Spring 2003

The Effectiveness of Monitor-Based Augmented Reality Paradigms for Learning Space-Related Technical Tasks

Christopher Michael Opalenik
Embry-Riddle Aeronautical University - Daytona Beach

The Effectiveness of Monitor-Based Augmented Reality Paradigms for Learning Space-Related Technical Tasks

By

Christopher Michael Opalenik

A Thesis Submitted to the Department of Human Factors and Systems Embry-Riddle Aeronautical University for the Degree of Master of Science in Human Factors & Systems
Embry-Riddle Aeronautical University
Daytona Beach, Florida
Spring 2003

The Effectiveness of Monitor-Based Augmented Reality Paradigms for Learning

Space-Related Technical Tasks

By

Christopher Michael Opalenik

This thesis was prepared under the direction of the candidate's thesis committee chair, Dennis A. Vincenzi, Ph.D., Department of Human Factors & Systems, and has been approved by the members of the thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors and Systems.

THESIS COMMITTEE:

Dennis A. Vincenzi, Ph.D., Chair

Steven Hall, Ph.D., Member

Sathya N. Gangadharan, Ph.D., Member

Anthony Majoros, Ph.D., Member

Shawn Doherty, Ph.D., Member

MS HFS Program Coordinator

Department Chair, Human Factors & Systems

Associate Chancellor for Academic Affairs

ABSTRACT

Author: Christopher Michael Opalenik

Title: The Effectiveness of Monitor-Based Augmented Reality Paradigms for Learning Space-Related Technical Tasks

Institution: Embry-Riddle Aeronautical University

Degree: Master of Science in Human Factors & Systems

Year: 2003

Currently today there are many types of media that can help individuals learn and excel in the on going effort to acquire knowledge for a specific trait or function in a workplace, laboratory, or learning facility. Technology has advanced in the fields of transportation, information gathering, and education. The need for better recall of information is in demand in a wide variety of areas. Augmented reality (AR) is a technology that may help meet this demand. AR is a hybrid of reality and virtual reality (VR) that uses the three-dimensional location viewed through a video or optical see-through media to capture the object's coordinates and add virtual images, objects, or text superimposed on the scene (Azuma, 1997).

The purpose of this research is to investigate four different modes of presentation and the effect of those modes on learning and recall of information using monitor-based Augmented Reality. The four modes of presentation are Select, Observe, Interact and Print modes. Each mode possesses different attributes that may affect learning and recall. The Select mode can be described as a mode of presentation that allows movement of the work piece in front of the tracking camera. The Observe mode involves information presentation using a pre-recorded video scene presented with no interaction with the work piece. The Interact mode allows the user to view a pre-recorded video scene that allows the user to point and click on the component of the work piece with a computer mouse on the monitor. The Print mode consists of printed material of each work piece component.

It was hypothesized that the Select mode would provide the user with the richest presentation of information due to information access capabilities helping to decrease work time, reduce the amount of error likelihood during usage, enhance the user's motivation for learning tasks, and increase concurrent learning and performances due to recall and retention. It was predicted that the Select mode would result in trainees that would recall the greatest amount of information even after extended periods of time had elapsed. This hypothesis was not supported. No significant differences between the four groups were found.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Several types of media are available that may assist society to learn and excel in the acquisition of knowledge and skills. Technology has advanced in a wide variety of areas including, but not limited to, the fields of transportation, information gathering, and education. Researchers no longer have just textbooks or reference manuals to help acquire answers to problems discovered on the job site or in an educational environment. Computer media is a highly expanding source of gathering, searching, and information processing in the 21st century that may help find solutions to knowledge acquisition and training problems. Researchers can now connect to the Internet and access textbooks and encyclopedias from across the world, talk online, and have web video conferences with experts answering questions on a wide variety of topics.

The need to learn using different cognitive media is in demand for maintenance, manufacturing, and inspection tasks. The media employed must be easily accessible and compatible with existing applications and equipment. Video-based media such as a video camera and VCR can be used for cognitive tasks. These types of media are incorporated as valuable tools for increasing monitor-based AR characteristics.

AR is a hybrid of reality and virtual reality (VR) that uses objects location through a video or optic see through media to capture an object(s) coordinates and add virtual images, objects, or text superimposed, attached to the scene (Azuma, 1997). By using both realities, a person can obtain the benefits of being part of the real scene along with spatial traits of objects or text information that can help the person learn, decrease costs, and increase equipment availability for manufacturing, maintenance, and inspection. This text information that is rendered is called an annotation, a graphical box with a leader line pointing to a reference point containing information pertaining to inspection, function, or terminology of a specific component of a work piece. Monitor-based AR uses a camera view of a scene and is displayed to the user with the annotations in a highly visual environment. The changing of the camera's position determines the object or annotations rendering of the images. The changing position enables graphic objects or annotations to be positioned in the video scene that is spatially registered to actual features of the real-world environment scene. The field of AR is advancing and is in demand for research in a wide variety of fields. These fields range from medical usage when performing surgery that requires precise and in depth training to perform life saving actions, to the military field by training soldiers to identify and engage potential hostile targets in an unfamiliar environment.

This literature review will provide information on how monitor-based AR can increase learning performances, decrease significant training time, and improve resource investment through four effects in cognitive psychology in the areas of information access, reduced error likelihood, enhanced motivation, and concurrent learning and performance. A discussion on the different types of AR systems and their characteristics will be explored to help illustrate the functionality and limitations of the system. Cognitive characteristics relating to learning will then be discussed by linking spatial and recall abilities with explanations of how AR defines the location of a real-world environment object and how to apply the location of the object in a non-virtual environment.

Optical and Video See - Through versus Monitor-Based Augmented Reality.

To capture and enhance actual views of the real environment using a video camera or optical display, researchers need to capture and control certain attributes of the current environment. These enhanced images can be viewed using a Head Mounted Display (HMD) to see computer generated objects or annotations superimposed on real-world images. There are two types of HMD displays that AR research is currently utilizing. The two types of HMD AR systems are optical see-through and video-see-through.

An optical see-through HMD uses the real-world environment seen through two half-transparent mirrors placed in front of the wearer's eyes (Fuchs & Rolland, 2001). The mirrors only cover the forward immediate field-of- view and do not cover peripheral vision outside the line of sight of the wearer. The images are then superimposed on the mirrors from a camera mounted on top of the HMD. This combines the images the wearer is seeing in the actual environment along with graphical depictions of data or imagery the HMD receives from the AR system.

Video see-through AR shares some similar characteristics but captures the real-world environment and displays the objects or annotations differently. The difference is the real-world view is captured through two miniature video cameras mounted on the HMD. The computer then generates annotations, and objects are then combined with the video representation of the world (Fuchs & Rolland, 2001). The major difference between the two types of HMD's involves the mechanics through which the wearer views the real-world environment, either through a video displayed image of the environment or the actual scene being viewed. Figure 1 illustrates the components of a typical AR System.

The other type of equipment that AR can use to manipulate scenes with computer generated objects and annotations are monitor–based Augmented Reality or AR. This system is similar with respect to the Video see-through HMD but it does not require the wearing of a HMD. The video camera is separate from the user located in a static position with no interaction from the user required to capture the real-world environment. Both the real-world environment and AR objects and annotations are displayed on a monitor screen. This gives the user a view of only what is shown from the point of reference of the camera.

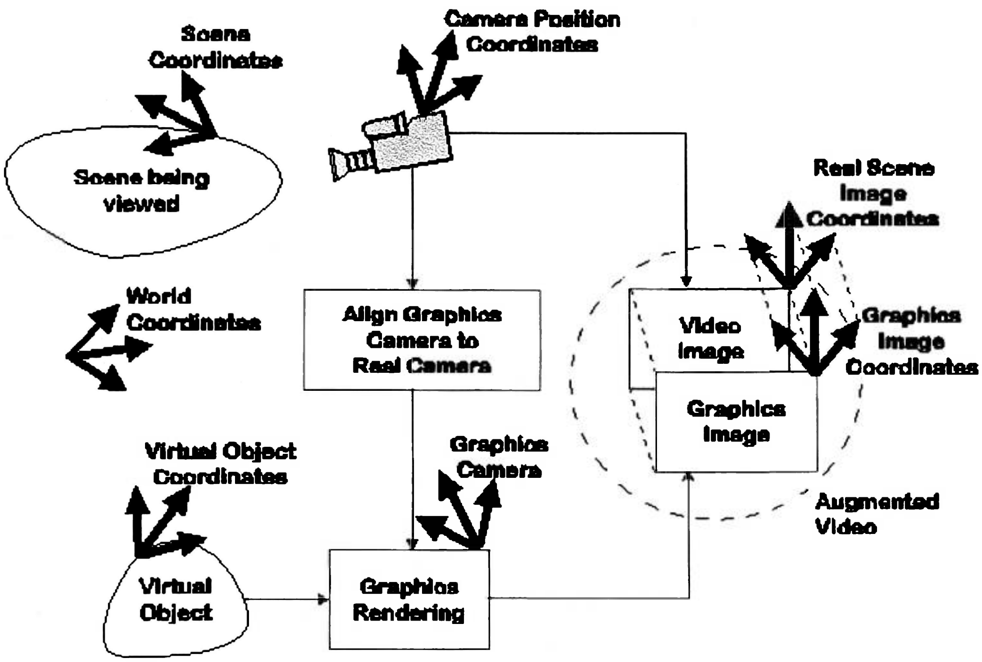

Figure 1. Components of a Typical AR System.

Even though the Optical and Video See-Through AR appears similar to monitor-based AR, there are some distinct differences. The following paragraphs discuss advantages and disadvantages between the HMD modes of AR and monitor-based AR with respect to system latency and field-of-view between the two types.

In order to create and display an AR object or annotation, several procedures need to be performed before the actual rendering of the scene can occur. The real-world environment scene needs to be registered and geometrically sized to render an object or annotation that is spatially correct within the environment. Movement of the tracking camera distorts the rendering of the object and annotations. To correct the distortions, the AR system needs to be geometrically calibrated between the tracking device and the tracking camera with reference to the real-world environment. This geometric calibration results in lag time, which is the largest source of registration errors in most systems (Fuchs & Rolland, 2001).

To decrease the amount of lag time for a system, some sacrifices must be made. The most significant AR characteristic that constantly changes is the image quality of the AR scene. As an AR system matures, researchers can compensate for the adjustments of lag time and image quality with faster and more efficient systems along with more advanced modes for image rendering

The most productive of the AR systems are the Video See-Through and the monitor-based AR systems. Both use video AR systems characteristics that capture real-world coordinates and process the coordinates to display the AR images. Since the video image is captured before the user can view them, researchers can delay the scene until the computer-generated AR images are rendered. This can reduce lag time because the manipulation of the real world scene is video based and not optical.

The Optical See-Through is quite different. There are no methods to induce artificial delays in the real-world environment scene. The Optical See-Through uses a semi-transparent mirror through which the user views the environment in its natural state instead of video feed. One way to reduce lag time is to limit the amount of head movement of the wearer to a defined boundary. The boundary could be a movement of 60 milliseconds for predicted tracking for a HMD AR system (Fuchs & Rolland, 2001).

