U.S. ARMY MOBILE AUGMENTED AND VIRTUAL REALITY TRAINING SYSTEMS FOR HANDHELD IED DETECTORS

[a]Dean Reed, [b]Latika Eifert, [c]Shane Reyolds, [d]Travis Hillyer, [e]Clive Hoayun

[a],[c],[e]Institute for Simulation and Training, University of Central Florida
[b],[d] Army Research Laboratory-Human Research and Engineering Directorate Advanced Training and Simulation Division, SFC Paul Ray Smith Center

[a]dreed@ist.ucf.edu, [b]latika.k.eifert.civ@mail.mil, [c]sreynold@ist.ucf.edu,[d]travis.r.hillyer2.civ@mail.mil, [e]choayun@ist.ucf.edu

ABSTRACT

The University of Central Florida's Institute for Simulation and Training (IST), and the U.S. Army's Research Laboratory (ARL) have collaborated on the creation of a suite of next generation mobile augmented reality (AR) and virtual reality (VR) applications. Our focus for this ongoing effort is full spectrum hand-held, mobile simulation-based training for advanced IED detectors. IST developed game engine based VR trainers capable of fully immersing the Soldier Trainees on low-cost mobile devices. A very advanced handheld AR trainer can convincingly emulate the proper motion required to employ the dual sensor detector in high fidelity virtual environments representing potential real operational environments, even to the representation of soil characteristics. This paper will discuss recent advancements, both hardware and software oriented, that enable the rapid deployment of high-quality end AR/VR training Apps. In addition, this paper discusses real world challenges associated with tackling complex training applications with mobile hardware.

Keywords: augmented reality, virtual reality, mobile embedded training

1. INTRODUCTION

Key enablers such as smart mobile devices with embedded GPUs and HMDs such as the Microsoft HoloLens have made low cost, mobile AR and VR training applications for the U.S. Army achievable. The cross-platform advanced game engines such as Unity have improved the fidelity of the trainee's experience. These tools continue to offer rapid and affordable development environments.

UCF IST has developed AR/VR Apps to support soldier IED training. Some of the Apps are in a ready to deploy state, while others are works in progress. We have achieved a diverse set of implementations that span the gamut of Milgram's Continuum.

2. AUGMENTED REALITY FOR TRAINING

IST and the U.S. Army partners at the Army Research Laboratories have implemented AR applications using several different technologies. Only a few years ago, it was noted that AR wearable HMDs and ancillary tracking systems were not sufficient to provide complete training systems (Stevens, Eifert 2015). The advent of powerful HMD systems such as the Microsoft HoloLens has enabled lightweight, ad-hoc functional on-demand training. IST developed a HoloLens application using a representation of a real life mine/IED device master-trainer. The authors used a motion capture system to collect the variety of critical motions that represent proper employment technique. The avatar can respond to voice queries and demonstrate complex tasks with the IED detector such as proper sweep techniques and can discuss and demonstrate calibration techniques in a variety of soil conditions supported by the trainer (Figure 1).

Figure 1: Microsoft HoloLens AR Instructor Application

A second AR application targeting the same IED sensor uses a distinctly different approach to achieve a very low cost and novel hardware surrogate trainer capable of delivering hands-on detector swing instruction. IST developers employed computer vision algorithms to enable a realistic mixed-reality application that allows trainees to deploy a low cost but convincing surrogate training system.

This AR simulation supports the capability to rapidly change the rendered visual environment. A simple change of the background fiducial image triggers a change in AR environment that also impacts the sensor performance and the simulation behaviour.

2.1. TRAINING COST SAVINGS

Current "real" IED detectors can easily exceed $20,000 per unit, but our AR training solution cost less than $500 per unit. IST combined modern 3D printing capability with high-performance tablet solutions to arrive at a very affordable system with proper weight form factor for the end training. A hybrid solution that adds our training device to a real IED detector was developed, but this only cost-effective when a supply of devices that are beyond repair is available. We removed the physical sensor head to implement the AR sensor head.

Figure 2: Low-Cost Surrogate (Left) Real Device (Right)

Another benefit of offering a low-cost approach is that it increases the accessibility of training. The expensive real/operational IED detectors can be difficult to obtain due to the perceived high dollar loss if a device is lost or damaged during training exercises. Our surrogate AR training system allows the user to practice in soil conditions that only exist in soil conditions from remote locales that would normally only be accessible via time-consuming training excursions or by building expensive infrastructure training ranges.

