



This is to certify that the
thesis entitled
COMPARATIVE EFFECTIVENESS OF AUGMENTED
REALITY IN OBJECT ASSEMBLY

presented by

Kwok H. Tang

has been accepted towards fulfillment
of the requirements for

M.S. degree in Computer Science &
Engineering

A handwritten signature in black ink, consisting of a large, stylized 'C' followed by a series of loops and a horizontal line at the end.

Major professor

Date 12/14/01

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

COMPARATIVE EFFECTIVENESS OF AUGMENTED REALITY
IN OBJECT ASSEMBLY

By

Kwok H. Tang

A THESIS

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

Department of Computer Science and Engineering

2001

Abstract

Comparative Effectiveness of Augmented Reality in Object Assembly

By

Kwok H. Tang

With all the speculations about the instructional capabilities of Augmented Reality (AR), there has been very little empirical research studying the actual effectiveness of AR as an instructional medium. The purpose of the research reported in this thesis is to explore the effectiveness of using AR in a computer assisted assembly task. Instructions for the assembly task are displayed in user's field of view and registered onto the workspace. Instructions are presented to the user as 3D objects superimposed on real objects to explicitly demonstrate the exact assembly step. Three other instructional media are compared and contrasted with the AR system: a printed manual, computer assisted instruction using a monitor-based display, and computer assisted instruction using a head-mounted display. Initial findings show that overlaying 3D instructions on the actual work reduce the error rate of an assembly task, particularly highly correlated and sequential errors. The result suggests that a part of the mental calculation of the assembly is offloaded to the computer since the system automatically calculates the position and orientation of the assembly part and overlay and provides an appropriate visualization according to the user's view.

Keywords: Augmented reality, computer assisted instruction, human computer interaction, usability study.

® Copyright December 11, 2001 by

Arthur Tang

All Rights Reserved

Acknowledgements

This thesis research cannot be made possible without the help of many people. To begin, I would like to thank Dr. Charles Owen for his supportive assistance throughout this research project. He is an excellent advisor and editor. His advices and comments has been very valuable. Next, I would like to thank Dr. Frank Biocca for introduced me into the area of HCI in virtual environment and the fascinating world of academic research in the first place. I would not have attempted graduate school without his advices and assistance. Next, I would like to thank Dr. George Stockman for his suggestions and comments on my research project.

I would also like to thank Dr. Weimin Mou for his support on the statistical analysis. A special thanks goes to Dr. Alex Terrazas for valuable discussion and review, and most importantly for sorting out some technical submission problem at the very last moment. Also, thanks to Dr. Duncan Rowland for reviewing the thesis and Dr Thomas Muth for allowing me to recruit participants in the experiment in his class. Thanks to Kenny Lee for setting up and trying out the stimulus materials, and all the other miscellaneous laboring jobs during the experiment. And finally, thanks to all the staff members in the M.I.N.D. Lab and the MET Lab for the assistances and suggestions that seem insignificant, but very important.

TABLE OF CONTENTS

List of Figures	viii
List of Tables	ix
List of Abbreviation	x
Chapter 1: Introduction	1
1.1. Research Problem	2
1.2. Research Contributions	3
1.3. Outline of the Thesis	3
Chapter 2: Augmented Reality Systems	4
2.1. Basic Components of an Augmented Reality System	4
2.1.1. See-through Head Mounted Display	4
2.1.2. Tracking System	8
2.1.2.1 Time-frequency Measurement Tracking	8
2.1.2.2 Spatial Scan Tracking	8
2.1.2.3 Inertial Sensing	9
2.1.2.4 Mechanical Linkages Tracking	9
2.1.2.5 Direct-field Sensing	9
2.2. Calibrations in Augmented Reality	10
2.2.1. Pointer Calibration	11
2.2.2. Workspace Calibration	11
2.2.3. Display Calibration	11
2.3. Error Evaluation for Calibration	12

Chapter 3:	Overview of Manufacturing Assembly	13
	3.1. The Importance of Manual Assembly in Manufacturing	13
	3.2. Issues in Manual Assembly	14
	3.3. Using Augmented Reality for Computer Assisted Instruction	15
Chapter 4	Methodology	17
	4.1 Hypothesis	17
	4.2 Method	18
	4.3 The Assembly Task	18
	4.4 Experimental Setups	19
	4.4.1 Treatment 1: Printed Media	20
	4.4.2 Treatment 2: Computer Assisted Instruction on LCD Monitor	21
	4.4.3 Treatment 3: Computer Assisted Instruction on See-through Head-mounted Display	21
	4.4.4 Treatment 4: Augmented Reality	22
	4.5 Participants	23
	4.6 Limiting Unwanted Variables	24
	4.7 Experimental Procedure	24
	4.8 Measurements	25
Chapter 5	Results	28
	5.1 Descriptive Statistics	28
	5.2 Effect of Time of Completion on Treatment Conditions	31
	5.3 Effect of Accuracy on Treatment Conditions	32

5.3.1	Effect of Total Errors on Treatment Conditions	32
5.3.2	Effect of Dependent Error on Treatment Conditions	33
5.3.3	Effect of Independent Error on Treatment Conditions	34
5.4	Effect of NASA TLX on Treatment Conditions	34
5.5	Effect of Spatial Cognitive Ability on Performance	34
5.5.1	Effect of Spatial Cognitive Ability on Total Error	34
5.5.2	Effect of Spatial Cognitive Ability on Dependent Error ..	35
5.5.3	Effect of Spatial Cognitive Ability on Independent Error	35
5.5.4	Effect of Spatial Cognitive Ability on Time of Completion	36
Chapter 6	Discussions and Conclusions	37
6.1	Effect of Information Overlay on Performance	37
6.2	Effect of Attention Switching and Mental Transformation Offloading on Performance	38
6.3	Effect of Treatment Conditions on Mental Workload	40
6.4	Effect of Spatial Cognitive Ability on Performance	40
6.5	Effect of Dependent Error in Augmented Reality	41
6.6	Effect of Attention Tunneling in Augmented Reality	42
6.7	Conclusion	42
Appendix	Procedural Steps for the Assembly Task	44
Bibliography	54

List of Figures

- Figure 2.1. Components of a typical AR system.
- Figure 2.2. Video see-through HMD.
- Figure 2.3. Optical see-through HMD.
- Figure 2.4. Transformations between coordinate systems.
- Figure 4.1. The completed assembly task.
- Figure 4.2. Treatment condition 1: printed manual.
- Figure 4.3. Treatment condition 2: CAI on LCD.
- Figure 4.4. Treatment condition 3: CAI on HMD.
- Figure 4.5. Treatment condition 4: AR.
- Figure 5.1. Bar chart on the average time of completion in each treatment conditions.
- Figure 5.2. Bar chart on the average number of dependent error, independent error, and total error in each treatment conditions.

List of Tables

Table 4.1	NASA TLX Rating Scale Defination.
Table 4.2	Table 6.2. Combination of pair wise comparisons between the 6 rating scales.
Table 5.1.	Descriptive statistics for time of completion in each treatment conditions.
Table 5.2.	Descriptive statistics for number of error in each treatment conditions.
Table 5.3.	Average score on Spatial Cognition Test in each treatment conditions.
Table 5.4	Average score on NASA TLX Rating in each treatment conditions.
Table 5.5.	ANOVA Post Hoc comparisons of time of completion on treatment conditions.
Table 5.6.	ANOVA Post Hoc comparisons of total error on treatment condition.
Table 5.7.	ANOVA Post Hoc comparisons of dependent error on treatment conditions.
Table 5.8.	Bivariate Correlation of Spatial Cognitive Ability and Performance.
Table 5.9.	Bivariate Correlation of Spatial Cognitive Ability and Dependent Error.
Table 5.10:	Bivariate Correlation of Spatial Cognitive Ability and Independent Error.
Table 5.11:	Bivariate Correlation of Spatial Cognitive Ability and Time of Completion.
Table 6.1:	Bivariate correlation analysis of Spatial Cognitive Ability and combined sample from treatment 1, 2 and 3.

List of Abbreviation

ANOVA	Analysis of Variance
AR	Augmented Reality
CAI	Computer Assisted Instruction
CCD	Charge-coupled Device
CRT	Cathode Ray Tube
DOF	Degree of freedom
dpi	dot per inch
FOV	Field of view
HDD	Head-down Display
HMD	Head-mounted Display
HUD	Head-up Display
LCD	Liquid Crystal Display
NASA	National Aeronautics and Space Administration
NASA TLX	NASA Task Load Index
PDA	Portable Digital Assistant
SPAAM	Single Point Active Alignment Method
VR	Virtual Reality

Chapter 1

Introduction

The term Augmented Reality (AR) is used to describe systems that combine computer generated environments with real environments. This combination might include the enhancement of an image with virtual annotations, the detection and amplification of soft sounds or those outside the normal range of hearing, or the use of haptics to increase the sensitivity of touch. Unlike Virtual Reality (VR), AR enhances the real environment rather than replacing it. In a typical AR system for augmented vision, a see-through head-mounted display (HMD) is used to composite computer generated graphics with the real environment. The superimposed graphics provide additional information to the user while the user is interacting with the real environment. AR technology has many potential applications, including computer assisted instructions (CAI), industrial training, computer-aided surgery, computer visualization, engineering design, interior design and modeling, and entertainment.

The idea of overlaying a computer generated synthetic environment over a real environment through an HMD dates back to Ivan Sutherland's idea of "*the ultimate display*" in 1965 [42, 43]. However, little research was done in this area until the last decade, when the tremendous advances in real time 3D graphics rendering, display technologies, motion tracking technologies and computer processing power solved many of the technical obstacles to the creation of practical systems. In 1990, researchers at The Boeing Company started a pilot project on using AR for wire bundle assembly [10, 29].