Another characteristic that all AR systems share is field-of view distortion. Optical see-through AR systems have the most difficult time with this environmental constant. With large binocular fields-of-view (FOV), the movement by the user needs to be minimized. The amount of FOV is very important to tasks that require any type of physical activities such as grabbing and moving objects. These types of activities increase the chances of AR distortion by allowing the scene to become out of the tracking cameras FOV. Optical see-through FOV is dependent on the size of the reflective mirrors. When performing a tangible task, the FOV will increase or decrease the amount of augmentation with the rendered scene. If the objects or annotations being manipulated with AR are in the field of view, then the AR system can render the images. The optical see-through AR system relies on two different kinds of variables. One is the three dimensional movement of the HMD. As the HMD moves in all three real-world environmental coordinate planes, the AR system renders visual objects or annotations to the user. The dimensions of the FOV will decrease the error of failing to produce the AR image. The other variable is movement of objects viewed with the HMD. If the object moves while the HMD is stationary, the AR image may not be rendered. Figure 2 illustrates typical components of an Optical See-Through AR system.

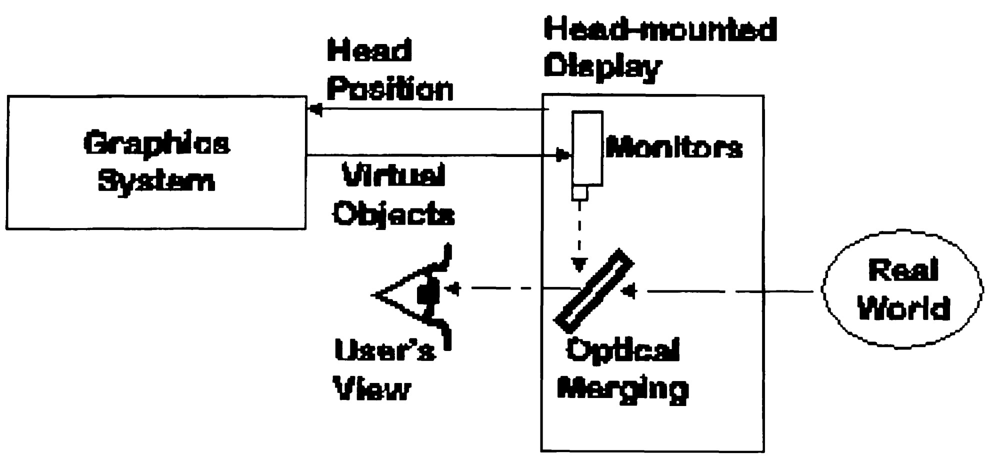

Figure 2. Optical See-Through AR System.

The Video See-Through AR has some different attributes compared to the HMD. One difference is the procedure for correcting the FOV of the tracking camera. A design matching the frustum of the wearer's eye with that of a camera is a challenging task (Fuchs & Rolland, 2001). This is not important for viewing distant objects but it is crucial when near-field visualization is used. The FOV needs to be narrow to decrease the amount of head movement as discussed with the Optical See–Through AR system. It also shares the same variable discussed above and shown in Figure 3.

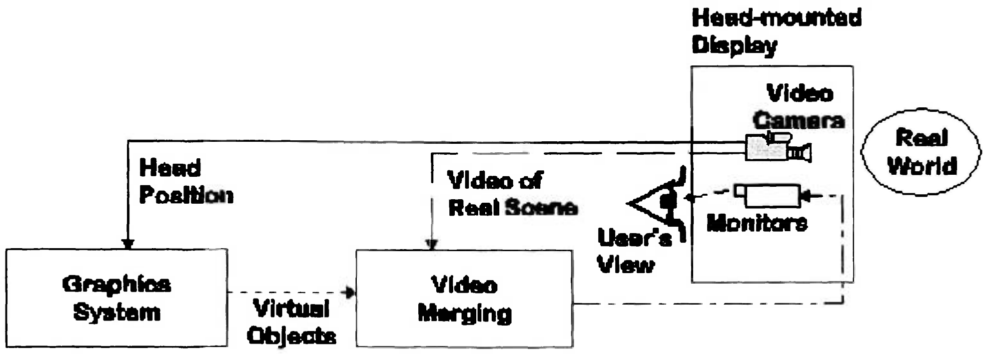

Figure 3. Video See-Through AR System

Monitor-based AR systems have the option to eliminate the movement of the tracking video camera discussed with both the Optical and Video See-Through AR systems. The monitor-based AR system has the capability to allow the tracking camera to be static or dynamic when rendering images during augmentation. The HMD AR system is constantly changing the FOV position, which increases other problems when projecting AR images. The FOV can be changed depending on the video camera's specification and static position. If the tracking camera were dynamic, then the tracking camera would have the same FOV distortion problems as the HMD AR system. Researchers can have better control over the position of the camera but will still have some error with FOV as shown in Figure 4.

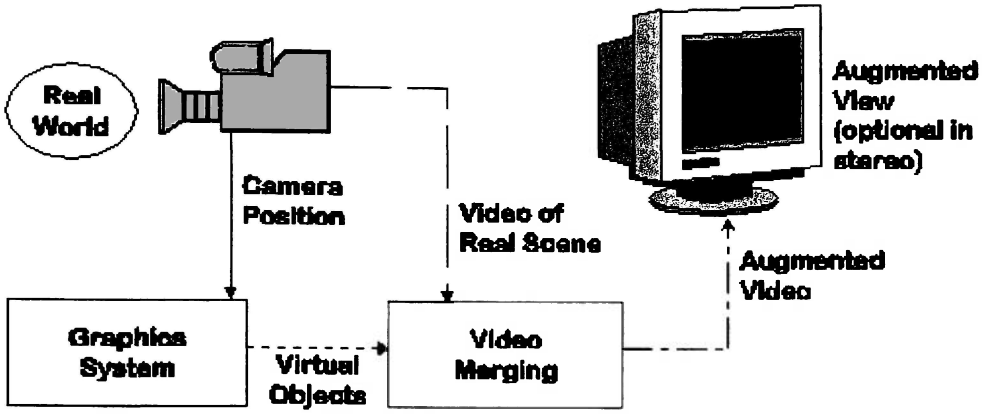

Figure 4. Monitor-Based AR System

Registering and Sensing Errors with Augmented Reality.

One of the most important goals for an AR system is to correctly render computer generated objects aligned with real-world environment coordinates in a scene. To correctly display the AR image in reference to the real world environment scene, the AR system must be registered with the coordinates of the real-world environment. The correct position of X, Y, and Z coordinates of virtual objects with respect to the real-world environment is known as registration (Holloway, 2001). The failure of the AR system to align virtual objects is called registration failure. Another problem with correct AR object overlaying is accurate tracking of the user's head or camera and then sensing the locations of other objects in the shared environment (Azuma, 2001). The requirement for correct and accurate long range sensors and trackers to report the current location of the user and the objects around the user in the real-world environment is called accurate sensing. This section will discuss about the problems with registration errors, how to control the tracking of the real-world environment scene, and how to produce accurate and concise AR objects and annotations with minimum distortion.

Perfect accurate registration of an object in AR is extremely difficult. Many types of variables can interfere with obtaining current real-world environment coordinates and referencing them with augmented objects. There are error models for AR that can help decrease error types and procedures on how to correct them. Researchers need to find out were the registration errors are created and then decide which ones can severely affect the AR system. Investigating the cause of distortion for each error found can be done mathematically with vectors and distant formulas. Results of registration error can be represented by a function of the level of accuracy created with correct registration. There is some AR software that can calibrate the AR system with reference to the real-world environment. All types of AR systems experience these problems.

The University of Washington's Human Interaction Lab (HIT) has developed several types of AR systems for various computer platforms, along with software the user can use to calibrate tracking cameras and HMD equipment to help decrease limitations and boundaries for pattern recognition. Calibration procedures referencing the real-world environment can help the AR system to eliminate most errors that can cause poor rendering.

Types of Registration Errors.

There are four different types of categories for registration errors. They are Acquisition/Alignment error, Head-Tracking error, Display error, and Viewing error (Holloway, 2001). The Acquisition/Alignment error processes data for virtual creation and then matches it with the real-world environment coordinates by overlaying the scene with text or 3D object rendering. An example would be the failure to calibrate the AR system effectively, which would result in an unsatisfactory rendering of objects and annotations. Head Tracking error is caused by the HMD tracking camera computing invalid real-world environment coordinates and then sending those coordinates to the AR system. The error source for Head Tracking error can be described as the increasing percentage of tracker delay, measurement error with static and dynamic movement of the HMD, or calibration of the AR system. Because AR systems are visual, errors in functionality of producing the image are common. Display error has many causes for improper rendering. Some can be optical distortions due to foreign objects between the visual capturing device and the environment. Other possibilities could be lateral color distortion, aligning of the object being recognized by the AR system to compute and display the designated object or annotation, or calibration of the tracking camera. Viewing error is caused by location and rotation of the viewer's line of sight. Non-calibration is the main cause of viewing error.

The largest source for registration error for all AR systems is fluxion of system delay. One attribute of AR systems is the expansion of different programming platforms that can utilize AR software. Every single type of computer hardware has limitations and advanced architecture design for speed and efficiency. AR system delay varies depending on the speed of the AR host computer and the amount of mathematical computation and manipulation needed to perform the required augmentation. System delay is more severe than Acquisition/Alignment error, Head-Tracking error, Display error, and Viewing error. There are six different kinds of delay sources for registration error (Holloway, 2001). The first delay source type is Tracker delay, which is the time needed to compute coordinates, mathematical calculations, and transpose coordinates with respect to the current real-world environment rendered. The next one is called Host-computer delay. This delay is caused by the computer that is receiving the data from the HMD or video along with host-based applications that run concurrently with the AR system. Image-generation delay deals with the rendering of the annotations and 3D objects during augmentation. Video sync delay and Frame delay all deal with the host computer's failure to send a steady video sync with the camera along with low frame delay making the augmented scene look choppy. If the display devices are either video or HMD, there could be additional delays inside the HMD or camera. The delays could be caused by conflicting resolution or lagging input signals from the video or HMD.

To help reduce the amount of registration error and position overlap of inconsistent AR annotations and objects, accurate sensing of real environmental reference points need to be optimized. Developing accurate sensing and tracking cameras such as HMD and video cameras are the largest concerns to be aware of when rendering. Currently there are no trackers that can produce a concise output for AR workloads. There are three intense demands from trackers and sensors. They are the amount of input variety and bandwidth, higher accuracy, and longer range for sensing patterns and objects (Azuma, 2001).

If the AR system does not render accurate objects or annotations, then the entire AR session becomes useless for learning tasks. AR systems currently are not being used for high fidelity scenarios that may physically harm a user if rendered images are incorrect. Dependable accuracy relies upon the ability of the tracker to process the pattern and display the annotation or object in the correct location. Trackers used in the past include mechanical trackers, magnetic trackers, ultrasonic trackers, video camera tracker, and optical trackers. Thus far, it appears that optical trackers are the most reliable because of rich high-resolution digital cameras, light sources, and real-time attributes. Video Camera are the second most reliable because of cost, usability, and maintainability. Other types of tracking equipment may increase accuracy towards almost perfect AR sessions in the future.

Limited range of a tracker may cause sensing problems when the AR system attempts to render an image. Range is defined as the distance from where the tracking camera is positioned to the location of the object being viewed either through a HMD, video camera, or monitor.

Because AR is visual, the range needs to be within a tolerant level. As distance for registering a pattern increases, the orientation error also increases displaying unrecognizable annotation and objects out of the field-of-view. One example that bests demonstrates this problem is using an AR system outdoor where the field of view is constantly changing and varies from sector to sector of the environment. AR is not at this advanced stage yet and will not be until all three sensing errors are controllable and maintainable.

Information Access.

