A Chromakey Augmented Virtual Environment for Deployable Training

Rudolph P. Darken, CDR Joseph A. Sullivan
Naval Postgraduate School
Monterey, California
darken@nps.navy.mil, jasullivan@nps.navy.mil

MAJ Mark Lennerton
U.S. Naval Academy
Annapolis, Maryland
mjlenn@usna.edu

ABSTRACT

In our attempts to construct virtual environments that simulate certain aspects of the real world, several significant technical shortfalls have limited our ability to elicit human behavior in the VE that approximates human behavior in the real world. Among these shortfalls are display issues and the insertion of the trainee in the training environment. Observable differences between the training environment and the real environment can severely limit the ability to induce stress and consequently to train complex tasks. These issues are exacerbated in a deployed setting where footprint matters most.

Our training problem involves skilled helicopter pilots and the ability to train at sea. The objective is to construct a training system suitable for deployed use that is low cost, small footprint, and that can be shown to have quantitative value as a training device. This paper will describe a mixed reality appended trainer solution that uses a Chromakey technique to mix the real environment inside the cockpit with the simulated environment outside the cockpit. This apparatus allows the near-field environment, including cockpit displays, maps, and controls, to "pass through" while the view out the window is replaced with a virtual simulation. Our prototype has been integrated with JSAF and interoperates with other simulators under development in our program. It will then describe an initial experiment that determined if experienced helicopter pilots could effectively navigate using this system and if their performance was reasonably similar to their performance in an actual helicopter. Results indicate that performance in the trainer does approximate actual performance within real world threshold values and that techniques for overland navigation used in actual flight also apply directly to navigation using the simulator. Future work includes adding NVG capabilities and further experimentation to determine the extent of training transfer possible with the system.

ABOUT THE AUTHORS

Rudolph Darken is an Associate Professor of Computer Science and a Technical Director of the Modeling, Virtual Environments, and Simulation (MOVES) Institute at the Naval Postgraduate School in Monterey, California. He is the Chair of the MOVES Curriculum Committee and directs the Laboratory for Human Performance Engineering. His research has been primarily focused on human factors and training in virtual environments with emphasis on navigation and wayfinding in large-scale virtual worlds. He is a Senior Editor of PRESENCE Journal, the MIT Press journal of teleoperators and virtual environments. He received his B.S. in Computer Science Engineering from the University of Illinois at Chicago in 1990 and his M.S. and D.Sc. degrees in Computer Science from The George Washington University in 1993 and 1995, respectively.

Major Mark Lennerton is an Instructor of Computer Science at the United States Naval Academy and a recent graduate of the Naval Postgraduate School. He was commissioned in 1990 as a U.S. Marine and underwent Naval Aviation Flight Training at NAS Pensacola later being designated a CH-53D Sea Stallion pilot. Additionally, he was designated a Forward Air Controller and deployed with the 13th MEU aboard ship to the Middle East in 1996. His research interests delve into providing compelling task oriented military training in deployable virtual environments. He received his B.S. in Computer Science from Keene State College in 1990 and an M.S. degree in Computer Science from Naval Postgraduate School in 2002.

CDR Joseph Sullivan is an instructor at the Naval Postgraduate School in the Department of Computer Science and deputy director of the Modeling, Virtual Environments, and Simulations (MOVES) Institute. CDR Sullivan graduated from Catholic University of America in 1986 with a Bachelor of Science in Computer Science, attended Aviation Officer Candidate School, and was commissioned in 1986. After completing his initial tour at Helicopter Antisubmarine Squadron Five (HS-5), CDR Sullivan was assigned instructor duty at Helicopter Antisubmarine Squadron Ten (HS-10). At HS-10, he trained student pilots for carrier-based helicopter missions, including low-level terrain navigation. After a tour on the USS New Orleans (LPH-11) as member of the Air Department, he was assigned to the Naval Postgraduate School where he earned a Masters of Science in Computer Science. His most recent operational tour was as Operations Officer at Helicopter Antisubmarine Squadron Eight (HS-8) where he was responsible for the training and readiness of all aircrew. As a member of the MOVES Institute, CDR Sullivan is pursuing a doctorate and leads all efforts in rotary wing training systems including prototype development and training assessment. CDR Sullivan developed the original Map Interpretation and Terrain Association Virtual Environment System (MITAVES) that was tested at HS-10 in 1997 and is the basis for all helicopter training systems efforts at the Naval Postgraduate School.

A Chromakey Augmented Virtual Environment for Deployable Training

Rudolph P. Darken, CDR Joseph A. Sullivan
Naval Postgraduate School
Monterey, California
darken@nps.navy.mil, jasullivan@nps.navy.mil

MAJ Mark Lennerton
U.S. Naval Academy
Annapolis, Maryland
mjlenn@usna.edu

INTRODUCTION

In our attempts to construct virtual environments that simulate certain aspects of the real world, several significant technical shortfalls have limited our ability to elicit human behavior in the VE that approximates human behavior in the real world. This is a particularly serious problem for training systems and one that merits our attention.

Our training problem involves skilled helicopter pilots and the ability to train at sea. This paper will describe a mixed reality solution that uses a Chromakey technique to mix the real environment inside the cockpit with the simulated environment outside the cockpit. It will then describe an initial experiment that determined if experienced helicopter pilots could effectively navigate using this system and if their performance was reasonably similar to their performance in an actual helicopter.

