Military Vehicle Training with Augmented Reality

Jonathan Brookshire, Taragay Oskiper, Vlad Branzoi, Supun Samarasekera, Rakesh Kumar

SRI International
Princeton, NJ
{jonathan.brookshire, taragay.oskiper, vlad.branzoi, supun.samarasekera, rakesh.kumar}@sri.com

Sean Cullen, Richard Schaffer

Lockheed Martin Mission Systems and Training
Burlington, MA
{sean.cullen, richard.l.schaffer}@lmco.com

ABSTRACT

In order to be effective in the field, the military trains warfighters to operate its many ground vehicles. The goals of training are for the warfighter to learn vehicle and weapon operations and dynamics (e.g., how the vehicle and gun turret work and "feel") in live tactical situations. Additionally, because many vehicles require multiple operators (e.g., a gunner and driver), team coordination is an important element of the tactical training.

The military employs both live and virtual reality training to achieve these goals. Live training, especially gunnery, requires significant facilities and range infrastructure and is also limited to specific sites due to safely restrictions. Such training events generally require travel/transportation to CTCs and ranges. Unfortunately, live training is expensive. In this paper, an augmented reality based vehicle training system is presented. The trainees are able to drive on physical terrain and engage virtual entities for tactical and gunnery training. By augmenting the real world using virtual entities and effects, along with existing training aids and devices, training anywhere and anytime is enabled.

The details of the vehicle-borne augmented reality system for augmenting both the driver's periscope and the gunner's remote weapon sight are presented. The system relies on inertial measurements, cameras, and GPS to provide jitter free, robust and real-time 6-DOF (degree of freedom) pose estimation. These poses are used to render synthetic targets (e.g., dismounts, technical, target) to the driver and gunner. An iPad style instructor interfaces controls the augmented engagement and provides student scores.

The system is evaluated on an Army Stryker vehicle operating in a real range. The consistency and quality of target insertions between the driver's three augmented periscopes and the gunner's augmented weapon sights are compared. The importance of each sensor is evaluated by removing its input and comparing.

ABOUT THE AUTHORS

Jonathan Brookshire is current a Senior Computer Scientist at SRI International, Princeton, New Jersey. He received his PhD from MIT in ECE. At SRI he is the principal engineer for projects on Augmented Reality for Stryker and Navigation for enabling SAR in GPS-denied environments. Prior to joining SRI, he was employed at iRobot.

Taragay Oskiper is a Senior Principal Research Scientist at SRI International, Princeton. He received his Ph.D in Electrical Engineering from Princeton University. He has over ten years' experience in developing vision-aided motion estimation and multi-sensor fusion algorithms for navigation and augmented reality for both video-see-through and optical-see-through platforms. He has acted as the lead algorithm developer for numerous augmented reality projects, most recently the Office of Naval Research AITT program, at Sarnoff and now SRI International Princeton.

Vlad Branzoi is a Computer Scientist at SRI International Sarnoff. He received his M.S. in Computer Science from Columbia University under Prof. Shree Nayar. Vlad Branzoi has over 10 years' experience in building novel sensors, integrated multi-sensor systems for training, robotics and mobile applications.

Supun Samarasekera is the Technical Director of the Vision and Robotics Laboratory at SRI International Sarnoff. He received his M.S. degree from University of Pennsylvania. Prior to joining SRI, he was employed at Siemens Corp. Supun Samarasekera has over 17 years' experience in building integrated multi-sensor systems for training, security & other applications. He has led programs for robotics, 3D modeling, training, visualization, aerial video

surveillance, multi-sensor tracking and medical image processing applications. He has received a number of technical achievement awards for his technical work at SRI.

Rakesh Kumar is the Director of the Center for Vision Technologies at SRI International, Princeton, New Jersey. Prior to joining SRI International Sarnoff, he was employed at IBM. He received his Ph.D. in Computer Science from the University of Massachusetts at Amherst in 1992. His technical interests are in the areas of computer vision, computer graphics, image processing and multimedia. Rakesh Kumar received the Sarnoff Presidents Award in 2009 and Sarnoff Technical Achievement awards in 1994 and 1996 for his work in registration of multi-sensor, multi-dimensional medical images and alignment of video to three dimensional scene models respectively. He received the University of Massachusetts Amherst School of Computer Science, Outstanding Achievement and Advocacy Award for Technology Development (2013). He was an Associate Editor for the Institute of Electrical and Electronics Engineers (IEEE) Transactions on Pattern Analysis and Machine Intelligence from 1999 to 2003. He has served in different capacities on a number of computer vision conferences and National Science Foundation (NSF) review panels. He has co-authored more than 50 research publications and has received over 50 patents.

Richard Schaffer is a Lockheed Martin Fellow and Principal Investigator at Lockheed Martin Mission Systems and Training (MST). He leads the Human Immersive Simulation Lab at MST's Advanced Simulation Centers. Richard received his S.B. degree from the Massachusetts Institute of Technology and has over 30 years of experience in modeling and simulation research and development. His areas of research have included distributed simulation, environment modeling, and immersive simulation. In 2010 he received the NTSA's lifetime achievement award.

Sean Cullen is a Sr. Staff Software Engineer at Lockheed Martin MST. He received his B.S. in Computer Science from Middle Tennessee State University. Sean Cullen has over 17 years' experience in military based modeling and simulation. He has been the Project Engineer on multiple augmented reality programs at Lockheed Martin and has extensive experience in 3D graphics.