Information acquisition is dependent on access to equipment and written documentation for inspection and cognitive tasks for a work piece. In the field of maintenance and inspection, there are many types of media that can help workers access information on a particular subject or task. In the aerospace industry, NASA may need to repair or replace faulty or damaged parts that require technical manuals and visual representation on the component to inspect, disassemble, and then replace the component. In order to accomplish this task, the worker would need to search through manuals, find the part, review the instructions and specifications pertaining to the component, and then apply the information to complete the task. The process of obtaining information involves document-related activities and visual recognition tasks that are done separately but in sequence. These searches can take hours depending on the amount of information the trainee has at his or her disposal. Monitor-based AR limits the time to find the materials and the specific information needed in training will decrease the investigation of the materials (Majoros & Neumann, 1998).

Reduced Error Likelihood.

Errors are likely to occur when using a medium for learning through visual or verbal actions. Constant iterations of tasks by a novice can produce cognitive errors from acquiring the information rapidly. This causes loss of recall and incorrect data retained that can result in errors when the information is needed in the real world environment. Monitor-based AR can reduce the potential for error when engaged in learning activities by recall and retention of tasks performed. For novices this is a constant problem until the information is mastered. The chance of errors is a function of the interaction of user's expertise and factors of situations such as a tasks environment (Majoros & Neumann, 1998). Utilizing monitor-based AR can reduce the amount of effort expended by the trainee in acquiring information by speeding up the transition of novice to expert when learning. AR complements human associative information processing and memory because virtual objects have a defined location and they can associate with features of the real world, not virtual (Majoros & Neumann, 1998). This can increase memorization and decrease stress.

An important attribute associated with monitor-based AR is the utilization of a work piece component's real-world environment coordinates to render objects and annotations for the user. AR then registers each component defined by the AR system in view. Experimental research with monitor-based AR tasks have discovered that by incorporating virtual objects into scenes, the objects would become part of the existing world scenes when attention is obtained and are spatially defined just like the components in the real-world environment scene.

Monitor-based AR uses spatial cognition to direct attention of the rendered object or annotation during learning tasks. This is an imaginary-related ability to know the location or layout of a physical environment (Majoros & Neumann, 2001). When an individual reads a manual on a mechanical work piece, the individual can often recall the approximate location of information within the documents in a two-dimensional world. But the person does not just learn the actual characteristics of the work piece. It helps to couple the spatial location and renders objects or annotations in the scene. Monitor-based AR uses this coupling effect by inserting the virtual objects and/or annotations in the real-world environment scene and the two become spatially defined within the same scene.

Past research involving text comprehension has indicated that readers retain some spatial information such as word location. In an experiment done by IBM Research Center and the Center of Research in Cognitive Psychology, (CREPCO), two types of text representation of the content and representation of the surface form of the text that includes the spatial location of the words was used (Baccino & Pynte, 1994). The purpose of the experiment was to investigate word locations and the retention of text content due to the memory of the reader. Other experiments have discovered a connection between recall of word location and recall of content (Baccino & Pynte, 1994). Monitor-based AR uses a similar type of spatial ability with annotations that contain information pertaining to a component of a work piece. A leader line connecting the annotation to the component establishes the position of the annotations rendered. This type of monitor-based AR composition will use recall of the annotation content coupled with the leader line pointing to the component. The annotation can contain information describing the function of the component and important information that may facilitate retention of spatial information based on a description of the location in the annotation and reference to which component the leader line is referencing.

Spatial information linking the component and annotations may increase learning ability and recall capability with respect to where the annotation points, and may facilitate retention of spatial information based on the description of the component being referenced. The amount of information that can be learned is based on how much material the person retains through exercises and learning procedures. The amount of retention depends on the level of recall of the person. Monitor-based AR may increase the amount of information recalled by an individual by linking spatial location and virtual object attributes (Majoros & Neumann, 1998). Experiments have proven that knowledge of location and memory of the presented content are not independent (Lovelace & Southhall, 1983). When a person commits visual text to memory, the location may help the person retain the information by having select words aid in the amount of retention achieved during the cognitive task. Monitor-based AR provides these conditions by presenting the real-world environment with text information together instead of separately.

For a person to memorize and retain information from text or abstract objects, the person relies on visual cues and relevant objects to retain information in long-term and short-term memory. Words and visual aids are not memorized in the same way (Dwyer & Melo, 1983). Memorizations of text documents alone are committed to memory by rehearsal when reading the text. There are no associations present that can help elaborate on what the document is trying to present. Visual representation of an object may help the person understand what is being described, but it may lack important information such as inspection criteria for a mechanical piece. By using specific visual material, a person can facilitate task completion and enhance recall. Monitor-based AR uses both visual and text attributes to enhance the relationship between the various attributes and information acquisition.

Enhances Motivation.

People do not like to spend time in a learning environment by sitting in a chair and reading a manual while viewing 2D pictures that are intended to represent an object for learning and performing a task. People are visually oriented and rely heavily on vision more than any other sense. AR uses a rich environment that uses sensory displays to help people acquire information in cognitive situations. Contrast and brightness captures attention and can increase the time span used for cognitive tasks. By using object and annotations, monitor-based AR can enhance the motivation of the trainee and increase the learning amount compared to a 2D medium like a textbook.

Fred Brooks labels monitor-based AR as a form of intelligence amplification (IA) because computers are utilized as a tool to increase the simplicity of tasks to be performed (Azuma, 1997). The computer-generated overlay created by AR is under program control. This means that the participant can view virtual objects or annotations generated by a computer when the computer scans a symbol embedded in the actual environment linked to a component by a leader line.

There are six potential uses for monitor-based AR program control. One is the designing of objects to increase the trainees' focus on attention during a learning session. It can enhance the trainees' ability to parse elements into method sets. Also it designs objects or annotations to be adjustable with respect to languages or distances from the person. The Individual can have objects or annotations rely on the operating conditions of the monitor-based AR session. Finally, it can improve the operator's monitor-based AR tool ability of discrimination by moving objects in the AR world. All of these controls are functional in a video-based AR environment (Majoros & Neumann, 1998).

The reason for increasing motivation is to assist in the learning and retention of a particular component of information, and apply the acquired knowledge in an actual real world work environment. Studies have found that as time spent on a task increases, the amount of information retained also increases for the participant (Brown, 2001). Monitor-based AR can increase the behavior of the participant by giving the participant the ability to determine the amount of time for retention of the information. In a classroom setting, a professor determines the amount of time for retention. The student, depending on their level of recall, may have a decreasing affect of the material they learned and therefore cannot access the information. AR gives the participant the freedom to learn information pertaining to a task and decrease pressure as a result of acquisition of the information being learned. The environment is a key factor in performing cognitive tasks. Attention of the participant can decrease if the learning environment contains attributes of distraction, decreased motivation, and may lead to direction of attention off-task (Brown, 2001).

Some learning tasks can only be performed at facilities that contain the media or text material required for the learning session. Monitor-based AR allows the environment to be controlled. Portability is a key factor in motivation because the participant can take the learning tool and perform tasks wherever he or she wishes. Monitor-based AR can be viewed with visual media equipment from a previously recorded session such as a VCR tape. This type of learning is called on-the-job (OTJ) training.

Concurrent Learning and Performance.

Today, researchers use learning techniques such as classrooms with textbooks and computer-based learning along with assigned tasks to help increase cognitive saturation. Companies use on-the-job learning to increase the level of learning by having experienced co-workers teach each other during cognitive sessions. Monitor-based AR can eliminate some forms of learning tasks and make some procedures unnecessary. Monitor-based AR allows the possibility to create different types of learning functions that are "just-in-time". Textbooks, videotapes, and other types of media can be outdated from a week to a year depending on the advancement of the participant being taught. With monitor-based AR, learning can be updated and expanded more quickly than with other types of media.

Visual stimulation plays an important part in teaching skills along with simple visual displays. An experiment designed to measure the effectiveness of augmented visual feedback for increasing landing skills indicates that augmented visual feedback speeds up the acquisition of landing and bombing skills for novice military pilots (Lintern, 1991). AR has the same attributes as augmented visual feedback does. Besides increasing the acquisition rate of task performance and learning, monitor-based AR can also help the participant's transfer of skills from learning sessions to the working environment to be more efficient and more stable.

Another important objective in AR is transfer of skills from monitor-based AR. The characteristics of a controlled work piece, the motive to activate the task, and knowledge about the function of the task are three attributes that affect transfer of skills in a learning environment (Lintern, 1991). To have a successful transfer of skills, the information that is learned needs to have certain attributes and operations. In learning environments, the need for similar operational tasks along with the quantity of fidelity is sometimes used to illustrate trainer design features (Lintern, 1991). Monitor-based AR can render annotations or objects in the actual real-world environment in which the work piece is found, or simulate an environment that contains the work piece through video in a similar transferable environment. The environment is similar to the exact place of the work piece therefore transfer of skill is increased. The transfer of skills is performed in monitor-based AR because of its on-the-job learning attribute.

Visualization ability is a key attribute when dealing with learning and monitor-based AR. Studies have been conducted to investigate whether visual learning will increase and enhance the ability of users to learn and recall the information presented in a cognitive session. Visualization ability is described as the ability to manipulate or transform images of spatial patterns into other types of arranged patterns. This is a very important predictor of learning and performance of visual based learning. People either have high visual abilities or low visual abilities (Bostrom, Davis, Olfman & Sein, 1993). Monitor-based AR adapts to these constraints depending upon the individual being trained and that upon which the user is being tested. Studies and experiments using high and low visualization abilities discovered high visual trainees could have the ability to take greater advantage of the monitor-based AR system and create mental representations of the AR session.

Researchers also have to explore the results pertaining to users with low visualization abilities with video-base AR. Studies and experiments have shown that users with low visualization abilities can eliminate or surpass users with high visualization abilities by increasing the learning methods and manipulating interfaces that the low visualization participant is using compared to the high visualization participants learning procedure. Based on these findings it can be said that the cognitive advantages provided by monitor-based AR can help increase learning for users with either high or low visualization abilities based on the flexibility of system control (Bostrom, Davis, Olfman & Sein, 1993).

The amount of time a participant spends on a visual pattern or component may reflect the total amount of retention of information the participant has in his or her visual memory. There are two types of visual memory. One is Short Term Visual Memory (STVM) and the other is Long Term Visual Memory (LTVM). STVM deals with an active control process referred to as visualization. LTVM reflects long-term knowledge of information acquired as processes of memorization (Avons & Phillips, 1980). The amount of time information is presented for a participant to retain increases the likelihood of acquisition and recall. Both LTVM and STVM are dependent on the time given for learning information. Studies show that as participants increase post-stimulus processing time, the more LTVM the participant uses. This provides evidence that rehearsal leads to an increase in LTVM. There are two types of rehearsal processes called maintenance rehearsal and elaborative rehearsal (Avons & Phillips, 1980). Maintenance rehearsal does not increase LTVM where as elaborative rehearsal does increase LTVM. AR increases LTVM by incorporating learning tasks with visual annotations that can be viewed at any stage of sporadic or constant elaboration. The amount of time the participant views the annotation increase LTVM and retention of the object or annotations is improved.

Monitor-based AR has many contributors towards cognitive tasks by improving training for maintenance procedures, increasing manufacturing products will less cost and more proficiency, and inspections with less time spend and more reliability on identifying potential problems of a component. The augmentations of objects or annotations can increase learning performances, decrease significant training time, and improve resources effects in four areas that were discussed. The four areas are information access, reductions in error likelihood, increased motivation, and concurrent learning and performances.