2.2. AR TRACKING DISCUSSION

The two approaches here, while both using AR, provide very different end applications. The HoloLens approach leverages real-time HMD sensor fusion to determine the correct AR viewing frustum. It uses time of flight sensors, as well as inertial sensors to calculate surfaces available for AR interaction. The onboard sensors are coupled with a form fitting graphics array to provide the rendering of the virtual graphics in the real world.

One identified problem with the HoloLens is the timeliness of the real-time environment mapping. It can take many seconds to acquire a fully realized mapping of the operational area depending on environmental conditions. As an example, we noticed that starting a brand new training exercise introduced in a new area crowded with people presented a challenge for the onboard sensors to realize the local coordinate system.

Conversely, we used a fiducial marker approach on the AR Swing trainer. Marker tracking coupled with internal IMUs available with the mobile device presents an excellent and affordable solution for this domain.

IST evaluated a surveyed a number of fiducial API solutions: Metai SDK, ARMedia, XZIMG, ARMedia, Wikitude, IN2AR, Obvious Engine, Vuforia and NyARToolkit. By trying each solution to compare the initial quality of pose estimation, we determined that the best contenders for our use case are Vuforia and Wikitude. A detailed comparison of the two led us to choose Vuforia over Wikitude because of how mature the Vuforia API is regarding 3D target recognition and independent object tracking. We created our IED swing trainer (Figure 3) with a generic interface to encapsulate the pose estimation software, so that we may rapidly change our solution to take advantage of newer or better solutions.

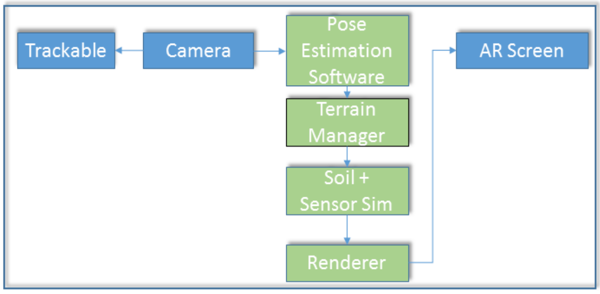

Figure 3: Fiducial Based AR Swing Trainer AR System Block Diagram

Fiducial markers have been successfully used in HMD based AR systems for over 15 years (Kato, Billinghurst 1999). As expected, our fiducial marker system works well within the constraints of the visual marker regions to provide solid pose estimation. Vuforia's extended tracking algorithms offered some improvements to the pure dead reckoning based positioning on sensor values, but we see significant tracking dropouts when we have no tracker images available to the onboard cameras. Sometimes, this can be overcome by cleverly populating our environment with tracking images that seem native to the training area. The non-uniform fiducial markers we employ are geotypical in appearance and offer a benefit in being natural to the end user application. Unfortunately, pre-positioning the markers relative to the physical camera on the mobile device can be error prone. The graph below (Figure 4.) illustrates tracking dropout, and shows a correlation between the distance of the fiducial and the AR camera speed when we lose tracking.

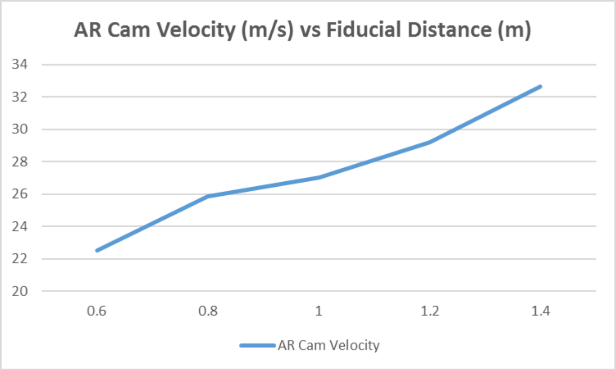

Figure 4: Distance from Fiducial and AR Velocity when tracking is lost.

Our experiment to find the AR pose accuracy leverages a high-quality camera mount with a built in levelling system for all degrees of freedom. We collected data at various distances from the center of our fiducial marker with a built-in grid for Cartesian correlation. A correlated grid incorporated into the 3D virtual environment is used to correlate and calibrate the measurement system. Figure 5 illustrates the experiment setup that we used to measure the Pose Estimation software. All lighting conditions and viewing angles to the fiducial were held constant. We collected a data sets at various heights (the y-axis) from fiducial marker to gauge a single variable change against the quality of position tracking.

