purchase decisions. AR-powered dynamic contextualization presents 3-D visualizations registered to actual products in the store and in the proper context for impulse decision making. The dynamic contextualization can be blended to reproduce endless consumption situations that can affect shoppers' perception of a brand and purchase decision.

Dynamic contextualization is made possible by AR technologies that modify the perception of the real world in real time [1]. Several empirical studies on the effectiveness of augmented reality technologies in human computer interaction provide evidence that augmented reality systems improve human performance. Tang et al showed statistical significance to support the hypothesis that augmented technologies improve the operational performance in instructing assembly tasks [24]. The

Archeoguide system is an outdoor augmented reality guide that offers personalized tours of archaeological sites. It uses augmented reality technologies to improve information presentation, simulate ancient environment, and visually recover destroyed sites [41]. Considerable active research in augmented reality includes a broad range of application domains. This thesis explores the technical feasibility and benefits of augmented reality for advertising and consumer experience.

### **3.2. THE PROMOPAD SYSTEM**

The PromoPad is a mediated experience device that provides an in-store virtual experience with 3-D product visualization. The system consists of a front-end client component and a back-end server component. The front-end component is a light-weight display device that slips into a cradle in a shopping cart. Tablet PC technologies are used to implement this front-end device. With a camera attached to the back of the Tablet PC, the client device is aware of the position and orientation of the shopper relative to the shopping cart and store shelves through the use of visual marker technologies. It is also capable of providing the shopper a see-through view of the shelves and additional information that is related to the items in the view. The back-end components consist of one or more servers that contain inventory databases, customer profiles and business logic, from which information in the databases is filtered and returned to the front-end component. The PromoPad employs augmented reality technologies and passes an augmented camera image from the rear of the Tablet PC to the display. Figure 8 shows a typical usage of the front-end component prototype.



Figure 8 Using the PromoPad in a store setting

The goal of this design is an intelligent shopping aid that provides shoppers automatic and meaningful help when needed, as well as minimizing human interference and effort. With wireless communication, a Tablet PC can have different modes for shoppers in different shopping situations. Planned shoppers may use a tablet PC to optimize their shopping routes in a store and to quickly find items they plan to buy. Bargain shoppers may use it to rapidly located sale items. Recreational or detail-oriented shoppers can use a Tablet PC to obtain product information that is not on the packaging. For example, augmented reality images of the winery and wine ratings or reviews can be displayed as a shopper inspects a bottle of wine. Content customization and personalization of a Tablet PC can greatly facilitate the convenience of all types of shoppers, enhancing the shopping experience. It is important to recognize that the vast majority of grocery and convenience store purchases are impulse purchases. Even slight improvements in marketing performance can result in massive increases in sales.

### **3.3. AUTOMATED CONTEXT-AWARE ASSISTANCE**

Using augmented reality in a shopping environment, the information that can be delivered to the user's attention can be vast. It can range from the introduction of a new product, a sales sign, or directions to a related product. Cluttering the user's view on the Tablet PC with large amount of information would be very easy. Thus, how to selectively display the most interesting and important information for each individual user becomes a major concern. The system must filter the information stream and provide relevant information that can be accommodated in the tablet display. For example, if the system chooses to flood the user with large amount of promotion information, price comparisons, and in-store advertising, then the system accomplishes little more than what could be accomplished by handing the customer a thick flier. The new capability of the PromoPad is that it can selectively display information that is related to the product under inspection and information that is tailored to individual needs. In other words, the information that is presented to the user is highly related to the context of the user, and the product under inspection. Three criteria are applied to determine the relevance of a piece of information to a specific user at a single point in the store:

1. User's location and orientation
2. User's previous shopping history and pattern
3. Product complementary relationship in the store database

We discuss the detail of these three criteria in this section.

### **3.3.1. USER'S LOCATION AND ORIENTATION**

The user's location and orientation determine what products the user is currently inspecting. When the consumer is using the PromoPad during her shopping trip, it is reasonable to assume that the position and the orientation of the Table PC when it is deployed are an approximation of the position and orientation of the consumer as well. A variety of AutoID systems are in development that will allow high-quality tracking of product relative to the PromoPad and knowledge of purchase (cart insertion) decisions.

With an in-store tracking system the PromoPad is aware of its 3-D position relative to store shelves and products. Considerable on-going research has been exploring the use of ultrasonic, RFID, infrared, and vision-based technologies to achieve location-awareness [18, 19]. The tracking method for such a system, however, is challenging. The quality of the tracking system directly determines the robustness and scalability of the whole system. The prototype system utilized a vision-based fiducial system and its improvement proposed by Owen, Xiao, and Middlin [23]. The system, a component of the ImageTclAR augmented reality development environment [72], is robust (high correlation) and fast (consistently under 2ms). The fiducial marker images serve as visual clues that accurately tell the system where the camera is pointed and what it is looking at. Figure 9 shows our experimental shelf with several fiducial images on the bottom. As a prototype, the vision-based system has limitations in the amount of store area that could be covered and the obtrusiveness of the marker images. A larger-scale system could be built based on ultrasonic or RF tracking technologies in combination with object recognition capabilities. As this thesis focuses on the human-computer information issues of a system as well as technical issues of registration and composition, larger-scale tracking solutions are considered beyond the immediate scope.



Figure 9 the experimental shelf with fiducial images

The location information required for the PromoPad is considerably more rigorous than that of traditional context-aware computing systems. Owen, et al. [14] discuss many issues relative to augmentation of imagery for AR applications such as the PromoPad. Augmented reality requires modification of the camera image. Achieving pixel-resolution registration of computer graphics with store shelf contents requires high-accuracy knowledge of the location and orientation of the PromoPad. Visual fiducial systems provide sufficient accuracy for high-quality image modifications.

With the tracking system, the PromoPad is aware of the 3-D position and orientation of the consumer relative to the product and store shelves. It then sends a query to the back-end server and displays feedback on the Tablet PC. For example, when the consumer is in the dairy products aisle, the server returns the promotional information for various milk brands.

### **3.3.2. USER PROFILE**

A user profile includes such data as brand preference, buying history, shopping pattern, and preference. User profiles also include individual and aggregate behaviors based on

shopping habits and demographics. Each time the consumer checks out, purchases are recorded in the store membership database. These systems are already common in many stores that feature *loyalty cards* and there is evidence that many consumers utilize these systems [73]. From loyalty card systems or future, automated variations, stores can create personal profiles based on the previous purchases that the consumer has made. For non-member consumers, a generic profile with demographic manipulations can be used.

The consumer will scan her member card or login in as a member before using the PromoPad. Based on history information, the system applies business logic at the database inquiry. The system is able to answer questions like “How likely is it that the customer buy a carton of milk on this visit?”, “How interested is this customer in some toys for 2-3 years old girls?”, “Will the customer like this brand of frozen pizza?” Carefully applying data mining techniques and planning business logic, the system can even predict more sophisticated conditions [74]. Answers to these questions help the system to predict whether or not the consumer will be interested in certain classes of information. If the answer is affirmative, then system will consider that the consumer is definitely interested in this information and delivers it to the consumer using store directions and emphasis of the product on the shelf. If the answer is moderately positive, then it can consider this information may trigger an *impulse purchase*. If the answer is strongly negative, then it interprets that the consumer doesn't like this information or related products, and hence the system will not bother the consumer at all.

### **3.3.3. PRODUCT CONTEXT**

Product context is the set of *complementary products* that are associated with the focal product or the product under inspection. For example, a golf club can be associated with

golf balls, hats, or shoes. A digital camera can be associated with a tripod, or pictures of a vacation. Products are perceived as more meaningful or even more valuable in context. A detailed description of complementary products is presented in Section 4.1.

### **3.4. TECHNICAL ISSUES**

To achieve the goal of a Tablet PC as a see-through augmentation device, several technical issues must to be addressed. First, the means of tracking the location context needs to be robust, scalable and stable. Second, the real image in the Tablet PC display should be adjusted to offer a true see-through view as if the Tablet PC display was transparent so that the device is well integrated with the environment and in harmony with real products. Third, the virtual objects should be accurately registered to the real image. Thus, the Tablet PC display can act as a ‘magic frame’ that allows the user to ‘look through’ the frame with additional information that cannot be seen otherwise. Finally, the system needs to be able to deal with a variety of different virtual and real composition methods, including overlays, occlusion, and diminishment.

#### **3.4.1. IN-STORE TRACKING**

The location of a shopper as a 3-D position and orientation relative to product and store shelves is acquired by an in-store tracking system. When the shopper is using the PromoPad, it is reasonable to assume that the position and the orientation of the Table PC are a good approximation of the position and orientation of the shopper. Considerable research has explored the use of ultrasonic, RFID, infrared, and vision-based technologies to achieve location-awareness (as reviewed in [18]). A variety of existing technologies can be scaled for this application to store-size volumes with large quantities of

PromoPads. In the experiments presented in this thesis, a vision-based fiducial system designed by Owen, Xiao, and Middlin is used [23]. Fiducials are markers that provide visual cues of the position and orientation of the user in a vision-based tracking system. As reported, this fiducial system is robust to partial occlusion and noise, computationally efficient, and scalable. Fiducial systems work well with inexpensive cameras and are easy to deploy. They are also a good match to the monitor-based augmented reality approach for this application [75].

Location information required for the PromoPad is considerably more rigorous than that required by traditional context-aware computing systems. Dynamic contextualization and augmented reality require modification of the camera image. Owen, et al. [14] discuss many issues relative to augmentation of imagery for augmented reality applications.

### **3.4.2. VIDEO SEE-THROUGH SYSTEMS**

The view as seen on the Tablet PC display is derived from the image captured by a camera mounted on the rear of the tablet. The view of the camera is in turn determined by the camera's intrinsic and extrinsic parameters. The intrinsic parameters of a camera describe how the camera will convert objects within the camera's field of view into an image. The extrinsic parameters describe the position and orientation of the camera in space. Figure 10 illustrates a perspective camera projection model. The optical axis, which is orthogonal to the retinal plane  $\mathfrak{R}$ , passes through the center of projection  $C$  and intersects with  $\mathfrak{R}$  at the principal point  $c$  on the image plane. The distance between the center of projection  $C$  and the retinal plane  $\mathfrak{R}$  is the camera focal length  $f$ . Let  $M$  denote the world coordinate of some point on the tip of the wine bottle. The corresponding point

$m$  on the retinal plane is the intersection of the line that passes through  $M$  and  $C$  and the retinal plane  $\mathcal{R}$ . Thus, intuitively, what the camera can see is the volume inside the infinite pyramid whose apex is  $C$  and the four lines that form the edges of the pyramid pass through the four corners of the retinal plane, as illustrated in Figure 10. A detailed derivation of the projection matrix that maps the world coordinate to the retinal plane coordinate can be found in computer vision books such as [76].

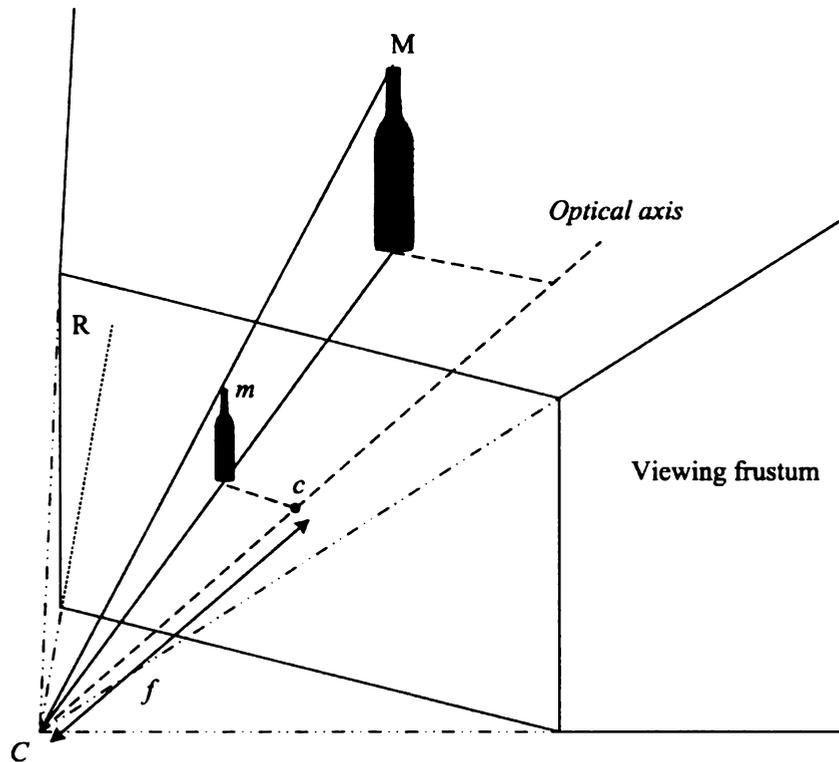


Figure 10 Perspective camera projection model

This pyramid is referred to as the *graphics frustum*. Graphics frustums for rendering are often truncated with near and far clipping planes, where the near clipping plane avoids rendering of objects too close to the camera or at the singularity point  $C$  and the far clipping plane avoids rendering objects so far away from the camera as to be considered no longer visible. The graphics frustum for the virtual elements has to match

the camera viewpoint and intrinsic parameters in order to mimic the view of the camera and have the virtual objects accurately registered with the camera image.

### 3.4.2.1 REGISTRATION

It is assumed that a calibrated camera, camera focal length  $f$ , center of projection  $C$ , principal point  $c$  and size and position of retinal plane  $\mathfrak{R}$  are known to the system. Before setting up the viewing frustum, the viewport rectangle needs to be set at the same resolution (viewport rectangle size in pixels) as that of the camera retinal plane in order to have the same view of the camera if the frustum is set according to the camera's intrinsic parameters. For example, if the camera resolution is 640 by 480 pixels, then the size of the viewport rectangle needs to be set as 640 by 480 as well. The parameters defining a viewing frustum that matches the camera point of view are  $(l, b, -n)$ , which specifies the 3-D coordinates of the lower left corner of the near clipping plane, and  $(r, t, -n)$ , which specifies the upper right corner of the near clipping plane [77]. The values of  $l, r, t$ , and  $b$  are as follows.

$$l = -n \cdot \frac{c_x}{f};$$

$$r = n \cdot \frac{w - c_x}{f};$$

$$t = -n \cdot \frac{c_y}{f};$$

$$b = n \cdot \frac{h - c_y}{f};$$

where  $(c_x, c_y)$  is the 2-D coordinate of principal point  $c$  in the retinal plane  $\mathfrak{R}$ ,  $w$  and  $h$  are the width and height of the retinal plane  $\mathfrak{R}$ ,  $f$  is the camera focal length. These

parameters much be measured in the same unit, usually pixels.  $n$  is the distance from the camera to the near clipping plane and is of the unit as  $l$ ,  $r$ ,  $t$ , and  $b$ . The virtual objects that are rendered in this frustum are well aligned with the camera image.

### 3.4.2.2 ZOOMING

Given this model, both the camera image of the real world and the virtual augmentations appear on the Tablet PC display with the camera point of view. It will provide users a more realistic view if the image can be properly zoomed and shifted as if it was seen from the user's point of view, yielding the effect of a 'magic frame' or magnifying glass. Moreover, the Tablet PC window is usually larger than the camera image and therefore able to accommodate more augmentations. Figure 11 illustrates the different viewing frustums caused by the camera and the user's point of view.

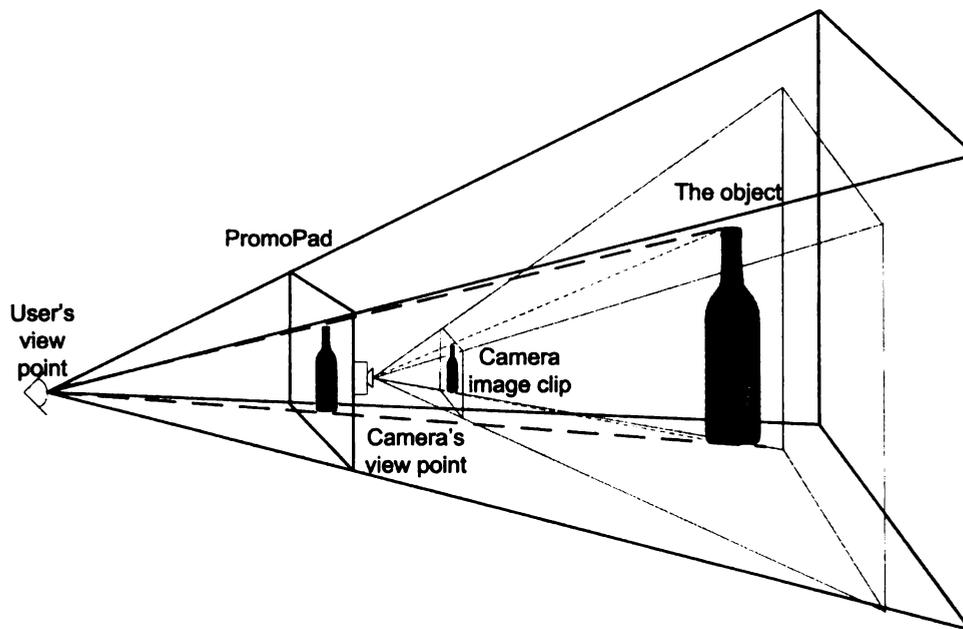


Figure 11 Perspective view of frustums

The construction of a frustum from the user's point of view has two steps. First, the viewport rectangle needs to be set as the size of the display window. In other words, if

the display is in full screen mode and the screen resolution is 1024 by 768, then the viewport rectangle has to be set as 1024 by 768 too.