In 1992, a research group at University of North Carolina, Chapel Hill started a research project to explore using AR in surgical settings. The project overlays 3D ultrasonic echography images on a patient to endow the doctor with “*xray vision*” to see into the patient’s body [3]. The Computer Graphics and User Interfaces Laboratory at Columbia University developed a prototype using AR to assist maintenance for a laser printer in 1993 [14]. In 1997 they also developed a system called “*The Touring Machine*”, a Mobile AR system that overlays tourist information onto the user’s view [15]. In 1998, the Massachusetts Institute of Technology Media Laboratory developed an AR application that enhances the game of billiards by calculating and overlaying strategic shots in a game [21]. In the same year, the Mixed Reality Laboratory in Tokyo, Japan developed an AR air hockey game, where two players hit a virtual puck with real mallets on a real table [32].

1.1 Research Problem

There have been many speculations about what AR can do, but there have been very few empirical research studies exploring the effectiveness of AR. Even though a number of AR prototypes and test-bed applications were developed in the last decade, they were mainly “*proof-of-concept*” applications or demonstrations. Currently there is a lack of theories and guidelines in computer-human interaction to support the design of this emerging environment. This thesis is an early attempt to study the effectiveness of using AR to create an interface that “assists the user’s memory for procedures (procedural memory) and context specific reference information (semantic memory)” [7]. This is a specific example of “*Intelligence Amplification*”, a term coined by Frederick

Brooks – to denote using the machine “*to couple the mind and the machine together with broad-band channels*” to increase human performance on specific tasks [9].

1.2 Research Contributions

The purpose of this thesis is to explore the effectiveness of using AR in a computer assisted assembly task. Information for the task is displayed in user’s view and registered in the workspace. Instructions can be presented to the user as 3D objects superimposed on real objects to demonstrate the exact direction explicitly. This research has produced 3 significant contributions in this area:

1. Support for the assertion that AR improves human performance in assembly tasks.
2. Provide theoretical basis for improved AR user interfaces.
3. Indication of some of the potential weakness of current AR systems.

1.3 Outline of the Thesis

The thesis presents the results of an empirical study of the effectiveness of an AR environment in a specific assembly task. Chapter 1 presents an introduction, and gives the general motivation to the research problem. Chapter 2 gives an overview of technologies used in AR and issues involved in the design of AR systems. Chapter 3 examines issues in manufacturing assembly, problems faced by system developers, and how AR can potentially solve some of these problems. Chapter 4 presents the methodology utilized to examine the research problem, and describes the design of the experiment. Chapter 5 presents the experimental results. Chapter 6 discusses experimental results, and presents the conclusions drawn from the experiment.

Chapter 2

Augmented Reality System

There are many methods for augmented human perception. This thesis focuses on the augmentation of human vision with a see-through head-mounted display. What AR attempts to do is to superimpose informative virtual environment over the user's field of view according to the position of the user and the direction the user looking. This chapter will explore the design issues in building an AR system.

2.1 Basic Components of an Augmented Reality System

A typical AR system consists of four components: HMD, tracking system, a computer, and software (Figure 2.1). The tracking system estimates the position and orientation of the user's head. This information is used to compute a viewpoint for graphics that will be displayed in the HMD. Tracking of the user's vision and HMD apparatus allows the system to render graphics that register with the real world as viewed through the semi-transparent display.

2.1.1. See-through Head Mounted Display

A see-through HMD is a device that combines virtual computer generated graphics with the real environment. There are two major types of see-through HMD for AR system: optical see-through and video see-through. A video see-through HMD consists of an opaque HMD and two small video cameras mounted on the outside of the HMD. Real time video streams from the two cameras are combined with computer-

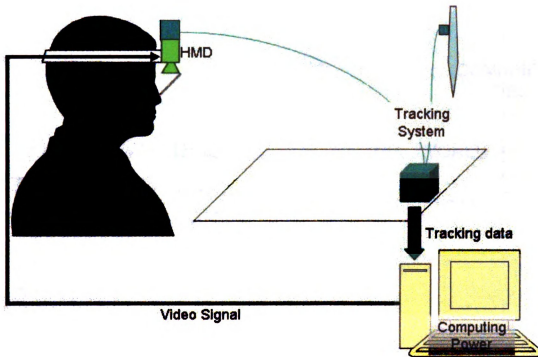


Figure 2.1. Components of a typical AR system.

generated graphics presented inside the opaque HMD (Figure 2.2). An optical see-through HMD overlays computer graphics on the visual environment using a partially transmissive half-silvered mirror (Figure 2.3). There is also a technology that uses high intensity light to paint images onto the retina of the user's eyes.

There are advantages and disadvantages to both types of displays. Video see-through displays position the two cameras as an approximation of the position of user's eyes. Consequently, the video streams being seen in the opaque HMD are displaced by the cameras position. This *eye-offset* problem can complicate tasks that require very

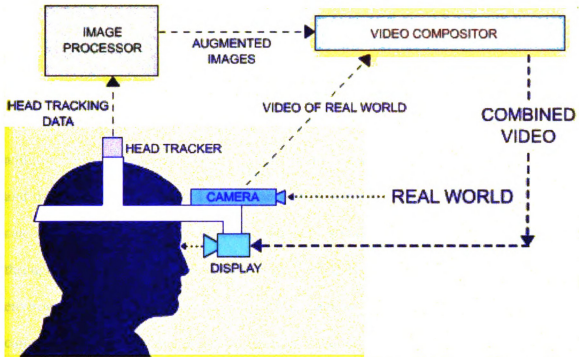


Figure 2.2. Video see-through HMD.

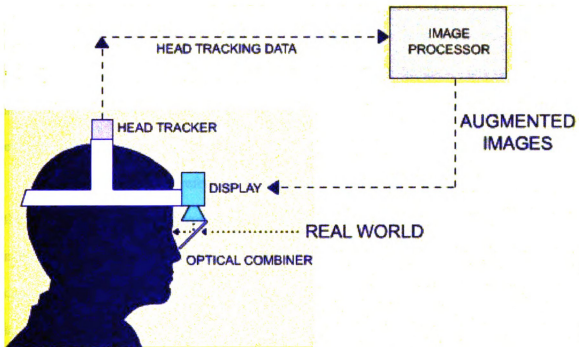


Figure 2.3. Optical see-through HMD.

accurate hand-eye coordination (e.g. Surgery) [6, 38]. This is not a problem for optical see-through displays.

Both optical and video see-through displays require rendering of the graphics in response to some method for tracking the head position and orientation in real time. In addition, video see-through displays require digitizing and re-rendering the video signals, and this usually adds at least 1/30 of a second of delay to the video stream. This *latency* can lead to unnatural hand-eye coordination or simulation sickness [36]. Video see-through displays limit the *resolution* and *field of view* for both the real and virtual environment to the resolution and FOV of the cameras and the display. With the current camera and display technologies, this limit is far inferior to the resolution of the human eye [40].

Video see-through uses video-mixing equipment to “*paint*” the virtual graphics onto the real environment [12], while optical see-through uses half-silver mirrors to optically combine the real and virtual environment. One of the disadvantages of optical see-through techniques is the real scene cannot be obscured by the virtual scene, and everything in the virtual environment looks semi-transparent.

Display calibration refers to the alignment between the virtual world displayed in the HMD and the physical world. Display calibration for video see-through HMD can be achieved using traditional camera calibration techniques. These calibration procedures can be performed once and reused. For optical see-through techniques, users are required to perform an online calibration procedure to determine the viewing parameters such as center of projection of the display, and geometric relation between the head tracker, eyes, and the display. Since these relationships vary among different users and are dependent

upon the worn position of the display, users are required to perform the calibration procedure every time before operation. Section 2.2.3 describes display calibration in more detail.

2.1.2. Tracking System

A tracking system is used in an AR environment to approximate the position on the user's head, and the direction the user is looking. According to Roland *et al.* [39], tracking technologies can be classified as (1) time-frequency measurement, (2) spatial scan, (3) inertial sensing, (4) mechanical linkages, and (5) direct-field sensing.

2.1.2.1 Time-frequency Measurement Tracking

Time-frequency measurement tracking systems measure the time and/or phase difference of pulsed signals traveling to at least 3 stationary points to determine the position and orientation of the source. Typical pulsed signals being used in time-frequency measurement tracking include ultrasonic, infrared laser-diode, and radio signals. This is, by far, the most precise measurement technique, but suffers from limitations due to occlusion and low update rate. Also ultrasonic signals are sensitive to noise from CRT (Cathode Ray Tube) sweep frequencies and disk drives, and tracker lag increases as the distances between the receivers and emitters increase.

2.1.2.2 Spatial Scan Tracking

Spatial scan tracking systems uses optical sensing devices, such as CCD (Charge-coupled Device) cameras, to scan for targets in a working volume and determine the

position and orientation. Examples of targets in the working volume include fiduciary marks, bar codes, and infrared light sources. Spatial scan tracking has a very good update rate, and could have, in principle, unlimited scalability. But these systems suffer from occlusion and optical noise.

2.1.2.3 Inertial Sensing

Inertial sensing trackers measure the change of momentum of the target to determine the position and orientation. These systems typically use sensing devices such as mechanical gyroscopes and/or accelerometers. Inertial sensing can operate without a source of reference, but suffers from accumulated error over time.

2.1.2.4 Mechanical Linkage Tracking

Mechanical linkage tracking physically links the target to a reference point. With an encoder attached to the linkages, the system uses the angular rotation measured by the encoder to determine the position and orientation of the target. Mechanical linkage tracking usually has a high accuracy and low lag, but usually with a limited working volume and range of motion.