Military Vehicle Training with Augmented Reality

Jonathan Brookshire, Taragay Oskiper, Vlad Branzoi, Supun Samarasekera, Rakesh Kumar

SRI International
Princeton, NJ
{jonathan.brookshire, taragay.oskiper, vlad.branzoi, supun.samarasekera, rakesh.kumar}@sri.com

Sean Cullen, Richard Schaffer

Lockheed Martin Mission Systems and Training
Burlington, MA
{sean.cullen, richard.l.schaffer}@lmco.com

INTRODUCTION

Like any complex task, mastering the operation of military vehicles requires training. This training often includes classroom learning where discussions and lectures provide a foundational understanding of the vehicle and tactics. Of course, in-vehicle training is essential for operators to learn, and develop the muscle memory for, vehicle and weapons control/dynamics. Further, tactics and team coordination are often practiced as part of the in-vehicle exercises. Although clearly an essential element, in-vehicle training can be costly and access to training ranges reduces training frequency. Additionally, when live munitions are required (e.g., during live gunnery training), necessary safety precautions can greatly add to the expense and infrequency.

As a result of the challenges associated with in-vehicle training, an array of solutions has been developed. Live training (see Table 1) takes place in a real vehicle on an equipped range. Often, targets pop-up at a set of fixed locations or travel along installed rails. Munitions are live, and safety is a key concern. The operator drives the real vehicle on real terrain and can experience the system's real dynamics.

Table 1. Comparison of different in-vehicle training approaches

	Environment	Vehicle	Targets	Weapons	Facility Requirements
Live-fire	Real	Real	Synthetic/Passive	Live	Equipped range, targets
Laser-based	Real	Real	Real	Simulated	Equipped range, Equipped vehicle, targets
Augmented reality	Real	Real	Synthetic/Responsive	Simulated	Equipped vehicle
Virtual reality	Synthetic	Synthetic	Synthetic/Responsive	Simulated	Simulator

Given the dangers and costs associated with live weapons, laser-based solutions (e.g., MILES) were developed. An advantage of these systems is that, by using a laser instead of live weapons, vehicles and dismounts can jointly train. A disadvantage of this approach is that weapons and targets must be equipped with the laser system, and the range must be outfitted if performance statistics are desired.

At the other end of the spectrum, virtual reality solutions provide a video game-style interface. The environment and all system dynamics are synthetic and a simulator is required. This simulator often includes some physical aspect of the vehicle (e.g., dashboard or instrumented weapon). An advantage of these systems is that any terrain, target, and scenario can be simulated. However, simulators are expensive and the synthetic vehicle dynamics do not

Figure 1. Periscope system installed on the Stryker

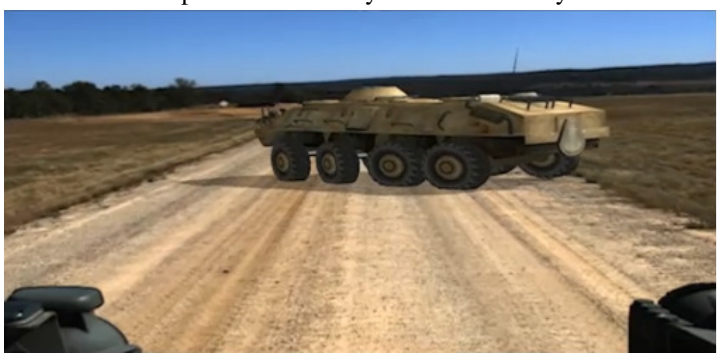


Figure 2. Augmented views of live video from the periscopes at Ft. Benning

often provide a substitute for the real thing.

This work focuses on an augmented reality (AR) alternative designed to address some of these shortcomings, as shown in Figure 2. Augmented reality uses a live video feed and overlays synthetic targets on top of that video feed. Using video game techniques, the targets are transformed to appear as if on the physical terrain in the video feed. As with live-fire and laser-based training, the environment and vehicle are real. Similar to virtual reality, the targets are simulated and can be made responsive to the trainees' actions. Weapons are simulated, and performance statistics can be gathered. As with all solutions, some infrastructure is required, but only the vehicle must be equipped with the AR system.

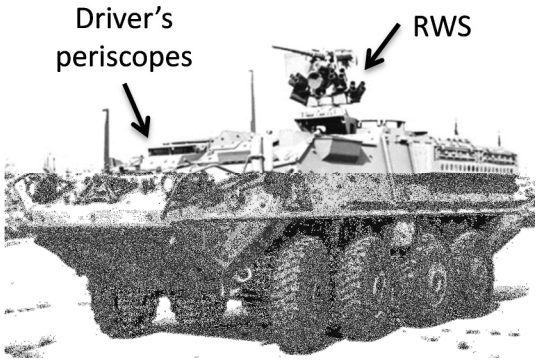


Figure 3. The driver's periscopes and the RWS display of the Stryker vehicle are augmented as a motivating example.

In this paper, the AR training system is described. As a motivating example, an Army Stryker vehicle is outfitted. In particular, augmented displays for the driver's three periscopes and the gunner's remote weapon station (RWS) fire control unit (FCU) are provided. All four AR units are synchronized and display the same augmented targets, allowing the driver and gunner to not only practice their individual functions but also their team coordination.