Limitations of Contemporary Displays

A common and largely unresolved issue in the design and implementation of VEs involves the choice of display technology and its suitability to the requirements of an application. Early VEs began with head-mounted displays (HMDs), but these were problematic for a variety of reasons. Early HMDs were very low resolution, often low field of view, had cables that the wearer had to avoid, and tended to be heavy and bulky to wear. These were their obvious drawbacks.

An aspect of the problem that received lesser attention was the issue of the representation of user's own body. In a typical HMD application, the user had no body representation except possibly a disembodied hand if a glove device was being used. This shortcoming was accepted at the time partly because it was not known how to reintroduce the body into the VE and partly because it was not known if this directly caused a performance deficit or if the deficit could be overcome in some way. If the tasks to be performed in the VE primarily involved distant objects, not being able to see one's hands or body did not seem to be a critical concern. However, when tasks involved near-field haptics or manipulation of objects within arms reach, an inaccurate or absent representation of the hands and arms tends to make the task all but impossible to perform.

When the CAVE was introduced in 1993 (Cruz-Neira, Sandin, & DeFanti, 1993), it addressed some of these issues. The user now had a body that was viewed to be in the virtual space. Most head-worn devices were either removed or at least minimized without giving up stereoscopy if required by the application. However, CAVEs and other projection-based technologies suffer on two important fronts; they are vastly more expensive then their HMD counterparts (e.g. the C6™ Immersive Virtual Reality System, MechDyne Corp. and Iowa State University), and although the user's body is now in the virtual space, interacting with near-field objects remains problematic. Since the user sees his actual hand and arm while interacting with purely virtual objects, occlusion is a serious problem. If the user places his hand behind a virtual object that is located within the confines of the projection surfaces, the hand can still be seen. Furthermore, if the objects are purely virtual, there is no way to provide haptic cues to assist the user in manipulating the object. The exception to this would be cases where physical objects are placed within the CAVE.

A desirable solution would allow distant-field virtual objects to remain virtual and near-field physical objects to remain physical. This would solve the near-field haptic and occlusion problem while facilitating the advantages of VE simulation and reconfigurability.

A related approach is "passive haptics" where a full occlusion display is still used but a physical surrogate is used for the virtual objects (Meehan, Insko, Whitton, & Brooks, 2002). When the user interacts with near-field objects, a virtual object is seen, but a physical object is touched. This can have a strong impact on cognitive involvement in the task (Unguder, 2001), but spatial correlation is a critical issue. The virtual and real objects have to line up accurately to achieve this end. Also, while this solves the haptic part of the problem, it does not solve the visual aspects of not seeing one's own body while manipulating an object.

System Requirements

These are exactly the issues we had to deal with in developing a VE training system for military helicopter pilots. Our solution had to be small – small enough to eventually fit on board a ship at sea. It had to be inexpensive – inexpensive enough that there could be many of them to network together for team training exercises. It had to be reconfigurable so that we could simulate any training area of interest and that we could conduct training exercises with a variety of other simulators. Most importantly, we had to show that within the constraints of the specific training tasks of interest to this system, human performance in the simulator approximated human performance in the aircraft. While this does not, in and of itself, confirm the utility of the system as an effective trainer, it is a necessary prerequisite to the training transfer studies that will follow.

A critical aspect of this is the presence of task-related stress. The real world tasks for which this trainer is intended are inherently stressful to execute for a variety of reasons. It is important to be able to induce this stress in the simulator as it directly affects both behavior and overall performance.

A typical task that we would be interested in training would involve a myriad of concurrent tasks competing for limited attentional resources. There would certainly be a complex overland navigation task to some target location via a series of intermediate checkpoints, usually involving multiple aircraft and the communications between them. This is a very stressful task because of time pressure and the consequences of failure. Most simulators fail in inducing this type of stress because the pilot is not completely "involved" in the simulated mission. The perception that this is a not merely a simulation but an actual flight is not achieved. This seems to be related to but not identical to what we would call "presence". A simulator that is capable of inducing stress in its users is desirable because we anticipate that it will allow a closer link between behavior and performance in the simulator to behavior and performance in the air which is an important goal of the system.

There are no tasks that we are interested in training for helicopter pilots and crews that do not involve navigation and spatial orientation. In all cases, the helicopter is moving over large areas and must be oriented to the environment in space and time for successful execution of any mission. This means that the navigator must be able to accurately correlate what is seen on a paper map to what is viewed from the cockpit. He must also do this while managing time and airspeed so that the aircraft arrives at the right location at the right time. It is not enough to avoid becoming lost. The aircraft must arrive neither too early nor too late while maintaining other important aspects of the task such as concealment and radio communications.

There are metrics in the air training community as to how well these tasks must be performed. However, these metrics tend to be qualitative in nature. Determining what "acceptable" performance is on an absolute scale is a difficult undertaking. Yet, the availability of experienced aviators, many with flight instructor experience, allowed us to develop metrics that could be linked to actual flight performance. A key aspect to our requirements was the need to be able to see cockpit gauges and paper maps. If we were to digitize everything and place it all in the virtual environment, not only would the near-field haptics problem be an issue, but readability of the map would be a factor as well. Our earlier implementations avoided this problem entirely by using either projection or conventional displays in a panoramic view (J. Sullivan, Darken, & McLean, 1998; J. A. Sullivan, 1998). The user would sit in front of the displays with a simplified control device and conduct the task of map interpretation and terrain association. The shortcomings of this approach are that it is still a relatively large footprint apparatus and also we found that it had very limited ability to induce cockpit stress.