The frustum that matches the user's point of view is defined as follows.

$$l = -\frac{W}{2 \cdot p_x}$$

$$r = \frac{W}{2 \cdot p_x}$$

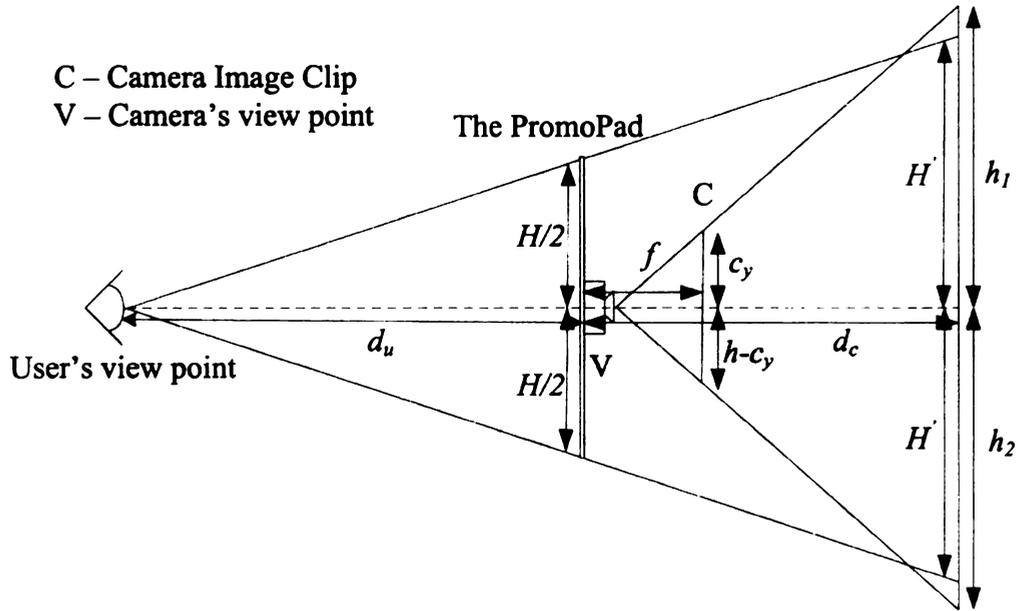
$$t = -\frac{H}{2 \cdot p_y}$$

$$b = \frac{H}{2 \cdot p_y}$$

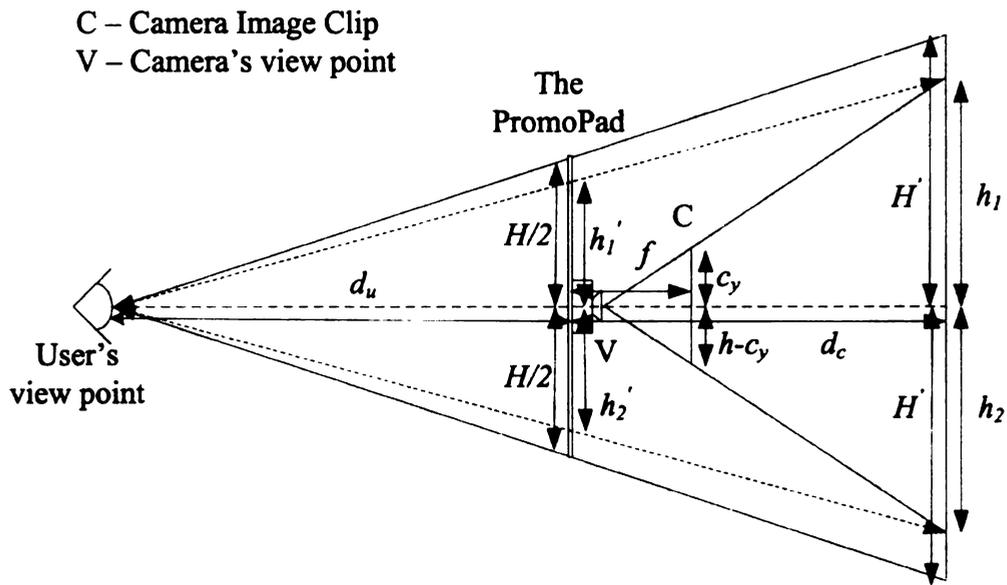
where  $W$  and  $H$  are the window size in pixels and  $p_x$ ,  $p_y$  are the number of pixels horizontal and vertical per measurement unit respectively. The distance to the near clipping plane is set to the distance from the user's point of view ( $d_u$ ) to the display window, which can be normalized in the application or obtained by existing eye-tracking systems [78]. This frustum model matches to the user's point of view.

The camera image has to be adjusted to match the frustum model from the user's point of view. Since the camera is rigidly attached at the central axis of the Tablet PC, it is reasonable to assume that the user's point of view and the camera's point of view are along the optical axis. Figure 12 shows the vertical 2-D view of the perspective view with two situations. If the camera captures a bigger view as shown in Figure 12a, then the mapping is just a truncation of the invisible part and projection to the display window. If the camera captures a smaller view as shown in Figure 12b, than the camera image will be mapped to the display window according to the proportion without any truncation.

However, a small area near the borders is out of sight of the camera and will not be displayed.



a. Camera captures bigger view



b. Camera captures smaller view

Figure 12 Vertical 2-D view of the perspective frustums

Let  $d_u$  be the distance between user's point of view and the tablet and  $d_c$  be the distance between the camera view point and the focal object. Then:

$$H' = \frac{H \cdot (d_u + d_c)}{2 \cdot d_u}$$

$$h_1 = \frac{c_y \cdot d_c}{f}$$

$$h_2 = \frac{(h - c_y) \cdot d_c}{f}$$

A comparison between  $H'$  and  $h_1, h_2$  indicates whether the image will be larger or smaller than the viewport on the display. If the image is larger than the display, then the upper  $h_1 - H'$  and lower  $h_2 - H'$  are truncated. If the image is smaller than the display,

the camera image maps to  $(-h_1', h_2')$ , where  $h_1' = \frac{h_1 \cdot d_u}{d_u + d_c}$ ,  $h_2' = \frac{h_2 \cdot d_u}{d_u + d_c}$ . The horizontal

projection can be calculated the same way.

Although there will be some blank area near the border in some cases, this is negligible in the limited depth of product shelves and can often be covered over with augmentations. Another limitation of this mapping is that it scales identically over the entire camera image, since the captured image is 2D. The zoom operation doesn't differentiate the depth of different objects in the camera image. However, the depth is correct for the focal object, which is the focus of the display in this application. Since the object and the tablet are simultaneously tracked, the depth, and therefore the zoom factor, is continuously varied to keep the focal object the correct size. Figure 13 shows the different effects of the Tablet PC window when the display is from the camera's point of view and adjusted to the user's point of view. A cereal box is behind the Tablet PC and

in the scene of the camera. The PromoPad captures the cereal box and augments the view with a nutrition bar and a piece of advertising information. Figure 13a shows the augmented view from the camera's viewpoint. The adjusted view is shown in Figure 13b with the effect of 'magic frame'.



a. display from the camera's point of view



b. display zoomed to user's point of view

Figure 13 Tablet PC displays from different viewpoints

### 3.4.2.3 COMPOSITION

At this point both the real image and virtual elements are properly aligned and adjusted to the user's point of view. The next step is to compose the real image and virtual elements to produce a final image that conveys dynamic contextualization. Augmenting context in the foreground is relatively straightforward since it does not involve any 'mixing' of the different sources, the virtual elements can be simply overlaid onto the camera image to provide augmented context.

Putting the augmentations in the background or immersing into the shelf display, however, is technical more challenging. The contour of the front objects needs to be determined and modeled using an *occlusion model* so that the front objects accurately occlude the virtual object in the background. The occlusion model is rendered in the graphics system as a transparent (invisible) object. Technically, this is accomplished by rendering only to the graphic system depth buffer, omitting actual rendering of pixels. It

is also important that occlusion models be rendered prior to any future occluded content. Since the invisible object is in front (virtually) of the background, the occlusion model creates a hole in the overlay image such that the underlying graphics show through without occlusion. An occlusion model simulates the occlusion that would have occurred had the real object been a virtual object in the graphics system. Figure 14 shows an occlusion model with a virtual sign at the back of real sauce cans. In an immersive setting, the depth of the virtual object needs to be compared with all of the real objects or other virtual objects that may occlude it. Diminishing context is achieved by augmenting over the competition with background or contextual settings of the focal product to yield a virtual diminished view.

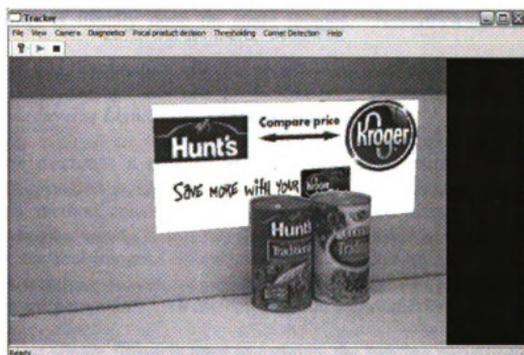


Figure 14 Occlusion model

### **3.4.3. REALTIME INVERSE LIGHTING**

In Augmented Reality (AR) systems, virtual objects are seamlessly integrated onto real scene in real time [1]. By “seamlessly”, we mean that the virtual objects are precisely registered with the real world in 3-D. This seamlessness is usually achieved by rendering the virtual objects with the same camera setting as that of the real world so that the virtual

objects are placed as if they were there. From a computer graphics point of view, rendering a model includes two aspects: geometric properties and illumination characters. Applied in AR, great amount of research effort has been conducted to discover the geometric properties of the scene, i.e. the camera settings and the pose of the virtual objects. However, the illumination characters have not obtained comparable attention as opposed to geometric properties, although illumination also plays a significant role in determining the quality of seamlessness and realistic. Furthermore, lighting is one of the physical context categorized in [2]. The virtual objects will certainly look faked if they do not exhibit the same illumination characters, such as highlight spot and shadows, as their surroundings. Retrieving lighting conditions from image, a problem referred to as *Inverse Lighting*, is usually done on static images. Nevertheless, unlike geometric properties, real life illumination conditions are usually dynamic and complex, which makes *Real time Inverse Lighting* enormously challenging.

This section discusses a new method of achieving common illumination in AR systems. This method consists two phases. First, the illumination parameters are retrieved from the real images. Second, the virtual objects are relit with synthetic light using the retrieved parameters. All the computation and rendering take place in real time. A non-linear least square fitting is used to estimate the lighting parameters from the real image. To overcome the performance limitation of non-linear least square fitting and achieve real time performance, some optimization methods are used. The problem of finding the lighting characters from image is called *inverse lighting* in the literature.

The contribution of this work is a purely software-based dynamic reverse lighting system for AR. Unlike most existing reverse lighting systems, either work on special

hardware or static images, our system can be deployed on commonly available hardware and give real time performance.

### **3.4.3.1 PRELIMINARIES**

This section describes the mathematical ground of the illumination problem in computer graphics.

*Assumptions:* Some simplifying assumptions are made to duplicate lighting conditions in real-time. First, it is assumed that there is only one point light source at a finite distance with no significant attenuation of light by the media. It is also assumed that some geometry and surface reflectance are known, and that the surface reflectance is Lambertian, meaning the surface illumination is scattered uniformly in all directions. The surface illumination is, then, independent of viewpoint. Lambertian illumination is the model commonly used for *diffuse* illumination in graphics systems. In many applications the geometry of the physical scene is already known and considerable research has been done to estimate the geometry from image, which is called “shape from shading” in the literature. A review of shape from shading algorithms can be found in [79]. Moreover, Dror, Adelson, and Willsky’s 2001 work [80], which investigates the possible solution of estimating surface reflectance from images under unknown illumination situations, makes it possible to assume that the surface reflectance is known.

### **3.4.3.2 LIGHTING MODEL**

Based on the assumptions of only one point light source and Lambertian surface reflectance, the lighting model is presented in Equation 1. Equation 1 outlines the color of each pixel in the image, determined by the color and position of the light source, the

normal of the point on the surface (geometry property) and the diffuse factor of the material (reflectance property).

$$C_P = L_D \cdot D \cdot \frac{\vec{N} \cdot (\vec{L}_p - \vec{P})}{|\vec{L}_p - \vec{P}|} \quad (1)$$

In this equation,  $C_P$  is the pixel color (triple),  $L_D$  is the light color,  $D$  is the diffusion factor, a constant ranging from 0 to 1.  $\vec{N}$  is the surface normal,  $\vec{L}_p$  is the light position and  $\vec{P}$  is the point on the surface.

Given an image, it is straightforward to acquire the color for each pixel. Based on the assumptions in the previous section, the unknowns of Equation 1 are  $L_D$  and the vector  $\vec{L}_p$ , which are the color and position of the light source respectively.

### 3.4.3.3 NON-LINEAR LEAST SQUARE ESTIMATION OF PARAMETERS

Non-linear least square estimation is used to estimate the unknowns from a set of sample points. For a given non-linear system,  $y_i = f(\vec{x}_i, \vec{\lambda})$ ,  $i = 1, 2, \dots, n$ ,  $n$  is the number of samples,  $\vec{x}$  is a set of known variables, non-linear least square estimation (also called non-linear least square fitting) solves these equations to find the values of  $\vec{\lambda}$ , which best satisfies this system of equations. All non-linear least square estimation methods are iterative. From an initial guess  $\vec{\lambda}_0$ , the estimation uses one of the descent algorithms to produce a series of vectors  $\vec{\lambda}_1, \vec{\lambda}_2, \vec{\lambda}_3 \dots$ , which, hopefully, will converge to the actual value of  $\vec{\lambda}$ . The details of non-linear least square methods and optimizations can be found in [81, 82].

The crucial step, which significantly affects the performance of a non-linear least square system, is the descent algorithm that determines how to make the series of  $\bar{\lambda}$  converge to the true value. A descent algorithm contains two parts: first, along which direction to update  $\bar{\lambda}$ ; Second, how much to update. Some descent algorithms work better if the initial guess is far away from the true value, while some have a better result when the initial guess is close to the true value; some converge faster while some are more conservative. The choice of descent algorithm requires considerable experimentation. The initial guess, also plays an important role in determining the accuracy and time complexity of the non-linear estimation. A poorly picked initial guess will not achieve real-time performance, cause the system converge to local minimum, or even worse, not converge at all. The strategies for choosing a descent algorithm and giving an initial guess that are used in this work are discussed in next section.

#### **3.4.4. PERFORMANCE ANALYSIS**

In order to achieve real-time performance, speed and accuracy are the two performance aspects that of the most concern in this application. The estimation needs to be accurate enough so that the synthetic light source lights up the virtual objects with illumination conditions common with that of the real scene. It is also necessary that the estimation is fast enough so that the virtual light and virtual objects can be merged into the real scene with no discernible delay. However, speed and accuracy are often competing goals. In order to improve one of them, it is inevitable to sacrifice the other one at some extent. The goal is to find the best possible balancing between these two aspects.

### 3.4.4.1 ACCURACY

This section addresses the facts that affect the accuracy of the estimation.

*Number and position of the sample points.* Since the number of sample points is usually larger than the number of unknowns, this system is over-determined. Nonetheless, since the measurement will inevitably include measurement errors, the more sample points available, the more information is known about the system, and therefore, the more accurate the results.

*The descent algorithm.* The descent algorithm determines the direction ( $\vec{h}$ ) along which to update the estimates and how much to move ( $\alpha$ ). Thus, the updated value is:  $\vec{\lambda}_{i+1} = \vec{\lambda}_i + \alpha \vec{h}$ . The procedure for finding  $\alpha$  is a line search. The algorithms implemented are the Newton method, and the Gauss-Newton method. It is not surprising that the latter gives much better results in terms of accuracy than the former one, since theoretically, the Gauss-Newton method has guaranteed convergence with line search, given two conditions that are affordable in this application [81]. Thus, the Gauss-Newton method was included in the final system.

*Position of the initial guess.* The position of the initial guess plays an important role in the accuracy of our estimation. For the first frame of a video sequence, there are no clues from previous experience. The frames thereafter can use the result of previous frame as the initial guess of the current frame. Thus, the quality of estimating the first frame is even more important since it affects the quality of whole video sequence. A poorly supplied initial guesses causes the system to converge to a local minimum or a saddle point, or even worse, not converge at all. Although the parameters of lighting conditions are unknown, some strategies can be applied to give a better initial guess.

First, the color of the light is related to the color of the brightest spot in the image and the material color of that spot. Second, the position of the light source is close to the brightest spot in the image. The brightest spot in the image, however, only gives the  $x$  and  $y$  coordinate. Based on the problem statement that the light source is close to the scene so that this work is worth investigating, we can supply the  $z$  coordinate of our guess with some value in between the scene and the camera. All of the above are trying to supply the estimator with the initial guess that is as close to the true value is possible.

*Parameter settings in non-linear estimation.* Some parameters in non-linear estimation, such as the maximum number of iterations in the descent algorithm and line search algorithm and the threshold of tolerable error also influence the accuracy of the result. The larger the maximum number and the smaller the threshold, the more accurate the result will be.

#### **3.4.4.2 COMPUTATIONAL COMPLEXITY**

Some of the factors that are mentioned in the previous section also affect the speed of convergence. For example, since non-linear estimation involves several iterations over the number of sample points, the more sample points there are, the longer that algorithm takes to converge. This is also true for the parameter settings of descent algorithm and line search algorithm and the algorithms themselves. Although the Gauss-Newton method converges on virtually all initial guesses, it has only linear convergence, while the Newton method has quadratic convergence.

One exception is the position of initial guess. The strategies of supplying an initial guess that is close to the true value not only improve the accuracy, but also reduce the time to converge. However, extra computation has to be done to get the closer initial

guess. Table 1 shows the time (in milliseconds) that it takes to converge for random initial guess (group A) and initial guesses with the strategies described in the previous section (group B). The average time it needs to converge in group A is much larger than that in group B (122.5 vs. 37.4). With the average of 122.5ms per frame, it is impossible to provide real time performance. In addition, the standard deviation of group A is also much larger than that of group B (149.2 vs. 1.5), which means that the resulting video of group B is much smoother than that of group A. So the additional computation needed to apply the strategies pays off.