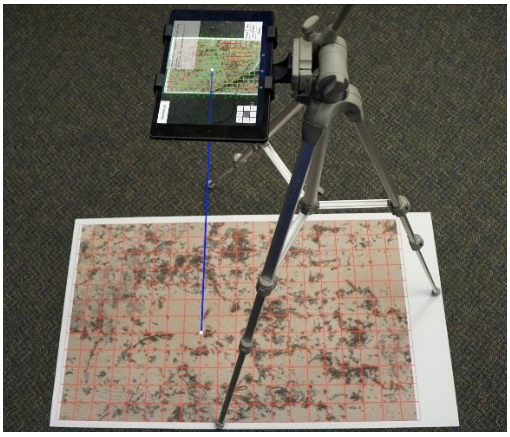

Figure 5. A System to Measure Pose Estimation Accuracy

Distance from Trackable (m)	Number of Samples (N)	Average Pose Estimation Error (cm)	Standard Deviation (cm)
0.8	147	2.7	0.49
1.0	147	3.6	0.81
1.2	147	5.7	1.73
1.4	123	7.3	1.61

Table 1. Average Pose Estimation Error

Table 1 shows that the accuracy of the pose estimation increases with distance from the trackable image. This conclusion can be attributed to per-frame image quality that decays with the focal distance capability of the on-board camera.

Due to the nature of the applicattion, we needed to characterized the tracking quality throughout the entire swing of the IED detector. Understanding the tracking error as it chages as you move relative to the fiducial marker is critical to presenting a quality AR based simulation. Initially we theorized that the fiducial measurement error was also related to the orientation of the marker with respect to the physical camera. The physical camera 16:10 ratio is a landscape field of view, but our marker is presented in a portrait mode. A second variation of data collection was conducted with the fiducial in an aspect ratio consistent with the physical camera. We simply rotated the marker 90 degrees to align the physical camera into proportion with the fiducial extents. The difference in landscape was only 3mm different than the results in portrait mode. This result speaks to the consistency in the Vuforia API performance.

When we overlay the accuracy over a representation of the features, we noticed the direct relationship with camera location and the pose estimation accuracy outcome. Figure 6 shows a map of pose estimation error relative to the real camera position. The yellow “+” symbols are individual tracking features discovered in the overall image. In our experiment, the Vuforia pose estimation improves when the physical camera is centered with respect to the tracking image.

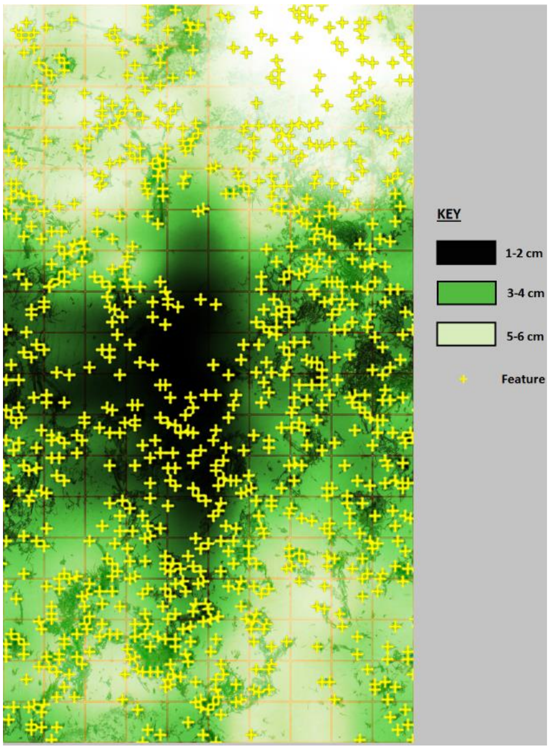

Figure 6. Accuracy Heat Map

For our use case, an important relationship exists between the distance from the trackable image, and the rate in which you can maintain quality tracking. In practice, a Soldier will travel at a suggested 1 m/s while covering a lane of up to 1.5m in width. Since our fiducial tracking system will need to be able to maintain our swing rate requirements while maintaining quality tracking we can see that we will need to maintain dense marker coverage within camera view.

2.3. AR FIDUCIAL FOR SOIL REPRESENTATION

With the AR Swing Trainer application, clever utilization of the Fiducial improves the application experience visually while providing a rapid way to modify the simulation environment. Our trackable images are correlated to geo-specific training environments, meaning that the regional soil conditions are changed with the visual tracker. Employment of scalable fiducials (Neumann, et al 1999), allows us to create convenient way of providing a "close in" view of the world or quickly change our trackable to a different dimension to achieve a zoomed out effect. We use a straight forward mapping approach to load a new terrain environment as soon as a new marker is detected. Since our markers are integrated using non-uniform high quality dense imagery, we have noted that pose estimation can be less consistent than employing the pre-made markers provided by the Vuforia software.