While it likely would be unusual for a typical user to view the world through a head-worn display with limited field-of-view, the population of interest to us, specifically military helicopter pilots, is special in that they are all extensively trained on the use of night vision goggles (NVGs) which are very similar in this way. The typical field-of-view of an NVG display is around 40º. The horizontal FOV on our HMD was 48º. Aviators are trained to change their visual scan pattern when in NVGs with a side-to-side motion to enlarge their practical FOV in spite of the narrow FOV of the display.

To minimize the footprint of the overall system, we decided to pursue an "appended" training device approach. An appended trainer is like an embedded trainer but it requires that additional equipment be used for the purposes of training while a pure embedded trainer is able to be used as a trainer without additional apparatus. In our case, we assume that we can tap the controls of the helicopter as a serial device input to the simulation. We add the graphics engine, tracking, HMD, and other training specific items.

RELATED WORK

The idea for a Chromakey augmentation to a conventional virtual environment display was originally developed for infantry training at the Southwest Research Institute (Wurpts, 2000). It has also been used for vehicle simulation, specifically for a HMMWV (High Mobility Multipurpose Wheeled Vehicle) simulator that included both driver and gunner stations.

As compared to a more conventional augmented reality implementation, where virtual stimuli are placed over the real world, the Chromakey Augmented Virtual Environment places the real world over the virtual. Using a real-time video mixing technique, the real world as viewed by a camera is "passed through" to the display while anything viewed as blue is replaced by the virtual environment.

For comparison, the MARS (Mobile Augmented Reality System) at Columbia University is an excellent example of virtual augmentation to a real world display (Höllerer, Feiner, Terauchi, Rashid, & Hallaway, 1999). In MARS, rather than use a camera, a see-through display is used that facilitates video overlays. There is no camera.

The Chromakey Augmented Virtual Environment is an augmented reality problem inside the cockpit (like MARS) and a virtual environment problem outside the cockpit.

IMPLEMENTATION

The implementation is initially only the right seat of a side-by-side dual pilot helicopter. It is intended to be generic to any helicopter of this type. We assume that the use of this trainer would be a mission-level trainer used by experienced pilots and crew. Therefore, full fidelity helicopter flight dynamics, platform-specific controls, and on-board displays are not necessary. The focus will always be on tasks other than actual flight. This is typical of any mission rehearsal type of simulator.

We began with an investigation of available HMDs and cameras. Since the primary cues involved in navigation are at a large distance from the observer, we decided to simplify the optics by developing a monoscopic display only. A stereoscopic display would require a second camera and would only benefit the near-field aspects of the task. This could be a topic for future work.

We used a Virtual Research V8™ HMD with 640x480 resolution at 48º horizontal by 36º vertical field-of-view (60º diagonal). The camera is mounted just above the eyes and in front of the tracking sensor. Consequently, the eyepoint is virtually translated approximately 3.5 inches up and 3.5 inches out from the face (see Figure 1). An issue of concern was what effects eye-lens displacement might have on user performance and adaptability to the environment. A hand-to-eye coordination test was conducted in our experiment to study this issue. It is possible to move the virtual eyepoint back to the eyes with the use of mirrors but there is a durability issue with that solution.

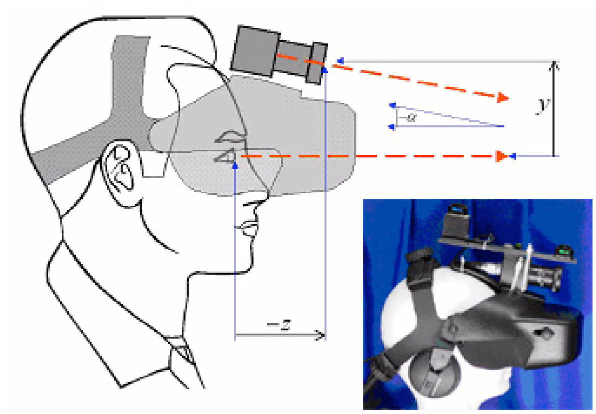

Figure 1. Eye-lens displacement

The camera lens is a fixed focal length (4mm) lens with depth of field within arm length. This allowed us to adjust focus and aperture settings as needed. Lowering the aperture to a lower f/stop allows more light to enter the camera but reduces depth of field. We used an Auto Gain Control (AGC) and Electronic Light Control (ELC) Panasonic camera. This allows the user to view all objects and displays of interest within the cockpit clearly. There are individual differences between users such that not all could see everything from their hands to the simulated gauges with reasonable clarity. Consequently, we observed some users leaning forward in the seat to be able to see the gauges more clearly. The field-of-view of the camera lens was compatible with the field-of-view of the HMD.

The blue cloth used was standard material developed specifically for Chromakey applications. Seams are carefully covered with blue cloth tape to hide them. Lighting is an important aspect of the system because it is critical that the variations in blue light that are seen by the camera are minimal. There can be no shadows on the blue screen. If the blue screen is not well lit or is lit with partial-spectrum lights, "ghosting" will occur in the mixing of the video signals. This appears as noise in the virtual component of the mixed video signal. What is desired is a clear virtual image wherever there is glass in the cockpit with clear edges to the real video from inside the cockpit. Poor lighting or a low quality mixer may cause an undesirable fuzzy edge at the seam or the virtual and the real.

We used flicker-free fluorescent lights pointed at the blue screen but away and shielded from the camera so that the camera could never look directly into a light.

