

Augmented Reality for Urban Skills Training

Dennis G. Brown¹
dbrown@ait.nrl.navy.mil

Roy Stripling¹
stripling@itd.nrl.navy.mil

Joseph T. Coyne²
jcoyne@sainc.com

¹U.S. Naval Research Laboratory

²Strategic Analysis, Inc.

ABSTRACT

Warfighters develop and maintain their skills through training. Since fully-manned live training in the real world is often too expensive (by many measures), scientists have developed many types of training systems ranging from classroom sessions to those using virtual reality. Recently, researchers have used augmented reality (AR) to insert virtual entities into the real world, attempting to create a low cost, repeatable, and effective substitute for fully-manned live training. However, very little evaluation of the effectiveness of AR for training has been performed.

We performed a pilot study to evaluate the use of wearable AR in teaching urban skills, specifically, room clearing in teams. Eight teams of two were briefed on room clearing techniques, given hands-on instruction, and then allowed to practice those techniques with or without the AR system. After this instructional period, subjects performed several room clearing scenarios against real people using infrared-based practice weapons that logged the number of hits on the subjects and the enemy and neutral forces. During these trials, a subject matter expert evaluated how well the subjects applied the room-clearing techniques.

In this paper, we describe the pilot study in more detail, including the hardware and software testbed, and then provide an analysis of the results of the pilot study.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/methodology; J.7 [Computer Applications]: Computers in Other Systems—Military;

Keywords: augmented reality, training, evaluation

1 INTRODUCTION

Modern wars are more often fought in cities than in open battlefields, and warfighter training has been updated to reflect this change. Military Operations in Urban Terrain (MOUT) training is an important component of a warfighter's initial and continued development. Much of this training occurs in purpose-built MOUT facilities, using simulated ammunition and half the team acting as the opposing forces (OPFOR). As an alternative, virtual reality (VR) training systems for MOUT operations are improving. Both of those training modes have several drawbacks. The MOUT facility training provides the trainee with a real-world experience, but there are manpower issues (must schedule two teams, or split one team so that half plays OPFOR), the exercise is not completely repeatable, and there are issues with the simulated munitions such as setup, injuries, and cleanup. In contrast, the VR training provides a safe, controlled, and repeatable training scenario, but it deprives the trainee of many real-world cues that are not yet simulated, requires special equipment that is not easily moved for the most immersive simulations, and does not allow completely realistic navigation through the environment.

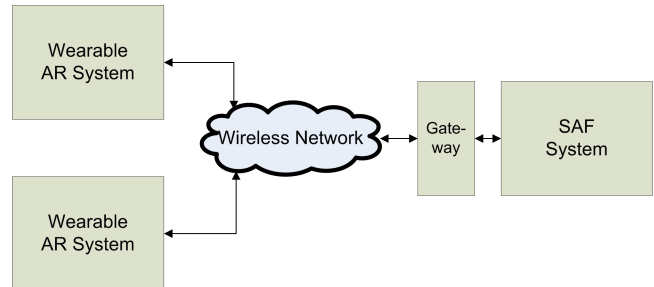


Figure 1: Major pieces of the testbed.

In an effort to create a training method that combines the control and repeatability of VR with the authenticity of the real world, we have researched and developed a prototype system that uses augmented reality (AR). Augmented reality technology adds computer-generated information to the real world. For training, animated three-dimensional computer-generated forces are inserted into the environment. The AR training system moves the repeatability and control of a VR system into a real-world training environment.

Other groups have considered the use of AR for MOUT training. A system presented by Small and Foxlin [6] allows trainees to practice close-quarters battles. MR MOUT [4] provides virtual targets in a realistic set in an example of mixed reality. VICTER [1] was built to fit within the limitations of the current Land Warrior system [3], replacing pieces of that system as necessary. The system described in this paper is the second generation of our own AR system for MOUT training [2].

Although several prototype systems have been built, very little evaluation of the effectiveness of AR for MOUT training has been performed. We ran a pilot study to evaluate the usefulness of wearable AR in teaching urban skills to teams, specifically, team room clearing. Participants, in teams of two, were briefed on room clearing techniques, then allowed to practice these techniques with or without the AR system, and finally evaluated in a simulated room clearing task, without AR, against real people acting as opposing forces. We will start by describing the evaluation testbed, then describe the evaluation process in more detail, and finally analyze the results.