A complete AR system has two main components: (1) a localization system which provides the pose of the vehicle and (2) a rendering system which displays the synthetic targets which would be visible at the current pose. Both qualitative and quantitative analysis of the system's performance is provided. Specifically, the consistency of the AR renderings between the periscopes and the FCU are examined. Because the quality of the renderings depends on the localization quality, localization against ground truth is compared. Further, the importance of each sensor and how it contributes to the final result is examined.

In the Technical Approach section, the system hardware is presented, and in the Algorithms section the navigation, rendering, and foreground obstacle modelling software components are discussed. Finally, the Vehicle Integration is discussed and results from experiments are presented.

TECHNICAL APPROACH

The objective of the system is to provide an AR, vehicle-borne training system on the Stryker vehicle. The three periscopes of the driver and the RWS FCU display of the gunner are augmented. In this section, the hardware system is described. Figure 4 depicts a system block diagram. The periscope-mounted system includes three sensor packages, one mounted immediately

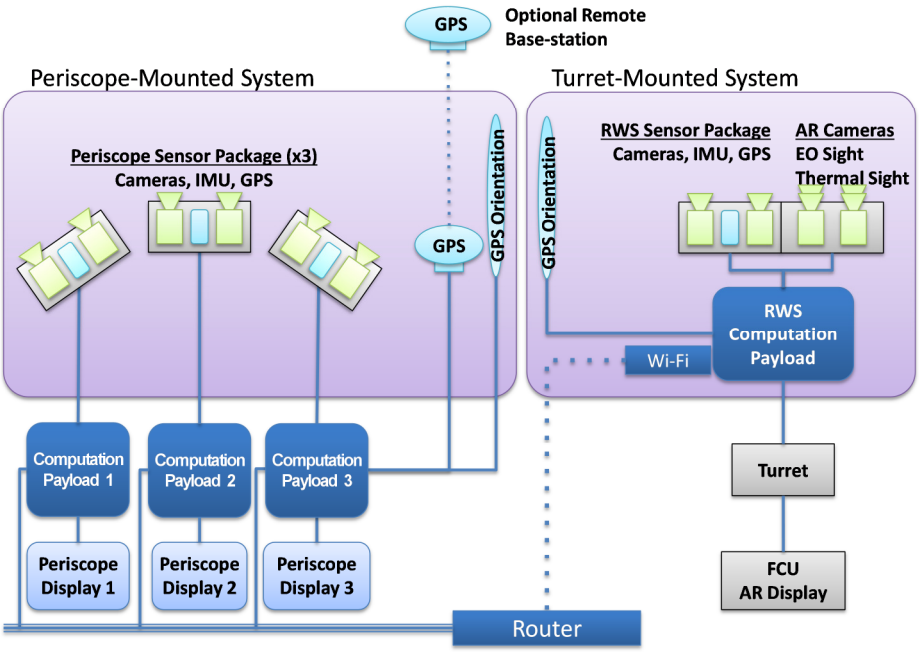


Figure 4. The system block diagram shows four basic AR sensor packages (periscope x3 and RWS) and the displays (three periscope displays and FCU display).

above each of the driver's periscopes. The RWS sensor package, along with the PC, is mounted on the turret, above the pan-tilt joints. For the RWS, the AR entities are rendered on the digital video; the video is then converted to analog and injected to be displayed on the real FCU.

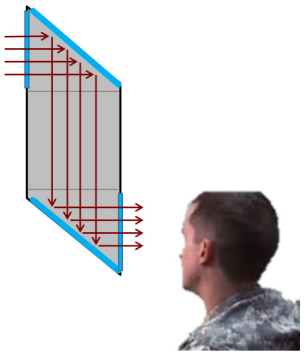


Figure 5. The original periscope reflects light to the driver.

Periscope Sensor Package

As suggested in Figure 5, the driver's three periscopes reflect light to the driver, while protecting him from direct enemy fire. In order to augment these views, a sensing package is mounted directly above the periscope and an easily removable LCD is installed over the periscope viewport (Figure 6). This sensing package consists of:

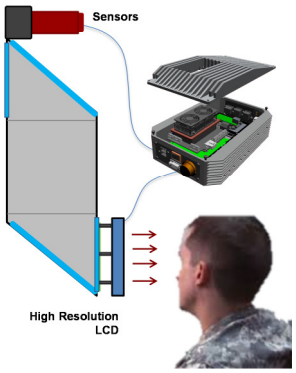


Figure 6. The augmented reality periscope projects the view from an adjacent camera to an LCD

1. **Two Allied Vision GT1920 GigE cameras.** The first "augmentation" camera is a color camera and located immediately above the periscope (as suggested in Figure 6). The second "navigation" camera is 30cm above the first camera and used as the primary navigation camera (see Algorithms). Both cameras run at 20Hz and have a horizontal field of view (FOV) of about 50 degrees.
2. **Microstrain 3DM-GX3 IMU.** The Inertial Measurement Unit (IMU) provides high frequency acceleration and rotational rate data which captures high speed movements.
3. **Custom trigger board.** The Arduino-based trigger board provides a 20Hz signal which synchronizes the cameras, IMU, and GPS.
4. **uBlox GPS receiver.** The uBlox GPS receiver is not strictly necessary, but provides a convenient way to provide a GPS timestamp with every trigger pulse. This allows the three periscopes to be synchronized without requiring a shared trigger.
5. **GPS heading receiver.** The SITEX is a dual-antenna marine GPS receiver which provides heading information, especially important when the vehicle is not moving. A single SITEX is shared between all periscopes.
6. **Differential GPS receiver.** The Trimble receiver integrates corrections from an optional, nearby base-station to provide centimeter-accurate latitude and longitude. A single Trimble system is shared by the periscopes.
7. **Computational payload.** All localization and rendering is performed in real-time on-board the vehicle using an Intel i7 3GHz small form-factor computer.