The tracking device is an InterSense™ IS600 Mark 2. Our initial implementation uses both head translation as well as rotation but we are in the process of studying whether or not head translation can be removed and what effects that might have on performance. While users do tend to move their head considerably while using the system, it is usually for the purposes of gaining information obtained via the camera, not the VE component, suggesting that translation may not be important. Miniaturization of the system would greatly benefit from the removal of translation because of the size of the ultrasonic transmitters and sensors that we are currently using. The rotation-only tracking devices are quite small and are therefore preferable to this implementation.

While the actual system would be used in a real helicopter, we are unable to fit an actual helicopter in our laboratory, so we constructed a surrogate platform with a seat, controls, and flat panel display for simulated gauges (see schematic diagram in Figure 2 and photo in Figure 3). The black frame behind the user is an approximation of the metal frame in an actual cockpit. We placed bluescreen material behind the frame and in front of the platform, everywhere that the cockpit would be glass. The platform is elevated so that we could simulate the "chin bubble" which is to the right and down from where the user sits. This is used for very low level flight, landing zone maneuvers, and any other task involving a requirement to see the ground directly below the aircraft.

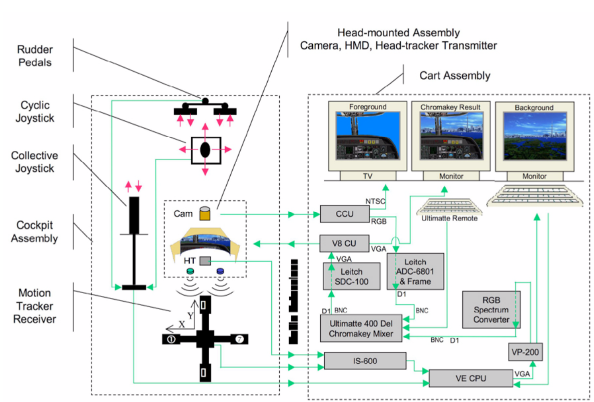

Figure 2. A schematic diagram of the apparatus.

The virtual helicopter is controlled via voice. For the purposes of our study, where we were not interested in this aspect of the interface, we controlled the helicopter in the same way that a real helicopter is controlled. The navigator, who is the user of the system, calls out verbal instructions to the pilot at the controls (PAC) who controls the aircraft with the keyboard. We do not yet use any speech recognition to control the interface. The PAC is not a part of our implementation so this role is filled by the proctor. However, for a single-user system, speech recognition would be needed.
The cockpit displays include an airspeed indicator, an attitude indicator, an altimeter indicating height above mean sea level (MSL), a turn rate indicator, a compass, and a vertical speed indicator. These are simulated on a flat panel display that is placed in the cockpit (see Figure 3). This is on a separate computer from the simulation for performance reasons. For our study, we added a physical clock for timing legs of flight. This was not on the computer screen but was an actual clock, similar to the type that is used in practice.

Figure 3. The apparatus with the user seated using a paper map.

What the users sees is a mixed view of both simulated and actual stimuli (see Figure 4). Everything inside of the cockpit is real. In practice, it would actually be the pilot's helicopter and not a mock-up like what we have in our laboratory. Everything outside of the cockpit is simulated based on the control inputs given by the user, his head movements, and the current status of others in the simulation.

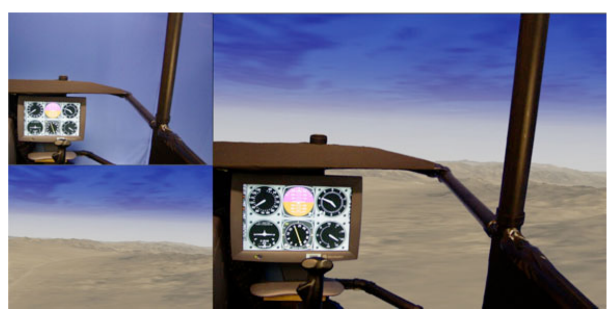

Figure 4. The top left is the camera-only view. The bottom left is the VE-only view. The right is the mixed view that the user sees via the HMD.

The mixing component we use is an Ultimatte 400-Deluxe Composite Video Mixer to which we have several converters of the various signals from the camera and the graphics engine so that we can properly mix and then display the resultant image in the V8 HMD with minimal latency. Latency inherent to the system is a factor of the tracker and rendering engine as opposed to the mixer which appears to introduce negligible delay.

In the schematic diagram in Figure 2, all components of the apparatus are depicted. There are three basic assemblies; the HMD assembly that includes the display, the camera, and the tracker, the cockpit assembly that includes the seat, controls, and flat panel gauge display, and the cart assembly that includes the graphics engine, mixing boards, and converters.

Because of the modular design of the hardware, the system is capable of operating with any type of simulation software desired. Our simulation was developed using Vega™ from MultiGen-Paradigm with several high fidelity terrain models of Ft. Irwin, California, and 29 Palms, California. Resolution of the terrain models had to be as high as possible because of the tasks we were interested in studying. If performance was not at the level we expected, we wanted to know if it was a factor of the apparatus, not a factor of the model fidelity. We used models that we had previously studied in other experimentation where we found that pilots were able to correlate a paper map to the virtual terrain. The models are based on Digital Terrain Elevation Data (DTED) and CIB imagery both of 10 meter resolution.

In its entirety, a Chromakey Augmented Virtual Environment costs under $50,000. We anticipate that the cost will drop as we are able to move the video processing portion of the pipeline to a video card instead of having a separate mixer for this purpose. The majority of the costs in our implementation are in the HMD and the video mixer.