2.1.2.5 Direct-field Sensing

Direct-field sensing trackers use magnetic field sensors to measure a static magnetic field to determine the position and orientation of the sensors. The source of the magnetic field can be generated artificially or Earth's natural magnetic field could be used. Direct-field sensing trackers are inexpensive, lightweight, compact, and can be used

without any pre-calibration. But they usually have a larger latency, smaller working volume, and suffer from magnetic interference from metallic objects such as iron and aluminum.

2.2 Calibrations in Augmented Reality

The tracking system only provides information about the position and orientation of user's head relative to the source of the tracking system. In order for the computer graphics to merge with the real world in a spatially meaningful way, a series of calibrations is required. "Calibration is the process of instantiating parameter values for 'models', which map the physical environment to internal representations, so that the computer's internal model matches the physical world" [27]. Typically, 3 calibration procedures are necessary to obtain the parameters of these geometric relations: pointer calibration, workspace calibration, and display calibration.

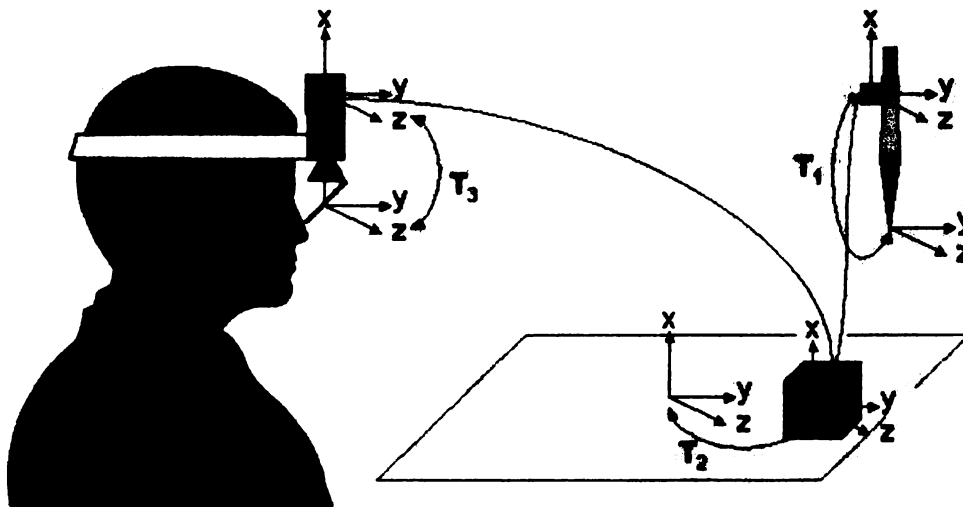


Figure 2.4. Transformations between coordinate systems.

2.2.1. Pointer Calibration

Pointer calibration determines a geometric transformation from the marker of a tracking system attached to the pointer to the tip of the pointer (Transformation T_1 in Figure 2.5). Pointer calibration is necessary because we need to pick points in the workspace to align with the virtual world in workspace calibration. The result of pointer calibration can be stored and reused as long as the marker is rigidly attached to the pointer.

2.2.2. Workspace Calibration

Workspace calibration is the alignment of the real world to the virtual world (Transformation T_2 in Figure 2.5). With a calibrated pointer, we can pick points in the real world and estimate a rigid body transformation or affine transformation to the equivalent points in the virtual world. This calibration can be stored and reused as long as the source of the tracking system remains stable relative to the workspace.

2.2.3. Display Calibration

Display calibration refers to a method to estimate the transformation that applies to the virtual object displayed on the HMD, so that the virtual object is registered with the real object (Transformation T_3 in Figure 2.5). Display calibration methods for video see-through HMD are being studied extensively in literatures on Camera Calibration in Computer Vision, such as [46], [2], and [24].

Azuma described a few methods to calibrate see-through HMD in [1]. One non-systematic calibration method for optical see-through HMD is to align a virtual object displayed in the HMD with a real object by moving the user's viewpoint until it "*looks*

correct'. This approach requires a "*skilled user*", and generally does not achieve robust results; registration becomes inaccurate when the user move away from the calibration point. Azuma also describes a more systematic method using a boresight alignment through a long pipe. Tuceryan and Navab developed an optical see-through calibration method called Single Point Active Alignment Method (SPAAM) [44]. This method uses a single point at a known location in the workspace to calibrate with crosshairs displayed in the HMD. This method is considered to be a more user-friendly method because using a single point for alignment simplified user interaction. Also, the user is not required to move the head to a fixed location and is free to move during the alignment.

2.3 Error Evaluation for Calibration

Since human performance and calibration error in AR is highly correlated, it is very important to get quantitative data of calibration error to evaluate human performance in AR. Calibration error evaluation for video see-thorough HMD can be done using traditional image-based methods [19]. For optical see-through HMD, this approach is not applicable since user's retinal images is not available. McGarrity *et al.* described an online calibration error evaluation method for see-through HMD that is capable of producing quantitative metric data [28].

Chapter 3

Overview of Manufacturing Assembly

One of the most exciting applications of AR is assembly and maintenance. In general, manufacturing processes consist of 4 series of operations: fabrication, assembly, inspection, and testing. This thesis only focuses on the assembly operation in a manufacturing process.

3.1 The Importance of Manual Assembly in Manufacturing

While many assembly operations are automated, there are still a significant number of assembly operations that cannot be done using automation and require a human assembler. Automated assembly is good for assembly tasks that have a well-defined location for acquiring and inserting parts, and for mass production manufacturing processes. For certain assembly processes, “people are good at assembly in spite of their lack of certain abilities. People use vision or, for occluded objects, special aptitude to get within range of an assembly task. They then use tactile sensing in coordination with movement to achieve the task” [37]. Also, in a market where customers are constantly changing what they want or for products that are highly customized, the cost for redesigning the automated processes can become substantial.

Manual assembly is typically used in manufacturing processes where automation is not cost-effective, products are highly customized, or processes cannot be done by automatic machineries (e.g. high quality soldering, parts that are fragile to machineries). Example products of these kinds of processes include aircraft, mainframe computers,

military equipment, rapid prototypes, medical devices, and National Aeronautics and Space Administration (NASA) contract works.

In the early 1990s, a new manufacturing conceptual framework, agile manufacturing, began to be employed widely. Agile manufacturing is a manufacturing operation that has the flexibility to change the manufacturing process quickly and efficiently to match rapid changes in market demands. Agile manufacturing has resulted in mass customization in small quantities of highly specialized products. It usually relies heavily on manual operations for flexibility.

3.2 Issues in Manual Assembly

One of the main problems in manual assembly is that expert assemblers are hard to train, particularly for assembling processes that requires problem solving skills. It usually takes months or even years for a novice assembler to develop expert knowledge for assembling processes that have high complexity. In some cases, even the experts need to refer to the instructional manual for procedures with high complexity, or procedures that are rarely performed.

In agile manufacturing, assemblers face the challenge of a continuously changing assembly process. It is impractical to retrain assemblers every time the assembly processes are changed. Assemblers need to be cross-trained to different assembly tasks so they have a deeper understanding of the process as a whole, and this trainings usually needs to be done on the job.

3.3 Augmented Reality for Computer Assisted Instruction

CAI is typically used in complex assembly tasks that involve a huge set of assembly instructions, so the assembler can pull out the appropriate instructions online when needed. However, the limited sensorimotor bandwidth (the amount of information flow between the human user and the computer) of current interfaces of computer and portable digital assistants (PDA) make them inadequate for hands-free operation and continuous data access with high interface-user information transfer rates. The limitation of sensorimotor bandwidth of modern computer interfaces (i.e., small screens, limited input/output options, etc.) makes it hard for the powerful multimedia computer to utilize its capabilities [5, 7].

In this research project, we present an AR system designed to guide and train assembly workers for assembly tasks of large complexity. This approach is very different from the traditional printed manual or online CAI approaches. In an augmented reality environment, 3D synthesized computer graphics are overlaid in the user's field of view. A study conducted by Haines, *et al.* [17] indicated pilots that use Head-up Display (HUD) have less head and eye movement when comparing with pilots that use Head-down Display (HDD) in the cockpit panels. By reducing head and eye movement and increasing "*eye-on-the-workspace*" time, user performance is expected to increase. By overlaying equivalent information on the work pieces in a spatially meaningful way, time for information searching in the instructional medium (e.g. printed manual, handheld display, machine display panel) is reduced.

By "seaming" the information to the real environment, AR technologies could be used "as a complement of human cognitive processes" [31]. Using AR as an instructional

medium can reduce the overhead of attention switching between the instructional media and the task. AR systems can also be used to augment human attention. Synthesized computer graphics are merged with the user's view, so attention can be caught by arrows, tags, highlighting the object with wire-frame, playing 3D animations, etc. Invisible objects that are blocked from view can also be indicated.

AR technologies can also facilitate on-the-job training. Human beings tend to memorize information better when they are docked to a space at the frame of reference of the real world. Demosthenes, a Greek orator born around 384 B.C., used a strategy, known as the Method of Loci, to memorize long speeches by mentally walking through one's house, associating each item of the speech to different spots or different objects in the house. In the field of neuroscience, there have been a number of theories suggesting that there is a strong relation between spatial location and working memory [33, 34]. Kirsh argued that "methods used to manage our space are key to organization of our thought patterns and behavior" [22]. By spatially relating pieces of information to physical objects and locations in the real world, AR provides a strong leverage of spatial cognition and memory [8].

Chapter 4

Methodology

This thesis hypothesizes that using AR in CAI expands human capability to absorb and process information. This chapter expands the hypotheses in detail, and explains the methodology used to investigate these hypotheses.