The amount of time during cognitive tasks when searching for specific information can decrease cost in training due to monitor-based AR's ability to present the information clear and precise. Monitor-based AR's ability to reduce potential errors during learning activities was also explained to be true by recall and retention of tasks performed. The amount of errors created when objects or annotation are renders to the scene, can be reduced by coupling spatial locations and rendering of objects or annotations in a scene. This is an important part in reducing error in any AR system. Using sensory displays for increased recall and retention can increase the level of motivation an individual has during cognitive tasks. Monitor-based AR uses concurrent learning by utilizing "just-in-time" training attributes along with transfer of skill during training.

This leads to multiple tasks such as training, recall, and retention of information by combining the actual work environment with cognitive tasks.

The differences in AR systems help distinguish which systems are more efficient and productive in certain environments. Monitor-based AR proved to be more efficient in spatial traits and cognitive tasks in a static real-world environment. Discussing the different types of registration errors and sensing error can help explain and prepare researchers in developing other AR experiments. All of these characteristics and conditions of monitor-based AR prove that by being part of the real learning scene, rendering of objects or annotations help in recall, spatial ability, and retention of knowledge will greatly facilitate the acquisition of knowledge in the individual.

STATEMENT OF THE HYPOTHESIS AND PREDICTIONS

Four groups will be set up, a video instructional group (Observe group), an interactive, video instructional group (Interact group), a monitor-based augmented reality instructional group (Select group), and a print-based instructional group (Print group), each utilizing a different mode of information presentation, but all being presented with the same pictorial views and information. The Observe, Interact, and Print groups represent the most common forms of media used in instructional presentation today. These groups will be used to examine the effectiveness of AR in presenting information for purposes of learning when compared to current methods.

Previous research conducted in the areas of recall, spatial ability, and learning suggests that a technology, such as AR would greatly facilitate the acquisition of knowledge in the learner. The AR system used by the Select group provides more sensory interfaces such as visual, spatial, verbal, tactile, and proprioceptive, than the other groups that utilize only one or two interfaces. This increase in modal interaction should create more memory traces and elaboration cues which will assist in acquisition, retention, and recall of knowledge.

Hypothesis

It is hypothesized that the instructional session will increase the amount of knowledge acquired concerning the work-piece as reflected on two recall tests, and that this improvement will differ across the four mediums of information presentation.

Prediction 1.

The Select group will achieve significantly higher test scores during a post-recall test condition than the Observe, Interact, and Print groups following the administration of a training session.

Prediction 2.

The Select group will achieve significantly higher test scores than the Observe, Interact, and Print groups during a long-term recall test condition administered one week after the training session.

METHOD

Participants

Participants were taken from the undergraduate population at Embry-Riddle Aeronautical University. The participants were randomly selected and a demographic questionnaire was administered. There were three phases to the experiment. They were the documentation session, the instructional session, and the experimental session. The documentation session involved a demographic survey, consent form, brief visuospatial test, and a self-screening vision test. The demographic survey consisted of questions taken by the participants relating to human behavior such as claustrophobic and motion sickness and classifications such as the participants degree, current year in college, gender, and other factors that may affect the results of the experiment.

A total of 96 participants were selected. There were four groups with 24 participants in each group. After the survey, each participant was tested to determine his or her visuospatial memory, and visual perception ability. The self-screening vision test was used for testing the participant's distance and reading power within visual distance of 20 inches from the position of the participant to the monitor. The spatialization ability test that was used was the Brief Visuospatial Memory Test – Revised. Each participant was evaluated based on his or her demographic survey, visuospatial memory test, and visual ability test.

Apparatus

The apparatus used for the treatment conditions consisted of a Silicon Graphics O2 Desktop CPU with an IRIX v 6.5 operating system (see Figure 5). A Toshiba Color Stream color television model number 27A41 was also used. The television had one S-Video Input, two Video-In and one Video-Out connections (see Figure 6). A JVC Super VHS player/recorder with one S-Video-In and one S-Video-out connection along with one audio/video in and out connections was used (see Figure 7). The video media device used was a Sony color video camera (Model: CCX-Z11) that fed the AR video to the CPU to display the images (Figure 7). A manually operated turning platform was used to rotate the object's Y-axis to display the dimensions of the work piece. The software that was used for AR functionality was ARToolKit v. 2.431 from the University of Washington. All of the apparatus functioned together for each work piece in the experiments (see Figure 8).

The self-screening vision tester was a measuring tool that contained an optical mirror that was capable of sliding closer to and farther from the participant's eye. The Brief Visuospatial Memory Test–Revised was a certified test from the Psychological Assessment Resources, Inc. by Ralph H. B. Benedict, PhD that measured the participant's visuospatial memory. The experimenter used a short version that consisted of four sub tests. The participants used a number two pencil and four sheets of blank paper to record the answers.

Figure 5. Display of the CPU and equipment

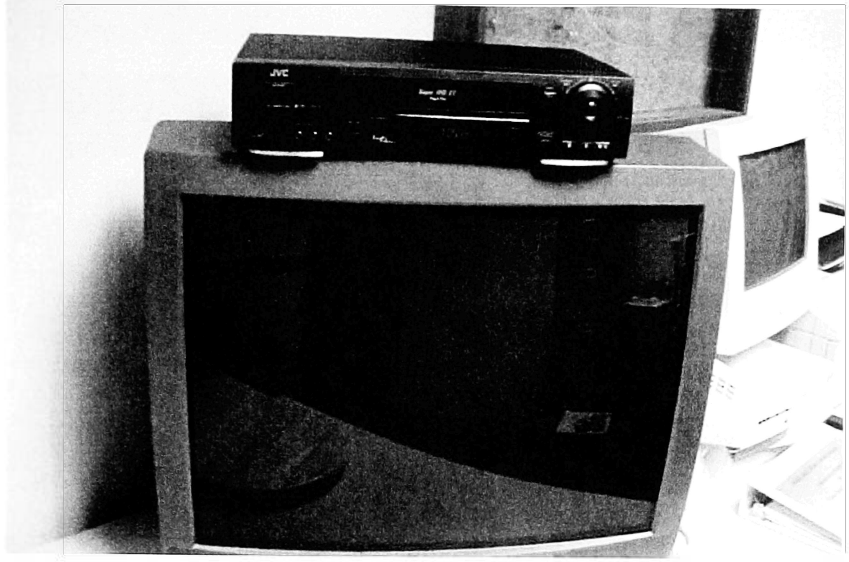

Figure 6. Television and VHS equipment for participant viewing

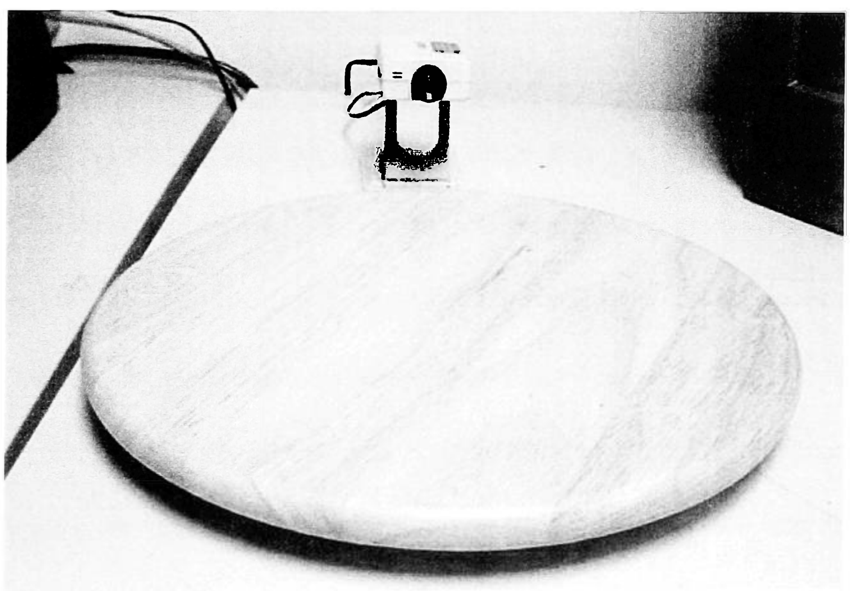

Figure 7. Camera for recognizing Augmented Reality patterns

Figure 8. Complete display of the equipment for Augmented Reality

Performance Measures

Data was collected to determine the effectiveness of monitor-based AR by conducting a training session and administering two tests to see if monitor-based AR would result in greater learning and recall with respect to inspection tasks and visual perception of each part of the work piece (solenoid valve). Each participant was administered a short term recall test and a long term recall test to determine the level of retention for each part over differing periods of time. The short-term recall test was administered immediately after the experimental session was finished. The long-term recall test was administered one week after the experiment was performed. Both tests consisted of ten multiple-choice questions, and five matching questions. The long-term recall test was given by the experimenter one week after the first test, and was administered by emailing a copy of the test to the participants. The participant would then email the completed long-term recall test to the experimenter the day it was received. Both tests were identical in content with respect to questions about inspection and visual perception information generated by the training session.

Design

The experimental design was a 4 x (2 x 24) mixed design. There was one between-subjects independent variable, the mode of information presentation. This variable was broken up into four factors, video-based presentation (Observe group), video-based interactive presentation (Interact group), augmented reality presentation (Select group), and text-based presentation (Print group). The second independent variable was a within-subjects variable, length of time between instructional session and recall test. There were two levels of this variable, short-term recall test, and a long-term recall test. There was one dependent variable, amount of information correctly recalled, measured by the percentage score of each of the two recall tests.

Procedure

Each participant completed the consent form, demographic survey, Brief Visuospatial Memory Test, and the visual perception test. The work piece used in the experiment was from NASA, Kennedy Space Center Florida, and was a Solenoid Valve that regulates Nitrogen flow on the Space Shuttle, (Figure 9).

The experimental treatment began following the completion of the visuospatial test. Participants were randomly assigned to a treatment group, and given instructions on how to use the equipment provided to their training group. The Observe group underwent video training, so they were given instructions on the use of the particular VCR with which they were provided. The Interact group underwent video-based interactive training. They were given instruction on the use of the computer to bring up text boxes explaining the work-piece functions as the video training ran on the computer monitor. The Select group underwent video-based AR training. They were given instructions regarding how to interact with the computer to find information on the functions of the work-piece. Lastly, the Print group was given print-based learning tools. They were given instructions on the nature of the text they were reading and appropriate illustrations.

The four groups then went through the instructional session, learning about the terminology, functions, and locations of the work-piece (Solenoid Valve supplied by NASA) and its components. The participants were then given a recall test to measure how much knowledge they acquired from the instructional session. This test was scored on a zero through one hundred percent scale, with one hundred percent being a perfect score, much like the scale found in academics.

Following the short-term recall test, a short interview was conducted to debrief the participants and record their opinions on the instructional mode they experienced. This concluded session one.

The last session, session two, was conducted exactly one week later. Participants were emailed the same recall test as in session one to measure how much information the participant retained after one week without any rehearsal. Participants emailed their answers back to the experimenter. The test was also scored on the same percentage scale as the test taken immediately after the instructional session. This concluded session two, and the experimental testing.