**Table 1 Convergence time (ms) comparison for group A (random initial guess) and group B (initial guess supplied with our strategies)**

|   | 1  | 2  | 3   | 4   | 5  | 6  | 7  | 8  | 9  | 10  | Avg.  | Std. Dev. |
|---|----|----|-----|-----|----|----|----|----|----|-----|-------|-----------|
| A | 93 | 53 | 539 | 106 | 49 | 92 | 71 | 65 | 30 | 127 | 122.5 | 149.2     |
| B | 38 | 36 | 36  | 40  | 35 | 38 | 37 | 37 | 39 | 38  | 37.4  | 1.5       |

### **3.4.5. WORKING SYSTEM**

This section discusses the experiment, experiment settings, and some efforts that have been made to improve performance and ensure convergence. Figure 15 is the flowchart of the working system. The system consists of an estimation stage and rendering stage. In the estimation stage, the lighting parameters are estimated from the original video image. In rendering stage, the virtual objects are added to the scene and are relit with the synthetic light.

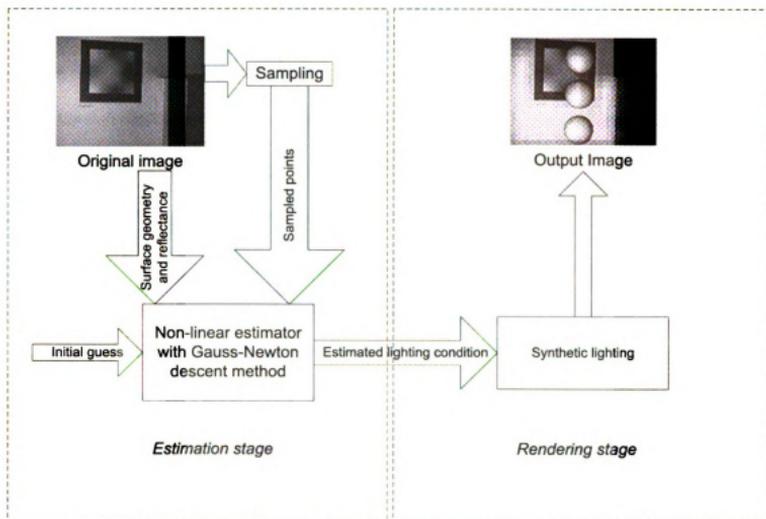


Figure 15 Flowchart of the working system

### 3.4.5.1 SYSTEM CONFIGURATION

The experiment was carried out with an easily configurable setup. A nearly dark room simulates the “only one light source” environment; a flashlight is employed as a single point light source; a white wooden block with known dimension and a white background makes up the scene to be lit up. The scene is captured by a Logitech web camera. All the computations are done on an HP Tablet PC with 1GHz Intel Centrino Mobile processor and 1.5GB RAM. .

Figure 16 shows a frame in the captured scene. The pattern at the upper left corner is a marker image that tells the system the position and orientation of the scene relative to the camera [23]. Combined with the known position and dimension of the block, the real world coordinate and normal for all points on the block or the background can be computed.

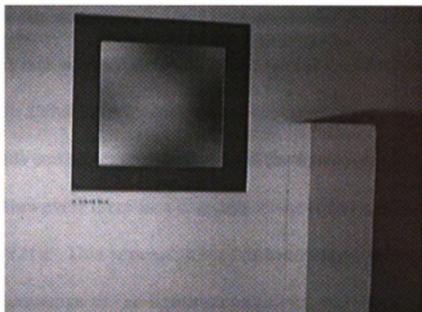


Figure 16 A frame from the captured video sequence

### 3.4.5.2 SAMPLING STRATEGY

The sampling strategy is a major element in determining the quality of this work, as discussed in the previous section. The more sample points, the more accurate result will be, albeit with a tradeoff in the form of increased convergence time. Suppose the resolution of the camera is  $m$  by  $n$ . There are potentially  $m*n$  sample points (pixels). It is neither necessary nor beneficial to feed all the sample points to the estimation procedure. The intensity of the pixels that are in the shadow is not yielded by direct lighting and involve more complicated lighting model, so those pixels are eliminated from consideration. For the remaining pixels, the following guidelines are applied:

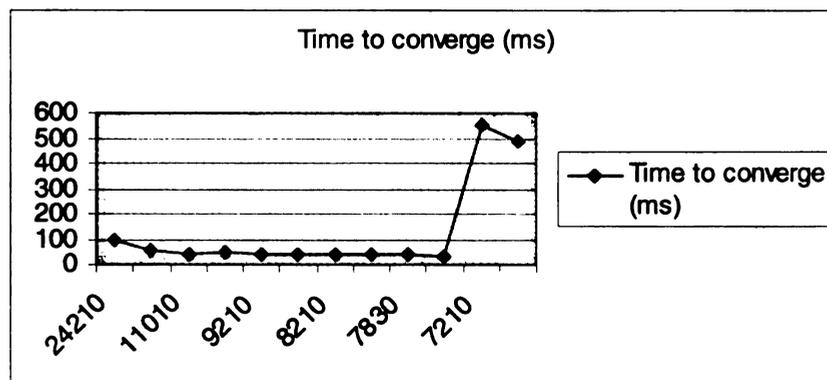
1. Sampling should include points from both the background and different facets of the local geometry.
2. Sampling should cover the largest possible area within the image.
3. Sampling should ensure that the range of the possible intensity is covered as much as possible.

According to the above guideline, the speed and accuracy of the estimation was tested on different numbers of sample points. The flashlight was fixed and the position was

measured to evaluate the accuracy. Table 2 lists the results of performance vs. number of points. Figure 17 gives a more intuitive view of the experiment results. When the number of points was reduced from 24210 to 7610, the performance gradually increased. When 35 milliseconds convergence was achieved, the system exhibited a satisfactory real time frame rate. However, there is a dramatic drop at the speed when the number of points decreases to 7210. This appears to be due to too few sample points and therefore insufficient sample coverage of the lighting conditions, making it more difficult for the non-linear estimation to converge in a short time.

**Table 2 Number of points versus performance**

| Number of points | Time to converge (ms) |
|------------------|-----------------------|
| 24210            | 94                    |
| 14210            | 55                    |
| 11010            | 42                    |
| 10250            | 49                    |
| 9210             | 42                    |
| 8490             | 40                    |
| 8210             | 38                    |
| 7940             | 37                    |
| 7830             | 38                    |
| 7610             | 35                    |
| 7210             | 553                   |
| 6810             | 490                   |



**Figure 17 Illustration of number of points versus performance**

### 3.4.5.3 JUSTIFICATION OF CONVERGENCE CONDITIONS

This section discusses the effort made to ensure the convergence of the non-linear system. It is very common that a non-linear system may not converge, or sometimes, converge to a local minimum.

First, the descent algorithm was implemented with a line search to enforce  $F(\bar{x}, \bar{\lambda}_{i+1}) < F(\bar{x}, \bar{\lambda}_i)$ , where  $F(\bar{x}, \bar{\lambda}_i) = \sum_{j=1}^n (f(\bar{x}_j, \bar{\lambda}_i) - y_j)^2$ . This prevents divergence and convergence to a maximum. It also reduces the possibility of convergence to a saddle point.

Second, by using the heuristics that we discussed in Section 3.4.5.2, the initial guess is close to the true value compared to random chosen points. This greatly raises the possibility of convergence to global minimum, which is the desired convergence point.

Third, the sampling strategies minimize the possibility of getting a singular Jacobian during the iterations of the estimation procedure, and then, the Gauss-Newton method is working.

### 3.4.5.4 RESULTS

The average converge time is 35 ms per frame, which gives real time performance with approximately 30 frame/second. The average distance of the estimated light source to the real light source is 0.64 inches. Figure 18 shows a frame (same as Figure 16) with a virtual ball and teapot lit up with the estimated light source. Of course, achieving 30 frames per second performance with just the inverse lighting solution does not necessarily imply a complete real time solution, as resources are also required for

tracking, rendering, and composition. However, it is clear that real time performance is possible and, indeed, practical for this solution.

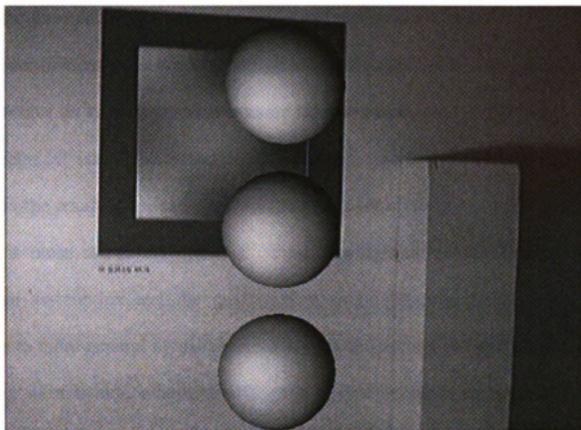


Figure 18 A frame with a virtual ball and teapot lit with estimated light source

### 3.5. SUMMARY

This chapter presents the concept of a shopping assistant that utilizes augmented reality technologies to provide personalized advertising and in-store shopping assistant based on dynamic contextualization long with technical details of the system design. The PromoPad system is a step towards ubiquitous computing in the highly lucrative grocery shopping segment. The development goal is to offer a pleasant and inviting shopping experience that is mediated by an augmented reality-based Tablet PC. Beyond traditional context awareness, this chapter developed the concept of dynamic contextualization, which suggests the modification of context to direct the interest flow of users. Dynamic

contextualization, the real-time modification of context, can be enabled by augmented reality technologies with augmentations and diminishments. Dynamic contextualization is based on, but extends beyond, the spatial and temporal context of the user. Location context, user context, and product context are integrated in this design to address the requirements of an intelligent context-aware shopping assistant.

The technical issues discussed in this chapter improve the realism of PromoPad system and the result is appealing. Nevertheless, a lot of work can be done in the future to make it more stable and general. First, the lighting model is restricted by the Lambertian assumption and the criteria of a single lighting source at infinity. An adaptation to more general lighting condition such as specular reflections, natural outdoor lighting, or normal indoor lighting will great widen the application of this technology. Possible approaches could include a signal processing [83] or statistical model.

Although the result of inverse lighting alone gives real-time performance, it is not included in the working PromoPad system due to the limitation of the assumptions and other computational cost such as tracking, rendering.

The design methodology of the PromoPad system can be extended to other circumstances such as tourism guides, training assistants, etc. Nevertheless, designers of other systems need to deliberately consider the context factors based on the requirements of an application domain.

Although this chapter has addressed several important issues in designing the PromoPad, it has not discussed the privacy issue in the project. The privacy issue arises when the retailers collect the consumption activities and try to predict the consumer's interest based on her previous shopping behavior. It is necessary to balance the tradeoff

between automation and privacy to meet the needs of both retailers and consumers. Consumers may be willing to sacrifice certain degree of their privacy in return for certain value, and retailers definitely should respect the privacy of their customers. The goal of this study is to maximize the automation and the privacy issue is beyond of the scope of this thesis.

## CHAPTER 4. DYNAMIC CONTEXTUALIZATION AND MARKETING PERCEPTIONS

A key capability of PromoPad is the ability to modify the perception of a product. As a store-based augmented reality device, PromoPad can change the context of a product. Key to the effectiveness of dynamic contextualization is the choice of the revised context. Ideally, context should increase the perceived value of products a store wishes to promote or, potentially, decrease the value of a product a store chooses not to promote. Surprisingly, both are the case. As an example, store-brand products are often more profitable than name-brand products. Hence, stores routinely use advertising that espouses the better value of store-brands, while often creating a setting that appears to decrease the perceived value of the same product from a major manufacturer. They don't seek to not sell the name-brand product at all (clearly, they could simple choose to not stock it), they seek to level the playing field a bit by balancing the perceived values.

This chapter examines the context of a *focal product*, a specific product of interest in the system. In PromoPad that is a product being displayed that the store has chosen to dynamically contextualize, effectively the *focal point* of the system. The product relationships discussed in this chapter are derived from existing studies into physical contextualization of products. This thesis examines the application of these techniques in a dynamic contextualization setting where less is currently known about effectiveness. Indeed, the value of these methods is the subject of empirical results presented in Chapter 5.

## 4.1. COMPLEMENTARY PRODUCTS

A complementary product is a product that enjoys an associative relationship with the focal product. By contextualizing the focal product with a matching product, image or symbol, the consumer's attitude toward the focal product can be influenced. Product contextualization can include functional, aesthetic, or sociocultural complements of the focal product [84].

*Functional complementary products* are products that can be consumed or utilized jointly in order to facilitate some operational relationship. For example, golf clubs can be functionally complemented by golf balls, bag, shoes, etc. A user purchasing hot dog buns is likely to purchase the hot dogs to place in them. Hence, functionally complementary products can have very close relationships that influence simultaneous purchase.

*Aesthetic complementary products* are products that are consumed because they form an inherently pleasant relationship with each other. Consumer motivation in using these products is the aesthetic pleasure derived from their juxtaposition. For example, a baroque painting in a baroque designed house is aesthetically complementary to the house. Aesthetic complementary is often highly subjective; hence it is not currently included in our experiment design, though use of experts may allow for aesthetic suggestions [85].

*Sociocultural complementary products* are group of products that involve consumption activities and/or products that hold little or no inherent relationship to each other, but are instead related through a sociocultural process of association and ascription of meaning. Groupings are valued for the ability to communicate social messages within a particular culture at a particular historic moment. For example, it is easy to socioculturally

associate BMWs with MBAs, Rolex watches, etc. Tie-dyed t-shirts are often socioculturally associated with patched blue jeans, army fatigue jackets etc. Table 3 lists some examples of product complementary as used in the base PromoPad evaluation products database.

**Table 3 Product complementary examples**

| <b>Focal Products</b> | <b>Functional Complementary</b>   | <b>Sociocultural Complementary</b>                           |
|-----------------------|---|--|
| Digital camera        | Photo papers, memory card, printer for digital camera, picture-editing software | Vacation package, plane ticket, ball park tickets            |
| PDA                   | PDA keyboard, PDA software, Wireless Internet access, memory                    | Tie, pen, cell phone, laser pointer pen                      |
| Perfume               | Body wash, deodorant, antiperspirant  | Jewelry, candles   |
| Pen                   | Notebook, highlighter, pencil jar   | Hair tie   |
| Candy bar             | Soda, popcorns, ice cream   | Ball park tickets, Big 'n' Tall clothes or shoes             |
| Wine                  | Wine stand, cork screw, glasses   | Crystal container, romantic dinner, travel package to winery |
| Shampoo               | Conditioner, hair dryer, hair gel, body wash                                    | Fruits, herbs  |
| Detergent             | Fabric softener, stain remover  | Glass cleanser, floor cleaner                                |

## **4.2. DYNAMIC CONTEXTUALIZATION OVERVIEW**

From the perspective of consumer psychology, dynamic contextualization with PromoPad can simulate an enhanced product experience. "Enhance" implies a combination of both direct experience and virtual experience. Traditionally, product experiences have been dichotomized as direct or indirect. *Direct product experience* is the unmediated interaction between consumers and products in full sensory capacity,

including visual, auditory, taste and smell, haptic and orienting [86]. Direct product experience is often obtained from personal inspection of a real product. *Indirect product experience* is experience gained through secondary sources such as advertising. Compared with indirect experience, direct experience is much richer for several reasons. First, product information is self-generated by the shopper and thus is more trustworthy. Second, the shopper can see, feel and touch a product and get input from multiple sensory channels. Third, the shopper can inspect a product in a sequence and pace of her choice and customize the information to her cognitive needs. However, direct experience from personal inspection is not perfect from a consumer learning perspective in that it is often limited to the product per se and it is not easy to incorporate external information such as the background, users and use scenarios of a product. This disadvantage can be overcome with virtual experiences as simulated in 3-D visualization.

The beauty of augmented reality is that it enables a shopper to inspect a product personally and at the same time, to view additional objects in 3-D visualization in the Tablet PC display. Objects in 3-D visualization are found to be able to generate a new form of mediated experience – *virtual experience* [10]. A virtual experience is a form of indirect experience, because both are mediated experiences [87]. However, virtual experience tends to be richer than indirect experience rendered by printed ads, television commercials, or even two-dimensional (2-D) images on the Web. Li, Daugherty, and Biocca indicate that virtual experience, as simulated in 3-D visualization, consists of more active cognitive and affective activities than 2-D marketing messages [88]. They attribute these psychological and emotional effects to the interface properties of 3-D advertising, as well as to the psychological sensation of presence.

### **4.3. DYNAMIC CONTEXTUALIZATION WITH AUGMENTED REALITY**

Dynamic contextualization is a process of contextual information rendering in multimedia form in response to cognitive needs of users when they are interacting with real objects in a changing physical environment. It is an extension of the concepts: *product contextualization* and *virtual product contextualization*. Researchers define product contextualization as the placement of the product in a particular setting that will resonate with the consumers and make clear that product's consumption practices [84]. Product contextualization is often seen in store displays and advertisement. In electronic commerce, product contextualization can be easily simulated with 3-D visualization, which can offer a variety of ways for the consumer to arrange a focal product with other complimentary products on the computer screen. Researchers use virtual contextualization to refer to the placement of complimentary products along with a focal product in 3-D visualization in order to affect the user's perception of the focal product [9]. For example, the user can arrange a set of furniture in different settings in 3-D on a website to select the preferable combination. Research demonstrated that virtual contextualization can lead to better consumer experience, brand attitude, and hence influences purchase intention [53].

Dynamic contextualization is theoretically superior to virtual contextualization in that it is a combination of both direct experience and virtual experience, resulting in an enhanced product experience. Augmented reality lies between the real world and completely virtual reality [13]. Users can add virtual objects to their perception of the real world to create an augmented reality. Although consumers can view various combinations of a focal product with different complimentary products in virtual

contextualization, their product experience is simulated and virtual in the sense that they have no direct contact with a real focal product. In dynamic contextualization using augmented reality technologies, consumers can inspect a real focal product in a virtual context that is simulated to meet their cognitive needs. Consumers can not only see the real product but also instantly access additional product information on the Tablet PC, such as complimentary products and background information of the focal product. Such an enhanced consumer experience in dynamic contextualization is even richer than merely a direct product experience.