Remote Weapon Station (RWS) Sensor Package

Generally, the RWS consists of a pan-tilt weapons mount on the vehicle's exterior, controlled from the FCU within the vehicle's cabin. The user interfaces to the FCU via push-buttons and control grip (joystick); a live video feed is displayed on the FCU monitor for the operator. Beneath the RWS weapon mount is the Sight Servo Assembly (SSA), connected by another rotational joint to correct for weapon elevation and parallax. The RWS' native sensor package includes an electro-optical (EO) visible light camera, long-wave infrared (LWIR) camera, and optional laser range finder (LRF).

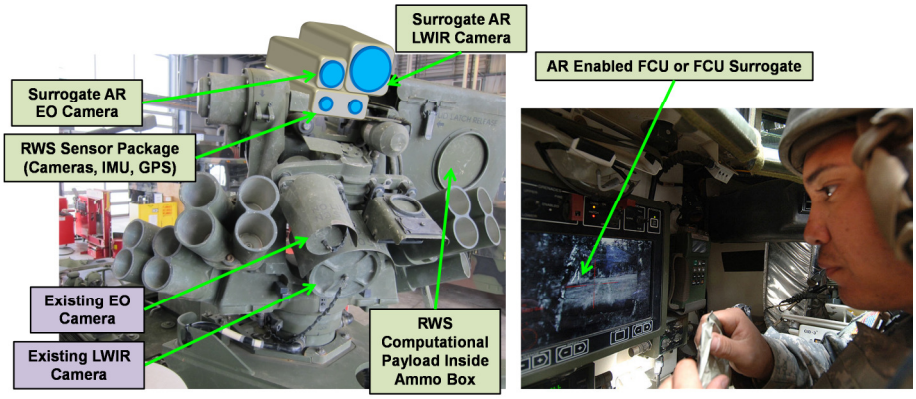


Figure 7. The RWS sensor package is installed on the gun mount and the computational payload replaces the ammo box.

The goal of the RWS augmentation is to render artificial entities on the variable zoom EO and LWIR cameras. As described in the Vehicle Integration section, the video feeds from the RWS' native sensor are interrupted and signals from surrogate EO and LWIR cameras are injected. (In future work, the native EO and LWIR sensors could be augmented, but using surrogate cameras allowed us to conduct experiments on a test vehicle when the Stryker was not available.) As a result, the sensor package for the RWS is similar to the periscope package described previously. The two significant additions are (1) a FLIR Tau2 640x480 LWIR camera and Ophir SpIR variable zoom lens, and (2) an Allied Vision GT1380C camera with a Fuji H22x11.5A-M41 zoom lens. Figure 7 (left) shows the installed RWS sensor package; on the right, the FCU monitor displays AR targets to the operator.