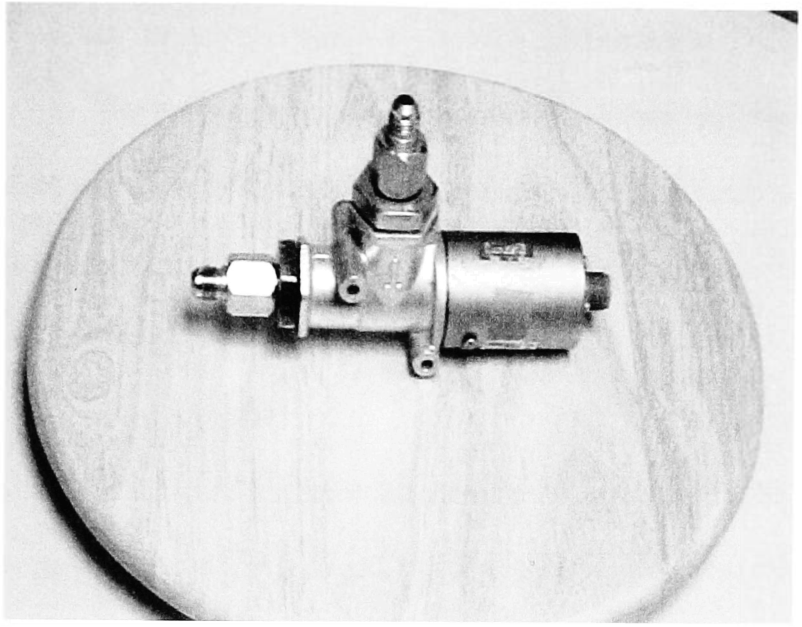

Figure 9. View of the Solenoid Valve on a turntable

Data Collection

Data was collected on the two tests administered during the course of the experiment. The test scores were analyzed using analysis of variance (ANOVA).

Independent Variables. The first independent variable for the experiment was mode of information presentation. The second independent variable was the length of time between the instructional sessions and long term recall. The length of time between instructional sessions was a within-subject variable.

Dependent Variable. The dependent variable that was measured was the amount of information recalled about the work piece. The scores from each of the two tests for each mode was measured and compared to each of the four modes.

Data Analysis. Experiment data was analyzed with an analysis of variance (ANOVA) controlling spatial ability to determine the paradigm that served as a tool for learning purposes.

RESULTS

The data from the short-term recall and long-term recall tests were collected from a total of 96 participants between the group modes (Print, Observed, Interact, and Select). Each group had 24 participants that successfully completed both tests. The data was analyzed using analysis of variance (ANOVA).

The main effect of mode of presentation was examined and the following results were found. The means for the Print, Observe, Interact and Select modes for the short-term recall test were 9.396, 9.683, 9.104, and 8.875 respectively. These means do not differ significantly when analyzed with a between subjects ANOVA, $F(3,92) = 0.127$, $p = .0.944$. Mode of presentation did not significantly affect amount of recall and are shown in Table 1.

The main effect of tests was examined and the following results were found. The means for the short-term and long-term recall tests were 9.500 and 8.729 respectively. The short-term recall and long-term recall tests were significantly different when analyzed with a within subjects ANOVA, $F(1,92) = 12.226$, $p < 0.001$ and are shown in Table 1. This indicates that the mean scores from the short-term and long-term tests had different amounts of recall from the time span of immediate testing and the experimental session to one week.

The effect of the interaction of mode of presentation and tests was examined and the following results were found. The mean for the Print, Observe, Interact and Select modes in were 9.833, 8.958, 9.292, 8.875 for the short term recall test, and 9.458, 8.750, 9.417, and 8.333 for the long term recall test, respectively. These means do not differ significantly when analyzed with an ANOVA, $F(3,92) = 0.408$, $p = 0.748$. The interaction of mode of presentation and tests did not significantly affect amount of recall shown in Table 1.

The amount of recall retained from each mode of presentation is dependent on the interaction between modes and their test scores, which showed to have not significance. The standard deviation between each of the modes of presentation for each test does not differ greatly. The mean differences between the two tests for each mode of presentation are almost parallel (Figure 10). This explains that the only significance found was between the two tests in each mode of presentation. The amount of recall from the short-term recall test to the long-term recall test in each mode was similar compared to all modes of presentation.

TABLE 1.

Mixed Design ANOVA

Source	DF	SS	MS	F	Sig
Mode of Presentation	3	6.604	2.201	0.127	0.944
Tests	1	28.521	28.521	12.226	0.001
Mode of Presentation X Tests (Interaction)	3	2.854	0.951	0.408	0.748

TABLE 2.

Descriptive Statistics

AR Group		Mean	Std. Deviation	N
Short term recall				
	Print	9.83	3.171	24
	Observe	9.29	3.057	24
	Interact	9.46	3.036	24
	Select	9.42	2.552	24
	Total	9.50	2.924	96
Long term recall				
	Print	8.96	3.483	24
	Observe	8.87	3.803	24
	Interact	8.75	2.908	24
	Select	8.33	2.929	24
	Total	8.73	3.259	96

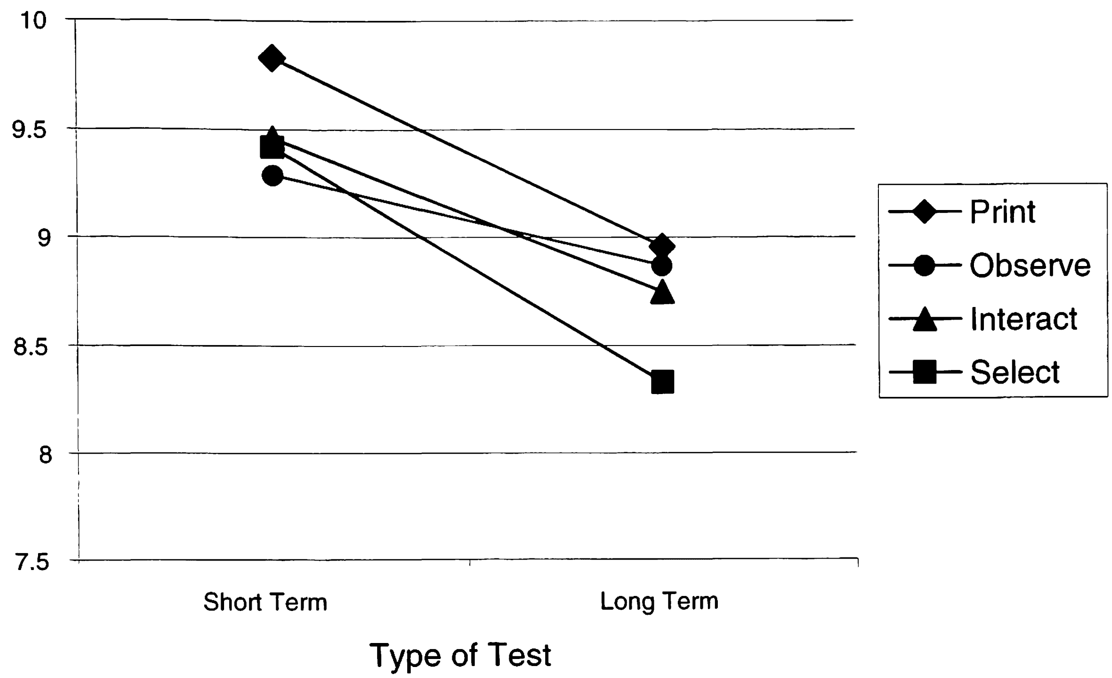

Figure 10. Test Mean Results for Short Term and Long Term Recall Tests for each mode condition.

DISCUSSION

The hypothesis stating the Select mode would result in greater recall of the annotations than any other mode of presentation was not supported. There were some confounds and uncontrolled variables that could have changed the data to produce unexpected results such as experiment procedures, presentation of each test and the AR environment itself. Some of the procedures performed during the experiment could have been done differently. The participants in each mode could have performed an inspection task with each component after the experiment session. The questions could have more spatial content to each component and an oral inspection test on each component. The experimenter could present the disassembled work piece to the participant and then have the individual answer questions verbally instead of written. This could help with spatial recall of the work piece because it allows more interaction with the participant. The AR environment itself could have increased amounts of unexpected data by unintentionally induced another training. The unexpected data could have come from the presentation time of the annotations from hardware limitations, experiment procedures, the population sample, or the different experimenters conducting and maintaining the experimental environment.

The two most important requirements for an AR system to display objects and annotations in a defined spatial area of a real-world environment are high registration and sensing requirements. One of the biggest problems with AR systems is the measurement of movement between the tracking camera and the rendered images of the registered position are fully displayed to the user known as system latency or lag time. If an AR system has a noticeable amount of lag time showing the annotation or objects, it may decrease the amount of recall and learning in the experiment. The SGI OS2 computer did produce lag time caused by rendering low frame rates and reaction time from the turntable operation, and processing time for the AR system to catch up. The rendered annotations were “swimming” around the scene in a way that did not seem to be apart of the real-environment when displayed to each individual participant. One way to decrease the amount of lag time is to use frameless rendering. Frameless rendering is a process that renders an image by continuously updating an image only when information is available instead by displaying the object or annotation per frame (Fuchs & Rolland, 2001). The ARToolKit v2.431 did not support frameless rendering during the experiment and was updating the rendered annotations per frame, which may have caused the scene to display improperly. Another possible way that lag time could have be decreased is to network one host computer and one image generating computer to help increase system processing speed and lessen the amount of processor and memory usage. The host computer can receive the real world coordinates from the tracking camera and have the registered environmental coordinated already cataloged in a database or data file. The image generator computer would take both the registered and real world coordinated from the host and display the object or annotation through a media for example a HMD or monitor. Parallel processing can be used to send data between each computer and control concurrent processing during AR sessions.

The annotations themselves were displayed in a field-of-view that was limited by what the tracking camera and resolution of the television could produce. The Sony camera could only record what was in view, and some of the work piece’s components could not be in the same view if the turn table was positioned in one spot along with the table turning in front of the camera. To compensate for this the participant had to move the tracking camera base to three positions to view a defined amount of annotations when the tracking camera scanned the environment. This additional interaction could have decreased the amount of learning because of additional learning skills needed for positioning the camera. The reason for switching the tracking camera to three positions was to bring the maximum amount of rendered component annotations to the participant's view. The annotations were colliding and the field-of-view limitations of the tracking camera created the three positions. The ARToolKit v2.431 didn't have any functionality that prevented the annotations from colliding making the text unreadable at times. Developing conditions inside ARToolKit v2.431 source code to register borders of the annotation and prevent overlapping was needed. The leader lines pointing to the components could have caused confusion by being too thin, not having enough contrast to be clearly seen by the participant, and not being positioned in a completely spatial correct manner.

The participants could have increased the amount of error in each of the four modes of presentation. The participants came from freshmen and sophomore classes that received extra credit for participation. Scores from the BVMT-R for each participant indicated that 24 out of the 96 did not have a 100% score with respect to spatial locations. This would have only affected the test scores if the percentages were lower and not spread out among the four modes of presentation. The motivation and desire to perform the experiment may have resulted in poor performance from the students. A sample pool of individuals who would have the desire to use the AR system could have increased their motivation to learn. The amount of time for the experiment was eight minutes long culminating in the short-term recall test. The amount of time may have needed to be increase for more acquisition of spatial recall for the Select and Interact modes to show some significance. The participant's learning curve for using the AR system could have affected the eight-minute experiment time to four to six minutes. Researchers could retest all the participants who performed the Interact and Select modes but with a different work piece and same experimental instructions and see how the test scores increased due to familiarity of the AR system.

The short term and long term recall tests were significantly different due to level of recall for each test in each mode of presentation. In each mode of presentation, there was a large percentage of the information retained, but the amount of recall was not significantly different when compared to each mode. The amount of information retained from the annotations decreased from short-term to long-term but when compared to the level of recall of all other modes, the amount of recall was almost parallel. The reason for the tests being significantly different was the difference in the amount of recall that was observed within one week between the tests. The decrease in information recalled shows a loss of information in each mode over time, but amount of information lost as a result of time was not significantly different when each mode was compared to other modes.