4.1 Hypothesis

Based on the discussion in section 3.3, the following hypotheses were generated:

- H1: Overlaying information in the user's view using a see-through HMD improves user performance on an assembly task by reducing head and eye movement.
- H2: Overlaying information in the user's field of view using AR in a spatially meaningful way improves user performance on the assembly task by reducing attention switching between the instructional media and the workspace.
- H3: By offloading the mental transformation tasks to the computer, subjects using AR instructions will perform better when comparing with subjects using traditional instructional media, where pictorial instructions need to be mentally transformed to the subjects' point of view.
- H4: Mental workload for the assembly task using traditional instructional media is higher than using AR instruction.
- H5: Individuals with better spatial ability will perform better in an assembly task based on traditional pictorial-based instruction.

4.2 Method

To test the hypotheses, an experiment was employed to compare the effectiveness of 4 different instructional media for an assembly task: a printed manual (treatment 1), CAI on a Liquid Crystal Display (LCD) monitor (treatment 2), CAI on a see-through HMD (treatment 3), and AR (treatment 4). The experiment uses a *between subject* design among the 4 treatment conditions. Subjects are required to complete the experimental assembly task according to the procedural instructions presented using the specific media as per the treatment condition. An assembly task made up of Duplo® is used in the experiment to minimize bias towards a population with expertise in a certain knowledge related to the assembly task.

4.3 The Assembly Task

The assembly task consists of 56 procedural steps. For each procedural step, subjects are required to acquire a part of a specific color and size from an unsorted part-bin and insert the part onto the current subassembly in a specific position and orientation according to the presented instruction. The assembly task is 3 dimensional in nature; some procedural steps subjects are required to put a part on top of parts that was previously inserted. Some of the procedural steps are correlated, so a mistake made in a previous step could potentially generate additional mistakes in the later steps. Figure 4.1 shows the completed assembly. The 56 procedural steps are shown in Appendix A.

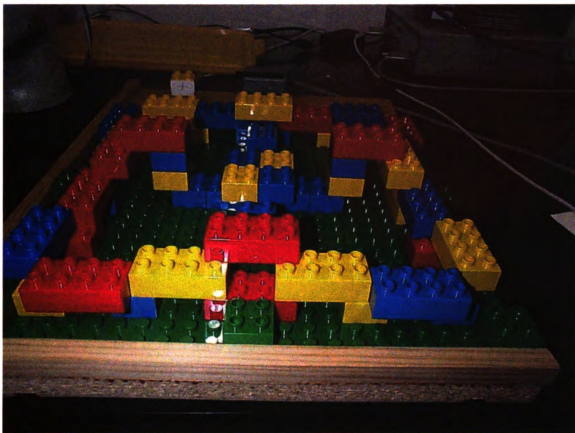


Figure 4.1. The completed assembly task.

4.4 Experimental Setups

Instructions for all 4 treatment conditions use pictorial representation, without any language. Appendix A shows the 56 procedural steps of the assembly task. The display resolution of all 4 treatments is set to 640 x 480 pixels, using 16-bit color. The graphics used in all 4 treatments are rendered using the ImageTclAR Toolkit developed by The Media and Entertainment Technologies Laboratory at Michigan State University [35]. In order to facilitate hands-free task engaged operation, subjects in the treatment 2 (CAI on LCD), treatment 3 (CAI on HMD), and treatment 4 (AR) used voice command to control the instructions. The voice command “*next*” prompts the instruction to the next

procedural step, while the voice command “*previous*” prompts the instruction to the previous procedural step. A human agent is used to interpret the voice command and control the instruction accordingly to ensure maximum accuracy on the voice recognition task. An audio signal is played to the user as a conformation of the voice command.

4.4.1 Treatment 1: Printed Media

The printed media is produced using a color solid ink printer with the resolution of 1200 dot per inch (dpi). The instructions are printed single sided, with one procedural step per page (Figure 4.2). Subjects are free to move the manual to anywhere in the workspace, or hold it in their hand during operation.

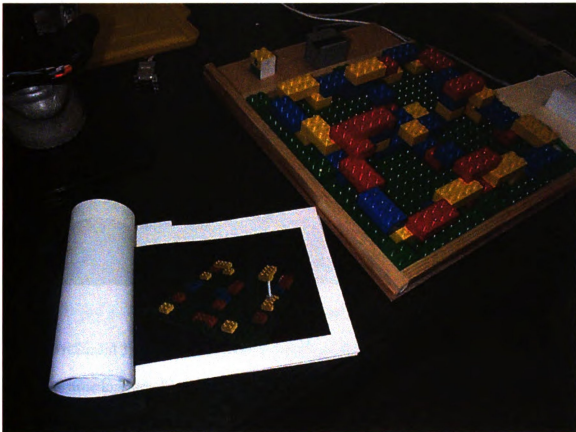


Figure 4.2. Treatment condition 1: printed manual.

4.4.2 Treatment 2: Computer Assisted Instruction on LCD monitor

Instructions are displayed on a laptop computer placed on the workspace (Figure 4.3). The size of the LCD monitor is 15-inch (diagonal), and the native resolution of the screen is 1400 x 1050 pixels. The pictorial instructions were displayed in full screen. Before the start of the experiment, subjects are free to adjust the brightness and orientation of the screen.

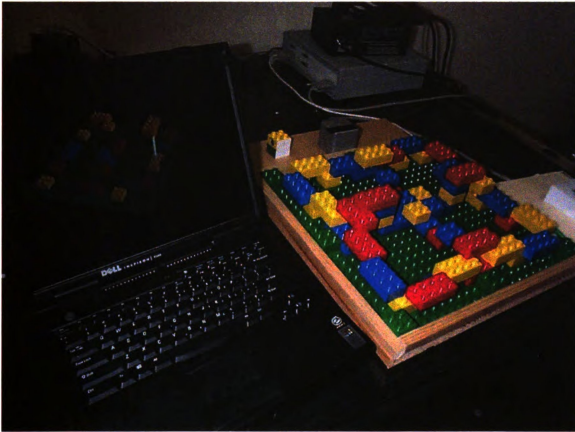


Figure 4.3. Treatment condition 2: CAI on LCD.

4.4.3 Treatment 3: Computer Assisted Instruction on See-through Head-mounted Display

Instructions are displayed on a see-through HMD. The see-through HMD used in the experiment is the Sony Glasstron LDI-100B (Figure 4.4). It has a native resolution of 832 x 624 pixels and a simulated 30 inches (diagonal) screen at 4 feet ahead.

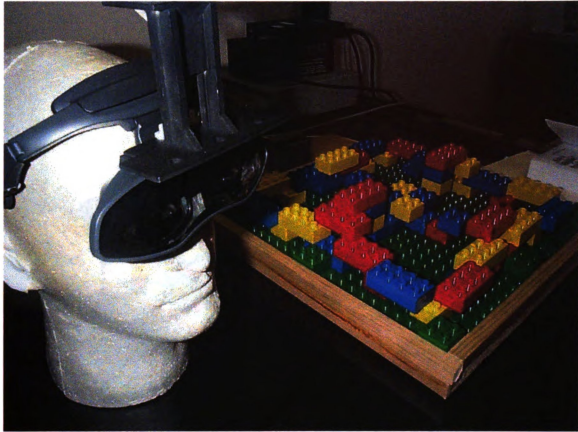


Figure 4.4. Treatment condition 3: CAI on HMD.

4.4.4 Treatment 4: Augmented Reality

Instructions are displayed in stereo using the Sony Glasstron LDI-100B. Head motion of the subjects are tracked using the Polhemus Fastrak® 6 DOF magnetic tracker. Stereo graphics are rendered in real time based on the data from the magnetic tracker, using a computer with dual Intel® Pentium® III Xeon™ 800 MHz processors, 512MB RDRAM® and a 3Dlabs Wildcat II 5110 graphic accelerator, running under Microsoft® Windows® 2000 Professional. The program is written using the ImageTclAR Toolkit [35]. The Toolkit uses a variation of the SPAAM algorithm for stereo display calibration. The calibration procedure will be described in Section 4.7.



Figure 4.5. Treatment condition 4: AR.

4.5 Participants

75 subjects were recruited in an introductory undergraduate class at a large midwestern university in the United States who volunteered to participate in the study for class credit. Subjects were from a general college student population with majors ranging from Information Technologies and Law and Policy, to Business Management and Media Arts. None of the subjects had previous experience in any AR environment. Subjects were randomly assigned to each treatment condition. The number of males and females was arranged to be distributed evenly among different treatments to control a possible gender effect to the experiment.

4.6 Limiting Unwanted Variables

In treatment conditions 3 and 4, instructions are presented to the subjects through a see-through HMD. Light from the real world will be attenuated and distorted by the half-silver mirror when entering the see-through HMD. The subjects' FOV is limited by the HMD (Horizontal FOV is about 28 degree for the Sony Glasstron HMD). And people generally feel uncomfortable with a load (the Sony Glasstron HMD weights about 120g) on ones head. These are factors that count as disadvantages to performance in treatment conditions 3 and 4. To eliminate these factors from the experiment, subjects in all treatments are required to wear the HMD during operation so that these variables remain constant among different treatment.

In treatment condition 4, subjects are required to perform a display calibration and error evaluation procedure that takes 8-12 minutes. This procedure generally is considered to be challenging for an untrained user, and can potentially induce fatigue and mental workload factors to the assembly task that affect subjects' performance. To eliminate these factors from the experiment, subjects in all treatment were required to perform the display calibration and error evaluation procedure so that these variables remain constant among different treatment.