Dynamic contextualization modifies the user's perception of the reality by either augmenting context or diminishing context. The latter is referred to as diminished reality in the literature [22].

#### **4.3.1. AUGMENTING CONTEXT**

By adding context to the focal product, PromoPad is able to give a consumer more information about the focal product than is possible in traditional media. Theoretically, the added context can be coupons, advertisements or complementing products as discussed in previous section. Based on the advertiser's needs, these pieces of information could be 2D pictures or 3-D objects that appear beside, in the foreground, or in the background of the focal product or immerse into the shelf display. It is actually possible to have content in the display with depths deeper than the physical shelf, allowing a virtual extension of the store space. Figure 19 illustrates the augmentation of a box of spaghetti with an image of cooked spaghetti with sauce.



Figure 19 Augmenting the box of spaghetti with cooked spaghetti and sauce

PromoPad can place information such as complementary settings of the product into the background of the focal product. Although it may not draw the consumer's active attention, the new information affects the consumer's attitude towards this product. The immersive setting will function in a similar fashion. Placing augmentations in the background or immersed into the layout is more technically challenging. The contour of the front objects needs to be determined and modeled using an occlusion model as discussed in Chapter 3 so that the front objects accurately occlude the virtual object in the background. In an immersive setting, the depth of the virtual object needs to be compared with all the real objects or other virtual objects that may occlude it. Figure 20 gives an example of augmenting the background. A comparison of store brand and name brand appears at the background.

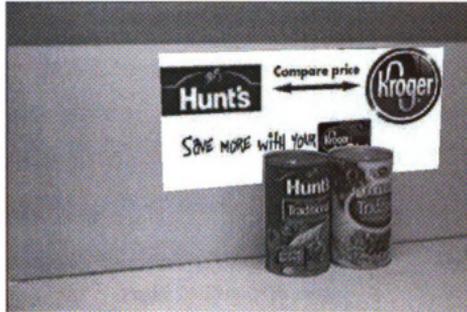


Figure 20 Augmenting the background

#### **4.3.2. DIMINISHING CONTEXT**

Whereas augmenting context highlights the focal product by delivering augmented virtual objects to the consumer, diminishing context emphasizes the focal product by hiding surrounding product items, most likely non-complementary products or competing brands. Figure 21 illustrates this idea by virtually removing the competition from the surrounding settings. Removing the competition gives more room to display information for the product that the retailer plans to introduce to the consumer or increase the sales volume at that period of time. It also allows the vender to specifically deemphasize a competing product.



Figure 21 Diminishing context

Both augmentations and diminishments allow retailers to apply business strategy and direct user's interests. Table 4 lists several possible examples of augmentations and diminishments to the focal products, which are listed in Table 3, other than coupons and sales offer.

**Table 4 Examples of augmentations and diminishments**

| <b>Focal Products</b> | <b>Augmentations</b>   | <b>Diminishments</b>                                     |
|-----------------------|--|--|
| Digital camera        | Picture slideshow, feature demonstration, accessories          | Outmoded models, security locks and latches, film camera |
| PDA                   | PDA keyboard, PDA software, Wireless Internet access, memory   | Security locks and latches, laptop computers             |
| Perfume               | Flowers, romantic pictures                                     | Disliked brands or scent of the consumer                 |
| Pen                   | Notebook, grade report, back to school picture                 | Crayon, scissors   |
| Candy bar             | Cartoon characters, ice cream                                  | Mint drops, energy bar                                   |
| Wine                  | Glasses, roses, picture of a grand banquet                     | All other than the bottle under inspection               |
| Shampoo               | Hair dryer, fruits, picture of model with beautiful hair       | Hair dye   |
| Detergent             | Picture of silk or wool, movie clip shows the effect after use | Unfavorable ingredient varieties                         |

\* This is determined by user profile, hence it is user dependent.

## **CHAPTER 5. EMPIRICAL STUDIES**

Existing research on product contextualization has assumed store settings and physical contextualization. Dynamic contextualization introduces many new capabilities for product marketing, but it must be shown that these new methods are as or more effective than traditional methods. Hence, several empirical studies were conducted to test the effectiveness and feasibility of dynamic contextualization based on PromoPad test-bed. This chapter discusses the methodologies, procedures, and data analysis of empirical studies.

### **5.1. INTRODUCTION**

Considerable research has been conducted on utilizing AR technologies in various application domains, including tourist guide [41, 45], assembly instruction [24, 39], and others [4]. Nevertheless, most of the work reported focuses on technical issues of presenting information in the form of computer generated imagery. The usability of these systems and their influence on users, however, is not as well addressed. Throughout the design and development of PromoPad, emphasis was placed on not only the technical issues related to using AR in public environments [7], but also the question of if the virtual experience [8] created by AR will have significant influence on consumers. Previously studies have shown that a virtual experience can enhance both consumer learning and experience [9], but these studies were conducted in very different environments, such as online commerce and the World Wide Web. PromoPad brings a

virtual experience to physical (brick and mortar) store settings using AR technologies. This is especially appealing to retailers who seek methods that allow them to manipulate consumers' interest by changing the setting in the system without having to physically change store signs or shelf layout. In addition to the labor saving, placing products in a 3D visualization results in more product knowledge, better brand attitude, and elevated purchase intention relative to traditional advertising [10].

Several empirical studies on the effectiveness of augmented reality technologies in terms of human computer interaction provide sufficient evidence that augmented reality systems improve operational performance in an instructed assembly task, training, and tourism guide [24, 89, 90]. While this previous work has focused on improving performance in terms of learning time or decreased process mistakes, the PromoPad project is focused on *influence*.

The concept of a computerized shopping assistant is not new. The PSA [70], MyGrocer [68], and Project Voyager [69] are prototype shopping assistants that provide product reviews, promotions, and pricing information, but are not augmented reality devices. PromoPad addresses different issues from a different point of view and hence, proposes different solutions. In addition to providing traditional shopping assistant information, PromoPad, powered with dynamic contextualization, also influences consumers' interests through the modified perception of the product in situ. User studies were conducted to justify this statement.

## **5.2. USER STUDY 1: AUGMENTING CONEXT**

Product contextualization is the placement of a product into a setting more conducive to purchase intent. A consumer is more likely to purchase a product in a display than one sitting among other products on the shelf. Traditional product contextualization is realized by shelf layout or store signs. Augmented reality allows for *dynamic contextualization*, the contextualization of a product on the AR display through the use of computer-generated augmentations [6]. In addition, dynamic contextualization associates focal product with complementary products in the form of computer-generated virtual objects. This experiment tests the effectiveness of augmenting context, as a part of dynamic contextualization.

With AR, store managers can create contextualization of products by modifying the system settings; there is no need to modify the physical shelf setting in different places. In addition, virtual contextualization can include animations or video and can be three-dimensional having a perceivable depth relative to the product.

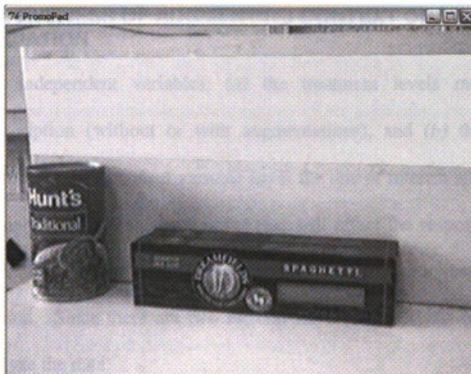
### **5.2.1. EXPERIMENT DESCRIPTION**

This experiment used boxed spaghetti and canned spaghetti sauce to test the effect of augmenting context. Spaghetti and canned spaghetti sauce are commonly available in grocery stores and exhibit a *functional contextualization* since they are usually consumed together as functional complements [6]. Figure 22 shows the experimental shelf with a box of spaghetti and Hunt's brand sauce. The pattern images are fiducial markers for the vision-based tracking system [23].



Figure 22 Spaghetti and sauce can

There are two levels of treatment: (i) without augmenting context and (ii) with augmenting context. Figure 23(a) illustrates the view in the PromoPad without augmentations. It is simply the video captured by the camera except for the virtual patches used to hide the fiducial markers. The participants who received treatment (i) saw this view. The participants who received treatment (ii) saw a view like Figure 23(b). The spaghetti is contextualized with an animated rotating sauce can in front of it and a virtual image showing a recipe of the spaghetti with Hunt's sauce. The question is if contextualization provides the consumers with such information that the Hunt's sauce is associated with the spaghetti, thereby boosting sales volume of Hunt's sauce to consumers who would also purchase spaghetti.



(a) View in the PromoPad without augmented context



(b) View in the PromoPad with augmented context

Figure 23 the view in the PromoPad for two treatment levels

### **5.2.2. METHODOLOGIES**

Two effects of augmenting context were tested in this experiment. One is the effectiveness on product connection; the other is the effectiveness on purchase intent.

### **5.2.2.1 EFFECTIVENESS OF AUGMENTING CONTEXT ON PRODUCT ASSOCIATION**

There are two independent variables: (*a*) the treatment levels mentioned in the Experiment Description (without or with augmentations), and (*b*) the familiarity of spaghetti as a product. Independent variable (*a*) is the one of interest and subject to test. Independent variable (*b*) is a nuisance factor that will affect the response but is known and controllable. There is one dependent variable (response), participants' perception of product connection. Since there are two factors, ANOVA (Analysis of Variance) [91] was used to analyze the data.

The study format consisted of a pre-experiment survey, utilization of the PromoPad system in the randomly selected treatment level, and a post-experiment survey. In the pre-experiment survey, subjects scored their familiarity on a five-point Likert scale where 1 refers to no familiarity and 5 refers to very familiar. The participants were randomly selected to receive a treatment level. After the use of the PromoPad, the participants were asked to complete a post-experiment survey where they scored their perceptions of the product association. The information collected in the pre-experiment helped to control the nuisance factor and therefore conduct a more accurate data analysis of the responses.

### **5.2.2.2 EFFECTIVENESS OF AUGMENTING CONTEXT ON PURCHASE INTENT**

There are two independent variables: (*a*) the treatment levels mentioned in the Experiment Description (without or with augmentations), and (*b*) the preference of spaghetti and sauce (doesn't like spaghetti, like spaghetti but prefer homemade sauce, and like spaghetti but prefer canned sauce). Again, independent variable (*a*) is the one of

interest and subject to test. Independent variable (*b*) is a nuisance factor that will affect the response but it is also known and controllable. There is one dependent variable (response), participants' purchase intent of Hunt's sauce quantified on the Likert scale. Again, the data was analyzed using ANOVA.

In the pre-experiment survey, subjects scored their preference on a five-point Likert scale where 1 refers to no interest and 5 refers to yes highly preferential. The participants were randomly selected to receive a treatment level. After the use of the PromoPad, the participants were asked to complete a post-experiment survey where they scored their purchase intent on a five-point Likert scale with 1 refers to no and 5 refers to yes. The information collected in the pre-experiment helped control the nuisance factor and therefore perform a more accurate data analysis of the responses.

### **5.2.3. PARTICIPANTS**

20 graduate and undergraduate students aged from 18 to 35 voluntarily participated in this study. All participants had no handicaps that limit their use of hands, arms, and eyes. All participants had no prior experience with AR systems.

### **5.2.4. PROCEDURE**

Participants entered the lab. They were given a brief description of the experiment and a demonstration of the PromoPad system by the experimenter. They were then asked to sign a consent form. Assuming they consented to the experiment, participants were then asked to complete a pre-experiment survey. Then a treatment was randomly chosen for the participant. For treatment (*i*) (without augmentations), the participant would use the PromoPad system with only a video display, no augmentations of the imagery exception the virtual background used to cover the fiducial images; the view in the PromoPad is as

shown in Figure 23(a). For treatment (ii) (with augmentations), the participant used the PromoPad system with augmented imagery, as shown in Figure 23(b). After the use of the PromoPad system, the participant is asked to complete a post-experiment survey which collects the responses.

### 5.2.5. DATA ANALYSIS

This section discusses detailed data analysis and statistical tests.

#### 5.2.5.1 PRODUCT ASSOCIATION

For the participants who received treatment (i) (without augmentations), score mean is 3, variance is 1.33, and median is 3. For the participants who received treatment (ii) (with augmentations), score mean is 4.2, variance is 1.067, and median is 4.5. Along with the histogram (Figure 24) and box plot (Figure 25), the effect is positive. Virtual contextualization boosts the consumers' perception of product association.

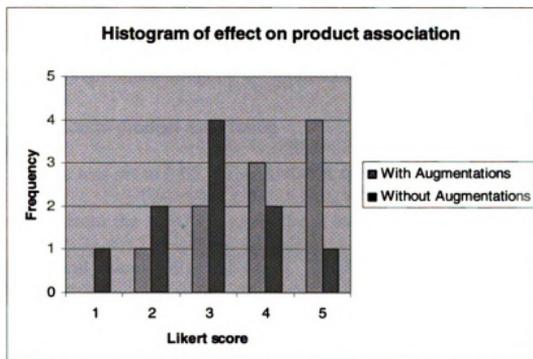


Figure 24 Histogram of effect on product association

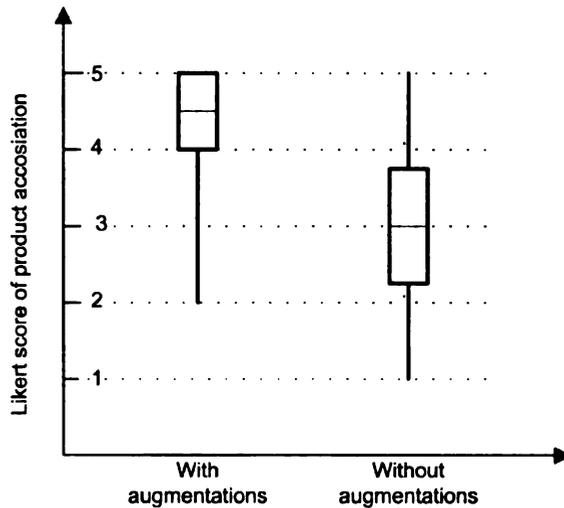


Figure 25 Box plot of effect on product association

The perception of whether the two products are functionally contextual depends on familiarity with the products. It is assumed that for those who are not familiar with the products, it is more difficult for them to associate these products than those who are familiar with the products. Hence, in this test, there are two factors: (a) with or without augmentations; (b) familiar or unfamiliar with the products, and one response (the level of functional complement the user perceive). The null hypothesis to be tested is:

$H_0$ : There is no significant effect of the treatment (with or without augmentations) on consumer's perception of product association

The critical value was set to 0.05. The ANOVA table of the test is presented in Table 5. As can be seen from the ANOVA table, the  $F$  statistics of the test is greater than  $F$  statistics of the critical value ( $169 > 161.4462$ ). The  $P$ -value is 0.0488, which is less than 0.05, the critical value. Hence there is statistically significant evidence to reject null hypothesis [91]. In other words, augmenting the context with connected products has a significant effect at influence consumers' perception of products complementary. Also, further observing the ANOVA table, the  $F$  statistics of the block is big, and the  $P$ -value

of the block is small, which means the participants' familiarity of the products also differs significantly. This justifies the presumption that the consumers' familiarity of the products is a nuisance factor that affects the response significantly.

**Table 5 ANOVA table for perception of product connection**

| Source of Variation | SS    | df | MS   | F   | P-value | F critical |
|---------------------|-------|----|------|-----|---------|------------|
| Factor a            | 1.173 | 1  | 1.17 | 169 | 0.048   | 161.4462   |
| Factor b            | 2.006 | 1  | 2.01 | 289 | 0.037   | 161.4462   |
| Error               | 0.006 | 1  | 0.01 |     |         |            |
| Total               | 3.19  | 3  |      |     |         |            |

### 5.2.5.2 PURCHASE INTENT

Another set of questions in the post-experiment survey asked for the purchase possibility of Hunt's sauce when purchasing spaghetti on a five-point Likert scale where 1 refers to not likely 5 refers to very likely. This quantifies the purchase intent.

For the participants that do not see augmentations, the score mean is 2.4, variance is 1.84, and median is 2.5. For the participants see augmentations, score mean is 3.4, variance is 1.64, and median is 4. Along with the histogram (Figure 26) and box plot (Figure 27), the effect is positive. Dynamic Contextualization increases purchase intent.

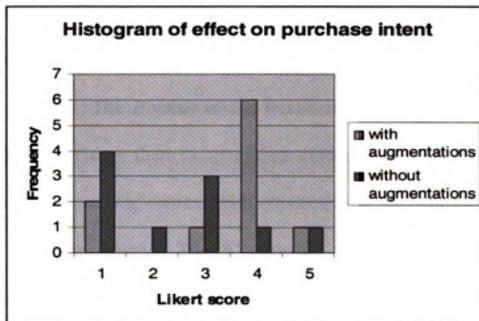


Figure 26 Histogram of effect on purchase intent

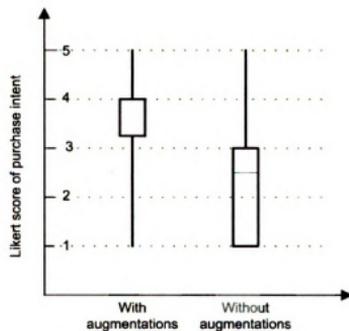


Figure 27 Box plot of effect on purchase intent

There are again two factors: (a) with or without augmentations; (b) the preference of spaghetti and sauce (doesn't like spaghetti, like spaghetti but prefer homemade sauce, and like spaghetti but prefer canned sauce). The response is the purchase intention quantified on the Likert scale. The null hypothesis to be tested is:

$H_0$ : There is no significant effect of virtual contextualization at consumer's purchase intent.