During the course of all 96 experiments, there were two different experimenters performing the Brief Visuospatial Memory Test – Revised, distributing the consent form and demographic form, and controlling the experimental session sporadically. This action was taken due to schedule conflicts with the participants and the chief experimenter. The junior experimenter was instructed and monitored by the chief experimenter for the first three to five sessions and then allowed the junior experimenter to continue by himself. The junior experimenter only collected the data and help run the experiment sessions. All of the data and statistical analysis was calculated and interpreted only by the chief experimenter. The inconsistency of not having one experimenter throughout all 96 experimental sessions may have contributed to the lack of significant data to support the hypothesis.

CONCLUSION

The results from this study indicate that the main effect for mode of presentation and the interaction of mode of presentation and tests did not produce any significant amount of difference in recall. The main effect for short –term and long-term recall tests was significant. Many questions were asked to investigate the potential effects they may have had on the results of the experiment. Even though no significant differences between any of the modes of presentation were found, the experiment has provided a first step toward better understanding and development of experimental environments and conditions to help investigate the potential of monitor-based AR. The effect of cognitive elaboration is a subtle effect that must be investigated with tightly controlled experiments. One of the causes for not having a significant amount of recall between the tests and no interaction within each group may have been because the experiment was not controlled properly. Cognition elaboration is very sensitive when an experiment's environment has variables that are changing sporadically during the experiment with out the experimenter being aware of it. To discover more potentials for monitor-based AR more research along with more control over experimental conditions, participant usage, and better AR software with more robust functionality.

REFERENCES

Azuma, Ronald T. (1997). A Survey of Augmented Reality.Huge Research Laboratories, 355-385

Azuma, Ronald T. (2001). Augmented Reality: Approaches and Technical Challenges. Fundamentals of Wearable Computers and Augmented Reality, 27- 59

Baccino, Thierry, and Pynte, Joel (1994). Spatial Coding and Discourse Models During Text Reading, Language and Cognitive Processes, 9(2), 143-155

Bostrom, Robert P, and Davis, Sidney A., and Olfman, Lorne, and Sein, Maung K. (1993). Visualization Ability as a Predictor of User Learning Success, International Journal Man-Machine Studies, 599-620

Brown, Kenneth G., (2001). Using Computers to Deliver Training: Which Employees Learn and Why, Personnel Psychology, 54(2), 271-296

Dwyer, Francis M., and Melo, Hermes De (1983). Effects of Mode of Instruction, Testing, Order of Testing, and Cued Recall on Student Achievement, The Journal of Experimental Education, 52 (1), 86-94

Fuchs, Henry, Rolland P., Jannick (2001). Optical versus Video See – Through Head – Mounted Displays, Fundamentals of Wearable Computers and Augmented Reality, 113 –153

Holloway, Richard L., (2001). Registration Error Analysis for Augmented Reality Systems, Fundamentals of Wearable Computers and Augmented Reality, 183 – 217

Lintern, Gavan ,(1991). An Introductional Perspective on Skill Transfer in Human-Machine Systems, The Human Factors Society, 33 (3), 251-266

Lovelace, Eugene A., and SouthHall, Stephen D. (1983). Memory for words in prose and their location on the page, Memory & Cognition, 11 (5), 429-434

Majoros, Anthony, and Neumann, Ulrich. (2001). Support of Crew Problem-Solving And Performance with Augmented Reality, The Boeing Company, 1-19

Majoros, Anthony, and Neumann, Ulrich. (1998). Cognitive, Performance, and System Issues For Augmented Reality Applications in Manufacturing and Maintenance, VRAIS '98, 1-9

Appendix A

ARToolKit v2.431 Verification Script

ARToolKit v2.431 Verification Script

ARToolKit has undergone many revisions just in the past year. New functionality has been added as AR technology has advanced. Conversely, other functions that are no longer needed have been deleted from the program. Therefore, the software is experimental by its very nature. This, and the fact that the software was downloaded from the Internet made it necessary to verify that functionality of this software was operational and correct, especially those functions critical to the completion of this research.

ARToolKit includes seven programs that assist the user in developing, and testing the ARToolKit installation and other user-created AR based applications. The functions include:

- SimpleTest: tests the ARToolKit, installation
- Exview: gives the OpenGL coordinates of the camera along three axes
- ModeTest: tests the linking of patterns and graphics
- Optical: switches the system calibrations between optical-based calibrations and video-based calibrations
- Mk_Patt: Creates new patterns and links them to graphics
- Calib_dist: calibrates camera lens distortion
- Calib_cparam: calibrates camera's focal point

Verification that the functions are operating properly was the first step in the experimentation. The functions were taken one by one and the results from running each executable program was compared against the expected results as specified by the user's manual, and the program's source code. After all functions have been verified operational and correct, ARToolKit will be ready to be incorporated into the experiment.

Verification for SimpleTest

SimpleTest (SimpleTest.exe) was the first function verified. This test shows whether or not the ARToolKit works properly. It uses one of the default computer graphics linked to one of the patterns included in the software package. If ARToolKit is installed properly with all programs and routines working correctly, then a blue cube will be overplayed on the real world video of the patt.Hiro pattern.

The verification procedure for SimpleTest is as follows:

1. The patt.Hiro pattern was found in the 'patterns' directory. This was printed out on a high quality laser printer for maximum resolution.

The pattern was attached to a flat, sturdy, plastic backing to keep the pattern from curling or folding during the testing. Figure 1 shows a sample pattern

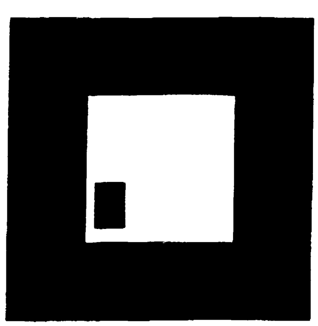

Figure 1. Sample Pattern

2. The camera was placed directly over the Hiro pattern (range = 2 3 inches). In the output window, ARToolKit showed the real-world video of the pattern and its surroundings, with the computer graphic of a blue cube projected over the inner square of the pattern. Figure 2 shows the graphic

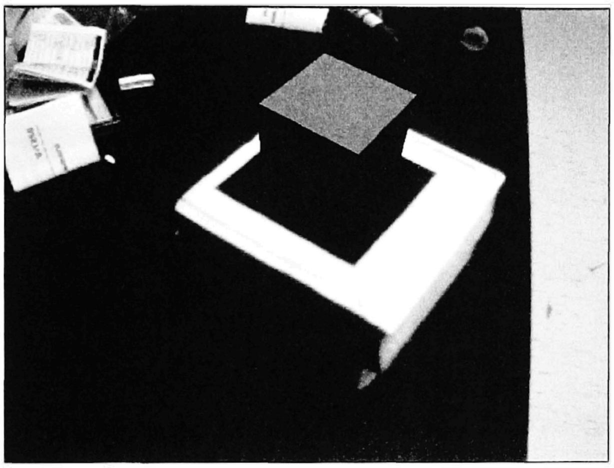
Figure 2. Graphic Picture

3. The pattern was elevated and rotated two different angles, the cube moved also, always maintaining its orientation with the pattern.

This test showed that ARToolKit was installed and configured correctly on the computer system. Furthermore, this test showed that the OpenGL and GLUT functions that drive ARToolKit were functioning properly.

<u>Verification for ModeTest</u>

ModeTest (ModeTest.exe) is an extension of the SimpleTest Program. It incorporates all four default computer graphics (sphere, cube, cone, and tours) linked to the four patterns included in the ARToolKit package. This program, in addition to testing the installation of ARToolKit, can be used as a shell to facilitate the construction of user-created applications. The source code for this program is also included allowing users to cut and paste code to create different applications. For verification purposes, the included patterns will used to validate that ARToolKit recognizes the pattern, interprets it correctly, and links the appropriate computer graphic with the real world video of the pattern. The procedure of verifying the ModeTest program is almost identical to that of the SimpleTest. (See Figure 1 and Figure 2)

The verification procedure for ModeTest is as follows:

1. The patterns patt.Sample1, patt.Sample2, and patt.Kanji were printed out and used in addition to the patt.Hiro pattern. The Sample1, Sample2, Hiro, and Kanji patterns were linked by the program to the default sphere, cube, cone, and tours graphics respectively.

2. Then each pattern was fastened to the plastic backing and placed underneath the camera.

3. The graphic that ARToolKit produced as an overlay of the pattern video was noted and the next pattern was placed under the camera. This was done for each pattern.

4. Then the procedure was carried out with two patterns together, then three, and finally all four patterns were placed under the camera at once.

ARToolKit correctly identified and linked the patterns to their appropriate computer graphics. This result was seen during singular and multiple pattern testing. With these results, it can be concluded that ARToolKit's pattern recognition and graphic display function, for single and multiple patterns, were operating correctly, Verification for Exview

The exview program (exview.exe) allows the ARToolKit user to view the position of the camera using the pattern as a target and as a reference point. From this reference point the program will plot the camera position using the X, Y, and Z-axes in the OpenGL plane, using the OpenGL coordinate system. The purpose of the program is to facilitate camera tracking when using head-mounted displays (HMD's) in AR.

The verification began with an examination of the documentation to determine the primary function, and also any options the program offered to the user. The program's source code was also examined. All functions were listed along with the desired result of their execution as referenced in the downloaded documentation and program source code.

The verification procedure for exview is as follows:

1. The program was opened, a simple double click on the exview.exe icon, which brought up a split screen.

The top window of the split screen showed a normal camera view, which performs ARToolKit's standard function of overlaying computer graphics on the video input from the camera. The bottom window is the main function of exview. This window shows a blue-green cube, which represents the camera, in an OpenGL plane. It also shows three pointers representing the orientation of the X, Y, and Z-axes in relation to the camera. Lastly, a bar below the window gives numerical coordinates for the camera along the axes.

This result matched the expected result. Therefore, the function was deemed operationally correct. Though this program will not be essential to this study, it may become very important in future, when this research turns from a stationary monitor-based AR system to using head-mounted displays.

Verification for Mk_Patt

The make new pattern program (mk_Patt.exe) is crucial to any experiment, or other uses, that ARToolKit is incorporated in. It allows the user to create any number of new patterns and link them to either the default graphics in ARToolKit, or graphics created by the user and stored in the software's graphics file.

The program's functionality was tested by creating four new patterns and linking them to the default graphics. To create new patterns, and have the program recognize them, the blankpat.gif file, found in the 'util' directory, was graphically edited to produce the preconceived pattern designs. Graphically editing the blankpat.gif file is the entire pattern design phase. After this was accomplished the new patterns were printed out with a high quality printer for maximum resolution.

The verification procedure for mk_Patt is as follows:

1. For the program to recognize the patterns, they were linked to the default graphics (sphere, cube, cone, and torus).

2. To accomplish this, the directory was changed to the 'bin' directory, and the command './mk_patt' was executed. The program asked for a camera parameter data file. The default path and filename prompted by the computer was used in this case.

3. Next, the camera was positioned directly overtop of one of the new patterns (range: 2-3 inches) until a green and red border appeared on the video output screen, in this case, a computer monitor. Figure 3 shows the red and green borders of a pattern.

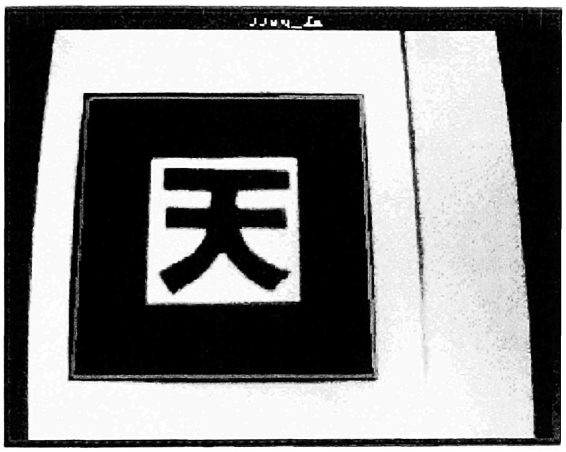

Figure 3. Camer Verification Borders

4. The left mouse button was clicked only when the red border is at the left and top of the pattern. Then, the patterns were given names, usually in the form: patt.*NAME*. For example, the four patterns just created were named patt.K, patt.C, patt.B, and patt.G.