Soil representation change causes the AR IED simulator to react differently as the user employs the device. We can simulate a variety of soil conditions: mineralization, density and humidity can be parametrically associated with an individual fiducial marker. Using this approach, the individual Soldier can practice in soil conditions that could only be achieved by building expensive ranges or travelling to the remote location where the real environmental conditions exists.

3. VIRTUAL REALITY APPLICATIONS

3.1. VR IED Trainer

We developed open-ended VR game engine-based software that draws on concepts borrowed from game design. Our VR environment is a high quality, third person view with an accurately modeled, real training range from Ft. Benning, GA called the McKenna MOUT facility.

This VR App offers the user multiple training scenarios, an advanced mode, and step-by-step guided tutorial for first-time users. This in-game tutorial style is consistent with modern commercial games that throw the player right into the action and point out the user controls as you progress through the introductory levels. This approach eliminates the need to train the user on how to use the training application. A "game within a game", or mini-game is included to implement precision sensor positioning.

Incorporation of a progressively more challenging set of IEDs that mimic actual training ranges was important to the training subject matter experts (SMEs). These SMEs proofed the simulated training range and offered constructive criticism to increase the training validity and constructive granularity of the overall training scenarios. The VR game is technically challenging, and has a very precise mini-game scenario that is embedded within the main vignette(s) to offer fine-grained sensor control and interaction. The responsiveness, competitive nature and visual quality of the environment contribute to the end users' enjoyment and by extent their willingness to participate in extensive training (Sveistrup, Thornton, Bruant et. al. 2004).

Soil Type: Medium
IEDs Detected: 0/2
Mode: MD Only

Figure 7. VR Training Scenario deployed on an Android Tablet

3.2. VR Hardware

IST and ARL have evaluated numerous consumer-off-the-shelf (COTS) candidate deployment targets for the VR HMD. We currently deploy the App to a tablet and phone form factor for common Android systems, but for immersion, we have created versions that run in HMDs as well. We have leveraged the Oculus Rift, HTC Vive, and the Samsung Gear VR to test the App in an immersive environment. An important benefit of an untethered VR solutions is the ability to train without an expensive personal computer. At the time of writing, the Google Daydream View and Samsung Gear VR are the most capable and offer the most mature API's for application development. UCF/IST chose to focus on the Samsung Gear VR for development since the Samsung "S" and "Note" lines of phones are the official devices used by large U.S. Army program of records such as Nett Warrior.

We chose the 3rd person perspective for our VR IED trainer as it offers best visual coverage of the simulated range environment. A trainee can observe the sweep speed of the Master Trainer Avatar as it navigates the environment.

The built-in controller on the side of the Samsung VR are assigned to fine tune VR camera controls while the Gear VR controller is assigned to the larger and more in-depth controls for Avatar locomotion and key IED detector functionality.

3.3. VR Game Environment

The real-world motion of sweeping a dual head sensor to reveal an (IED) is complex and precise. The timing of the sweeping motion coupled with the forward velocity are intentionally synchronized to provide optimal lane coverage with the onboard sensors. If the motion of the sensor sweep is too fast, slow, wide or shows abnormal tilting, called bowing or cupping, then it can induce safety issues by providing negative training.

Since the ARL supplied a subject matter expert (SME), we were able to capture and render the Master Trainer avatar in the mine detection training application performing the proper sweep techniques. This includes the techniques used to determine the outline pattern of a potential IED. This precision animation would take several weeks to animate using key-frames and would be error-prone. IST leveraged a commercial motion capture (MoCap) system and captured the animations with a real world SME, a Sergeant assigned to the ARL. We used a passive MoCap system that allowed to capture the motions of the Soldier and the detector. In this way were able to use the most accurate motion data possible. The motion data from the Soldier also takes into account the weight of the Detector, which is fairly significant, as the Soldier is sweeping with the unit.

We employed photogrammetry to implement an ultra-high fidelity representation of the IED detector. The IED detector complete with individual buttons was captured into a very dense point cloud consisting of over 19,000,000 individual points. Seventy-two individual camera positions were used as we spun the real IED detector to collect the surface mesh and photorealistic textures. Our initial ultra-high resolution model weighs in at over 80,000 polygons.