The ANOVA table of the test is shown in Table 6. As we can see from the ANOVA table, the  $F$  statistics of the treatment is greater than  $F$  statistics of the critical value ( $21.72778 > 18.51276$ ). The  $P$ -value of the treatment is 0.043072, which is less than 0.05, the critical value. Hence there is significant evidence to reject the null hypothesis. In the other words, augmenting context with connected products has a significant effect on consumers' purchase intent. The  $F$  statistics and  $P$ -value of the participants' preference show that participants' preference also has a significant impact on their purchase intent.

**Table 6 ANOVA table for consumer's purchase intent**

| Source of Variation | SS     | df | MS     | F      | P-value | F critical |
|---------------------|--------|----|--------|--------|---------|------------|
| Treatments          | 3.1248 | 1  | 3.1248 | 21.728 | 0.043   | 18.512     |
| Blocks              | 7.9156 | 2  | 3.9578 | 27.520 | 0.035   | 19.000     |
| Error               | 0.2876 | 2  | 0.1438 |        |         |            |
| Total               | 11.328 | 5  |        |        |         |            |

### **5.2.6. EXPERIMENT SUMMARY**

From the above data analysis, there is statistically significant evidence that virtual contextualization using AR technologies can boost consumers' perception and attitude towards the products and hence impulse purchase intention.

### **5.3. USER STUDY 2: DIMINISHING CONTEXT**

In addition to augmenting context with complementary products, dynamic contextualization is also capable of manipulating consumers' interests by highlighting one product using augmented imagery and/or virtually removing competitive products.

### 5.3.1. EXPERIMENT DESCRIPTION

In this experiment, three bottles of wine are used to test the effectiveness of manipulating consumers' interests using AR. Figure 28 shows the shelf setting for three commercially available wines. The augmented image in PromoPad virtually removes the Meridian and Beringer wines in order to promote Yellow Tail. Other than the augmentation, the presentation of the three wines was identical.

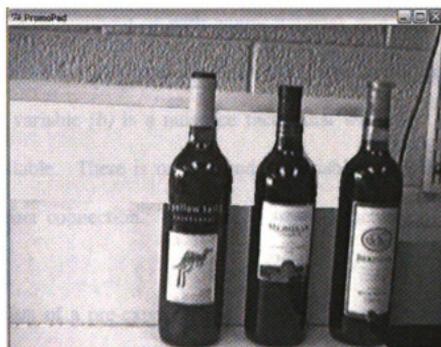


Figure 28 Wines

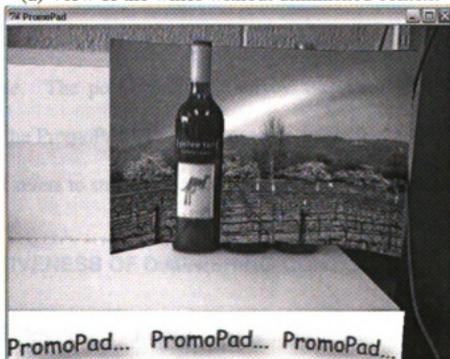
In a traditional store setting, if a store manager wishes to promote one product, he/she has to commit noticeable store space surrounding the product or move promoted products to a specific aisle placement. Powered with AR technologies, this promotion can be easily realized by changing the configuration at the server end. In addition, the signs can be dynamic, including video or animations, which are superior to physical store signs.

Again, participants are randomly selected to use the PromoPad with augmented imagery or without augmented imagery. The virtual imagery in this experiment setting is the virtual 'store sign' to manipulate users' interests. Without augmented imagery, the

participants will see the exact video coming through the camera as if they were looking at the shelf directly without the PromoPad, except for the background image to hide the fiducial images, as shown in Figure 29(a). In the augmented imagery, the participants see a virtual vineyard image behind the Yellow Tail wine bottle in addition to the incoming video. This virtual vineyard image appears to hide the other two wines as if they were removed from the shelf, as shown in Figure 29(b). By this virtual excision, the store manager attempts to attract more attention to Yellow Tail wine—the promoted product.



(a) View of the wines without diminished context



(b) View of the wines with diminished context

Figure 29 Two levels of treatment with wines

### **5.3.2. METHODOLOGIES**

Two effects of dynamic contextualization on manipulating consumers' interests were examined in this experiment. One is the effect on consumers' perception of product promotion status; the other is the effect on purchase intent.

#### **5.3.2.1 EFFECTIVENESS OF DIMINISHING CONTEXT ON PRODUCT PROMOTION STATUS**

There are two independent variables: *(a)* the treatment levels mentioned in the Experiment Description (without or with augmentations), and *(b)* the prior knowledge and experience with wines. Independent variable *(a)* is the one of interest and subject to test. Independent variable *(b)* is a nuisance factor that will affect the response but is known and controllable. There is one dependent variable (response), the participants' perception of product connection. Since there are two factors, ANOVA was used to analyze the data.

The study consists of a pre-experiment survey, the use of PromoPad system, and a post-experiment survey. The pre-experiment survey asks for the participant's prior knowledge and experience with wines, which is presume to be a nuisance factor that will affect the response. The post-experiment survey asks their perception of promotion status after using the PromoPad. The response was quantified as a score on a five-point Likert scale with 1 refers to strong negative and 5 refers to strong positive.

#### **5.3.2.2 EFFECTIVENESS OF DIMINISHING CONTEXT ON PURCHASE INTENT**

There are two independent variables: *(a)* the treatment levels mentioned in the Experiment Description (without or with augmentations), and *(b)* the preference of wines.

Independent variable (*a*) is the one of interest and subject to test. Independent variable (*b*) is a nuisance factor that will affect the response but is known and controllable. There is one dependent variable (response), participants' perception of product connection. Since there are two factors, ANOVA was used to analyze the data.

The study consists of a pre-experiment survey, the use of PromoPad system, and a post-experiment survey. The pre-experiment survey asks for the participant's preference with wines. The post-experiment survey asks their purchase intent after using the PromoPad. The response was quantified as a score on a five-point Likert scale where 1 refers to strong negative and 5 refers to strong positive.

### **5.3.3. DATA ANALYSIS**

This section presents detailed data analysis and statistical tests of user study 2.

#### **5.3.3.1 EFFECTIVENESS OF DIMINISHING CONTEXT ON PRODUCT PROMOTION STATUS**

Awareness of promotion is a valued consequence. Customers aware of a promotion are more likely to purchase the promoted product. Hence, one question is if participants think there is promotion for Yellow tail wine? For the participants that did not see augmentations, the Likert score mean is 2.7, variance is 2.23, and median is 2.5. For the participants seeing augmentations, the Likert score mean is 4, variance is 2, and median is 4.5. These results, illustrated by the histogram (Figure 30) and box plot (Figure 31), indicates a positive effect on participants' perception of the product's promotion status.

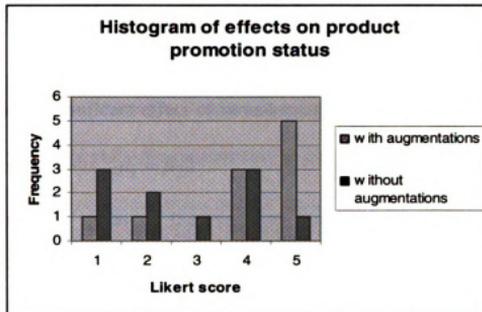


Figure 30 Histogram of effects on product promotion status

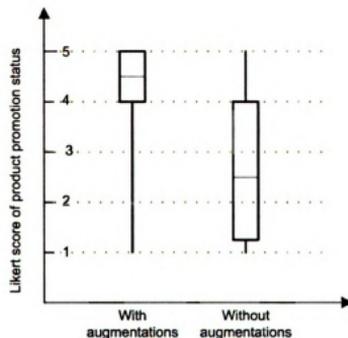


Figure 31 box plot of effects on product promotion status

An ANOVA was conducted to justify the observation of the data. Three levels of prior knowledge and experience with wines were set: 1. little experience; 2. average experience; 3. much experience. Thus there are two factors: (a) with or without virtual imagery and three blocks; (b) the experience to wines (1. little experience; 2. average experience; 3. much experience). The response is the participants' score on a five-point

Likert scale with 1 toward negative and 5 towards positive. The null hypothesis under test is:

$H_0$ : There is no significant effect of virtual imagery at consumer's perception of the product's promotion status. ( $\mu_1 = \mu_2$ )

$\mu_1$  and  $\mu_2$  are the mean of the responses under the two levels of factor (a), respectively.

Table 7 shows the ANOVA table of this analysis. Surprisingly, the  $F$  statistics and  $P$ -value indicate that neither the treatment nor the block have a significant impact on consumers' perception of the product promotion status.

**Table 7 ANOVA table for perception of wines**

| Source of Variation | SS     | df | MS     | F      | P-value | F critical |
|---------------------|--------|----|--------|--------|---------|------------|
| Factor (a)          | 0.1157 | 1  | 0.1157 | 0.109  | 0.7724  | 18.512     |
| Factor (b)          | 2.370  | 2  | 1.1852 | 1.1179 | 0.4721  | 19.000     |
| Error               | 2.120  | 2  | 1.0601 |        |         |            |
| Total               | 4.6065 | 5  |        |        |         |            |

Further analysis of the data was conducted with two separated tests on the two factors. A two-sample  $t$  test on factor (a) and a three level single factor ANOVA on factor (b).

A two-sample  $t$  test was run on factor (a). For without virtual imagery, the sample mean is 2.7, variance is 2.233; for with virtual imagery, the sample mean is 4, variance is 2. The  $t$  statistics of the two-sample is -1.998 ( $t$  critical value is 1.73.  $t_0 < -t_{critical}$ ) and the one-side  $P$ -value is 0.03. It indicates the treatment differs significantly.

On the other hand, for the three levels of factor (b), the ANOVA shows the  $F$  statistics is 3.5, ( $F_{critical} = 3.685$ ), the  $P$ -value is 0.06. So it can be concluded that there is no

significant impact of factor (*b*). This indicates that the presumption of prior experience with wines as a nuisance factor was not necessary.

Based on the two separated tests, it is concluded that diminishing context increases the consumers' perception of the promotion status.

### 5.3.3.2 EFFECTIVENESS OF DIMINISHING CONTEXT ON PURCHASE INTENT

The next question is to be analyzed is if augmentations boost purchase intent. For the participants that do not see augmentations, the Likert score mean is 2.2, variance is 1.511, and median is 2. For the participants see augmentations, the Likert score mean is 3.4, variance is 2.489, and median is 4. Examining only the histogram (Figure 32) and box plot (Figure 33), it seems hard to draw any conclusion. Although the mean and median score for with augmentations are higher than those without augmentations, the variance of the subjects with augmentations is greater too.

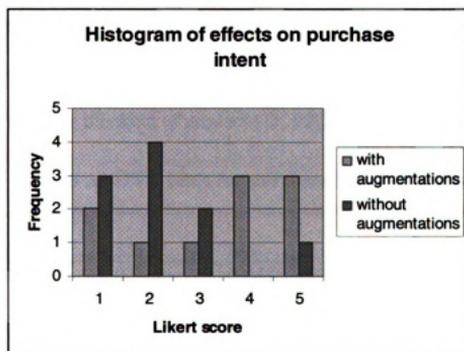


Figure 32 Histogram of effects on purchase intent

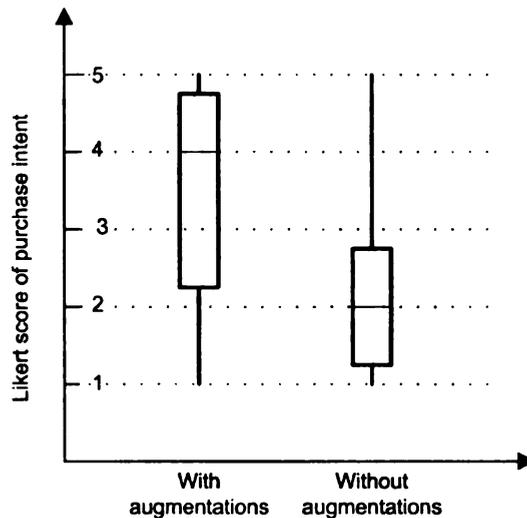


Figure 33 Box plot of effects on purchase intent

A more accurate ANOVA was performed with two factors: (a) with or without virtual imagery; (b) preference of wines (don't like wines, average, wine lovers)

From the ANOVA table (Table 8) of the test on the effect of virtual imagery to purchase intent, it can be seen that the F statistics and P-value for the treatments shows no significant influence of consumers' purchase intention, while the blocks do make a significant difference. The reason for this is that people's preference for this kind of product (wine) is hardly influenced by a promotion. The promotion is more likely to influence people who like this kind of product [11, 92].

**Table 8 ANOVA table for purchase intent of wines**

| Source of Variation | SS            | df       | MS     | F      | P-value | F critical |
|---------------------|---------------|----------|--------|--------|---------|------------|
| Treatments          | 0.0416        | 1        | 0.0416 | 0.392  | 0.594   | 18.512     |
| Blocks              | 7.9422        | 2        | 3.9711 | 37.422 | 0.026   | 19.000     |
| Error               | 0.2122        | 2        | 0.1061 |        |         |            |
| <b>Total</b>        | <b>8.1961</b> | <b>5</b> |        |        |         |            |

From the above data analysis, a conclusion can be drawn that, consumers' interest can be manipulated by virtual imagery, but the purchase intention highly depends on the property of the products.

#### **5.4. USER STUDY 3: FUNCTIONAL COMPLIMENTARY**

The concept of a functional complement as a marketing methodology is based on modifying the environment of a focal product to affect the consumer's attitude toward the focal product through the association of other products that have a functional relationship with the product being marketed. This experiment examines the effectiveness of using augmented reality technologies to artificially create functional complementary relationships for products. Functional complementary associates the focal product with products that can be consumed jointly in order to facilitate some operational relationship. For example, digital cameras are functionally complemented by tripods, memory sticks, etc. Functional complementary products can have very close relationships that influence simultaneous purchase [7]. Involvement is a term that describes the time, thought, energy, and other resources people devote to the purchase process [93]. Most of time, involvement is represented by the price of the product. In this experiment, we test the effect of functional contextualization using AR technologies, which will be referred to as virtual functional contextualization in the rest of the text, on a high involvement product (digital camera) and a low involvement product (wine). Table 9 lists the focal products and their functional complements that are conducted in the experiment. These products and their functional complementary were selected from a survey conducted by the Department of Advertising.

**Table 9 Experiment scenario**

| Focal Products | Functional complementary |
|----------------|--------------------------|
| Digital Camera | Tripod                   |
| Wine           | Glasses                  |



**Velbon**

Velbon Tripod EI Carnage Carbon Fiber 504

Includes 3-way Panhead (PH-250B)  
4 Section w/ split center column  
Neoprene grips and rubber feet  
Effortless Lever Lock System  
Tripod case

**\$259.95**



**Quantaray QSX 2001 UT Tripod**

- 3-way panhead
- metal leg brace with lock
- tension adjustment
- Holds up to 4.4 lbs
- Lifetime warranty

**\$19.99**

(a) Tripod – high involvement

(b) Tripod – low involvement



**SWAROVSKI Crystalline Red Wine Pair**

Clear crystal filled stems feature a beautifully faceted clear crystall base

**\$268.00 per pair**

Crafted in lead-free crystalline blown glass, the Rigoletto collection presents clean, simple shapes with no tarnishes, tints or imperfections. Dishwasher safe.

**\$6.30 each**



(c) Glasses – high involvement

(d) Glasses – low involvement

Figure 34 Functional complementary of camera (tripod) and wine (wine glasses)

For each focal product, there is one high involvement functional complementary product and one low involvement functional complementary product associated with it.

The involvement of the complementary product is a characteristic of an association that indicates if the associated product raises or lowers the combined value of the association. As an example, a high quality tripod may infer that the digital camera, the focal product, is an expensive, high quality camera and would be classified as high involvement. On the other hand, a low cost tripod may infer that the digital camera is a cheap, amateur one. Involvement is expected to modify a consumer's attitude toward the focal product. Figure 34 shows the virtual functional complementary products for digital camera (tripod) and wine (glasses).

This experiment was implemented using the ImageTclAR augmented reality development environment [72]. Participants were randomly selected to receive one of the two treatment levels. Table 10 lists the experiment settings for each treatment.

**Table 10 Experiment settings for each treatment**

| Treatments  | Experiment setting for digital camera    | Experiment setting for wine   |
|-------------|--|-------------------------------|
| Treatment 1 | Digital camera / high involvement tripod | Wine/high involvement glasses |
| Treatment 2 | Digital camera/low involvement tripod    | Wine/low involvement glasses  |

Figure 35 shows the experiment settings of this user study. Figure 35 (a) and Figure 35 (b) are the original shelf setting with a digital camera and bottle of wine respectively. Figure 36 is the view in PromoPad that the participants who were randomly determined to receive high involvement complementary treatment would see with (a) for digital camera and (b) for wine. Figure 37 is the view in PromoPad that the participants who were randomly determined to receive low involvement complementary treatment would see with (a) for digital camera and (b) for wine.



### **5.4.1. METHODOLOGIES**

This experiment is testing the effectiveness of virtual functional contextualization on both high involvement product and a low involvement product, so there are two sets of data with the same methodologies to analyze.

For the high involvement product (digital camera), the independent variable is the treatment level the participant receives (high involvement complementary or low involvement complementary). The dependent variable is the participant's rating of the focal product. With one two-level independent variable, a two-sample  $t$  test gives a good data analysis [91].

For the low involvement product (wine), the independent variable is the treatment level the participant receives (high involvement complementary or low involvement complementary). The dependent variable is the participant's rating of the focal product. Again, a two-sample  $t$  test will be used to analyze the data.

### **5.4.2. PARTICIPANTS**

12 graduate and undergraduate students aged from 18 to 35 participated in the study voluntarily. All participants have no handicaps to limit their use of hands, arms and eyes. All participants had no prior experience with AR systems.