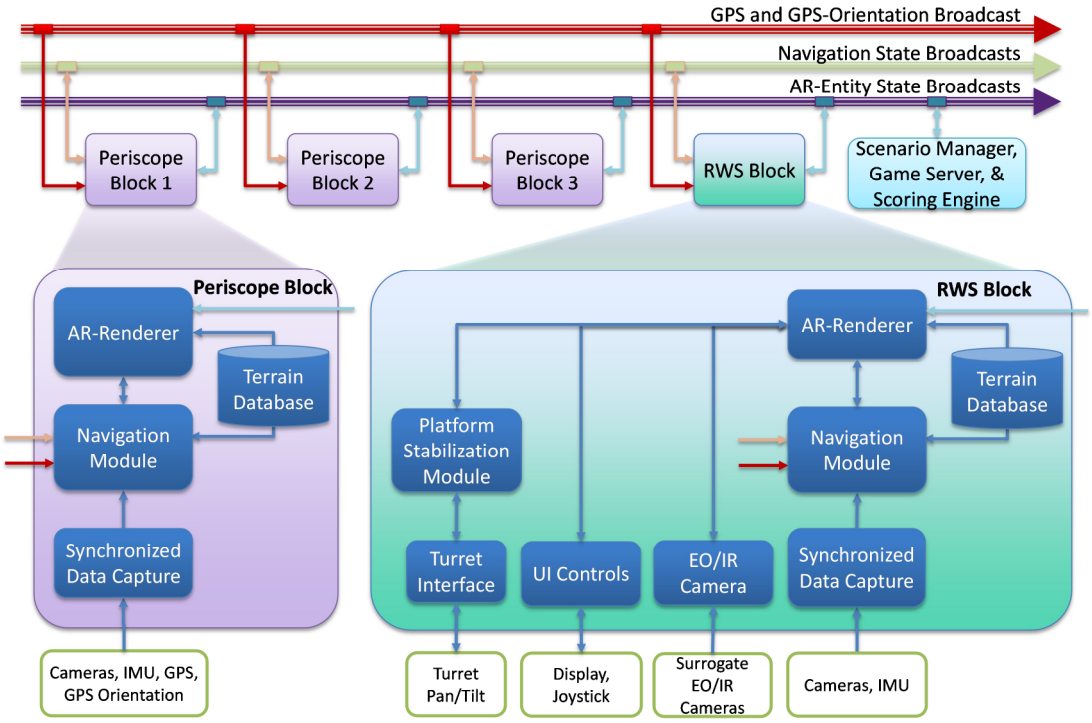


Figure 8. Algorithm data flow

ALGORITHMS

The block diagram in Figure 8 illustrates the different system algorithms. Each Periscope Block is identical and similar to the RWS Block. The four blocks are connected over a single Ethernet backbone, over which differential GPS, orientation GPS, navigation data, and AR-entity status (e.g., position of moving objects) is shared. Each block consists of an AR-Renderer which renders the AR targets visible from the current pose, provided by the Navigation Module. A Terrain Database provides the Navigation module with a reliable mean sea level height. It is also used by the AR-Renderer to generate targets that "sit" on the ground, hide behind terrain, and can follow terrain features as they drive. The terrain map is particularly important for generating targets that are realistic and can appear from behind terrain features during training.

The RWS Block also has several unique modules which allow it to interface with the RWS and FCU. In particular, the signals to the EO camera, LWIR camera, LRF and gun are decoded. This allows the augmentations to be displayed on the real FCU. In the following sections, the Navigation Module and the AR-Renderer are detailed.

Navigation Module

Past experience with AR (Oskiper, 2011) suggests that in order to produce a compelling AR experience, the jitter and latency must be tightly controlled. Jitter refers to how a rendered entity "jumps around" on the screen. Jitter of more than 1 pixel is easily detectable by the human eye. Latency refers the delay between AR target movement and movement of the environment. When latency is more than 1 frame (~1/20 of a second), the targets appear to float

above and below the ground and lag the motion caused by the vehicle. Jitter and latency are caused by noise in the pose estimates from localization and latency in the pose calculation, respectively.

The Navigation Module builds on previous work (Oskiper, 2012) which uses an error-state (indirect) Extended Kalman Filter (EKF) to fuse measurements from the IMU, cameras, and GPS. Many filtering solutions require an estimate of the platform dynamics to propagate the state forward. The error-state (EKF) does not require these unknown and difficult to estimate dynamics. This is because the filter estimates not the vehicle state, but rather the error between the IMU and the vehicle state. This formulation is advantageous because, first, it preserves the high-frequency IMU motions -- especially appropriate for vehicles which vibrate when idle and travel over rough terrain. Second, the highly non-linear vehicle state is replaced by the more linear error state, and more easily estimated by an EKF.

The reader is referred to (Oskiper, 2012 and Oskiper, 2011) for complete details, but an overview of the indirect EKF is provided here for completeness. The filter provides 6-DOF pose estimates for navigation by generating relative visual measurements at the feature track level and marginalizing out the 3D feature points, obtained via multi-view triangulation, from the measurement model. This reduces the state vector size and makes real time implementation possible by keeping computational complexity linear in the number of features. The algorithm incorporates two cameras (both employed in monocular fashion) and additional global measurements in the form of global heading from the SITEX GPS.

The total (full) states of the filter consist of the IMU location $\mathbf{T}_{IG}$, the gyroscope bias vector $\mathbf{b}_g$, velocity vector $\mathbf{v}_{IG}$ in global coordinate frame, accelerometer bias vector $\mathbf{b}_a$ and ground to IMU orientation $\mathbf{q}_{GI}$, expressed in terms of the quaternion representation for rotation (Kuipers, 1998). Hence, the total (full) state vector is given by

$$\mathbf{s} = [\mathbf{q}_{GI}^T \quad \mathbf{b}_g^T \quad \mathbf{v}_{IG}^T \quad \mathbf{b}_a^T \quad \mathbf{T}_{IG}^T]^T .$$

During filter operation, ground to IMU pose $\mathbf{P}_{GI}$ is predicted prior to each update instant by propagating the previous estimate using all the IMU readings between the current and previous video frames via IMU mechanization equations. After each update, estimates of the errors (which form the error-states of the filter) are fed-back to correct the predicted pose before it is propagated to the next update and so on.

In this work, the new SITEX heading sensor is integrated. This sensor uses two GPS receivers to measure the sensor's absolute heading, a direct (albeit rotated) measurement of $\mathbf{q}_{IG}$. The innovation associated with this global measurement can be calculated by converting $-\mathbf{q}_{GI} \otimes \mathbf{q}_{SI}$ to R_{SG}, where $\mathbf{q}_{SI}$ is the calibrated SITEX reference frame expressed in the IMU frame and R_{SG} is the rotation matrix representing the orientation of the SITEX in the ground frame. The expected heading is then $\gamma = \text{atan}\left(\frac{z_0}{z_2}\right)$, where z is the third column of R_{SG}. The heading is not calculated using a rotation matrix to Euler angle conversion, because such a conversion would give incorrect results. This is because the physical SITEX measures a heading by projecting the two receivers onto a 2D plane and neither of the other two orientations are considered.

Rendering

The Rendering subsystem creates a virtual world that closely represents the real world around the vehicle. Accurate terrain elevation data is used to generate a 3D polygonal model of the area and the tracking system informs the renderer where in the virtual model the vehicle is placed. An accurate camera pose and corresponding imagery captured from both the electro-optical (EO) and infrared (IR) cameras form the basis of the augmented imagery presented to the trainee.