2 EVALUATION TESTBED

The evaluation testbed assembled for this project consists of two wearable AR systems, wide area indoor tracking, the Army's One-SAF to drive the computer-generated forces, and wireless networking to tie the systems together. Figure 1 illustrates this simple testbed.

2.1 Wearable AR System

Participants wore a backpack loaded with commercial off-the-shelf (COTS) hardware that uses GOTS (government off-the-shelf) software. This wearable AR system is driven by a high-end laptop

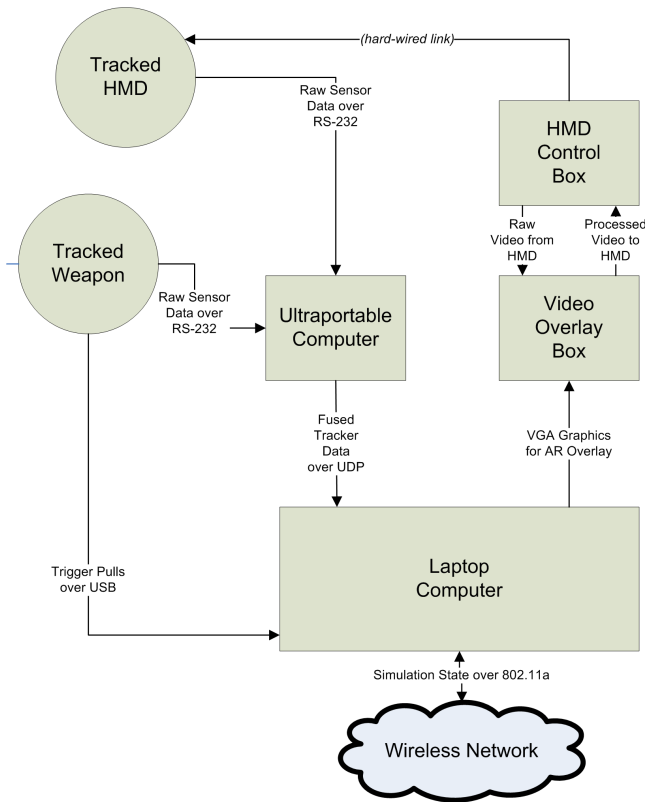


Figure 2: Components on the backpack.

computer with accelerated graphics. The system runs Lockheed Martin's ManSIM software. This software can be classified as a "first-person shooter" type of application, however, it is oriented toward training rather than entertainment. The wearable system also includes a video see-through head-mounted display (HMD) connected to a video overlay box. The computer generates graphics that go to the overlay box as a VGA signal; the box combines the VGA signal with the composite video output of the HMD, and feeds the combined signal back to the display to give the augmented view. We chose to use the external box rather software-based overlay to reduce the lag, since users walk through the building wearing the system, and too much lag in the display may be unsafe. The HMD and a handheld weapon proxy are both tracked by a vision- and inertia-based tracking system, and there is a second ultra-portable computer on the backpack to handle the tracking data fusion duties. Figure 2 illustrates these components and their various connections.

2.2 Computer-Generated Forces

The computer-generated forces used in the AR practice sessions were driven by the US Army's OneSAF Testbed Baseline Semi-Automated Forces (OTBSAF) system. OTBSAF connects through a gateway to a local instance of the Run-Time Infrastructure (RTI) to which the backpack systems also connect. The mobile users are reflected in real time in OTBSAF as friendly forces, and the computer-generated forces respond appropriately. These responses are sent to the backpack to control the visualizations of the computer-generated forces in the HMD.

2.3 Test Area Model

One of the most important tasks was building the model of the test area. This model needed to be very precise as it was used to calcu-



Figure 3: Model of test site, showing evaluation and training areas.

late occlusions of the computer-generated forces as they hid behind walls and beside windows. The occlusion model was used in the same way that ARQuake [5] used an occlusion model, basically, drawing the model in black, and allowing the real world view to be seen through the black pixels on the display. The test area was surveyed with a laser-based theodolite and that data was used in a common 3D modeling program to build the occlusion model. This data was also used to build the "floorplan"-style model that the SAF system uses to calculate visibility and paths. Figure 3 shows the floorplan of the test site.