### **5.4.3. PROCEDURE**

Participants entered the lab and were asked to sign a consent form after a brief introduction of the PromoPad system given by the experimenter. Participants were then asked to complete a pre-experiment survey, which collects the prior knowledge. Then a treatment was randomly chosen for the participant. For treatment 1 (low involvement complementary), the participant used the PromoPad system with augmentations of low

involvement complementary product as shown in Figure 37. For treatment 2 (high involvement complementary), the participant used the PromoPad system with augmentations of high involvement complementary product as shown in Figure 36. After the use of the PromoPad system, the participants were asked to fill a post-experiment survey, which collects the response, i.e., the participant's perception for the focal product. The responses were quantified as scores on a five-point Likert scale where 1 towards low value and 5 towards high value.

#### **5.4.4. DATA ANALYSIS**

This section discusses the data analysis of the effect of virtual functional contextualization.

##### **5.4.4.1 EFFECT OF VIRTUAL FUNCTIONAL CONTEXTUALIZATION ON HIGH INVOLVEMENT PRODUCT**

First to be presented is the data analysis of the effect of virtual functional contextualization on a high involvement product, i.e. digital camera. For the participants who received low involvement virtual complementary, the score mean is 2.83, variance is 0.57, median is 3. For the participants who received high involvement virtual complementary, the score mean is 3.67, variance is 0.27, median is 4. This preliminary analysis indicates a positive effect of high involvement complementary compared to low involvement complementary. The histogram (Figure 38) and box plot (Figure 39) of the rating as shown in Figure 39 illustrate the positive effect as well.

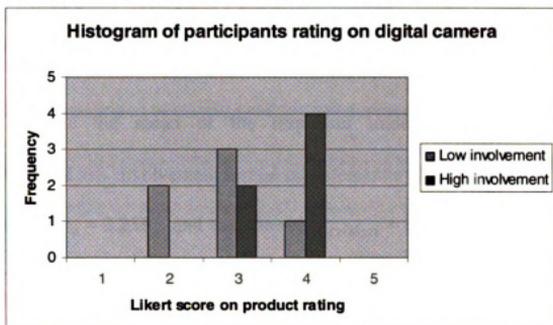


Figure 38 Histogram of rating on digital camera with two levels of complementary involvement

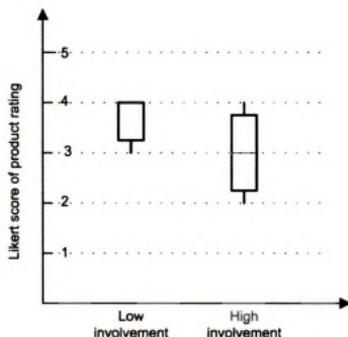


Figure 39 Box plot of rating on digital camera with two levels of complementary involvement

A more accurate statistical analysis was applied to justify the preliminary observation. A two-sample  $t$  test was done on the scores of the two levels of involvement. The null hypothesis is:

$H_0$ : There is no significant effect of levels of involvement on consumer's rating of the product (camera). ( $\mu_1 = \mu_2$ )

The alternative hypothesis is:

$H_1$ : Consumers' who receive high involvement complementary rate the product higher than those who receive low involvement complementary ( $\mu_1 > \mu_2$ )

$\mu_1$  and  $\mu_2$  are the mean of the responses under high involvement virtual complementary and low involvement virtual complementary, respectively.

The  $t$  statistics  $t_0 = 2.236$ , and the one-side  $t_{critical} = 1.833$ , so  $t_0 > t_{critical}$ , and P-value (0.026) less than the critical value (0.05), which meets the criteria to reject the null hypothesis. Thus it can be concluded that the statistical test supports the preliminary observation and levels of virtual complementary have a statistically significant effect on the consumers' perception of high involvement product.

Additional evidence supporting the above conclusion is the difference between the rating of the focal product and complementary product. As can be seen in Figure 40, the difference between the ratings for tripod and camera from a single participant is no greater than 1, which means that the participant's perception of focal product (camera) and virtual functional complementary (tripod) are highly correlated.

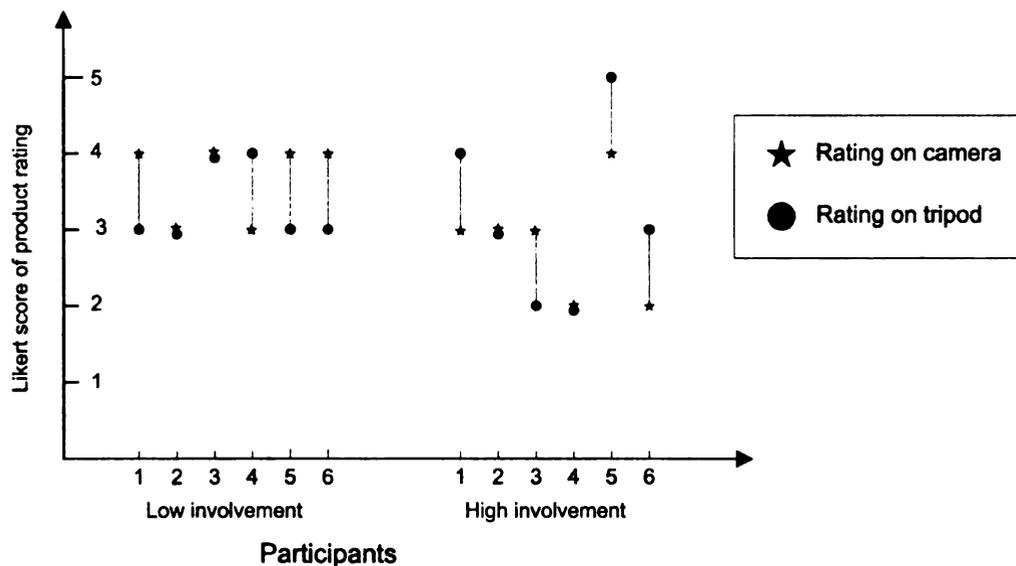


Figure 40 Participants rating on cameras and tripods in pair

A two-sample  $t$  test on the difference between the ratings of tripod and camera from two treatments with null hypothesis:

$H_0$ : There is no significant effect of levels of involvement on the difference of consumer's rating of the virtual functional complementary (tripod) and the focal product (camera) ( $\mu_1 = \mu_2$ )

The alternative hypothesis is:

$H_1: \mu_1 \neq \mu_2$

The  $t$  test shows that for high involvement, the mean is 0.33, variance is 0.667, for low involvement, the mean is -0.33, variance is 0.667,  $t_0 = 1.414$ ,  $t_{critical} = 1.81$ ,  $t_0 < t_{critical}$ , and P-value is 0.188. This is statistical evidence for accepting the null hypothesis. Hence, it can be concluded that there is no significant effect of levels of involvement on the difference of consumers' rating of the virtual functional complementary and the focal product. The consumers' rating of the focal product is highly correlated with the rating of the virtual functional complementary. In another words, different levels of virtual functional complementary has significant effect in determining consumers' perception of the focal product.

#### **5.4.4.2 EFFECT OF VIRTUAL FUNCTIONAL CONTEXTUALIZATION ON HIGH INVOLVEMENT PRODUCT**

The result of virtual functional contextualization on a low involvement product (wine) is satisfying as well. The perception of the product (wine) was quantified on the rating of a five-point Likert scale with 1 refers to low quality and 5 refers to high quality. For the low involvement, median is 3.5, mean is 3.389, variance is 0.6667. For the high involvement, median is 4.5, mean is 4.389, variance is 0.667. High involvement

complementary is 1 higher than low involvement complementary on average. The preliminary examination of the data suggests that the levels of virtual complementary involvement have a significant effect on consumers' perception of the focal product. The histogram (Figure 41) and box plot (Figure 42) also give a good observation of the preliminary result.

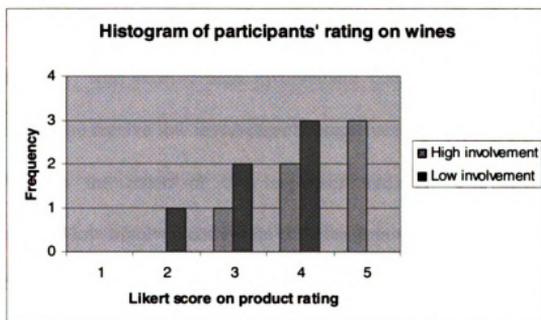


Figure 41 Histogram of rating on wine with two levels of complementary involvement

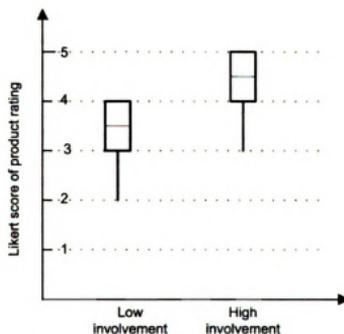


Figure 42 Box plot of rating on wine with two levels of complementary involvement

Using a more accurate statistical analysis to justify the preliminary observation, a two-sample  $t$  test was conducted on the scores of the two levels of involvement. The null hypothesis is:

$H_0$ : There is no significant effect of levels of involvement on consumer's rating of the product (wine). ( $\mu_1 = \mu_2$ )

The alternative hypothesis is:

$H_1$ : Consumers' who receive high involvement complementary rate the product higher than those who receive low involvement complementary ( $\mu_1 > \mu_2$ )

$\mu_1$  and  $\mu_2$  are the mean of the responses under high involvement virtual complementary and low involvement virtual complementary, respectively.

The  $t$  statistics  $t_0 = 2.262$ , and the one-side  $t_{critical} = 1.833$ , so  $t_0 > t_{critical}$ , and P-value (0.025) is less than that critical value (0.05), which meet the criteria to reject the null hypothesis. Thus it can be concluded that the statistical test supports the preliminary observation and levels of virtual complementary have significant effect on the consumers' perception of low involvement product.

Again, the difference of ratings of wine and virtual glasses made by each participant is no greater than 1, as shown in Figure 43. A more accurate statistical test shows that there is no significant effect (P-Value = 0.2729) of levels of involvement on the difference of consumers' rating of the virtual functional complementary and the focal product. The consumers' rating of the focal product is highly correlated with the rating of the virtual functional complementary. In another words, different levels of virtual functional complementary has significant effect in determining consumers' perception of the focal product.

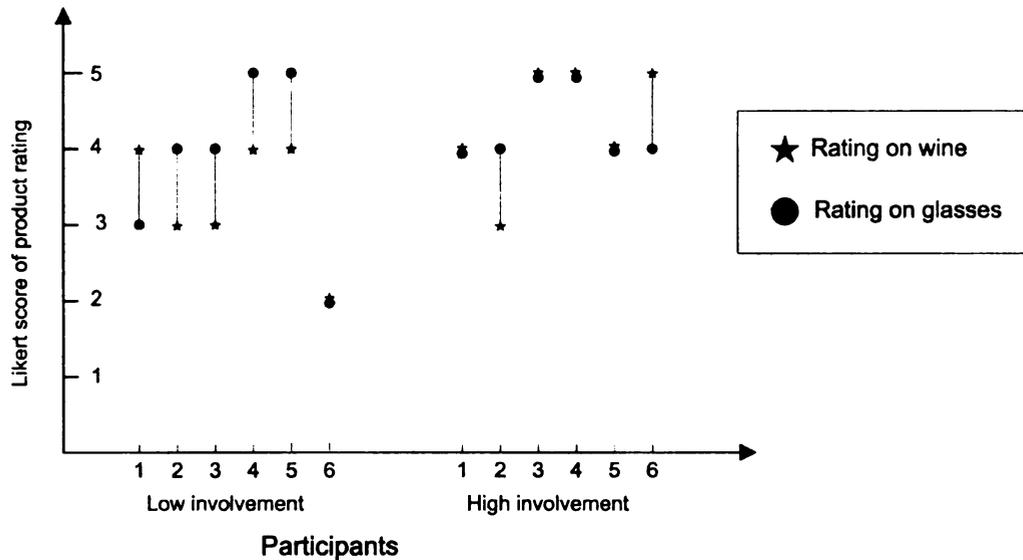


Figure 43 Participants rating on wine and glasses in pair

## 5.5. USER STUDY 4: 3D VIRTUAL CONTEXT

This study tests the effectiveness of 3D virtual context compared to 2D virtual context. A 3D virtual experience has been proven to lead to better understanding of the product and elevated purchase intent [10]. Theoretically, AR is a good medium for realizing 3D virtual experience. This study tests the effectiveness of 3D virtual experience using AR-based virtual contextualization versus 3D virtual contextualization.

### 5.5.1. EXPERIMENT DESCRIPTION

The experiment setting consists two bottles of wines. One wine is virtually contextualized with a 2D image of a wine dinner setting as shown in Figure 44(a). Another is virtually contextualized with several 3D objects in 3D to make up a wine dinner setting as shown in Figure 44(b).



(a) One wine is contextualized with a 2D image



(b) Another wine is contextualized with some 3D objects

Figure 44 Virtual context

The participants observed both wines using PromoPad and then ranked their perception of wines on a five-point Likert scale with 1 towards negative and 5 towards positive.

### **5.5.2. METHODOLOGIES**

The independent variable is the form of virtual context, 2D or 3D. The dependent variable is participants' response on the likableness of the wines, quantified on a five-point Likert scale, with 1 towards not likable, and 5 towards likable.

The study format consisted of a pre-experiment survey, utilization of PromoPad system, and a post-experiment survey.

### **5.5.3. PARTICIPANTS**

6 graduate students aged from 22 to 35 voluntarily participated in this study. All participants had no handicaps that limit their use of hands, arms, and eyes. All participants had no prior knowledge and experience of wines.

### **5.5.4. PROCEDURE**

Participants entered the lab. They were given a brief description of the experiment and a demonstration of PromoPad system by the experimenter. They were then asked to sign a consent form. Assuming they consented to the experiment, participants were then asked to complete a pre-experiment survey before they used the system. After the use of PromoPad system, the participants were asked to complete a post-experiment survey which collects the responses.

### **5.5.5. DATA ANALYSIS**

The mean of the scores on the wine with 3D virtual context is 4.33, median is 4.5, and variance is 0.6667. The mean of the scores on the wine with 2D virtual context is 3, median is 3.5, and variance is 1.6. Figure 45 shows the scores on likableness from each

participant. 5 out of 6 participants score the wine with 3D virtual context higher than the wine with 2D virtual context.

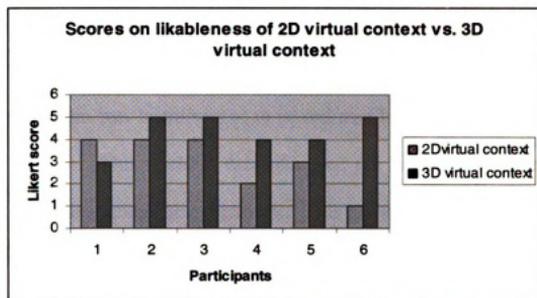


Figure 45 Scores on likableness

The box plot shown in Figure 46 illustrates the distribution of the scores. The majority of the scores on 3D virtual context are discernibly higher than the majority of the scores on 2D virtual context.

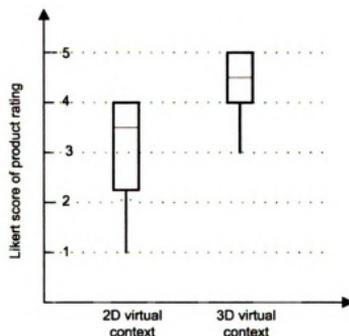


Figure 46 Box plot of likableness

From the above observation, it comes to the conclusion that 3D virtual context has a positive effect in boosting consumers' attitude towards one product. The two-sample *t*

test statistically justified this observation with a P-value equal to 0.029. This showed that 3D virtual context has statistically significant effect against 2D virtual context.

## **5.6. USER STUDY 5: USAGE PATTERN ANALYSIS**

The PromoPad system was instrumented to record user behavior in the form of tracking data relative to time. This tracking data indicated the location and orientation of the PromoPad at all times during the experiment. This allowed for analysis of users' usage behavior pattern when using PromoPad system with augmented imagery and without augmented imagery. 7 graduate and undergraduate students, aged from 18 to 35, participated this study voluntarily. They were given a brief introduction of the system by the experimenter. They were randomly chosen to use PromoPad system without augmented imagery first or use PromoPad system with augmented imagery first.

### **5.6.1. TIME PATTERN**

Some interesting observations were seen of the time that the participants spent using the system. Three timestamps were recording during the use of PromoPad system:

1. Start tracking time: this is the moment that the camera captures one of the fiducial images and starts tracking.
2. Effective in use time: this is the amount of time that the system is in use and the camera is capturing one of the fiducial images, i.e., the tracking system is in effective use.
3. Total time: this is the amount of time that the participant uses the system, from when the application starts to the application terminates.

Figure 47 shows the start tracking time when using the system with or without augmentations. The start tracking time for with augmentations is considerably smaller than that without augmentations. Augmentations give a visual clue of the successfulness of using the system and help attract the participant's attention to the focal objects.

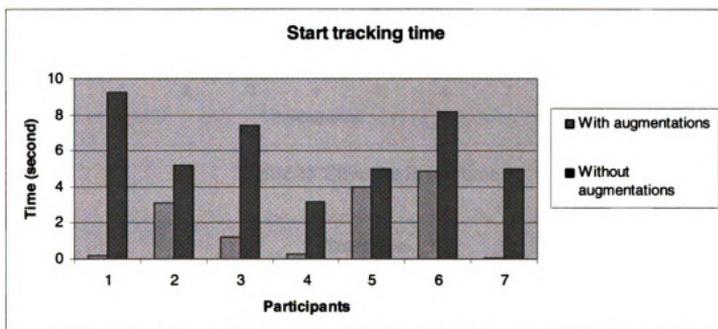


Figure 47 Start tracking time

Figure 48 shows the effective in use time with and without augmentations. Figure 49 shows the total time of using the system with or without augmentations. Both times with augmentations are about 2-4 times longer than the times spent without augmentations. Augmented imagery attracts participants' attention longer to the focal objects.