Figure 8. High Resolution Photogrametric Model

Some modern mobile devices have very capable 3D graphics capabilities. We have targeted devices with high-quality graphics processing such as the NVidia Tegra chipset and the Samsung Exynos SoC solutions. While these mobile solutions offer higher polygonal rendering capabilities than previous generations, we still have to consider optimization techniques to preserve high frame rate and battery life. As an example, we applied a series of decimation methods to reduce rendering requirements for our high resolution IED detector model. Our ultra-high resolution model shown in Figure 8, was reduced from 80,000 to 12,000 polygons to facilitate loading into the desktop implementation of the VR software. Upon initial import, the 12k high resolution model caused our frame rate to drop below 10 fps on the mobile devices. To run efficiently within our game engine on the mobile device, while maintaining a large geospecific training environment, we decimated the IED model to around 2,000 polygons.

The VR training environment is a geospecific implementation of the McKenna MOUT site in Ft. Benning, GA. It represents a complex urban environment with a 2 km surrounding area represented at an average of 0.3 m post spacing resolution. A portion of the geospecific source was derived from airborne LIDAR sensors. Measurement for individual buildings and their accurate interiors were collected on site.

In the VR training game level, one of the important variables left to the trainee is the pacing speed that he selects for the Master Trainer Avatar. In the real world, doctrine calls for 1 m/s total area clearance, which is a slow and steady rate. In real world operation, incorrect swing rates can contribute to missing a concealed IED. Application users have the ability to outpace recommended search rates, and receive either real-time feedback or receive a full after action review (AAR) that offers a correlated speed vs. lane area that was missed during the clearance exercise. A user can replay his motions, the game engine renders the actual lane coverage, and then highlights the regions that were left uncleared. A summary or rollup score of IEDs located vs. the number of total IEDs in the area is displayed for user feedback.

REFERENCES

Stevens, J.A., 2015. Augmented Reality Technology in U.S. Army Training (WIP). 2015. The Society for Modeling and Simulation, Proceedings of the Summer Simulation Conference #88.

Kato, A.H. and Billinghurst M., 1999. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. (IWAR '99) Proceedings. 2nd IEEE and ACM International Workshop on

Neumann U., You S., Cho Y., Lee J., Park J., 1999. Augmented Reality Tracking in Natural Environments, Mixed Reality - Merging Real and Virtual Worlds, Ohmsha and Springer-Verlag, 101-130

Sveistrup H., Thornton M., et al. 2004. Outcomes of Intervention Programs using Flatscreen Virtual Reality. Proceedings of the 26th Annual International Conference of the IEEE EMBS.

AUTHORS BIOGRAPHY

Dean Reed (B.S. in Computer Science, UCF 2000) is a Senior Associate for Simulation with the Institute for Simulation and Training (IST) of the University of Central Florida. Mr. Reed is a veteran of the U.S. Army, leads a team of developers at IST. He has worked on a vast array of projects under the auspices of the University including NASA Vision Spaceport. He is currently managing team efforts directed at evolving future training ranges on behalf of the U.S. Army.

Latika (Bonnie) Eifert (M.S. in Computer Engineering, UCF 2003). As Science and Technology Manager at the U.S. Army Research, Development and Engineering Command (RDECOM), Army Research Laboratory, Human Research Engineering Directorate, Simulation and Training Technology Center (ARL-HRED ATSD) located in Orlando, Florida, Ms. Eifert manages several projects associated with simulation and training. She is also supporting the Defense Advanced Research Project Agency (DARPA) by managing research program efforts.

Shane Reynolds is a graduate of UCF in Digital Media. He has specialized in developing compelling 3D content and mobile game engine development for over 10 years. Mr. Reynolds is a veteran of the U.S. Air Force. Shane is a Research Associate at the Institute for Simulation and Training where he has been a faculty member since 2008. His primary activities involve research and integration of modern technologies to train dismounted Soldiers at the squad level. Currently, his focal areas are technologies involving virtual reality, augmented reality, and photogrammetry.

Travis Hillyer is a Science & Technology Manager at the US Army Research Laboratory, Orlando, FL. He is a Mechanical Engineer by training and has been supporting research efforts in Live Training for ARL and DARPA for the past two years. The focus of his efforts with the Army Research Lab have been on characterizing and modeling free-space optical communication systems for improved Live Training realism and advanced ballistic simulation systems.

Clive Hoayun is a Computer Scientist and graduate of the University of Central Florida. While pursuing a Master's degree of Computer Science, he works with the IST working on a variety of U.S. Army projects. He excels at creating novel mobile applications and has interests in evolving compelling AR apps. Clive is a veteran of the U.S. Marine Corps.