4.7 Experimental Procedure

The experiment began after the participants read and signed a consent form indicating their voluntary participation in the experiment. The participants were first briefed about the whole experimental procedure. After that, they were instructed about

the display calibration procedure. The display calibration procedure involves aligning 9 crosshairs for each eye presented in the HMD sequentially (18 crosshairs total) to a crosshair located in the middle on the workspace. After completing the display calibration procedure, the experimenter explained the graphical metaphors used in the instructions, and in treatment 2, 3, and 4, the voice command used to control the instructions. They then entered the pretest environment and performed the training assembly task. Errors made in the pretest by the participants were explained after the participants finished the pretest, and participants were asked if they feel comfortable in performing the assembly, and if they want to repeat the pretest to get more familiar with the environment. When participants felt comfortable with the pretest environment, they were allowed to proceed to the main test environment. Participants were asked to perform the task in the main experiment as fast and as accurate as possible, and any question the subjects had were answered at that time. The participants then completed the assembly task. Immediately after the experiment, participants completed the post-test questionnaires, which includes the NASA TLX rating, demographic information, and the spatial ability test. After the participants completed the questionnaires, they were thanked and debriefed.

4.8 Measurements

Performance: Performance of the subject is defined as time of completion and the accuracy of the assembly task. Accuracy is measured in number of errors the subject made in the assembly task, where error is defined in a particular assembly step as: (1) a part is inserted at the wrong location, (2) a part is inserted with the wrong orientation, (3)

a part with the wrong color is inserted, (4) a part with the wrong size is insert, (5) a part is missing, and (6) an extra part is inserted.

Spatial ability: Spatial ability of subjects was measured using the mental rotation test [13]. The test includes two timed tests (3 minutes each) assessing 3D rotation of drawings of 42 pairs of cubes. 3 sides of each cube are visible, and the subject is to mentally rotate one or both cubes to determine if they are the same.

Mental Workload: Subjective measurement of mental workload on the assembly task of the subjects is collected using the NASA Task Load Index (NASA TLX) [18]. Subjects rate each of the 6 categories as shown in Table 6.1 based on their experience on the assembly task, using a 20 point scale. And then they were asked to do a pair wise comparison about which category is more important correspond to the assembly task among the 15 combinations as shown in Table 6.2. A mean weighted workload score can then be calculated by adding up on the rating multiplied by its respective weighting for each category.

Mental Demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculation, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Frustration Level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Table 6.1. NASA TLX Rating Scale Definition.

Mental demand vs. Physical demand	Mental demand vs. Temporal demand	Mental demand vs. Effort	Mental demand vs. Frustration level	Mental demand vs. Performance
	Physical demand vs. Temporal demand	Physical demand vs. Effort	Physical demand vs. Performance	Physical demand vs. Frustration level
		Temporal demand vs. Effort	Temporal demand vs. Performance	Temporal demand vs. Frustration level
			Effort vs. Performance	Effort vs. Frustration level
				Performance vs. Frustration level

Table 6.2. Combination of pair wise comparisons between the 6 rating scales.

Chapter 5

Results

A total of 75 subjects participated in the experiment, 18 in treatment condition 2, 19 in each of treatment conditions 1, 3 and 4. The average age of the participants is 20.63. 21 (28%) of the participants are female, and 54 (72%) are male. An alpha level of .05 (2-tailed) was used for all statistical tests.

5.1 Descriptive Statistics

Table 5.1 and Figure 5.1 illustrate the mean time for completing the assembly task in seconds. They demonstrate that treatment 4 (AR) has the shortest time of completion among the 4 treatment conditions, while treatment 1 (printed manual) has the longest time of completion.

Treatment Condition	N	Mean (seconds)	Median (seconds)	Std. Dev.
1: Printed Manual	19	864	847	289.61
2: CAI on LCD Display	18	686	716	158.29
3: CAI on HMD	19	668	687	211.74
4: AR	19	651	609	174.31

Table 5.1. Descriptive statistics for time of completion in each treatment conditions.

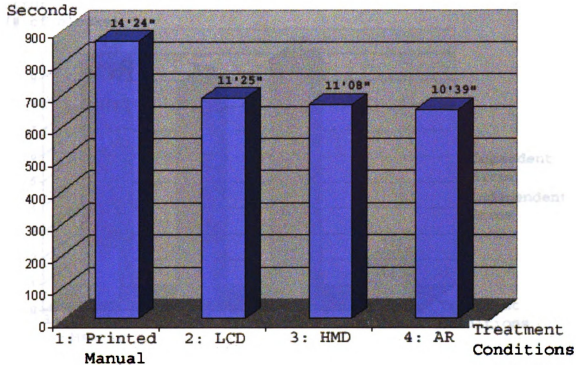


Figure 5.1. Bar chart on the average time of completion in each treatment conditions.

Table 5.2 and Figure 5.2 shows the average number of errors for the assembly task in number of steps. The total number of steps of the assembly task is 56. Two classes of errors are defined: dependent error and independent error. Dependent error is an error that is related to another error made previously in the assembly sequence. Independent error is an isolated error that does not related to a previous step. The statistics show that treatment condition 4 (AR) has significantly lower error rates in all categories. It also shows that a majority of errors in treatment 4 are independent errors, whereas treatment 1, 2 and 3 exhibit a majority proportion of dependent errors.

Treatment Condition	Average total error (# of steps)	Average dependent error (# of steps)	Average independent error (# of steps)
1: Printed Manual	9.37	7.21	2.16
2: CAI on LCD Display	8.44	6.17	2.28
3: CAI on HMD	9.50	7.11	2.39
4: AR	1.63	0.21	1.42

Table 5.2. Descriptive statistics for number of error in each treatment conditions.

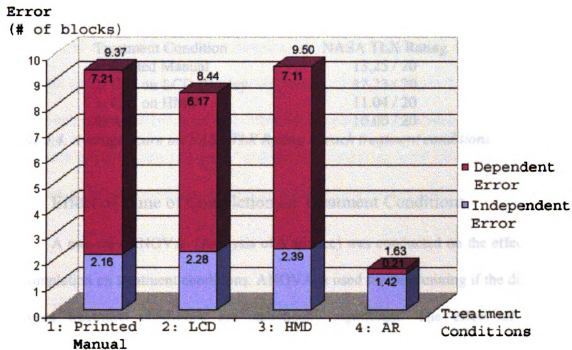


Figure 5.2. Bar chart on the average number of dependent error, independent error, and total error in each treatment conditions.

Table 5.3 shows the mean score of the spatial cognition test and the NASA TLX rating. The statistics show that subjects in treatment 1 have the highest mental workload, where subjects in treatment condition 4 have the lowest mental workload. It also shows that subjects among 4 treatment conditions have about the same mean in spatial cognition abilities.

Treatment Condition	Spatial Cognition Test
1: Printed Manual	26.95 / 42
2: CAI on LCD Display	28.22 / 42
3: CAI on HMD	26.00 / 42
4: AR	28.11 / 42

Table 5.3. Average score on Spatial Cognition Test in each treatment conditions.

Treatment Condition	NASA TLX Rating
1: Printed Manual	13.25 / 20
2: CAI on LCD Display	12.23 / 20
3: CAI on HMD	11.04 / 20
4: AR	10.00 / 20

Table 5.4. Average score on NASA TLX Rating in each treatment conditions.

5.2 Effect of Time of Completion on Treatment Conditions

A one-way ANOVA (Analysis of Variance) was conducted on the effect of time of completion on treatment conditions. ANOVA is used for determining if the differences between treatment conditions are statistically significant. The effect of time of completion depending on treatment conditions is statistically significant, $F(3, 71) = 3.75$, $p = .015$. Post Hoc comparisons were further conducted using the Bonferroni Method to obtain all possible pair wise comparisons among treatment conditions. The results are shown in Table 5.5.

(I) Setting	(J) Setting	Mean Difference (I-J)	Std. Error	Sig.
1	2	178.03	70.80	.085
	3	173.37	69.84	.092
	4	212.95	69.84	.019
2	1	-178.03	70.80	.085
	3	-4.66	70.80	1.000
	4	34.92	70.80	1.000
3	1	-173.37	69.84	.092
	2	4.66	70.80	1.000
	4	39.58	69.84	1.000
4	1	-212.95	69.84	.019
	2	-34.92	70.80	1.000
	3	-39.58	69.84	1.000

Table 5.5. ANOVA Post Hoc comparisons of time of completion on treatment conditions.

The analysis shows that there is a statistically significant effect between treatment conditions 1 and 4 ($p = .019$). The effect between treatment conditions 1 and 2 and

treatment conditions 1 and 3 trends toward significance ($p = .085$ and $.092$ respectively). But there is no significant effect between treatment conditions 2 and 3 ($p = 1.000$), treatment conditions 2 and 4 ($p = 1.000$), and treatment conditions 3 and 4 ($p = 1.000$). The results of the ANOVA analyses show that treatment conditions 2, 3 and 4 have a significant improvement in time of completion comparing with treatment condition 4. However, there is no statistically significant effect between treatment conditions 2, 3 and 4.

5.3 Effect of Accuracy on Treatment Conditions

5.3.1 Effect of Total Errors on Treatment Conditions

A one-way ANOVA was conducted on the effect of total error rate on treatment conditions. The effect of total error depending on treatment conditions is statistically significant, $F(3, 71) = 4.41$, $p = .007$. Post-Hoc Comparisons were further conducted using the Bonferroni Method to obtain all possible pair wise comparisons among treatment conditions. The results are shown in Table 5.6.

(I) Setting	(J) Setting	Mean Difference (I-J)	Std. Error	Sig.
1	2	.92	2.65	1.000
	3	-.68	2.61	1.000
	4	7.74	2.61	.025
2	1	-.92	2.65	1.000
	3	-1.61	2.65	1.000
	4	6.81	2.65	.073
3	1	.68	2.61	1.000
	2	1.61	2.65	1.000
	4	8.42	2.61	.012
4	1	-7.74	2.61	.025
	2	-6.81	2.65	.073
	3	-8.42	2.61	.012

Table 5.6. ANOVA Post Hoc comparisons of total error on treatment condition.