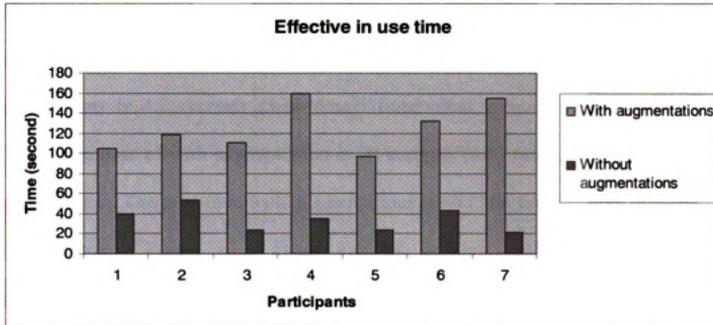


Figure 48 Effective in use time

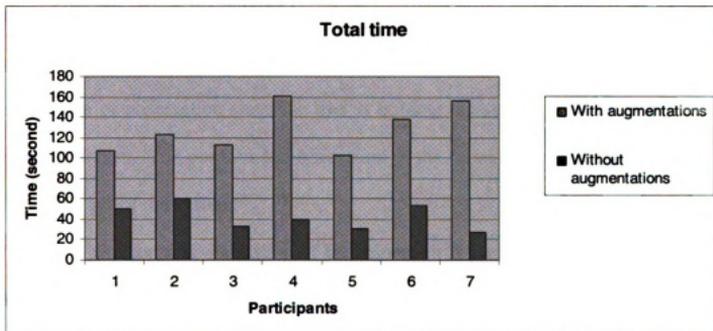


Figure 49 Total time

Table 11 Summary of time pattern

|          | With augmentations      |                           |                | Without augmentations   |                           |                |
|----------|-------------------------|---------------------------|----------------|-------------------------|---------------------------|----------------|
|          | Start tracking time (s) | Effective in use time (s) | Total time (s) | Start tracking time (s) | Effective in use time (s) | Total time (s) |
| Median   | 1.21                    | 118.34                    | 122.544        | 5.209                   | 35.419                    | 40.109         |
| Mean     | 1.949                   | 125.312                   | 128.348        | 6.174                   | 34.523                    | 42.198         |
| Variance | 4.027                   | 583.067                   | 549.834        | 4.581                   | 143.853                   | 157.43         |

Table 11 summarizes the time usage pattern between the use of the system with augmentations and without augmentations. On average, the start time with augmentations is 4.2 seconds (3.17 times) faster than without augmentations. The effective use time of with augmentation is 90 seconds (3.63 times) longer than without augmentations. The total time with augmentations is 86 (3.04 times) longer than without augmentations. Vision is a major part of learning [94]. Computer generated augmentations increases the attention span of users to focal objects and hence the users can better understand the focal objects.

### **5.6.2. MOVEMENT PATTERN**

The movement of the camera during the use of the system is actually the movement of user's points of observation since the user is holding the system and observing through the system screen. The position and orientation of the camera were recorded every 50 ms during each use of the system.

With the position and orientation of the camera, it is possible to project the points of observation onto the shelf back panel. The shelf back panel is a 40 \* 12 panel. Two focal objects were placed at (10, 0) and (30, 0), respectively. The observation points are plotted with virtual augmentations in Figure 50 and without virtual augmentations in Figure 51. In Figure 50, there are observable clusters close to the places where the focal objects were placed. In Figure 51, observation points are randomly spread out. With the use of virtual augmentations, the users' attention is more attracted to the focal objects, and hence, enhance the learning of the focal objects.

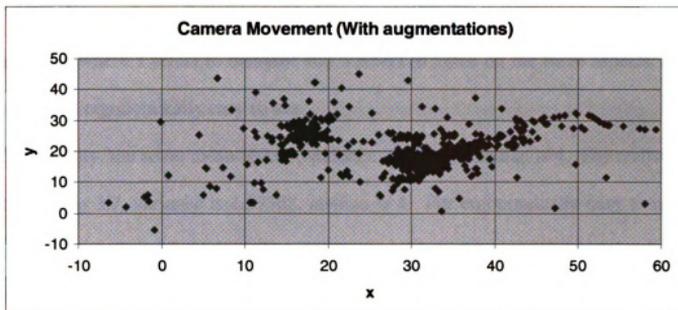


Figure 50 Camera movement on shelf back panel (with augmentations)

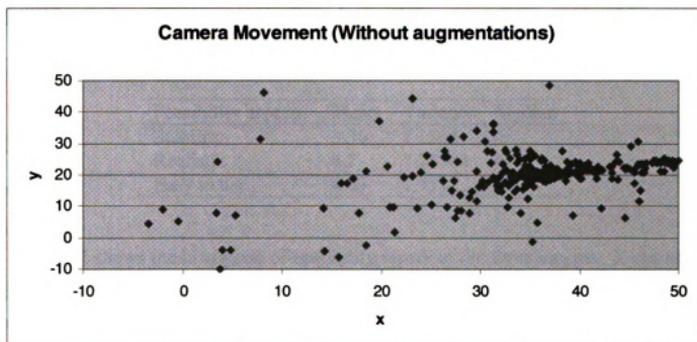


Figure 51 Camera movement on shelf back panel (without augmentations)

## 5.7. USER STUDY 5: FEASIBILITY ANALYSIS

20 graduate and undergraduate students aged from 18 to 35 voluntarily participated in this experiment to test the feasibility of using the PromoPad by a novice user. All of the participants have no prior experience with AR systems.

After using the PromoPad system, they were asked to score the system on a five-point Likert scale where 1 refers to disagree and 5 refers to agree on the three aspects: stable, realistic, and ergonomically easy to use.

For stability, the score mean is 3.65, variance is 1.1868, median is 4. For realistic, the score mean is 3.7, variance is 1.27368, median is 4. For ergonomically easy to use, the score mean is 4.05, variance is 1.3131, median is 4. All of the means are above average, and medians are at 4 out of 5, giving a very positive result. The analysis shows that novice users can master the system at first use. Table 12 lists the feasibility analysis summary.

**Table 12 Summary of feasibility analysis**

| Feasibility aspects | Mean | Variance | Median |
|---------------------|------|----------|--------|
| Stability           | 3.56 | 1.1868   | 4      |
| Realistic           | 3.7  | 1.27368  | 4      |
| Easy to use         | 4.05 | 1.3131   | 4      |

Figure 52 shows the histogram of feasibility scores of the three aspects. It can be seen that most participants are satisfied with the system performance and it is feasible to use AR by novice users. The box plot shown in Figure 53 also indicates that the majority of participants rating the feasibility aspects above average.

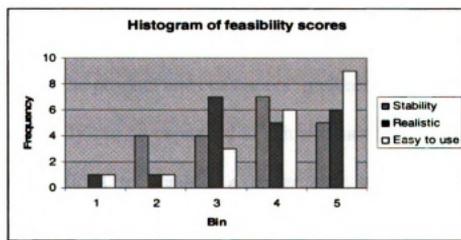


Figure 52 Histogram of feasibility scores

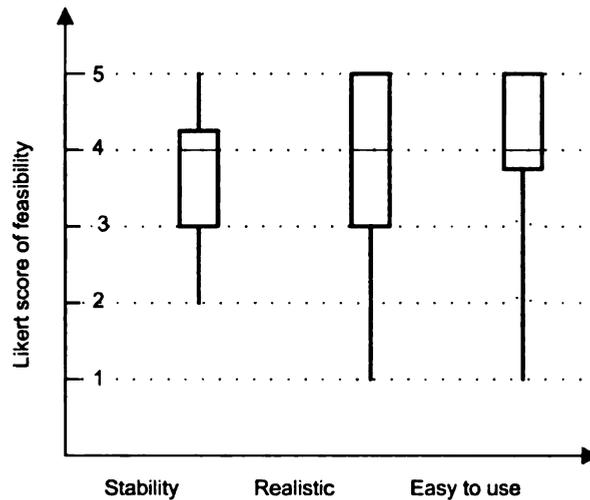


Figure 53 Box plot of feasibility scores

## 5.8. SUMMARY

This work conducts several user studies to evaluate the shopper influence that can be achieved by using augmented reality in a store setting. From the data analysis it can be concluded that AR technologies can influence consumers' perception of the products, and hence impulse purchase intent in a more flexible and dynamic manner compare to traditional in-store advertising mediums. Furthermore, since virtual imagery that promotes one product is more explicit than traditional virtual contextualization, the influence of the former is more significant than the latter. Dynamic contextualization draws attention to the complementary product and potentially increases purchase intent. The virtual promotion sign draws attention but the consumers' purchase intent depends more on the property of the product and consumers' preference.

Nevertheless, this study can be expanded to achieve more accurate analysis. First, it would be greatly beneficial to conduct a larger user study in an actual shopping environment, with participants from different ages, occupations, and education

backgrounds. This would yield a more accurate assessment of the influence of AR virtual settings and address more feasibility issues.

Second, from a statistical point of view, the larger the sample size, the more precise the test. This study includes some subjective questions to the participants. Although a pre-experiment survey was used to minimize the variability caused by participants' subjectivity, larger sample sizes will definitely help to cancel out this variability.

This work can also be extended to other application domains, where assistant information can guide or influence user's interaction with the environment for better experience, such as tourist guide, education, and learning.

## **CHAPTER 6. SUMMARY AND FUTURE WORKS**

This thesis examines the technical issues of employing context-aware computing in augmented reality, with the emphasis on dynamic contextualization using augmented reality technologies. Empirical studies demonstrate that dynamic contextualization has a statistically significant positive influence on users' perception of the focal objects, and attitude towards the focal objects.

A complete survey of the state of art of context-aware computing in conjunction with augmented reality technologies was presented. This survey listed research activities being carried out in the area of augmented reality, context-aware computing, the conjunction of these two, and AR oriented context-aware models. Nevertheless, it is still worthwhile and necessary to investigate more research effort in this area. This thesis is concerned with how context-aware computing can improve the experience of using AR systems, and how dynamic contextualization, enabled AR technologies, can influence users perception, attitude, and decision making. Clearly there are many additional AR application areas where the use of context-aware computer technologies will be a major benefit.

The work in this thesis has been focused on the specific application area of product marketing in a brick and mortar store setting. A shopping environment has a wide variety of contextual settings that can be used to test the new concept of dynamic contextualization and can be easily extended to other application domain. The design focused on how to effectively use and modify the context setting of a focal object to draw

users' attention, manipulate users' interests, and influence users' perception. The technical issues discussed in this thesis are approaches to align with the physical context, including video see-through models and inverse lighting models. The design principles are not limited in a shopping environment. They can be extended to other application domains such as education, training, gaming, instructions, and so on.

The effects of dynamic contextualization were tested in several user studies. These user studies tested the effectiveness of dynamic contextualization on influencing users' perception and decision making on virtual contextualization, object highlighting, virtual functional complementary, respectively. The experimental results and data analysis showed:

1. Augmenting context has a positive effect on influencing users' perception and purchase intent.
2. Diminished context has positive effect on highlighting a focal product.
3. Virtual functional complementary has a positive effect on influencing consumers' perception of the focal product.
4. 3D virtual context is more appealing than 2D virtual context.

Usage pattern analysis revealed some interesting observations of users' usage pattern of using AR-enabled systems as compared to non-AR system. The analysis showed that users tend to use AR-enabled systems longer than non-AR systems, and focus more on focal objects. How much of this effect may be due to novelty is yet to be determined. This usage pattern leads to a better learning experience and understanding of the focal objects [94].

The feasibility analysis showed that subjects were satisfied that using an AR-enabled system is realistic, stable, and ergonomically easy to use. Thus it appears to be feasible to employ such a system in public environment.

Future works can be conducted based on the accomplishments of this thesis. First, more participants from various education backgrounds, ages, occupations can be involved. User studies can be conducted in a real store setting. This will demonstrate a more accurate assessment. Second, the study can be made more general so that it can be easily adapted to different application domains. Third, more psychology expertise will help analysis the user behavior and thus, improve the system design. Finally, this system can be technically improved to be more stable and scalable to a real public environment setting.

## BIBLIOGRAPHY

- [1] Azuma, R.T., *A Survey of Augmented Reality*. In *Presence: Teleoperators and virtual Environments*, 1997. 6(4): p. 355-385.
- [2] Chen, G. and D. Kotz, *A Survey of Context-Aware Mobile Computing Research*. 2000, Department of Computer Science, Dartmouth College.
- [3] Korkea-aho, M., *Context-Aware Applications Survey*. 2000.
- [4] Azuma, R., et al., *Recent Advances in Augmented Reality*, in *IEEE Computer Graphics and Applications*. 2001. p. 34-47.
- [5] Mackay, W.E., *Augmenting Reality: A new paradigm for interacting with computers*. 1996: ORSAY-CEDEX, FRANCE.
- [6] Zhu, W., et al., *Personalized In-store E-Commerce with the PromoPad: an Augmented Reality Shopping Assistant*. *The Electronic Journal for E-Commerce Tools & Applications*, 2004. 1(3).
- [7] Zhu, W., et al. *Design of the PromoPad: an Automated Augmented Reality Shopping Assistant*. in *AMCIS 2006 SIGHCI minitrack on Human Cognition in Computing*. 2006. Acapulco, Mexico.
- [8] Li, H., T. Daugherty, and F. Biocca, *The role of virtual experience on consumer learning*. *Journal of Consumer Psychology*, 2002.
- [9] Li, H., T. Daugherty, and F. Biocca, *Characteristics of Virtual Experience in Electronic Commerce: A Protocol Analysis*. *Journal of Interactive Marketing*, 2001. 15(3): p. 13-30.
- [10] Li, H., T. Daugherty, and F. Biocca, *Impact of 3-D Advertising on Product Knowledge, Brand Attitude, and Purchase Intention: The Mediating Role of Presence*. *Journal of Advertising*, 2002. XXXI(3): p. 43-57.
- [11] Armata, K., *Signs that Sell*. *Progressive Grocer*, 1996. 17(21).
- [12] Baird, K.M., *EVALUATING THE EFFECTIVENESS OF AUGMENTED REALITY AND WEARABLE COMPUTING FOR A MANUFACTURING ASSEMBLY TASK*. 1999.

- [13] Milgram, P. and F. Kishino, *A Taxonomy of Mixed Reality Visual Displays*. IEICE Transactions on Information Systems, 1994. **E77-D(12)**.
- [14] Owen, C.B., et al., *Augmented imagery for digital video applications*, in *CRC Handbook of Video Databases*. 2003, CRC Press LLC.
- [15] Foley, J.D., et al., *Computer Graphics: Principles and Practice in C*. Second ed. 1995: Addison-Wesley Publishing Co.
- [16] MacIntyre, B., E.M. Coelho, and S.J. Julier. *Estimating and Adapting to Registration Errors in Augmented Reality Systems*. in *IEEE Virtual Reality Conference 2002*. 2002. Orlando, Florida.
- [17] Azuma, R., *Tracking Requirements for Augmented Reality*, in *Communications of the ACM*. 1993. p. 50-51.
- [18] Hightower, J. and G. Borriello, *A Survey and Taxonomy of Location Systems for Ubiquitous Computing*. 2001, University of Washington, Computer Science and Engineering.
- [19] Bishop, G., B.D. Allen, and G. Welch, *Tracking: Beyond 15 Minutes of Thought*. 2001.
- [20] Ipiña, D.L.d., P.R.S. Mendonça, and A. Hopper, *TRIP: A Low-Cost Vision-Based Location System for Ubiquitous Computing*, in *Personal and Ubiquitous Computing*. 2002. p. 206 - 219.
- [21] Cho, Y., J. Lee, and U. Neumann. *A Multi-ring Color Fiducial System and An Intensity-invariant Detection Method for Scalable Fiducial-Tracking Augmented Reality*. in *IEEE International Workshop on Augmented Reality*. 1998.
- [22] ARToolkit, [http://www.hitl.washington.edu/research/shared\\_space/download/](http://www.hitl.washington.edu/research/shared_space/download/).
- [23] Owen, C.B., F. Xiao, and P. Middlin. *What is the best fiducial?* in *The First IEEE International Augmented Reality Toolkit Workshop*. 2002. Darmstadt, Germany.
- [24] Tang, A., et al. *Comparative Effectiveness of Augmented Reality in Object Assembly*. in *Proceedings of ACM CHI '2002*. 2002. Darmstadt, Germany.
- [25] Tönnis, M., et al. *Experimental Evaluation of an Augmented Reality Visualization for Directing a Car Driver's Attention*. in *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*. 2005. Vienna, Austria.

- [26] Bonanni, L., C.-H. Lee, and T. Selker. *Attention-based design of augmented reality interfaces*. in *CHI '05 extended abstracts on Human factors in computing systems*. 2005. Portland, OR, USA.
- [27] Biocca, F., et al. *Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms*. in *Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI'06)*. 2006. Montréal, Québec, Canada.
- [28] Biocca, F., A. Tang, and D. Lamas. *Evolution of the mobile infosphere: iterative design of a high information-bandwidth, mobile augmented reality interface*. in *The International Conference on Augmented, Virtual Environments and Three-Dimensional Imaging, ICAV3D'2001*. 2001. Mykonos, Greece.
- [29] Schilit, B. and M. Theimer, *Disseminating Active Map Information to Mobile Hosts*, in *IEEE Network*. 1994. p. 22-32.
- [30] Schilit, B., N. Adams, and R. Want. *Context-Aware Computing Applications*. in *IEEE Workshop on Mobile Computing Systems and Applications*. 1994. Santa Cruz, CA, US.
- [31] Pascoe, J. *Adding Generic Contextual Capabilities to Wearable Computers*. in *2nd International Symposium on Wearable Computers*. 1998.
- [32] Dey, A.K. and G.D. Abowd. *Towards a Better Understanding of Context and Context-Awareness*. in *the Workshop on The What, Who, Where, When, and How of Context-Awareness, as part of the 2000 Conference on Human Factors in Computing Systems (CHI 2000)*. 2000. Hague, Netherlands.
- [33] Feiner, S., B. MacIntyre, and D. Seligmann, *Knowledge-based Augmented Reality*, in *Communications of the ACM*. 1993. p. 53-62.
- [34] Seligmann, D.D. and S. Feiner. *Automated generation of intent-based 3D Illustrations*. in *ACM SIGGRAPH Computer Graphics*. 1991. Las Vegas, Nev.
- [35] Friedrich, W. *ARVIKA – Augmented Reality for Development, Production and Service*. in *International Symposium on Mixed and Augmented Reality (ISMAR'02)*. 2002. Darmstadt, Germany.
- [36] <http://www.arvika.de/www/e/home/home.htm>, *ARVIKA Home Page*.
- [37] <http://www.boeing.com/defense-space/aerospace/training/instruct/augmented.htm>, *Boeing's instructional systems*.