3 EXPERIMENTAL DESIGN

This purpose of this pilot study was to measure the usefulness of AR at the application level and to set the stage for future work. Two conditions were evaluated: training with AR and without AR. Eight individuals grouped into four teams were tested for each condition, for a total of sixteen individuals in eight teams. This study was approved by the NRL Institutional Review Board.

3.1 Instruction

Each trial contained an instructional period and an evaluation period. During the instructional period, the team learned basic room clearing techniques. The doctrine for room clearing contains several specific techniques for entering a room, holding a weapon, and working as a team. First, the subject team watched an eight-minute video explaining the basic techniques used for room clearing. Next, the subjects were shown the techniques, in the practice area, by a subject matter expert (SME), for fifteen minutes. Finally, the subjects donned the AR backpacks and were allowed to practice room clearing techniques for fifteen minutes in the practice area. Subjects in both the AR and non-AR conditions were free to practice as they saw fit, but they were encouraged to perform several repetitions of clearing all of the rooms. In the AR condition, as a team started each new repetition, we would load a new SAF scenario, placing stationary but reactive enemy and neutral forces in the environment.

Subjects in both conditions wore the AR backpacks because we wanted to make sure the weight and bulk of the backpack system did not negate any possible positive effects of AR training. In the non-AR condition, the CGFs were simply not mixed in with the real-world video. The backpack was built on a strict budget, and some tradeoffs were made—for example, sacrificing a small, light computer in order to procure highly-accurate trackers, and using a heavyweight general-purpose simulation program rather than building a new single-purpose application from scratch that could run on an embedded PC. Thus, even with today’s technologies, the backpack could be much smaller, and in the future, will be even more compact, and we believe the bulk of a future system will not have a negative effect on users.

3.2 Evaluation

After the instructional period ended, the subjects were moved to another part of the test site to be evaluated. Here, participants performed in six room-clearing scenarios against real people. Each scenario had enemy and neutral forces in different positions. As in the training period, these forces were stationary and defended a particular corner of a room. The subjects and the people playing the enemy and neutral forces traded fire using “laser-tag-style” weapons. This weapon system counts the number of hits on the subjects and on the enemy and neutral forces. The participants once again wore the AR backpacks, however, this time it was solely for tracking and logging the user’s actions—the HMDs were raised above the subjects’ heads so that they did not occlude the natural sight abilities of the subjects.

3.3 Confounding Issues

We identified several issues going into the study that would ultimately affect the results, but we were not able to solve them, due to time, budget, institutional constraints, or the limitations of today’s hardware. These issues include:

- **Training and evaluation venues.** We trained and evaluated the users through several scenarios in the same two sets of rooms. In actual room clearing tasks, infantry will approach an unfamiliar set of rooms, clear them, then move on to yet another unfamiliar set of rooms. Thus, our subjects had the advantage of being able to create a plan for each scenario because they knew the layout of the rooms. One possible solution is to use cubicles to create different sets of “rooms” for each trial, but we were not able to use such a facility for this study. Another problem with our set of rooms used in the evaluation scenario is that, as seen in figure 3, four rooms have doorways clustered together, creating a very dangerous task for novice room clearer, as enemies had sight lines across many rooms.
- **Short training time with the AR system.** Each trial lasted around two hours, which is a lot to ask of any volunteer. At the same time, we felt that we could not remove any of the training or evaluation steps. As a result, the AR backpack training was set to fifteen minutes, or less than half of the total instruction period.
- **No feedback provided to the subjects during the AR training sessions.** The subjects knew when they shot another force, or when they were shot, but otherwise, were not told how well they were applying the room clearing techniques during the training sessions. Although it can be considered a control that both the AR and non-AR groups had no feedback, we failed to harness one power of a wearable AR system, which is the ability to provide immediate feedback tailored to a specific user. However, enhancing the system to support that capability was beyond the scope of this work.