EVALUATION

To evaluate the system as to its feasibility as a trainer and the accuracy of the stimuli as compared to a real world environment, we conducted an experiment to study if experienced helicopter pilots could navigate a complex route in the simulator and if their behavior and performance could be compared to real world performance.

We first completed a cognitive task analysis (CTA) of low-level overland helicopter navigation to determine what stimuli were necessary to the task and which were not. The results of the CTA are not described here but were used directly in the design of the simulator.

Method

Fifteen experienced military pilots took part in the study. All were male with a minimum of 8 years flying experience. Some had been flight instructors but all were completely familiar with all aspects of low level overland navigation from flight planning and preparation to actual flight. Dynamic prioritization of competing cockpit tasks is a critical quality required of all helicopter pilots and was in part, the interest of this study.

The basic procedure included an initial questionnaire followed by map preparation for the route of flight. Then we administered a series of pre-flight physiological tests immediately followed by the simulated flight. The procedure concluded with a series of post-flight physiological tests and an out brief performance evaluation session.

The initial questionnaire included both demographic information as well as a probing of parameters they felt were important in conducting an “acceptable” overland flight. We were interested in how much variation there was in our subjects’ evaluation of flight performance (be it their own or that of anyone else). We showed each participant a series of map images each with a fictitious intended and actual flight path. Participants were asked to indicate whether or not each fictitious flight was “acceptable” or “not acceptable”. This provided a baseline estimation of what baseline performance should be in our simulator.

The next phase was flight planning and preparation. Participants were given all the necessary resources (e.g. pens, tape, measuring tools) that would typically be used for map preparation. They were given a clean map and were able to mark it or alter it in any way they wished. There was no time limit on this phase of the experiment.

When they were ready to begin the actual flight, we first completed a series of physiological tests. The first test was visual acuity for which we used a modified Snellen chart. Then we performed a Dvorine pseudo-isochromatic color test to determine the presence of color blindness. Lastly, we performed a simple hand-to-eye coordination test. The participant sat three feet from the proctor. A soft ball was tossed to the participant who had to catch it with one hand or both hands. We distinguished between a clean catch where there was no discernible fumbling or dropping of the ball, a fumble, where the ball was caught but not cleanly, or a drop, where the ball was actually dropped. Before the flight began, we performed each of these tests twice. The first test was without the HMD apparatus (unhooded baseline condition). Then the participant put on the HMD apparatus and the tests were repeated. This is important because vision through the camera is degraded from normal vision in terms of acuity, FOV, and color resolution. The hand-to-eye coordination test is included to test the effects of the eye-hend displacement factor. We anticipated that there would be no fumbles or drops in the baseline test but that there may be some of both immediately after donning the HMD because adaptation has not taken place yet.

Figure 5. (left) View of the hooded phase of the hand-to-eye coordination test. (right) View of the same phase from the HMD.

To become accustomed to the interface, each participant was given a three minute exposure to a separate environment where they could fly at will anywhere they wished. They were asked to practice rolling in and out of turns and directing the pilot to specific features in the environment. This was done solely to acclimate them to the interface.

To begin the flight, the virtual helicopter was placed at a known location and was already in the air. During the flight, the participant was asked to listen for and acknowledge radio calls for his aircraft. These were mixed in a pre-recorded audio track with other "noise" signals that were to be ignored. Every two minutes, the participant was asked to mark their current position and orientation on the map with a tick mark. This would later be used to determine their estimated location as opposed to their actual location. The participant was given 30 minutes in which he was to complete as much of the course as possible.

The post-flight questionnaire determined if certain features were observed, both in the air and in the terrain. We also asked about comparisons between the simulated flight and actual flight in terms of cockpit management, interface, and fidelity of the simulation.

Since it was not feasible to obtain actual flight performance data for each of our participants, we had to determine some other way of linking performance in the simulator to actual flight performance. What we did was repeat the "flight path evaluation" task described earlier. Initially, we used fictitious flights in order to gain an understanding of what constitutes "acceptable" performance. Here, we used the actual flight paths of the participants in the study. They were not told that these were actual flight paths of other participants.

In summary, we collected data in the form of questionnaires and evaluation forms, recorded simulated flight data (position, orientation, and time), the participant's map, debrief, and peer evaluations.

Results

We were first interested in how pilots prepare for overland flights in actual practice. Even if the simulator can be shown to be effective, if it does not fit within operating procedures, it will not be used.

What we found was that there is a slight difference in preparation methods depending on whether the pilot is at sea or is shore-based (see Figure 6). On shore, pilots rely slightly more on map study than when they are at sea. This may have to do with availability of satellite imagery at sea rather than an actual preference. The only currently available simulation tool is TOPSCENE which was not preferred primarily due to access limitations. PFPS (Portable Flight Planning Software) is a planning tool, not a real-time simulator.

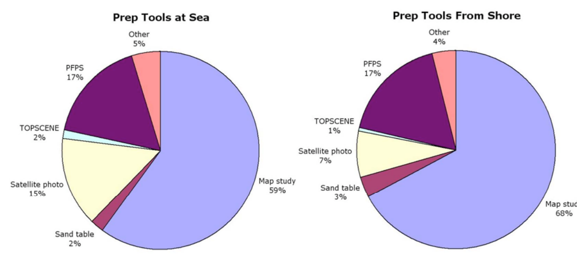

Figure 6. Preferences for flight preparation tools for low-level terrain navigation.