- [38] SCHWALD, B., et al. *STARMATE: Using Augmented Reality technology for computer guided maintenance of complex mechanical elements*. in *e2001 Conference*. 2001. Venice, Italy.
- [39] Schwald, B. and B.d. Laval, *An Augmented Reality System for Training and Assistance to Maintenance in the Industrial Context*. *Journal of WSCG*, 2003. 11(1).
- [40] Goose, S., et al., *Speech-Enabled Augmented Reality Supporting Mobile Industrial Maintenance*, in *IEEE Pervasive Computing*. 2003. p. 65-70.
- [41] Vlahakis, V., et al., *Archeoguide: An Augmented Reality Guide for Archaeological Sites*, in *IEEE Computer Graphics and Applications*. 2002. p. 38-50.
- [42] Vlahakis, V., J. Karigiannis, and N. Ioannidis, *Augmented Reality Touring of Archaeological Sites with the ARCHEOGUIDE System*, in *Cultivate Interactive*. 2003.
- [43] Hollerer, T., et al., *Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System*. *Computers & Graphics*, 1999. 23(6): p. 779-785.
- [44] Hollerer, T., S. Feiner, and J. Pavlik. *Situated Documentaries: Embedding Multimedia Presentations in the Real World*. in *ISWC '99 (International Symposium on Wearable Computers)*. 1999. San Francisco, CA.
- [45] Bederson, B.B. *Audio Augmented Reality: A Prototype Automated Tour Guide*. in *the ACM Human Computer in Computing Systems conference (CHI'95)*. 1995.
- [46] Dieter, S. and W. Daniel. *A Handheld Augmented Reality Museum Guide*. in *Proceedings of IADIS International Conference on Mobile Learning 2005 (ML2005)*. 2005.
- [47] State, A., et al. *Case Study: Observing a Volume Rendered Fetus within a Pregnant Patient*. in *IEEE Visualization 1994*. 1994. Los Alamitos, CA.
- [48] State, A., et al. *Technologies for Augmented Reality Systems: Realizing Ultrasound-Guided Needle Biopsies*. in *ACM SIGGRAPH, Computer Graphics 1996*. 1996. New Orleans, LA.
- [49] Devernay, F., F. Mourgues, and È. Coste-Manière. *Towards Endoscopic Augmented Reality for Robotically Assisted Minimally Invasive Cardiac Surgery*.

- in *International Workshop on Medical Imaging and Augmented Reality (MIAR '01)*. 2001. Shatin, N.T., Hong Kong.
- [50] Billingham, M., H. Kato, and I. Poupyrev, *The MagicBook—Moving Seamlessly between Reality and Virtuality*, in *IEEE Computer Graphics and Applications*. 2001.
- [51] Wang, T., et al. *A Simulation and Training Systems of Robot Assisted Surgery Based on Virtual Reality*. in *International Workshop on Medical Imaging and Augmented Reality (MIAR '01)*. 2001. Shatin, N.T., Hong Kong.
- [52] Daugherty, T., H. Li, and F. Biocca, *Experiential commerce: A summary of research investigating the impact of virtual experience on consumer learning*, in *Online Consumer Psychology: Understanding and Influencing Consumer Behavior in the Virtual World*, R. Yalch, Editor. 2005.
- [53] Host, B. *The Impact on Consumer Behavior by Virtual Reality: Survey in the German Furniture Market*. in *Proceedings of the 2001 Experiential E-commerce Conference*. 2001. East Lansing, MI.
- [54] Wierzbicki, R.J. and K. Margolf, *Affordable Virtual Reality Content as a Marketing Instrument in Small and Middle Enterprises*.
- [55] <http://www.pvimage.com>, *Princeton Video Image*: Lawrenceville, New Jersey.
- [56] <http://www.dynamicdigitalvirtualreality.com/virtual-reality.html>, *Dynamic Digital Advertising*.
- [57] Benford, S. and L. Fahlen. *A Spatial Model of Interaction in Large Virtual Environments*. in *the Third European Conference on computer Supported Cooperative Work*. 1993. Milan, Italy.
- [58] Julier, S., et al. *Information Filtering for Mobile Augmented Reality*. in *International Symposium on Augmented Reality 2000*. 2000. Munich, Germany.
- [59] Seligmann, D.D. and S. Feiner. *Specifying composite illustrations with communicative goals*. in *the 2nd annual ACM SIGGRAPH symposium on User interface software and technology*. 1989. Williamsburg, Virginia.
- [60] Seligmann, D.D. and S. Feiner. *Supporting interactivity in automated 3D illustrations*. in *the 1st International Conference on Intelligent User Interfaces*. 1993. Orlando, Florida, United States.

- [61] Rohn, E., *Predicting Context Aware Computing Performance*, in *Ubiquity - An ACM IT Magazine and Forum*. 2003.
- [62] Hirsh, H., C. Basu, and B.D. Davison, *Learning to Personalize -Recognizing patterns of behavior helps systems predict your next move*, in *Communications of the ACM*. 2000. p. 102-106.
- [63] Shardanand, U. and P. Maes. *Social Information Filtering: Algorithms for Automating "Word of Mouth"*. in *1995 Conference on Human Factors in Computing Systems (CHI'95)*. 1995. Denver, Colorado.
- [64] Barkhuus, L. and A. Dey. *Is Context-Aware Computing Taking Control Away from the User? Three Levels of Interactivity Examined*. in *UBICOMP 2003, 5th International Symposium on Ubiquitous Computing*. 2003.
- [65] Rekimoto, J. and K. Nagao. *The World through the Computer: Computer Augmented Interaction with Real World Environments*. in *Symposium on User Interface Software and Technology (UIST'95)*. 1995: ACM Press.
- [66] Rauterberg, M., T. Mauch, and R. Stebler. *The Digital Playing Desk: a Case Study for Augmented Reality*. in *5th IEEE International Workshop on Robot and Human Communication*. 1996. Tsukuba, Japan.
- [67] Rosenholtz, R., et al. *Feature Congestion: A Measure of Display Clutter*. in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2005. Portland, Oregon.
- [68] Kourouthanassis, P. and G. Roussos, *Developing Consumer-Friendly Pervasive Retail Systems*, in *PERVASIVEcomputing*. 2003. p. 32-39.
- [69] Chan, W., *Project Voyager: Building an Internet Presence for People, Places, and Things*, in *Media Laboratory*. 2001, Massachusetts Institute of Technology: Cambridge, MA. p. 57.
- [70] Asthana, A., M. Cravatts, and P. Krzyzanosowski. *An indoor wireless system for personalized shopping assistance*. in *IEEE Workshop on Mobile Computing Systems and Applications*. 1994. Santa Cruz, California: IEEE Computer Society Press.
- [71] Armata, K., *Progressive Grocer*. 1996. **75**(10): p. 21.
- [72] Owen, C.B., A. Tang, and F. Xiao. *ImageTclAR: a blended script and compiled code development system for augmented reality*. in *STARS2003, The International*

*Workshop on Software Technology for Augmented Reality Systems*. 2003. Tokyo, Japan.

- [73] Mauri, C., *Card loyalty. A new emerging issue in grocery retailing*. Journal of Retailing and Consumer Services, 2003. **10**(1): p. 13025.
- [74] Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. 2001: Springer-Verlag.
- [75] Tuceryan, M., et al., *Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System*. IEEE Transactions on Visualization and Computer Graphics, 1995. **1**(3): p. 255-273.
- [76] Shapiro, L.G. and G.C. Stockman, *Computer Vision*. 1st edition (January 23, 2001) ed. 2001: Prentice Hall.
- [77] Neider, J., T. Davis, and M. Woo, *OpenGL Programming Guide*. 1994: Addison-Wesley Publishing Company.
- [78] Haro, A., M. Flickner, and I.A. Essa. *Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance*. in *Proceedings IEEE CVPR 2000*. 2000. Hilton Head Island, South Carolina.
- [79] Zhang, R., et al., *Shape from Shading: A Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999. **21**(08): p. 690-706.
- [80] Dror, R.O., E.H. Adelson, and A.S. Willsky. *Estimating surface reflectance properties from images under unknown illumination*. in *the SPIE 4299: Human Vision and Electronic Imaging IV*. 2001. San Jose, CA.
- [81] Madsen, K., H.B. Nielsen, and O. Tingleff, *METHODS FOR NON-LINEAR LEAST SQUARES PROBLEMS*. 2004, Informatics and Mathematical Modelling, Technical University of Denmark.
- [82] Frandsen, P.E., et al., *UNCONSTRAINED OPTIMIZATION*. 2004, Informatics and Mathematical Modelling, Technical University of Denmark.
- [83] Ramamoorthi, R. and P. Hanrahan. *A Signal-Processing Framework for Inverse Rendering*. in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. 2001: ACM Press.

- [84] Englis, B.G. and M.R. Solomon, *Using Consumption Constellations to Develop Integrated Marketing Communications*. Journal of Business Research, 1996. 37(3): p. 183-191.
- [85] *Queer eye for the straight guy*. 2003, Bravo Television Network.
- [86] Gibson, J.J., *The Senses Considered as Perceptual Systems*. 1966: Boston: Houghton Mifflin.
- [87] Heeter, C., *Interactivity in the Context of Designed Experience*. Journal of Interactive Advertising, 2000. 1(1).
- [88] Li, H., T. Daugherty, and F. Biocca, *The role of virtual experience on consumer learning*. Journal of Consumer Psychology, 2003. 13(4).
- [89] Boud, A.C., et al. *Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks*. in *1999 International Conference on Information Visualisation*. 1999.
- [90] Vlahakis, V., et al., *Archeoguide: An Augmented Reality Guide for Archaeological Sites*. IEEE Computer Graphics and Applications, 2002. 22(5): p. 38-50.
- [91] Montgomery, D.C., *Design and Analysis of Experiments*. 5th ed. 1997: Wiley.
- [92] Bell, D.R., R.E. Bucklin, and C. Sismeiro, *Consumer Shopping Behaviors and In-Store Expenditure Decisions*. 2000.
- [93] Kennedy, S.H. and D.R. Corkindale, *Managing the advertising Process*. 1976, Lexington, MA: Saxon House/Lexington Books.
- [94] Felder, R.M. and B.A. Soloman, *Learning Styles and Strategies*.

## Appendix A. SUMMARY OF SURVEY QUESTIONS

### A.1 PRE-EXPERIMENT SURVEY QUESTIONS

1. Are you familiar with the term “augmented reality”?  
Don't Know      1     2     3     4     5     Very Familiar
2. Have you utilized an augmented reality system before?  
 Yes                       No
3. Do you usually shop for grocery?  
Never                      1     2     3     4     5     Very often
4. Do you like spaghetti?  
Not at all                1     2     3     4     5     Very much
5. Do you like canned sauce or homemade sauce when having spaghetti?  
Canned sauce      1     2     3     4     5     Homemade sauce
6. Are you familiar with wines?  
Not familiar          1     2     3     4     5     Very familiar
7. Describe your experience with wines.  
No experience      1     2     3     4     5     Enthusiast
8. Please check the wines that you are familiar with.  
 Beringer  
  
 Yellow Tail  
  
 Almaden  
  
 Meridian  
  
 Folonari
9. Please check the digital camera models that you are familiar with  
 Olympus FE-115  
  
 Sony Cyber-shot DSC-P30  
  
 Canon Powershot SD450  
  
 Nikon Coolpix 5700

Canon EOS Digital Rebel XT SLR

Kodak EasyShare C310

Polaroid FineShot 450

10. Please indicate your knowledge of digital photography on the following scale:

No knowledge    1     2     3     4     5     Expert knowledge

11. How often do you use a digital camera with a tripod?

Never                    1     2     3     4     5     Always

12. How would you describe your knowledge in wine?

No knowledge    1     2     3     4     5     Expert knowledge

13. How familiar are you with the Wine Spectator rating system for wines?

Not familiar        1     2     3     4     5     Very familiar

## **A.2 POST-EXPERIMENT SURVEY QUESTIONS**

1. Do you think the spaghetti is consumed together with Hunt's sauce?

Not at all            1     2     3     4     5     Yes

2. How likely would you purchase Hunt's sauce if you would purchase the spaghetti?

Not at all            1     2     3     4     5     Certainly

3. How likely would you purchase the spaghetti?

Not at all            1     2     3     4     5     Certainly

4. How likely would you purchase Hunt's sauce?

Not at all            1     2     3     4     5     Certainly

5. Would you try the spaghetti recipe?

Not at all            1     2     3     4     5     Certainly

6. How likely would you purchase the Yellow Tail?

Not at all            1     2     3     4     5     Certainly

7. Do you think there are other wines than the Yellow Tail?

Not at all            1     2     3     4     5     Certainly

8. Do you think there is promotion for the Yellow Tail?

Not at all            1     2     3     4     5     Certainly



Low Quality    1     2     3     4     5     High Quality

14. On each of the following scales, how would you rate the wine that you have just seen?

Bad                    1     2     3     4     5     Good

Unlikable            1     2     3     4     5     Likable

Low Quality        1     2     3     4     5     High Quality

Undesirable        1     2     3     4     5     Desirable

Common             1     2     3     4     5     Distinctive

Worthless           1     2     3     4     5     Valuable

Inferior             1     2     3     4     5     Superior

Not consider purchasing    1     2     3     4     5     Would consider purchasing

Not likely to buy    1     2     3     4     5     Likely to buy

15. Do you agree or disagree the following statement?

The system is stable.

Disagree            1     2     3     4     5     Agree

The view is realistic.

Disagree            1     2     3     4     5     Agree

The system is ergonomically easy to use.

Disagree            1     2     3     4     5     Agree

### **A.3 SUMMARY OF DATA ANALYSIS**

This section summarizes the data analysis of most the responses to the questions in the post-experiment surveys as listed in Table 13 and Table 14. The most interesting responses were discussed elaborately in Chapter 5

**Table 13 Summary of data analysis 1**

|  | Median | Mean | Variance | P-value | significance |
|--|--------|------|----------|---------|--------------|
| <b>Question: Do you think the spaghetti is consumed together with Hunt's sauce?</b>              |        |      |          |         |              |
| With AR  | 4.5    | 4.2  | 1.067    | 0.048   | Yes          |
| No AR  | 3      | 3    | 1.33     |         |              |
| <b>Question: How likely would you purchase Hunt's sauce if you would purchase the spaghetti?</b> |        |      |          |         |              |
| With AR  | 4      | 3.4  | 1.64     | 0.043   | Yes          |
| No AR  | 2.5    | 2.4  | 1.84     |         |              |
| <b>Question: How likely would you purchase the spaghetti?</b>                                    |        |      |          |         |              |
| With AR  | 3      | 3    | 2.22     | 0.44    | No           |
| No AR  | 2.5    | 2.9  | 2.32     |         |              |
| <b>Question: How likely would you purchase the Hunt's sauce?</b>                                 |        |      |          |         |              |
| With AR  | 3      | 2.8  | 1.95     | 0.03    | Yes          |
| No AR  | 1      | 1.6  | 1.6      |         |              |
| <b>Question: How likely would you purchase the Yellow Tail?</b>                                  |        |      |          |         |              |
| With AR  | 4      | 3.4  | 2.489    | 0.39    | No           |
| No AR  | 2      | 2.2  | 1.511    |         |              |
| <b>Question: Do you think there is promotion for the Yellow Tail?</b>                            |        |      |          |         |              |
| With AR  | 4.5    | 4    | 2        | 0.03    | Yes          |
| No AR  | 2.5    | 2.7  | 2.23     |         |              |
| <b>Question: How likely would purchase the Meridian?</b>   |        |      |          |         |              |
| With AR  | 2      | 2.2  | 1.956    | 0.27    | No           |
| No AR  | 2      | 1.9  | 0.544    |         |              |

**Table 14 Summary of data analysis 2**

|   |     |       |        |       |     |
|---|-----|-------|--------|-------|-----|
| Question: On the following scale, how would you rate the tripod that you have just seen? Low Quality 1 2 3 4 5 6 High Quality         |     |       |        |       |     |
| Low involvement   | 3   | 3.167 | 1.367  | 0.379 | No  |
| High involvement  | 3   | 3.333 | 0.2667 |       |     |
| Question: On the following scale, how would you rate the digital camera that you have just seen? Low Quality 1 2 3 4 5 6 High Quality |     |       |        |       |     |
| Low involvement   | 3   | 2.83  | 0.567  | 0.026 | Yes |
| High involvement  | 4   | 3.67  | 0.267  |       |     |
| Question: On the following scale, how would you rate the wine glasses that you have just seen? Low Quality 1 2 3 4 5 6 High Quality   |     |       |        |       |     |
| Low involvement   | 4   | 3.889 | 1.367  | 0.18  | No  |
| High involvement  | 4   | 4.306 | 0.2667 |       |     |
| Question: On the following scale, how would you rate the wine that you have just seen? Low Quality 1 2 3 4 5 6 High Quality           |     |       |        |       |     |
| Low involvement   | 4.5 | 4.389 | 0.667  | 0.025 | Yes |
| High involvement  | 3.5 | 3.389 | 0.667  |       |     |