- **Hardware glitches.** There were some intermittent problems with the systems that were out of our control, such as one eye going black in the HMD, or the tracker getting confused and temporarily flying the subject and/or weapon hundreds of meters away. We asked subjects to watch out for these problems and report them to us, so that we could fix the system and let them continue the training. We also measured the amount of time fixing problems and gave the subjects that much more time to practice. Each team suffered one to two of these episodes, which may have distracted them enough to affect the quality of training.
- **Subject pool.** We used subjects who had no formal training in room clearing techniques, however, the subjects had varying degrees of experience with paintball, laser tag, computer games, and augmented reality.
- **Inaccurate weapons for evaluation.** The weapons we chose for the evaluation were consumer-grade and based on infrared senders and receivers. The senders had a fairly wide angle, allowing subjects to be sloppy and still register hits, and allowing unwanted hits (such as friendly fire) to happen more frequently than if more accurate weapons had been used.
- **Differences between training and evaluation weapons.** During the AR training phase, the subjects saw a graphical weapon that was superimposed over the handheld weapon proxy. Subjects were to aim and shoot the graphical weapon, ignoring the proxy. Thus, they trained on one weapon and were evaluated on another. Both weapons acted as similarly as we could specify (in this case, one shot—any shot—on a force is a kill), but still, they were different.
- **Unnatural appearance of the computer-generated forces.** The CGFs, although registered and occluded, still didn’t look realistic—they had constant lighting unrelated to the actual real lighting, and sometimes had a ghostly appearance due to the video mixing hardware we chose.

4 ANALYSIS AND CONCLUSIONS

4.1 Measures

The subjects were evaluated using two basic measures. The first measure is objective and is based on survival and shots on enemy and neutral forces during each scenario. The raw data was taken straight off of the weapons system after each scenario and applied using the formula $team\ performance = ((number\ of\ surviving\ team\ members) + (0.5 * number\ of\ hostiles\ killed) + (0.1 * number\ of\ neutrals\ still\ alive)) / maximum\ possible\ score$. This formula was created with the input of our subject matter expert, taking into account the priorities of a military force: survive and achieve the objective. The division by the maximum score gives a normalized value between 0 and 1.

The second measure is subjective. Our SME followed the subjects during each scenario and rated the subjects on a scale of 1 to 5 for each of these attributes: aggressiveness, movement, security, communication between teammates, and coordination between teammates. These categories describe the fundamental skills one should learn through this training, but we had no objective way to measure them. During the trials, the SME did not know whether the subjects trained with or without AR.

4.2 Results

We found no significant difference between the performances of subjects using AR and those not using AR. Using the team perfor-

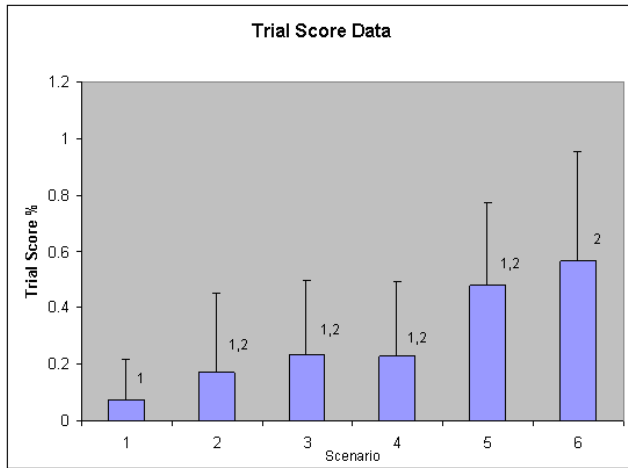


Figure 4: Mean scores for all users for each evaluation scenario. Error bars are standard deviation. Numbers indicate the Student-Newman-Keuls group to which data for the scenario belongs.

mance metric described above on the objective measures, the AR subjects had a mean score of 0.25 versus 0.35 for the non-AR subjects. The data were analyzed using a AR (2) x Scenario (6) repeated measures Analysis of Variance (ANOVA), which gave the values $F(1, 6) = .749, p = .420$.