To determine what characteristics an "acceptable" flight might have, we used the results of the initial fictitious flight evaluation and questionnaire. Quantifying flight performance is extremely problematic. In order to assess flight performance, we must know if (A) the pilot is on course and knows he is on course, (B) the pilot is off course but thinks he is on course, (C) the pilot is off course, knows he is off course, but does not know how to recover, (D) the pilot is off course, knows he is off course, and knows how to recover, or (E) the pilot is on course but thinks he is off course. To do this, we need to know not only where the aircraft is throughout a flight but also where the pilot thought he was. The initial evaluation task only showed the track of a fictitious helicopter. It did not give any indication of where the pilot thought he was. Therefore, we report only the qualitative characteristics from the questionnaires regarding the relative importance of different aspects of the navigation task (see Figure 7).

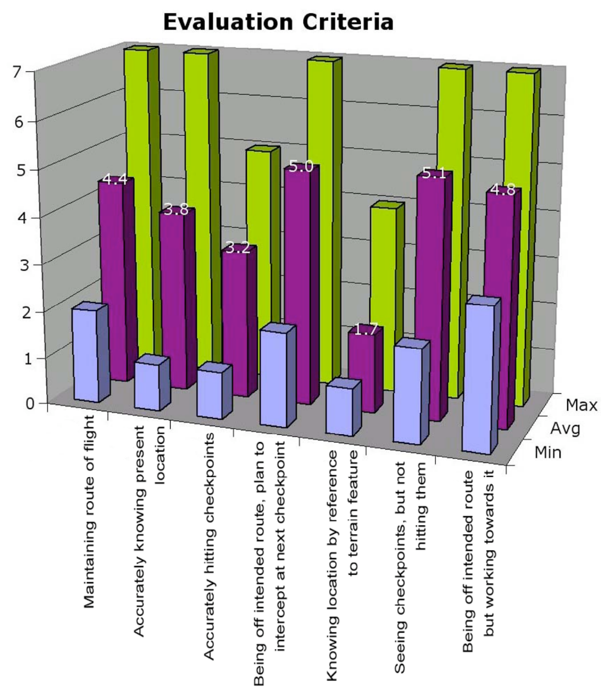

Figure 7. The relative importance of individual aspects of the overland navigation task.

The most important aspects of navigation performance seem to be (1) Seeing checkpoints, but not necessarily hitting them, (2) Being off of the intended route of flight but intended to intercept at the next checkpoint, and (3) Being off of the intended route of flight but working towards it. Clearly, military pilots understand that an expectation of always being on route is unreasonable. The most critical aspects of navigation have to do with being able to identify when you are off course and knowing how to recover. The key is at the checkpoints. Participants indicated that "reasonable proximity" to a checkpoint is approximately 260 meters. This is the only quantitative measure (other than time) that would translate to the evaluation of a flight path. Therefore, it would appear to be inappropriate to judge navigation performance on flight path alone.

The physiological test results were as expected (see Figure 8). The unhooded baseline performance on the hand-to-eye test was almost perfect. When they first put on the HMD, their performance dropped significantly. They adapt over the course of the simulated flight so that their post-exposure performance improved. Finally, when they take off the HMD, their performance returned almost to its original level, but not quite, suggesting that full re-adaptation had not yet occurred. Relative measurements for color perception and Dvorine are similar but are not affected as strongly.

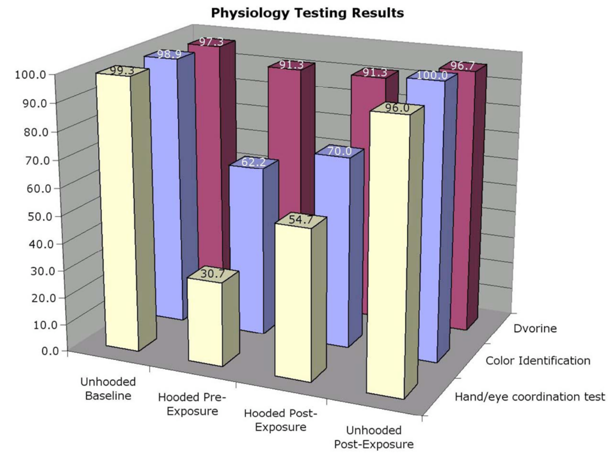

Figure 8. Results of the physiological tests.

The results of the simulated flights are presented in tabular form in Table 1. Recall that during the flight, subjects marked their maps every two minutes as to where they believed themselves to be and what direction they were heading. The quantitative results as measured in the simulation are in the first three columns. Distance is the average distance the participant was off-route at the time of a map check. Heading is the average error in degrees at the map checks. Minimum distance from flown path to checkpoint indicates how close the participant flew to the intended route measured at the checkpoints. As discussed earlier, quantitative measures do not capture the essence of "good navigation performance" as defined by our pilot participant group. What is more interesting is the last four columns. These are based on the peer assessment portion of the experiment. Given a threshold of 260 meters for each checkpoint (as established by the earlier survey), the number of acceptable checkpoint proximities is listed in the fourth column. Out of a possible nine checkpoints, the average across the group was 3.2 indicating that performance was comparable to real flight. On a 7 point scale, the participants' ability to maintain an acceptable route over the duration of the flight was an average of 3.6. Similarly, their ability to estimate their location on 2-minute interval map checks was an average of 3.6 on the same 7 point scale. Clearly, the data suggests that the task is neither too easy nor too hard as compared to actual in-flight navigation. The last column indicates the number of evaluators who rated the participant's overall route as "acceptable". A perfect score would be 14 (participants did not rate their own performance). The average across participants was 10.6.