The analysis shows that there are statistical significant effects between treatment conditions 1 and 4 ($p = .019$) and conditions 3 and 4 ($p = .012$). The effect between treatment conditions 2 and 4 trends toward significance ($p = .073$). But there is no significant effect between treatment conditions 1 and 2 ($p = 1.000$), treatment conditions 1 and 3 ($p = 1.000$), and treatment conditions 2 and 3 ($p = 1.000$). The results of the ANOVA analyses show that treatment conditions 4 has a significant improvement in total error comparing with treatment condition 1, 2 and 3. However, there is no statistically significant effect between treatment conditions 1, 2 and 3.

5.3.2 Effect of Dependent Error on Treatment Conditions

A one-way ANOVA was conducted on the effect of error rates of dependent error on treatment conditions. The effect of dependent error depending on treatment conditions is not statistically significant, $F(3, 71) = 4.68$, $p = .005$. Post-Hoc Comparisons were further conducted using the Bonferroni Method to obtain all possible pair wise comparisons among treatment conditions. The results are shown in Table 5.7.

(I) Setting	(J) Setting	Mean Difference (I-J)	Std. Error	Sig.
1	2	1.04	2.30	1.000
	3	-.53	2.27	1.000
	4	7.00	2.27	.017
2	1	-1.04	2.30	1.000
	3	-1.57	2.30	1.000
	4	5.96	2.30	.070
3	1	.53	2.27	1.000
	2	1.57	2.30	1.000
	4	7.53	2.27	.009
4	1	-7.00	2.27	.017
	2	-5.96	2.30	.070
	3	-7.53	2.27	.009

Table 5.7. ANOVA Post Hoc comparisons of dependent error on treatment conditions.

The analysis shows that there are statistical significant effects between treatment conditions 1 and 4 ($p = .017$) and conditions 3 and 4 ($p = .009$). The effect between treatment conditions 2 and 4 trends toward significance ($p = .070$). But there is no significant effect between treatment conditions 1 and 2 ($p = 1.000$), treatment conditions 1 and 3 ($p = 1.000$), and treatment conditions 2 and 3 ($p = 1.000$). The results of the ANOVA analyses show that treatment conditions 4 has a significant improvement in dependent error comparing with treatment condition 1, 2 and 3. However, there is no statistically significant effect between treatment conditions 1, 2 and 3.

5.3.3 Effect of Independent Error on Treatment Conditions

A one-way ANOVA was conducted on the effect of error rates of independent error on treatment conditions. The effect of independent error depending on treatment conditions is not statistically significant, $F(3, 71) = .967, p = .413$.

5.4 Effect of NASA TLX on Treatment Conditions

A one-way ANOVA was conducted on the effect of the NASA TLX rating on treatment conditions. The effect of NASA TLX depending on treatment conditions is statistically significant, $F(3, 71) = 6.26, p = .001$.

5.5 Effect of Spatial Cognitive Ability on Performance

5.5.1 Effect of Spatial Cognitive Ability on Total Error

A bivariate correlation analysis was conducted on the effect of spatial cognitive ability on total error. The results are as shown in Table 5.8.

Treatment Condition(s)	Pearson Correlation	Sig.
1	-.121	.621
2	-.350	.155
3	-.182	.457
4	-.270	.263

Table 5.8. Bivariate Correlation of Spatial Cognitive Ability and Performance.

The results show that there is no statistical significant correlation between spatial cognitive ability and total error in all treatment conditions.

5.5.2 Effect of Spatial Cognitive Ability on Dependent Error

A bivariate correlation analysis was conducted on the effect of spatial cognitive ability on dependent error. The results are as shown in Table 5.9.

Treatment Condition(s)	Pearson Correlation	Sig.
1	-.051	.835
2	-.409	.092
3	-.101	.682
4	.198	.417

Table 5.9. Bivariate Correlation of Spatial Cognitive Ability and Dependent Error.

The results show that there is no statistical significant correlation between spatial cognitive ability and dependent error in all treatment conditions.

5.5.3 Effect of Spatial Cognitive Ability on Independent Error

A bivariate correlation analysis was conducted on the effect of spatial cognitive ability on independent error. The results are as shown in Table 5.10.

Treatment Condition(s)	Pearson Correlation	Sig.
1	-.369	.120
2	-.001	.998
3	-.394	.095
4	.198	.417

Table 5.10. Bivariate Correlation of Spatial Cognitive Ability and Independent Error.

The results show that there is no statistical significant correlation between spatial cognitive ability and independent error in all treatment conditions.

5.5.4 Effect of Spatial Cognitive Ability on Time of Completion

A bivariate correlation analysis was conducted on the effect of spatial cognitive ability on time of completion. The results are as shown in Table 5.11.

Treatment Condition(s)	Pearson Correlation	Sig.
1	1.000	.263
2	-.493	.038
3	-.182	.457
4	.198	.417

Table 5.11. Bivariate Correlation of Spatial Cognitive Ability and Time of Completion.

The results show that there is no statistical significant correlation between spatial cognitive ability and time of completion in treatment conditions 1, 3 and 4, and there is a statistically significant effect in treatment conditions 2.

Chapter 6

Discussions and Conclusions

This chapter explores the experimental findings in relationship to the stated hypotheses. It investigates the implications of the results to the theoretical model, and provides further insight into the influence of AR in human performance and perception.

6.1 Effect of Information Overlay on Performance

Hypothesis 1 states that overlaying information in user's view using a see-through HMD improves user's performance on the assembly task by reducing head and eye movement. This hypothesis suggests that the performance of subjects, in terms of time of completion and accuracy, in treatment conditions 3 and 4 is expected to be better than treatment conditions 1 and 2. Even though there are statistical significant advantages in time of completion and accuracy in condition 4 comparing with conditions 1 and 2, there is no significant advantage in time of completion in condition 3 comparing with conditions 2, and no significant advantage in accuracy in condition 3 comparing with conditions 1 and 2. Therefore, this hypothesis is not supported.

In treatment conditions 1, 2 and 3, it is a common practice that the subjects count the number of bumps from the edge of the Duplo® base plate to determine the exact position of the part to be inserted. Some subjects in treatment condition 3 also reported that it is hard to perform counting on the instructions since they cannot touch the instructions physically. Some of the responses from subjects in treatment condition 3 stated that the overlaid instructions interfered with the workspace and it was hard to see

the workspace clearly. Other stated that the workspace was interfering with the overlaid instructions and it was hard to read the instructions clearly. Studies of HUDs for automobile drivers suggested that symbology placed within a 5 degree radius of the fovea is annoying to drivers [20, 41].

The Sony Glasstron HMD projects a simulated 30 inches (diagonal) screen at 4 feet ahead. The distance between the subject's head and the top of the workbench is approximately 1.5 feet. So the projected image in the HMD appears to be under the table. Some of the subjects in treatment condition 3 reported that it is hard to adjust the focus on a point under the workbench. A subset of the subjects moved their heads up and looked at a plain background on the wall when they read the instructions to solve the visual cluttering and/or focusing problem. This portion of subjects gained no advantages from increasing "eye-on-the workspace" time by overlaying of information.

This result suggested that overlaying information in the central vision area of the user's view does not facilitate improvement in human performance. Based on the limitations of FOV and resolution of the current HMD technologies, only a very limited amount of information can be placed outside of the central vision area of a user.

6.2 Effect of Attention Switching and Mental Transformation Offloading on Performance

Hypothesis 2 states that overlaying information in user's view using AR in a spatially meaningful way improves user performance on the assembly task by reducing attention switching between the instructional media and the workspace. Hypothesis 3 states that by offloading the mental transformation tasks to the computer, subjects using AR instructions will perform better when comparing with subjects using traditional

instructional media, where pictorial instructions needed to be mentally transformed to the subjects' point of view. These two hypotheses suggest that the performance of subjects, in terms of time of completion and accuracy, in treatment condition 4 is expected to be better than treatment conditions 1, 2 and 3. There is statistically significant advantage in accuracy (in total error and dependent error) in condition 4 comparing with condition 1, 2, and 3. But there is no statistical significant advantage in time of completion. Since there is a strong correlation between time of completion and accuracy by nature (e.g. the faster you go, the more mistake you make), having advantage in one category and having the same performance in another category would be considered an advantage in overall performance. Therefore, these hypotheses are supported.

There is extensive research in the field of ergonomics of HUDs for aircraft pilots concerning switching attention among information sources and the real environment. [4, 25, 26, 30] reported evidences that optically overlaid information cannot be processed in parallel. [16, 23, 45] reported that there is a time cost associated with the cognitive switching among the information displayed in HUD and the surrounding environment. In AR, synthetic computer graphics are registered with the real world, and they appear to be a part of the world. It eliminates the cognitive load of switching attention across information displayed and the working environment. However, there is no literature the author aware of about how computer-assisted mental transformation of pictorial diagram affects user performance. It is a general presumption that computer assistance in the mental transformation task may result in improvement in performance. It is not certain how these two factors contribute to user task performance; i.e. which factor contributes

more to improving task performance. More research is needed to determine the contributions of these factors.

6.3 Effect of Treatment Conditions on Mental Workload

Hypothesis 4 states that mental workload of the assembly task using traditional instructional media is higher than using AR instruction. This hypothesis suggests that the NASA TLX in treatment condition 4 is expected to be lower than treatment conditions 1, 2 and 3. The lower NASA TLX rating in condition 4 is statistically lower relative to conditions 1, 2 and 3, which indicates subjects' mental workload in condition 4 is lower than in condition 1, 2 and 3. Therefore, this hypothesis is supported.

6.4 Effect of Spatial Cognitive Ability on Performance

Hypothesis 5 states that individuals with better spatial ability will perform better in an assembly task based on traditional pictorial-based instruction. This hypothesis suggests that the score of the spatial cognition test in treatments 1, 2 and 3 is expected to be correlated to the performance of subjects. However, there is no statistical significant correlation between the score of spatial cognition test and performance of subjects in all treatment conditions. Therefore, this hypothesis is not supported.