Next, we looked at the team performance measurements between the six training scenarios for all subjects combined. Here we found a significant different between scenarios across all subjects ($F(5, 29) = 3.302, p = .018$). This finding indicates a steep learning effect during the evaluation scenarios. Figure 4 illustrates this effect. This result suggests that the subjects still had much to learn following the instructional period (with or without AR). Also, the subjects' performance may have improved as a results of increased familiarity with the room layout, as it was the same for each scenario (only the locations of the enemy and neutral forces changed).

Finally, we looked at the interaction between the AR condition and the training effect. In this case, $F(5, 29) = .381$ and $p = .858$. The AR condition does not seem to have a significant effect on the increase in performance.

For the subjective measures (team communication, team coordination, aggressiveness, movement, and security), again, we saw no significant differences between the AR and non-AR conditions. Table 1 shows the mean scores and ANOVA results for each measure. Once again, there is no strong effect of using AR or not on the scores. The ANOVA results, for each subjective measure, between scenarios and for the interaction between AR/non-AR and each scenario, mimic those shown above for the objective measures, and will not be listed here.

Measure	Mean non-AR	Mean AR	F	p
Team Communication	2.88	3.17	$F(1, 6) = .219$.656
Team Coordination	2.79	2.83	$F(1, 6) = .000$.991
Individual Aggressiveness	3.02	2.52	$F(1, 14) = 1.272$.278
Individual Movement	2.67	2.07	$F(1, 14) = 2.993$.106
Individual Security	2.50	2.20	$F(1, 14) = 1.300$.273

Table 1: Means and ANOVA values for subjective measures

5 FUTURE WORK

This study was a basic pilot study using a minimal number of users and just two conditions: training with and without augmented reality. One way to continue this work is to set up several experiments of a smaller scope that look at particular aspects of the use of AR for training. These shorter studies would allow more subjects through and would help us refine the larger experiment. We could look at comparing AR training to training using live targets or static targets, for certain tasks simpler than room clearing, to help determine in which cases AR training is effective. We could also look at varying certain attributes within the AR condition to help narrow down exactly what qualities and features are necessary in an AR system for training. For example, is the video-based display the best, the best bang for the buck, or inadequate?—How badly can the tracking degrade before the training transfer effect is reduced?—and so on. Finally, we can consider how an interactive AR system can provide immediate feedback to the user, possibly from an on-board application, or by providing a communications channel with an instructor who can watch many trainees at once. The results from these simpler experiments would help us refine the main experiment, which we would then like to rerun with many more subjects.

6 ACKNOWLEDGEMENTS

The following individuals contributed heavily to the success of this work: Dan Afergan, Elliott Cooper-Balis, Bryan Hurley, Zina LaBudde, and K.C. Pfluger. This work was funded by the Office of Naval Research under contract N0001405WX20394.

REFERENCES

- [1] P. Barham, B. Plamondon, P. Dumanoir, and P. Garrity. Victor: An embedded virtual simulation system for land warrior. In *Proc. of the 23rd Army Science Conference*, 2002.
- [2] Dennis G. Brown, Yohan Baillot, Simon J. Julier, Paul Maassel, David Armoza, Mark A. Livingston, and Lawrence J. Rosenblum. Building a mobile augmented reality system for embedded training: Lessons learned. In *Proc. of Interservice/Industry Training, Simulation, and Education Conference*, 2004.
- [3] Natick Soldier Center. Operational requirements document for land warrior, 2001. URL: http://www.natick.army.mil/soldier/WST/LW_ORD.PDF.
- [4] Charles E. Hughes, Christopher B. Stapleton, J. Michael Moshell, and Paulius Micikevicius. Challenges and opportunities simulating future combat systems via mixed reality. In *Proc. of the 23rd Army Science Conference*, 2002.
- [5] W. Piekarski and B. Thomas. Arquake: The outdoor augmented reality gaming system. *ACM Communications*, 45:36–38, 2002.
- [6] Daniel Small and Eric Foxlin. Augmented reality training system for close quarters combat, 2004. demonstrated at the ISMAR 2004 (International Symposium on Mixed and Augmented Reality).