Table 1. Performance results of all participants.

Subject	Location, Direction, & Proximity Result Averages				Peer Evaluations		
	Position & Orientation Differences		Checkpoint Proximities				
	Distance (Meters)	Heading (Deg)	Min Distance from Flown Path to Checkpoint (Meters)	Number of Acceptable Checkpoint Proximities *4	Average Proximity of Paths Score *1	Average Location Estimation Score *2	Total Acceptable Overall Performance Ratings *3
001	1,124.9	14.4	378.1	4	2.4	1.9	14
002	965.5	21.7	326.7	5	1.8	2.1	14
003	2,543.6	32.9	1346.0	1	6.3	5.5	0
004	2,403.2	30.0	1279.9	4	4.9	4.9	5
005	1,521.1	26.8	924.1	4	3.8	3.5	12
006	4,586.1	28.1	1416.3	1	6.6	6.4	0
007	1,435.5	17.1	739.3	3	3.5	3.4	13
008	1,872.5	18.6	575.2	3	3.4	3.6	14
009	1,132.1	12.1	533.8	4	2.7	2.5	14
010	1,054.9	21.7	231.9	5	1.8	2.2	14
011	1,677.9	19.2	473.1	4	2.3	2.8	14
012	1,133.9	19.4	558.4	4	3.7	3.1	14
013	1,587.6	30.7	576.9	2	4.7	3.8	7
014	880.1	7.1	481.2	3	2.6	2.1	14
015	1,172.2	18.2	579.4	1	4.1	3.5	10
Averages->	1,672.7	21.2	694.7	3.2	3.6	3.4	10.6

*1. The subject's ability to maintain a flight path in acceptable proximity to the intended path.
- Rated using a 1 to 7 scale, '1' indicating highly acceptable while '7' indicates not acceptable.
- This criteria is independent of the following criteria, meaning the proximity to the intended flight path was evaluated independently of whether or not the subject knew where they were.

*2. The subject's ability to correctly estimate their location.
- Rated using a 1 to 7 scale, '1' indicating highly acceptable while '7' indicates not acceptable.
- This criteria is independent of the preceding criteria, meaning the accuracy of the position estimation was to be evaluated independently of whether or not the subjects were on the intended route.

*3. The overall performance.
- Rated the overall performance as acceptable or not acceptable ('A' or 'N')

*4. Number of checkpoints flown within the subject pool's 260 meter threshold as established by question #22 of the questionnaire.

We also looked at visual scan patterns in the simulator as compared to real world flight (see Figure 9). Even if performance is comparable, if scan patterns are significantly different, this could be the cause of a reverse training effect. The simulator seems to cause a slight bias towards a forward view. Since there was no take-off or landing portion to the task, there was no reason to use the chin bubble. Other data are comparable between the real world and the simulator.

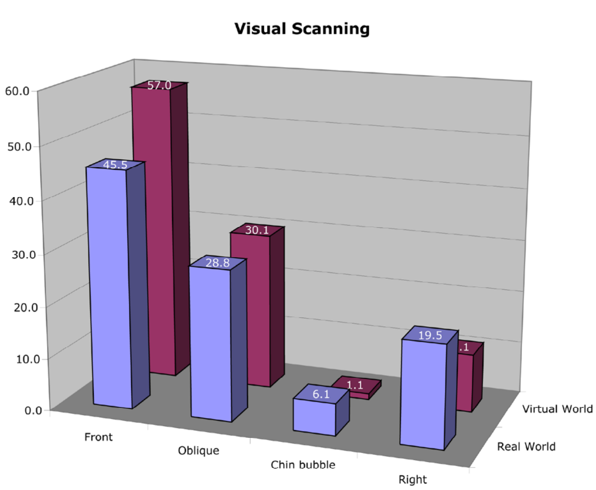

Figure 9. A comparison of visual scan patterns.

Discussion

We learned from this experiment that assessing navigation performance for this task is exceptionally difficult. We have worked extensively on land navigation in the past (Banker, 1997; Darken & Banker, 1998; Goerger, 1998). The advantage there is in our ability to directly measure real world performance. Here, we had to be creative in the way we determined if the trainer was simulating the correct aspects of the task. By having the participants assess performance, we were able to determine their tolerance for acceptable navigation and we could then apply this measure to their own performance. This gives us a baseline with which to compare performance in the simulation. It appears from our results that performance is where it should be – neither too high nor too low. We would have been alarmed if performance had been excellent as much as we would be if performance was dismal. It was average. Also consider that our participant pool, while expert pilots, were all in a non-flying status which also accounts for somewhat lower performance than might be expected by an active participant pool.

The post-exposure questionnaires show that overall, participants rated the simulation favorably. They believe it to be comparable to other visual simulators and that it reasonably simulates the complex task of overland navigation. It received low marks for viewing the map and cockpit gauges. This is a factor of the focal length of the lens we use. Participants commonly lean forward to view the gauges and they will tend to hold the map close to their face to read their markings. These are artifacts of the current technology that we can work to improve in the next generation.