5. After all four patterns were read into the computer the right mouse button was clicked to exit.

6. The patterns were moved to the 'Data' directory by typing 'mv patt.*NAME* Data\patt.*NAME*'.

7. The directory was then changed to the 'Data' directory.

8. Next, using a text editor, the 'Object Data' file was edited. The format of the patterns included in the file was followed.

9. Then, the total number of patterns was changed by changing the integer number at the top of the file. Also, the name of the graphics files were changed to link patterns K, C, B, and G to graphics of the sphere, cube, cone, and torus, respectively.

10. These edited files did not need to be recompiled. The directory was changed to the 'bin' directory, and the mode test function was run.

The mk_patt function worked perfectly. ARToolKit recognized each new pattern, and the correct graphic was portrayed on the monitor for each of the four patterns. With these results, the function was deemed operational and correct.

Verification for Optical

The optical.exe function is intended for use with optical or video-based head mounted displays. It allows the user to switch between these two modes, which is important as the camera and computer calibrations are different for each mode.

This function will not be used in this experiment as the subjects will be using a monitor-based AR system. Therefore, we do not have the need for this function, or, more importantly, we don't have the proper equipment to adequately test the optical executable file.

Verification for Program calib_dist

There are two programs that are included in ARToolkit 2.413 that are responsible for camera calibration. The values are stored in the camera parameter file called camera_para.dat. The toolkit comes with default values that are standard for all types of cameras. The camera calibration routines that are used are calib_dist and calib_cparam. Each routine requires a pattern that comes with the toolkit. The patterns are pdf format and need to be printed out for the calibration. The program calib_dist is used to calculate and measure the image center point and camera lens distortion.

The pattern must be on a sturdy piece of cardboard or poster board to give the pattern flat surface for calibration. The calib_dist program must be run first because the calib_cparam requires x, y, and z coordinates and other parameters to complete full camera calibration. Figure 4 shows the pattern

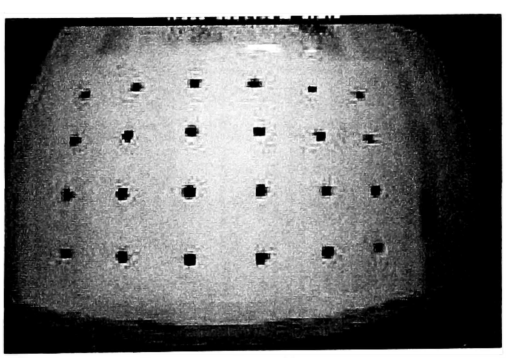

Figure 4. Calibration Pattern

The verification procedure for calib_dist is as follows:

1. Executed the program and captured image perpendicular to pattern by left clicking mouse.

2. After capturing image I then clicked with my mouse every dot starting with the top left to the right for every column.

1	2	3	4	5	6
7	8	9	10	11	12
3	14	15	16	17	18
19	20	21	22	23	24

3. During every dot clicked a red cross would be displayed on the dot and the window would verify each dot clicked on. After all the dots are clicked, the left button is pressed to store and unfreeze the camera. Figure 5 shows the procedure

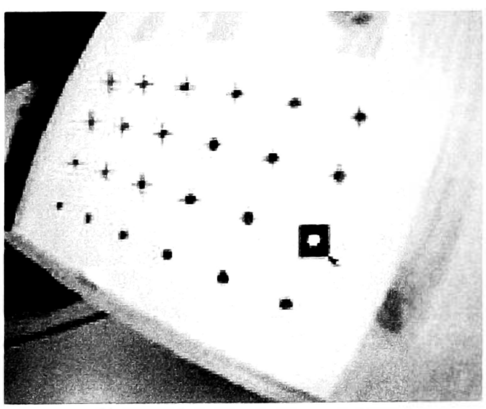

Figure 5 Calibration Sample Shot

4. The following process requires five to ten screen captures along with different screen shot angles with differences in pitch, roll and yaw. But all screen shots must display all 24 dots clearly and an estimate of 500 mm from the pattern. Figure 3.3 display two different screen shots.

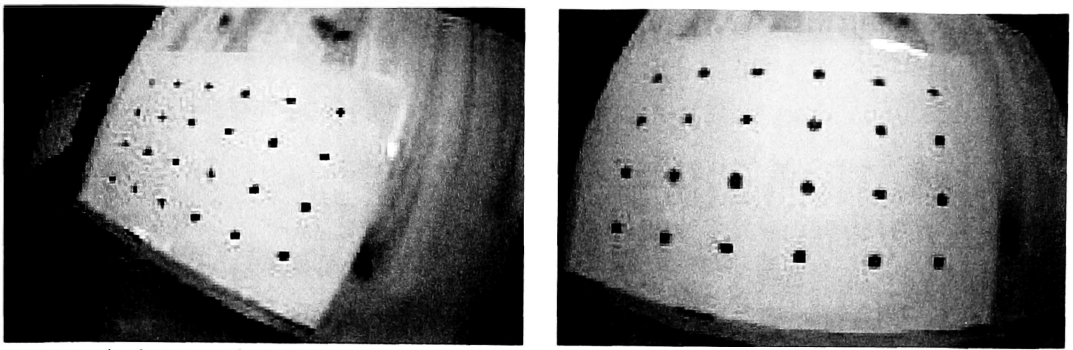

Figure 6. Calibration Sample Angle Shots

5. After the five screen shots, click on the left mouse button to perform calculations. The process took three to four minutes. Then the screen displayed four values. Center x, Center y, Dist Factor, and size adjustment. These values were recorded.

6. The program also allowed a visual review of each of the screen shots taken along with red lines that were connected through each point on the pattern. Figure 7 shows the results.

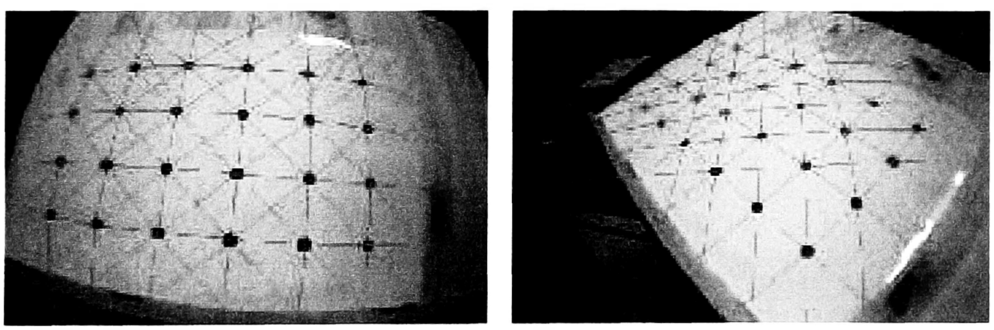

Figure 7. Calibration Procedure Screen Shots

The procedure for calib_dist did calculate and displayed values for each of the screen shots and coordinates needed for calib_cparam. I didn't discover any functional or systematic problems with the program. I did notice a difference in the 2.11 version of the SGI manual. It didn't describe the size adjust variable outcome. The problem was not with the program, but with the discrepancy with the manual. The final result with the calibrated camera will only be verified until the completion of calib_cparam.

Verification for Program calib_cparam

The other program to be verified is calib_cparam. This program uses a pattern that has seven horizontal and nine vertical lines. The calib_cparam program calculates the camera's focal point and other system parameters.

The pattern must be on a sturdy piece of cardboard or poster board to give the pattern flat surface for calibration. Figure 1.1 shows the patterns. The calib_cparam program will use the calib_cparam x, y, and z coordinates, dist facto, and size adjusts numbers to complete the camera calibration. Figure 4.1 shows the pattern

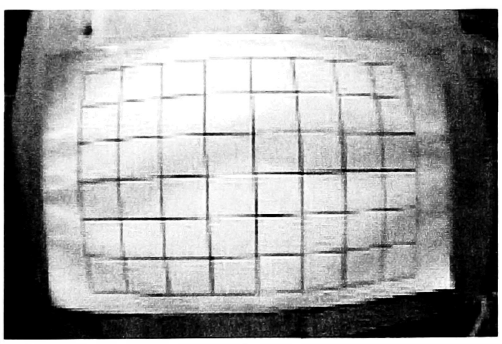

Figure 8. Verify Calibration is Correct

The verification procedure for calib_cparam is as follows:

1. Executed the program and captured image perpendicular to pattern by left clicking mouse. The window will ask for the four values computed from calib_dist. All values were entered.

2. A horizontal line appears on the captured screen. After capturing image I then use the arrow keys on the keyboard 9.2)

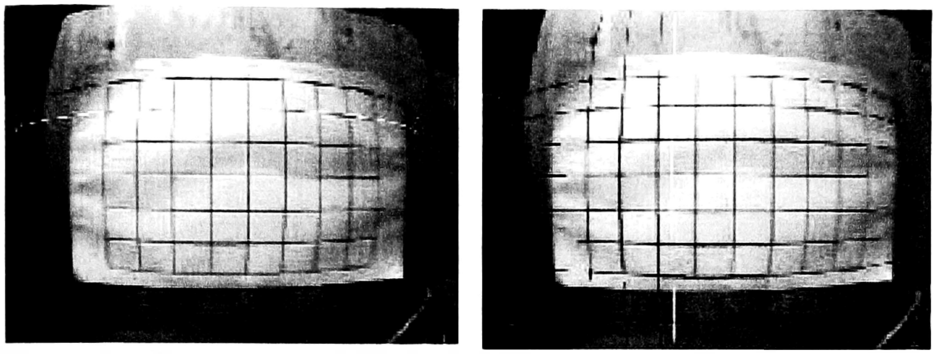

Figure 9. Horizontal and Vertical Calibration Procedure Screen Shots

3. During every horizontal line clicked, the line is displayed on the pattern After all the horizontal lines are set on each of the grid, the left button is pressed to store and unfreeze the camera. The ordering of the lines must be top to bottom.

4. The following process requires five to ten screen captures along with different screen shot angles with differences in pitch, roll and yaw. But all screen shots must display all 24 dots clearly and an estimate of 10 inches from the pattern.

5. The process is then repeated for each vertical line on the pattern. Figure 10 shows the vertical lines.

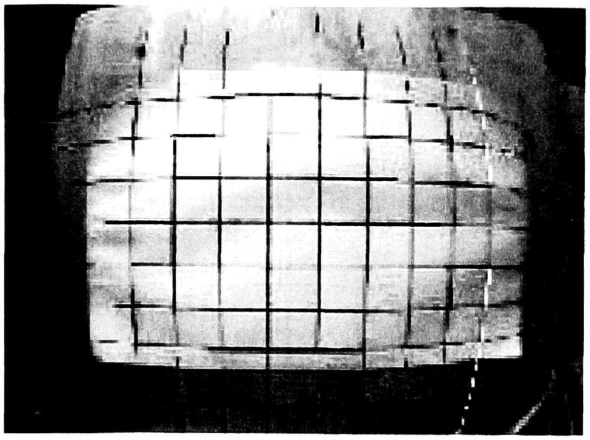
Figure 10 Repeat Vertical Line Calibrations

6. A file name was requested to save the new calibration values.

7. The new file is now in the bin/Data director under the ARToolkit2.431.

The calib_cparam program performed exactly what the instruction described. To verify the new camera parameter, I then executed the simpleTest program. The tracking and movability of the simpleTest pattern increased. Before we could only produce the OpenGL image if the camera was no less than 45 degrees to recognize the pattern. Now we can have the camera at an angle 10 degrees or greater to produce an image. There are no measurable values for the calibration except by comparing the default parameter values from ARToolKit download to the new parameters calculated. This verification was only done the minimum five times. I suggest redoing the calibration and capture each pattern eight to ten times.