For each frame, the system first copies the video imagery as it was received from the camera. Then, the system uses the camera pose and projection matrix (representing the current

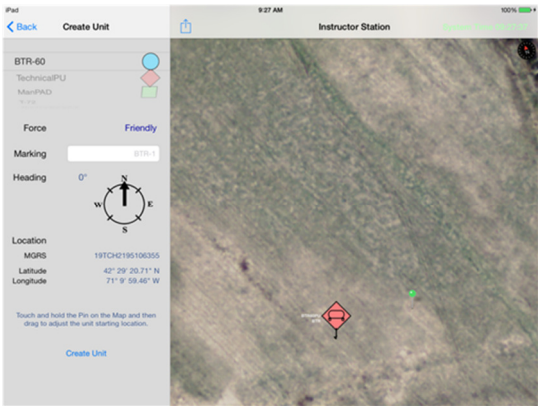


Figure 9 - Instructor Tablet

camera settings) to render the 3D terrain model into a depth buffer. Next, any virtual entities/effects are rendered to the scene with the depth buffer ensuring items blocked by terrain features will be occluded. The edges of virtual entities are slightly blurred with the camera imagery. This avoids the virtual entities appearing artificially sharp against the very good, but less than perfect, real world background generated by the real system's sensors and display. The rendering system is built using the Unity 3D game engine. This allowed for the easy inclusion of animated characters and shadow effects that add to the realism of the generated scene.

Instructor Tablet / Scenario Capabilities

Since this system was designed to facilitate gunnery training, the software allows for creation of scenarios that resemble live fire training events. The system has the ability to simulate BTRs, T-72s, Technical Pickups, and dismounted enemy combatants. The entities are simulated using the Unity game engine and can be commanded to move from point to point while taking damage as appropriate. The vehicles have the ability to drive out of and into virtual fighting positions that conceal the vehicle from the Stryker. If targets are not engaged within the specified period of time, they can return to the fighting position. This allows the instructor to control target exposure times in a way similar to those utilized in gunnery tables on live-fire ranges. Destroyed vehicles can be configured to remove themselves from the scene or to remain, depending on the instructor's intent. Instructors oversee training from the Instructor tablet that provides system monitoring/control as well as scenario authoring capabilities. The Instructor tablet connects to the system via Wi-Fi, and the scenario is simulated directly on the rendering computer.

Foreground Obscurations

As shown in Figure 10, the view from the periscopes is partially occluded for some configurations of the Stryker. Here, the armor obscures the bottom of the camera's view; an artificial target rendered in this area would incorrectly appear to be in front of the armor. The solution is to calculate a foreground mask and apply this mask to the rendered entities. In this way, targets are occluded by the armor, just as are real objects.

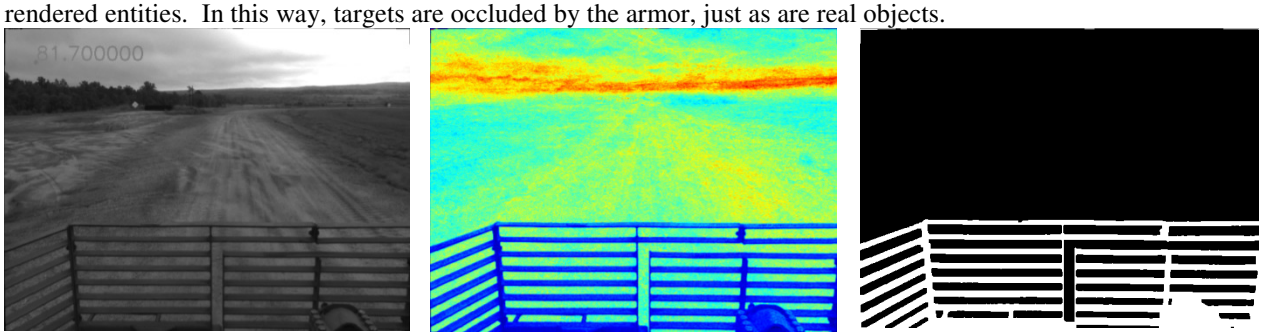


Figure 10. The view from the periscopes on some Strykers is partially obscured by the armor (left). A foreground mask is calculated (middle) and thresholded (right) during a calibration procedure where the Stryker is driven around.

It is not practical to manually create the foreground mask, as the armor is complex and slightly different on each vehicle. Instead, data is collected during a calibration procedure by simply recording the video while driving. Then, in an offline step, a background detection algorithm produces a foreground mask (Figure 10, right). An OpenCV (Bradski, 2000) background detection algorithm, $\mathrm{bgs}(f)$, is used which accepts an input frame, f, and returns a binary mask highlighting pixels different from the last frame. This algorithm is suitable for background detection; that is, it removes parts of the scene which do not change. However, it is desired to remove parts of the frame that do change. Thus, integration and thresholding are performed as $\sum_f \mathrm{bgs}(f) > t$. An example of the integrated image is shown in Figure 10 (middle). The threshold, t, is manually selected to produce the final mask shown in Figure 10 (right).

VEHICLE INTEGRATION

The system is installed on both a Stryker vehicle at Ft. Benning, Columbus, GA (Figure 11, top) and on a surrogate test vehicle (Figure 12) for testing. The periscopes are standalone and simply bolt to the vehicle. Power conditioning, displays and computation are mounted inside the vehicle. The RWS sensor package integrates with the FCU (see Figure 8) to tap into the control grip, EO camera, LWIR camera, and LRF messages. The RWS sensor

package injects analog video back into the FCU, bypassing the real EO and LWIR camera feeds. In this way, the driver simply looks at displays in front of his periscopes and the gunner uses the actual FCU.