Most importantly, participants reported that the simulator is capable of approximating cockpit stress and workload. This is partly due to the similarities to the real world task, but also because of the "cockpit management" skills that were required such as map folding and turning, writing, and visual scanning. While some participants reported average or lower than average stress during this task, we noted that most participants performed poorly on the secondary task of recognizing their radio calls and identifying specific features on the simulated terrain (e.g. buildings, vehicles, etc.).

While we did not use a standard simulator sickness questionnaire, we did ask about symptoms of simulator sickness in the questionnaire. Some early symptoms were reported but overall, simulator sickness did not interfere with participant performance.

CONCLUSIONS

The Chromakey Augmented Virtual Environment appears to be a viable approach to appended training for helicopter simulation – possibly for vehicle simulation in general. We believe the primary advantages of this apparatus are (1) its ability to simulate a stressful environment, (2) its close approximation of the real world task, (3) its small footprint, (4) its low cost, and (5) its potential for use in a deployed setting.

The apparatus has been used as one of several interfaces to VEHELO, the helicopter simulation portion of the Virtual Technologies and Environments (VIRTE) program at the Office of Naval Research. It has been integrated with AAAV and LCAC vehicle simulators (developed by Southwest Research Institute and Lockheed Martin respectively) in a common synthetic battle space compatible with JSAF. VEHELO will soon be tested in a full transfer of training experiment with the cooperation of the H-60 community at Naval Air Station North Island in San Diego, California. A set of pilot trainees will receive exposure to the simulator before their navigation flight where they must perform a complex overland navigation task. We will then capture data in the air using GPS and voice communications for comparison with a second group that does not receive the trainer but instead receives only the existing conventional syllabus.

The next steps for the Chromakey Augmented Virtual Environment are to expand on the mission profile. Currently, it is only capable of an isolated navigation simulation. However, to a helicopter pilot, navigation is done to achieve some other goal such as insertion or extraction or combat search and rescue. Extending the system to be able to train combined arms types of tasks or other missions requiring complex coordination are particularly of interest. We must also further investigate the lenses we use and the head mounted display.

We have prototyped a Night Vision Goggle (NVG) simulation (see Figure 10) and have evaluated it using the same procedures described in this paper. We were able to compare a simulation based on the physical attributes of materials in the terrain to one that approximated the way light reflects in the scene. We found that performance in the physics simulation was poorer to that of the non-physics simulation and that our NVG-expert pilots believed that the approximated simulation was adequate for the needs of their task. This will require further study, but if we can show that for some class of tasks, physically-based simulation for NVG is not required, this will greatly lower the costs of these simulators.

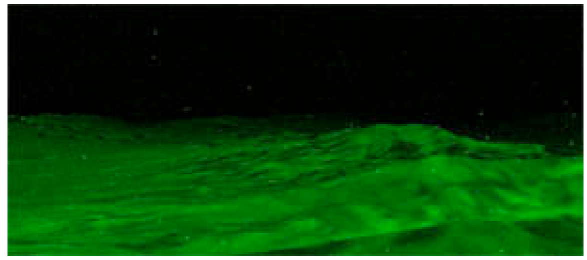

Figure 10. A scene from the NVG simulation.

We believe the ideas presented here can be extended to other types of simulation as well. By comparing what is achievable using the Chomakey technique as compared to more conventional simulation, we can better learn what aspects of tasks are suitable for training in simulation, and which are to be avoided. It is critical that we not "oversell" our technology but rather do the necessary experiments so that we can know what technologies best fit the training requirements. This is the goal of the VIRTE research program, and the goal of our research as well.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the military officer students at the Naval Postgraduate School who volunteered their time to participate in the study. This research is sponsored by the Virtual Technologies and Environments (VIRTE) Program at the Office of Naval Research.

REFERENCES

Banker, W. P. (1997). *Virtual Environments and Wayfinding in the Natural Environment.* Unpublished Masters, Naval Postgraduate School, Monterey, CA.

Cruz-Neira, C., Sandin, D., & DeFanti, T. (1993). Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. *Computer Graphics*, 135-142.

Darken, R. P., & Banker, W. P. (1998). *Navigating in Natural Environments: A Virtual Environment Training Transfer Study.* Paper presented at the IEEE Virtual Reality Annual International Symposium, Atlanta, GA.

Goerger, S. R. (1998). *Spatial Knowledge Acquisition and Transfer from Virtual to Natural Environments for Dismounted Land Navigation.* Unpublished Masters, Naval Postgraduate School, Monterey, CA.

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., & Hallaway, D. (1999). Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System. *Computers and Graphics, 23*(6), 779-785.

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2002). Physiological Measures of Presence in

Stressful Virtual Environments. *ACM Transactions on Graphics, 21*(3), 645-652.

Sullivan, J., Darken, R., & McLean, T. (1998, June 2-3, 1998). *Terrain Navigation Training for Helicopter Pilots Using a Virtual Environment.* Paper presented at the Third Annual Symposium on Situational Awareness in the Tactical Air Environment, Piney Point, MD.

Sullivan, J. A. (1998). *Helicopter Terrain Navigation Training Using a Wide Field of View Desktop Virtual Environment.* Unpublished Masters thesis, Naval Postgraduate School, Monterey, CA.

Unguder, E. (2001). *The Effects of Natural Locomotion on Maneuvering Task Performance in Virtual and Real Environments*. Unpublished Masters thesis, Naval Postgraduate School, Monterey, CA.

Wurpts, M. (2000, July 10-14, 2000). *Application of a Video See-Through Head-Mounted Display to Infantry Training.* Paper presented at the IMAGE 2000, Scottsdale, AZ.