All executable programs and the functions and subroutines associated with them have been found to be operating correctly according to their intended purpose as stated in the downloadable documentation and the source code. ARToolKit will be used to develop one of the experimental treatments later in the study.

Appendix B

AR Instructional Session Script

Observe Mode Instructional Session

Below is a list of instructions that you need to follow to perform the experiment. Please follow each direction precisely and to it is fullest. After reading, the instructions please ask questions if you do not understand them.

- You will have 8 minutes to watch the video presented to you.
- The video itself is not 8 minutes long but you can use the Fast Forward, Rewind, Pause, and Play button through the experiment.
- The starting position will be 0:00 on the VCR and will go to 4:30. Please do not go past each of the two numbers during the experiment.
- Please watch and view the contents of the video.
- The experimenter will stop the video after the 8 minutes has expired. Any questions?

Print Mode Instructional Session

Below is a list of instructions that you need to follow to perform the experiment. Please follow each direction precisely and to it is fullest. After reading, the instructions please ask questions if you do not understand them.

- You will have 8 minutes to view each of the 10 handouts.
- Read the text and study the picture
- You may flip back and forth through the 10 pages as many times as possible
- The experimenter will ask you to stop after the 8 minutes has expired. Any questions?

Select Mode Instructional Session

Below is a list of instructions that you need to follow to perform the experiment. Please follow each direction precisely and to it is fullest. After reading, the instructions please ask questions if you do not understand them.

- You will have 8 minutes to do the entire experiment.
- You will be able to move the turntable left and right with the work piece on it along with the camera to view each pattern..
- The experimenter will explain the three position that the camera will be moved to
- You will move the turntable and the camera for each of the patterns to be displayed on the television.
- Please watch and view the contents on the television.

Position 1 is for the following patterns

Patterns 7 5 J F

Position 2 is for the following patterns

Patterns 6 G

Position 3 is for the following patterns

Patterns R L 4 2

Interact Mode Instructional Session

Below is a list of instructions that you need to follow to perform the experiment. Please follow each direction precisely and to it is fullest. After reading, the instructions please ask questions if you do not understand them.

- You will have 8 minutes to do the entire experiment.
- The video itself is not 8 minutes long but you can use the Fast Forward, Rewind, Pause, and Play button through the experiment.
- You will use a VCR remote to Play, Fast Forward, Rewind, and Stop the video tape
- The starting position will be 0:00 on the VCR and will go to 3:10. Please do not go past each of the two numbers during the experiment.
- While watching the video on the Television screen you will click and hold down the left mouse button to have information appear on the screen. Depressing the left mouse button will have the information disappear on the screen.
- Please watch and view the contents on the television.

Appendix C

Brief Visuospatial Memory Test – Revised Instructions

1) **Say before slide 1 is shown:**

Turn to page T-1 in your booklet. I am about to show you a slide that has six figures on it. I want you to study the figures so that you can remember as many of them as possible. You will have just 10 seconds to study the entire display. I will present the figures right here *(examiner points to screen where figures will be displayed).* After I take the display away, try to draw each figure exactly as it appeared and in its correct location on the page. Are there any questions about these instructions?

2) Reread and clarify directions as much as necessary.

3) Show slide 1 for exactly 10 seconds. **Do not begin timing until subjects are scanning stimulus.**

4) **Say after 10 sec. period:**

Now draw as many of the figures as you can in their correct location on the page.

Give them as much time as they need.

5) **After all subjects are finished say:**

That was fine. Now I would like to see whether you can remember more of the figures if you have another chance. I will present the display again for 10 seconds. Try to remember as many of the figures as you can this time, including the ones you remembered on your last attempt. Try to draw each figure precisely and in its correct location. Are there any questions about these instructions?

6) Reread and clarify directions as much as necessary.

7) Show slide 2 for exactly 10 seconds. **Do not begin timing until subjects are scanning stimulus.**

8) **Say after 10 sec. period:**

Now draw as many of the figures as you can in their correct location on the page.

Give them as much time as they need.

9) **After all subjects are finished say:**

That was fine. Now I would like to see whether you can remember more of the figures if you have another chance. I will present the display again for 10 seconds. Try to remember as many of the figures as you can this time, including the ones you remembered on your last attempt. Try to draw each figure precisely and in its correct location. Are there any questions about these instructions?

10) Reread and clarify directions as much as necessary.

11) Show slide 3 for exactly 10 seconds. **Do not begin timing until subjects are scanning stimulus.**

12) Say after 10 sec. period:

Now draw as many of the figures as you can in their correct location on the page, and write down the time when you are finished.

Give them as much time as they need.

13) After all subjects are finished say:

Try not to forget the display because I may ask you to remember the figures later.

14) Take response booklets from subjects. Pass out reading task. Have subjects engage in predominantly verbal tasks (reading task?) for 20 minutes.

15) Pass out response sheets **opened to page DR.**

16) Say:

Remember the figures I showed you before? I want to see how many you can remember now. I know it sounds difficult, but try to draw as many of the figures as you can in their correct location on the page. Remember try to draw them accurately. When you are finished write the time down on your sheet. Just do the best you can.

Appendix D

Demographic Survey

Embry-Riddle Aeronautical University
Graduate Thesis Experimental Demographic Survey

Please fill in the scantron sheet correctly and accurately with a No 2 pencil. For questions one and two, please write on the appropriate space at the top of the scantron sheet. Please remember the assigned number at the top right hand corner of this sheet. Please put this number on the scantron labeled Period. If selected for the experiment we will track you as this number for your own privacy. We ask for complete honesty for all questions. None of this information will be viewed to any outside students or faculty members for the survey.

1. Name

2. Today's date

3. Home Phone Number

4. Age
   A) 17 to 20 years
   B) 21 to 25
   C) 26 to 30
   D) 31 and over

5. Current year in College
   A) Freshmen
   B) Sophomore
   C) Junior
   D) Senior
   E) Senior extended

6. What is your major?
   A) Engineering
   B) Aeronautical Science
   C) Communications
   D) Human Factors
   E) Aviation Maintenance

7. Sex
   A) Male
   B) Female

8. How often do you use the computer daily?
   A) 10 to 8 hours
   B) 7 to 5 hours
   C) 4 to 2 hours
   D) 1 hour or less

9. Do you get motion sickness while looking at a fixed point, such as a computer screen?
    A) Yes
    B) No

10. Are you claustrophobic, if put in a small room?
    A) Yes
    B) No

11. Do you normally eat 3 meals daily?
    A) Yes
    B) No

12. Did you eat breakfast/ lunch (if applicable)/ dinner (if applicable) today?
    A) Yes
    B) No

13. How much did you sleep last night?
    A) Less than 4 hrs.
    B) 4 – 5 hrs.
    C) 6 – 7 hrs.
    D) 8 – 9 hrs.
    E) 10 or more hrs.

14. Are you currently taking any medication that affects your, attention and/or concentration?
    A) Yes
    B) No

15. Do you know specifically what a mechanical valve is
    A) Yes
    B) No

Appendix E

Consent Form

Embry-Riddle Aeronautical University
Graduate Thesis Experimental
Participant Consent Form

Date: ____________________

When I sign this statement, I am giving my informed consent to the following basic considerations:

I understand clearly the procedures to be done, including any that might be experimental. The participant will complete the assigned experiment and execute all procedures set in the given program.

I understand clearly any discomforts and/or risks that might be associated with this research project. The Experiment will be done in a controlled environment. The participant will be in a closed room with only the testing equipment

I understand clearly any benefits anticipated from this research project. Each participant will receive bonus points at the completion of the experimental trials added to their class with permission from their instructor.

I understand that provisions have been made to protect my privacy and to maintain the confidentiality of data acquired through this research project. All participants' results will be referenced with a four-digit code that will only be referenced by the experiment team and not by any outside students or professors in and outside Embry-Riddle Aeronautical University.

The experimenter, Chris Opalenik (756-6956), has offered to answer my questions about the procedures. He can also be contacted for further information about this research project. The principal investigators/ supervisors are Dr. Sathya N. Gangadharan (226- 7005), and Dr. Dennis A. Vincenzi (226-7035).

I understand clearly that I may withdraw at any time from this research project without penalty or loss of benefits to which I am otherwise entitled.

I am not involved in any agreement for this project, whether written or oral, which includes language that clears the institution from alleged fault or guilt. I have not waived or released the institution or its representatives from liability for negligence, if any, which may arise in the conduct of the research project.

I, the person signing below, understand the above explanations. On this basis, I consent to participate voluntarily in the Effectiveness of an Augmented Reality Training Paradigm.

____
Signature of person giving consent

____
Signature of principal investigator

Appendix F

Recall Test

I. In the following questions below, choose the correct inspection task for each of the Solenoid Valve components. Each question will be testing position of the components, appearance characteristics, and mechanical attributes

1. The Inlet Fitting Assembly Nozzle connects to
    a. The Inlet and Outlet Fitting Assembly Bolt
    b. The Inlet Fitting Assembly Base
    c. The Body Valve
    d. The Inlet and Outlet Fitting Assembly Bolt and The Inlet Fitting Assembly Base

2. The Solenoid has how many threaded parts?
    e. only one small threaded end
    f. one small threaded and 1 large threaded end
    g. no threaded ends
    h. 2 identical threaded ends

3. The Spring is
    i. 1 inch long with multiple coils
    j. 1/2 inch long with only 4 coils
    k. not a true spring by definition
    l. 2 inches long

4. The Plunger
    m. connects to the Outlet Fitting Assembly Bolt
    n. is spherical in shape and has a metal ring around the end
    o. has a plastic O ring around its middle and is cylindrical
    p. part does not exist

5. The Outlet Fitting Assembly Bolt is female threaded and goes into the
    a. Outlet Fitting Assembly Base
    b. Inlet Fitting Assembly Base
    c. Plunger
    d. Solenoid

6. What type of shape is the Spacer?
    a. Rectangular and two inches long
    b. Spherical and 1/2 inch long
    c. Oval and 1/2 inch long
    d. Cylindrical and 2 inches long

7. The Inlet Fittng Assembly Base has how many threaded ends?
    a. 1 male threaded end
    b. 1 female threaded end
    c. 2 male threaded ends
    d. 2 female threaded ends

8. The shim goes around the
    a. large male threaded end of the Solenoid
    b. body valve
    c. Inlet Fitting Assembly Base
    d. Inlet Fitting Assembly Bolt

9. The Body Valve has how many connections
    a. 2 male and one female
    b. 3 male and no female
    c. 3 females and no males
    d. No connections

10. The Inlet Fitting Assembly Bolt is female threaded and goes into the
    a. Inlet Fitting Assembly Base
    b. Outlet Fitting Assembly Base
    c. Plunger
    d. Solenoid

II. Match up the correct part from the diagrams below for each of the components by writing the number on the diagram corresponding to the part. (One diagram will be attached to the test with blanks for the subject to write the letters in)

11. Solenoid
12. Spring
13. Plunger
14. Shim
15. Inlet Fitting Assembly Base