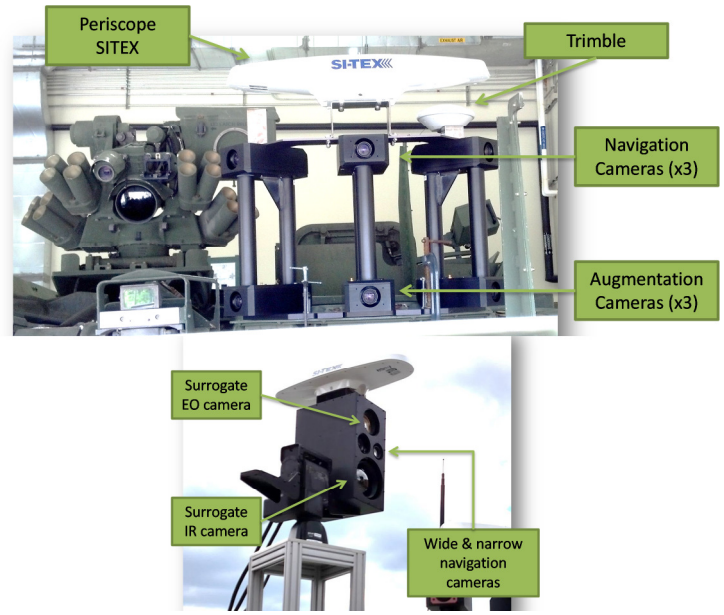


Figure 11. The periscope system installed on the Stryker (top) and a close-up of the RWS sensor package (bottom)

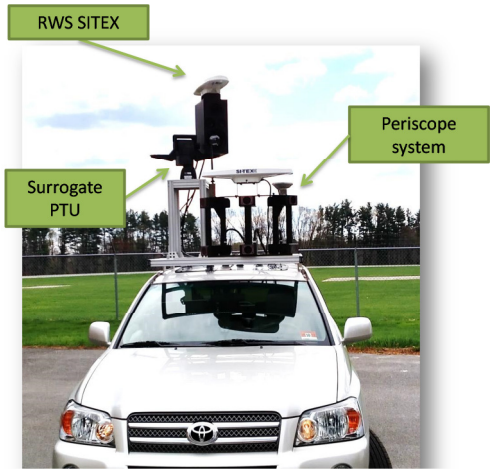


Figure 12. To test the system, a surrogate vehicle is used

EXPERIMENTS & RESULTS

Figure 13 shows augmented views from the RWS. Here, the background video is in grayscale and highlights the augmented entities. For the left view, the camera zoom is set to ~3 degrees horizontal FOV; the right view is set to ~60 degrees. In future work, the full variable zoom capability of the camera will be supported.

In order to qualitatively test the consistency of the renderings between the narrow FOV RWS, wide FOV RWS, and periscope cameras, an AR target is placed in the scene and viewed from three different cameras. The first row of Figure 14 shows the operator's view of the same AR vehicle. The second row shows post-processed views (digitally zoomed) of the periscope and wide FOV RWS views for comparison. The RWS cameras (b,c) agree well, within a pixel of the wide FOV camera. There is some difference noticeable between the narrow RWS (b) and periscope cameras (a). However, this 40 pixel shift (half the radar dish) is actually correct for a target at 250m and results from the displacement between the periscope and RWS cameras (about 0.75m).

Navigation

The ability to accurately render targets depends on the accuracy of the localization. As discussed in the Technical Approach section, a variety of sensors for localization are employed. Here, the contribution of two new sensors, the differential and heading GPS systems, are analyzed. Data over the three routes shown in Figure 15, ranging from ~640km to ~3km, is collected. All routes start and stop at the same "parking spot." As expected, good general agreement between the raw differential GPS and the estimated path (which itself uses differential GPS) can be noted.

Ideally, it is desired to compare the estimated localization with known ground truth at every time. Lacking such universal ground truth, however, error is measured when ground truth is available. First, note that all three paths start and stop at nearly the same point; thus, ground truth indicates that the start and stop locations should be at the same point. Second, the

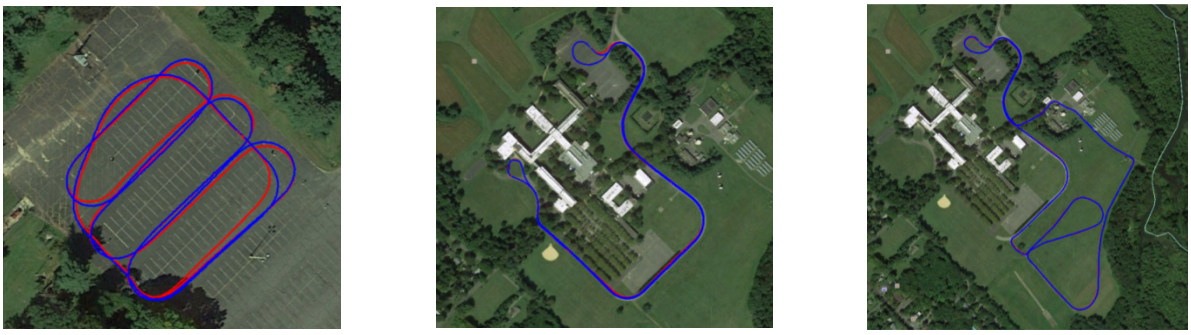
Figure 13. Augmented views from the RWS on the surrogate test vehicle

paths (a) and (b) travel on relatively flat, paved roadways. Thus, ground truth indicates that the vertical travel should be small. As shown in Table 2, less than a meter of error at the end of the loops and less than two degrees of heading error is experienced.