It is possible that a Type II error occurs in this correlation analysis. The correlation analysis might miss a small effect due to an insufficient sample size (about 19 in each treatment conditions). A bivariate correlation analysis was repeated with a combined sample from treatment 1, 2 and 3. The results are shown in table 6.1.

	Correlation analysis of spatial cognitive ability for combined sample from treatment condition 1, 2 and 3
Total error	$r(56) = -.207, p = .127$
Dependent error	$r(56) = -.164, p = .187$
Independent error	$r(56) = -.291, p = .030$
Time of completion	$r(56) = -.251, p = .062$

Table 6.1. Bivariate correlation analysis of Spatial Cognitive Ability and combined sample from treatment 1, 2 and 3.

The analysis shows that the correlation between independent error and spatial cognitive ability is statistically significant, and the correlation between time of completion and spatial cognitive ability trends toward significance. This result is contradictory to the original assertion about the falsity of hypothesis 5. A larger sample size is necessary in order to make an accurate assertion on hypothesis 5.

6.5 Effect of Dependent Error in Augmented Reality

In Section 5.1, it is noted that the number of dependent errors in treatment condition 4 is much lower than the other 3 treatment conditions. This may be due to the fact that determining position and orientation from pictorial diagram drawn from the author's perspective is a primitively hard task. Human beings tend to approximate the position and orientation using fixations and landmarks already in place. By overlaying the instruction in the exact position of the part at the location where it is to be inserted, AR not only reduces the cognitive workload to locate the position and orientation at the workspace from the instructional media, but also eliminates some dependency among procedural steps.

6.6 Effect of Attention Tunneling in Augmented Reality

It is observed that the rate of subjects correcting a mistake made in previous assembly steps in treatment condition 4 is much lower than in treatment condition 1, 2 and 3. This observation is coherent with a phenomenon called attention tunneling (also referred to as attention capture and cognitive capture in some literature). Attention tunneling refers to the phenomenon that attention is focused on the area cued, at the cost of other areas. Dopping-Hepenstal reported that “military pilots fixated more frequently on information presented on a HUD at the cost of scanning the outside scene” [11]. Yeh, *et al.* reported that “cueing aided the target detection task for expected targets but drew attention away from the presence of unexpected targets in the environment” [47]. Attention tunneling can reduce user performance and generate potentially hazardous scenarios. Yeh *et al.* recommended that the designer of such cueing systems more carefully evaluate operator reliance on automation.

6.7 Conclusion

The results of this research project support that AR improves human performance and relieves some of the user mental workload. The feature of overlaying and registering information on the workspace in a spatially meaningful way in AR allows it to serve as an effective instructional media. However, the limitations in the current display and tracking technologies are the biggest obstacles preventing AR from being realistic in practical uses. There is also a psychological implication in the phenomenon of attention tunneling which could possibly reduce human performance. AR system designer needs to

leverage the potential power of AR carefully in order to design a system that achieves an overall improvement of performance.

Appendix A

Procedural Steps for the Assembly Task



Step 1



Step 2



Step 3



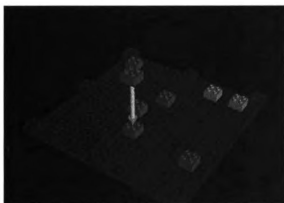
Step 4



Step 5



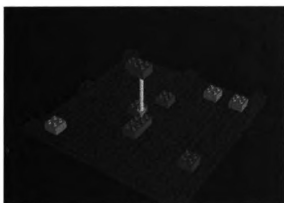
Step 6



Step 7



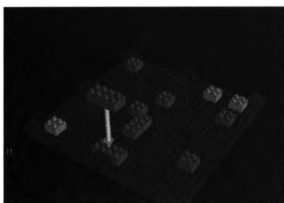
Step 8



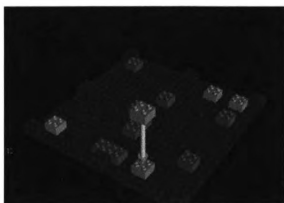
Step 9



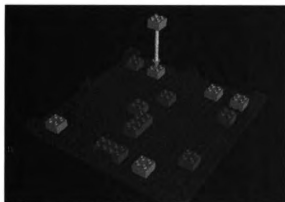
Step 10



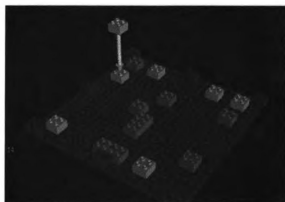
Step 11



Step 12



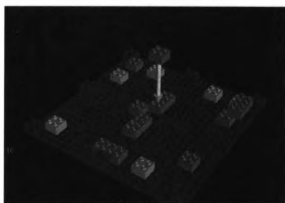
Step 13



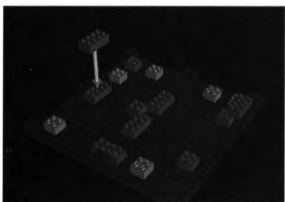
Step 14



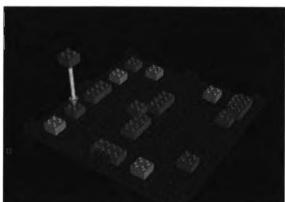
Step 15



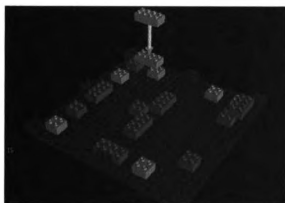
Step 16



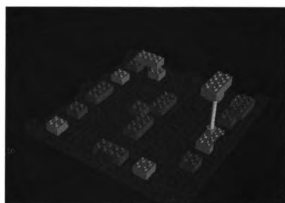
Step 17



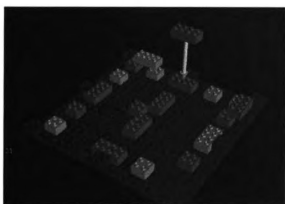
Step 18



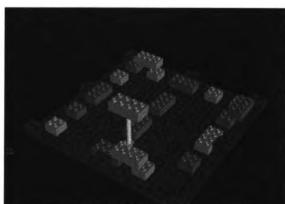
Step 19



Step 20



Step 21



Step 22



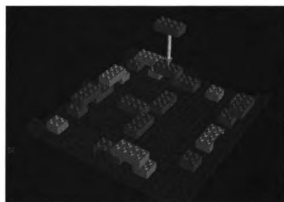
Step 23



Step 24



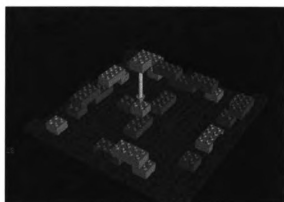
Step 25



Step 26



Step 27



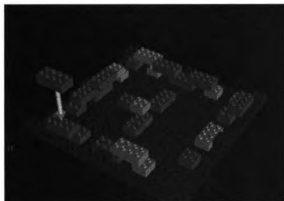
Step 28



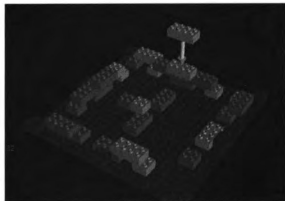
Step 29



Step 30



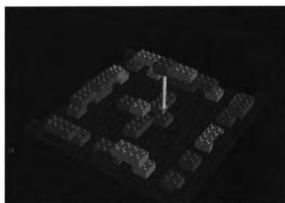
Step 31



Step 32



Step 33



Step 34



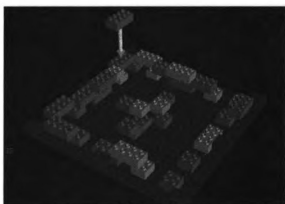
Step 35



Step 36



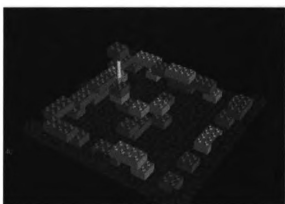
Step 37



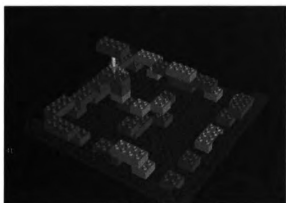
Step 38



Step 39



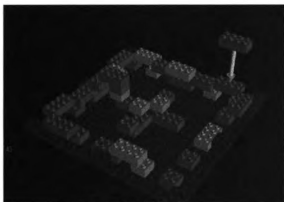
Step 40



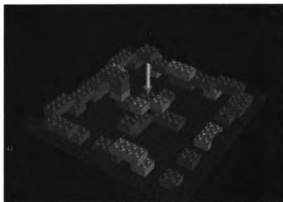
Step 41



Step 42



Step 43



Step 44



Step 45



Step 46



Step 47



Step 48



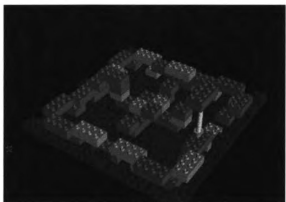
Step 49



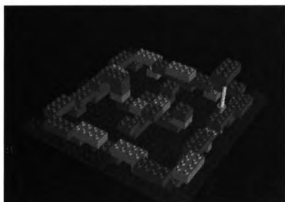
Step 50



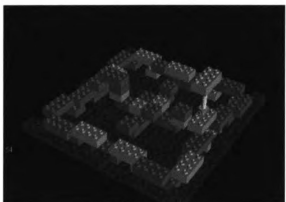
Step 51



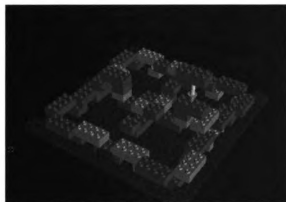
Step 52



Step 53



Step 54



Step 55



Step 56

Bibliographies

1. Azuma, R. (1995), *Dissertation: Predictive Tracking for Augmented Reality*. 1995, University of North Carolina, Chapel Hill: Chapel Hill. p. 262.
2. Bajura, M. (1993). Camera Calibration for Video See-Through Head-Mounted Display. Technical Report TR93-048. Chapel Hill: Department of Computer Science, University of North Carolina, Chapel Hill.
3. Bajura, M., H. Fuchs, and R. Ohbuchi (1992), *Merging Virtual Reality with the Real World: Seeing Ultrasound Imagery Within the Patient*. IEEE Computer Graphics, 1992. 26(2): p. 203-210.
4. Becklen, R. and D. Cervone (1983), *Selective looking and the notice of unexpected events*. Memory & Cognition, 1983. 11: p. 601-608.
5. Biocca, F. (2000). *Human-bandwidth and the design of Internet 2 interfaces: Human factors and psychosocial challenges*. In *Internet2 Socialtechnical Summit*. 2000. Ann Arbor, MI.
6. Biocca, F. and J. Rolland (1998), *Virtual Eyes Can Rearrange Your Body: Adaptation to visual displacement in see-through, head-mounted displays*. Presence, 1998. 7(3): p. 262-277.
7. Biocca, F., A. Tang, and D. Lamas (2001). *Evolution of the mobile infosphere: Iterative design of a high-information bandwidth, mobile augmented reality interface*. In *Euroimage 2001: International Conference on Augmented, Virtual Environments and 3D Imaging*. 2001. Mykonos, Greece.
8. Biocca, F., A. Tang, D. Lamas, J. Gregg, R. Brady, and P. Gai (2001). How do users organize virtual tools around their body in immersive virtual and augmented environment?: An exploratory study of egocentric spatial mapping of virtual tools in the mobile infosphere. Technical Report East Lansing: Media Interface and Network Design Labs, Michigan State University.
9. Brooks, F.P.J. (1996), *The Computer Scientist as Toolsmith II*. Communications of the ACM, 1996. 39(3): p. 61-68.
10. Caudell, T.P. and D.W. Mizell (1992). *Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes*. In *International Conference on System Sciences*. 1992. Kauai, Hawaii.
11. Dopping-Hepenstal, L.L. (1981), *Head-up displays: The integrity of flight information*. IEE Proceedings Part F, Communication, Radar and Signal Processing, 1981. 128(7): p. 440-442.

12. Edwards, E.K., J. Rolland, and K.P. Keller (1992). *Video see-through design for merging of real and virtual environment*. In *IEEE Virtual Reality Annual International Symposium*. 1992. Seattle, WA.
13. Ekstrom, R.B., J.W. French, H.H. Harman, and D. Derman (1976), *Manual for kit of factor-referenced cognitive tests*. 1976, Princeton, NJ: Educational Testing Service.
14. Feiner, S., B. MacIntyre, and D. Seligmann (1993), *Knowledge-based Augmented Reality*. Communications of the ACM, 1993. 36(7): p. 52-62.
15. Feiner, S., B. MacIntyre, H. Tobias, and A. Webster (1997). *A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment*. In *International Symposium on Wearable Computers*. 1997. Cambridge, MA.
16. Fisher, E., R. Haines, and T. Price (1980). Cognitive issues in head-up displays. Technical Report 1711. Moffett Field: NASA Ames Research Center.
17. Haines, R., E. Fischer, and T. Price (1980). Head-up transition behavior of pilots with and without head-up display in simulated low-visibility approaches. Technical Report 1720. Moffett Field: NASA Ames Research Center.
18. Hart, S.G. (1987). *Background Description and Application of the NASA Task Load Index (TLX)*. In *Department of Defense Human Engineering Technical Advisory Group Workshop on Workload*. 1987. Newport, RI.
19. Holloway, R.L. (2001), *Registration Error Analysis for Augmented Reality Systems*, in *Fundamentals of Computers and Augmented Reality*, W. Barfield and T.P. Caudell, Editors. 2001, Lawrence Erlbaum Associates, Publishers: Mahwah, NJ. p. 183-217.
20. Inzuka, Y., Y. Osumi, and K. Shinkai (1991). *Visibility of head up display for automobiles*. In *35th Annual Meeting of the Human Factors Society*. 1991.
21. Jebara, T., C. Eyster, J. Weaver, T. Starner, and A. Pentland (1997). *Stochastic: Augmenting the Billiards Experience with Probabilistic Vision and Wearable Computers*. In *International Symposium on Wearable Computers*. 1997. Cambridge, MA.
22. Kirsh, D. (1995), *The intelligent use of space*. Artificial Intelligence, 1995. 73(1): p. 31-68.
23. Larish, I. and C. Wickens (1991). *Attention and HUDs: Flying in the dark?* In *Society for Information Display International Symposium Digest of Technical Papers XXII*. 1991.

24. Lenz, R.K. and R.Y. Tsai (1988), *Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1988. 10(5): p. 713-729.
25. McCann, R.S., D.C. Foyle, and J.C. Johnston (1993). *Attentional limitations with head-up displays*. In *Seventh International Symposium on Aviation Psychology*. 1993. Columbus, OH.
26. McCann, R.S., J.M. Lynch, D.C. Foyle, and J.C. Johnston (1993). *Modeling attentional effects with head-up display*. In *Human Factors and Ergonomics Society 37th Annual Meeting*. 1993.
27. McGarrity, E. and M. Tuceryan (1999). *A Method for Calibrating See-through Head-mounted Displays for AR*. In *2nd IEEE International Workshop on Augmented Reality (IWAR 99)*. 1999. San Francisco, CA.
28. McGarrity, E., M. Tuceryan, C. Owen, and N. Navab (2001). *A new system for online quantitative evaluation of optical see-through augmentation*. In *IEEE and ACM International Symposium on Augmented Reality*. 2001. New York, NY.
29. Mizell, D.W. (2001), *Boeing's Wire Bundle Assembly Project*, in *Fundamentals of Wearable Computers and Augmented Reality*, W. Barfield and T.P. Caudell, Editors. 2001, Lawrence Erlbaum Associates, Publishers: Mahwah, NJ. p. 447-467.
30. Neisser, U. and R. Becklen (1975), *Selective looking: Attention to visually specified events*. Cognitive Psychology, 1975. 7: p. 480-494.
31. Neumann, U. and A. Majoros (1998). *Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance*. In *IEEE VRAIS '98*. 1998. Atlanta, GA.
32. Ohshima, T., K. Satoh, H. Yamamoto, and H. Tamura (1998), *AR2 Hockey system: A collaborative mixed reality system*. Transactions of the Virtual Reality Society of Japan, 1998. 3(2): p. 55-60.
33. O'Keefe, J. and J. Dostrovsky (1971), *The Hippocampus as a spatial map*. Brain Res, 1971. 34: p. 171-175.
34. O'Keefe, J. and L. Nadel (1978), *The hippocampus as a cognitive map*. 1978, Oxford: The Clarendon Press.
35. Owen, C. (2001). The ImageTclAR Augmented Reality Development Environment. "<http://metlab.cse.msu.edu/imagetclar>".
36. Pausch, R., T. Crea, and M. Conway (1992), *A Literature Survey for Virtual Environments: Military Flight Simulator Visual Systems and Simulator Sickness*. Presence, 1992. 1(3): p. 344-363.

37. Redford, A. and J. Chal (1994), *Design for Assembly, Principles and Practice*. 1994, London, England: McGraw-Hill Book Company.
38. Rolland, J., F. Biocca, F. Barlow, and A. Kancherla (1995). *Quantification of Adaptation to Virtual-Eye Location in See-Thru Head-Mounted Displays*. In *IEEE VRAIS '95*. 1995. Research Park Triangle.
39. Rolland, J., L. Davis, and Y. Baillet (2001), *A Survey of Tracking Technology for Virtual Environments*, in *Fundamentals of Wearable Computers and Augmented Reality*, W. Barfield and T.P. Caudell, Editors. 2001, Lawrence Erlbaum Associates, Publishers: Mahwah, NJ. p. 67-112.
40. Rolland, J. and H. Fuchs (2001), *Optical versus Video See-Through Head-Mounted Displays*, in *Fundamentals of Wearable Computers and Augmented Reality*, W. Barfield and T.P. Caudell, Editors. 2001, Lawrence Erlbaum Associates, Publishers: Mahwah, NJ. p. 113-156.
41. Sojourner, R. and J. Antin (1990), *The effects of a simulated head-up display speedometer on perceptual task performance*. Human Factors, 1990. 32(3): p. 329-339.
42. Sutherland, I.E. (1965). *The ultimate display*. In *IFIP Congress*. 1965.
43. Sutherland, I.E. (1968). *A Head-mounted Three Dimensional Display*. In *Proceedings of AFIPS Conference*. 1968.
44. Tuceryan, M. and N. Navab (2000). *Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR*. In *IEEE and ACM International Symposium for Augmented Reality*. 2000. Munich, Germany.
45. Weintraub, D.R., R. Haines, and R. Randle (1985). *Head-up display (HUD) utility II: Runway to HUD transitions monitoring eye focus and decision times*. In *Human Factors Society 29th annual meeting*. 1985.
46. Weng, J., P. Cohen, and M. Herniou (1992), *Camera calibration with distortion models and accuracy evaluation*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1992. 14(10): p. 965-980.
47. Yeh, M. and C.D. Wickens (2000). *Attention and Trust Biases in the Design of Augmented Reality Displays*. Technical Report ARL-00-3/FED-LAB-00-1. Savoy, IL: Aviation Research Lab, University of Illinois, Urbana-Champaign.

MICHIGAN STATE LIBRARIES



3 1293 02318 251