Next, how the differential and heading GPS affect the system is analyzed. In Table 3, the difference is shown between a system configuration with all sensors and with certain sensors removed or replaced. Although the heading sensor can be removed, some estimate of heading is needed to initialize the filter. For example, in previous work (Oskiper, 2012), a manual landmark procedure was used. In order to test the system without the heading sensor, the first measurement from the SITEX is used; after that first measurement, no further heading measurements are fused. A Trimble differential GPS is used, and to test the system without it, a lower-quality, standard GPS, the XSens is substituted. The degradation in Test 1 shows that the SITEX affects the position estimate by about 0.52m; it is not surprising that an improved heading helps the position estimate, as they are correlated. In Test 2, the Xsens is used instead of the Trimble; as a result, the difference is about 2m. In Test 3, the XSens and lack of SITEX result in a difference of about 2.13m, the majority of which appears to be attributable to the missing Trimble.

(a) Periscope 1 Camera **(b) Narrow FOV RWS Camera** **(c) Wide FOV RWS Camera**

Figure 14. The top row shows the operator's view of the same AR target from three different cameras. The bottom row shows digitally zoomed views of the (a) and (c) cameras for comparison with (b). The difference between (a) and (b) is due to the physical camera offset for a 250m target.



Path (a), 640m Path (b), 2.6km Path (c), 3.0km

Figure 15. Blue shows the differential GPS path, and red shows the estimated path using all sensors (including differential GPS). Paths (a) and (b) are on improved roadways; path (c) is partially off-road with elevation changes.

Table 2. Localization errors at loop closure and versus assumed level ground

Path	2D Translational Loop RMSE (m)	Heading Loop Error (deg)	Vertical RMSE (m)
(a)	0.37	-1.85	1.13
(b)	0.60	0.94	1.87
(c)	0.74	-0.37	N/A

Table 3. Root mean squared difference in position between filter output with all sensors and partial sensors

Path	Test 1 One Heading Measurement & Differential GPS (m)	Test 2 Heading GPS & Standard GPS (m)	Test 3 One Heading Measurement & Standard GPS (m)
(a)	0.44	1.89	1.96
(b)	0.86	2.03	2.36
(c)	0.28	2.03	2.07
Mean	0.52	1.98	2.13

In Table 4, the differences are shown for the same experiments for rotational error. In Test 1 and Test 3, a difference of 1.49 degrees and 1.83 degrees are present when the SITEX is removed. This is not surprising, as this sensor provides a heading estimate. The quality of the GPS position also affects heading as evident in Test2; 1.07 degrees of difference can be seen when the Xsens is used.

Table 4. Root mean squared difference in orientation between filter output with all sensors and partial sensors

Path	Test 1 One Heading Measurement & Differential GPS (deg)	Test 2 Heading GPS & Standard GPS (deg)	Test 3 One Heading Measurement & Standard GPS (deg)
(a)	2.18	1.42	2.82
(b)	1.48	1.33	1.56
(c)	0.82	0.46	1.11
mean	1.49	1.07	1.83

Finally, the same data set from Figure 14 is used and the jitter and drift of the AR insertion over the 500 frame sequence are observed. The tank jittered and drifted less than 10 pixels for the narrow FOV camera, corresponding to an angle accuracy of less than 0.05 degrees.

CONCLUSIONS

The augmented reality solution for training on vehicles has been discussed. Future work will continue to focus on the RWS, integrating the full variable zoom and focus capabilities of the EO/LWIR cameras. The work will include a variable-zoom calibration technique and variable zoom and focus rendering techniques. With that capability, the gunner will be able to engage virtual targets using the same camera controls as during live-fire training.

ACKNOWLEDGEMENTS

The material presented in this paper is based upon research supported by U.S. Army Project: Augmenting the Long-Range Sights & Periscope Sights on Army Vehicles for Embedded Training under Contract W15QKN-13-C-0083. The views, opinions, or findings contained in this report are those of the authors and should not be construed as an official Department of the U.S. Army position, policy, or decision unless so designated by other official documentation.

REFERENCES

Defense Science Board Report on Technology and Innovation Enablers for Superiority in 2030, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, Washington, D.C., October 2013, pages 60-65.

Reitmayr, G. (2006) & Drummond, T. Going Out: Robust Model-based Tracking for Outdoor Augmented Reality. In *International Symposium on Mixed and Augmented Reality*.

Bradski, G. (2000). OpenCV. *Dr. Dobb's Journal of Software Tools*.

Kuipers, J. (1998). *Quaternions and Rotation Sequences.* Princeton University Press.

Oskiper, T., Chiu, H., Zhu, Z., Samarasekera, S., & Kumar, R. (2011). Stable vision-aided navigation for large-area augmented reality. *IEEE Virtual Reality Conference.*

Oskiper, T., Samarasekera, S., & Kumar, R. (2012). Multi-sensor navigation algorithm using monocular camera, imu, and gps for large scale augmented reality. *IEEE ISMAR.*