

Experiences and Observations in Applying Augmented Reality to Live Training

Jon Barrilleaux

Peculiar Technologies • 3800 Lake Shore Ave., Oakland, CA 94610 • www.augsim.com

Abstract

Between 1992 and 1993 a series of live demonstrations were conducted that introduced the concept of Augmented Reality (AR) to live training involving ground combat vehicles. Since these early attempts, concepts for introducing virtual entities into a live training exercise, and for combining live and simulator based training in the same exercise have evolved and matured. Designs for wearable configurations are also on the drawing board, which would allow live infantry training. Besides the obvious problems of size, weight, and cost, there are many issues left to be addressed. The most challenging one was and still is accurate tracking, whether in the confines of a tank turret or building, or outdoors in a cluttered urban setting or a hilly forest.

Live Demonstrations

Between 1992 and 1993 a series of three high visibility Augmented Reality / Seamless Simulation demonstrations were conducted. Augmented Reality allowed vehicle crews to see virtual vehicles and weapon effects. Two-way Seamless Simulation allowed live and virtual vehicles to interact in real-time on the same battlefield. The work was primarily sponsored by STRICOM, the Army's command for simulation and training. The author of this paper served as the chief architect and technical manager.

The first demonstration was performed at Fort Hunter-Liggett, CA and was presented live at the 1992 I/ITSEC conference in San Antonio, TX. The second demonstration was performed at Fort Knox, KY for the 1993 February AUSA conference. The final and most ambitious effort was performed at Fort Knox, KY and was presented live at the 1993 May AUSA conference in Orlando, FL.

Figure 1 shows the 1993 May AUSA demonstration configuration. Live instrumented vehicles in the field, an M1 tank and a LOSAT anti-tank missile launcher, performed combat maneuvers against virtual attackers while accompanied by virtual platoon members. The live vehicles were located on the St. Vith range at Fort Knox in Kentucky. These were the platoon leaders for an M1 and a LOSAT platoon. Manned simulators

located at the Mounted Warfighting Testbed (MWTB) at Fort Knox provided the rest of the M1 platoon. A manned simulator located on the conference floor in Orlando, FL filled out the LOSAT platoon. The opposing force (OPFOR) of approximately 50 vehicles was provided by Computer Generated Forces (CGF) generated at the MWTB. Stealth and map displays at the conference showed the battle to the audience from a virtual world perspective. Simultaneously, live video from the St. Vith range showed the battle to the audience from a real world perspective.

Two-way communication of real-time simulation data between live and virtual players occurred through a series of data links. The live vehicles were linked to the range tower via a Time Division Multiple Access (TDMA) radio network, with a dialup line relay to the MWTB. Simulation data was exchanged in a custom compact protocol with overload management in the software to accommodate the limited bandwidth of the link. A translator at the MWTB connected the field network with their Distributed Interactive Simulation (DIS) network, which was in turn linked to the Orlando DIS network via a two-way long-haul network.

The technical groundwork laid by these demonstrations included:

- On-board real-time simulator graphics generation and tracker processing.
- Position and orientation sensing via a strap-on Inertial Measurement Unit (IMU).
- Articulation sensing via vehicle 1553 data bus and strap-on optical and ultrasound sensors.
- Two-way DIS networking in the field via TDMA.
- High-resolution terrain database correlated with ground truth.
- Basic packaging, cabling and weatherproofing for strap-on equipment on live vehicles.

Critical technical issues not addressed:

- Lack of integrated displays in the live vehicles — only a gun sight view of virtual terrain and entities was feasible at the time.
- Lack of head tracking to support integrated displays for all visual operational modes.

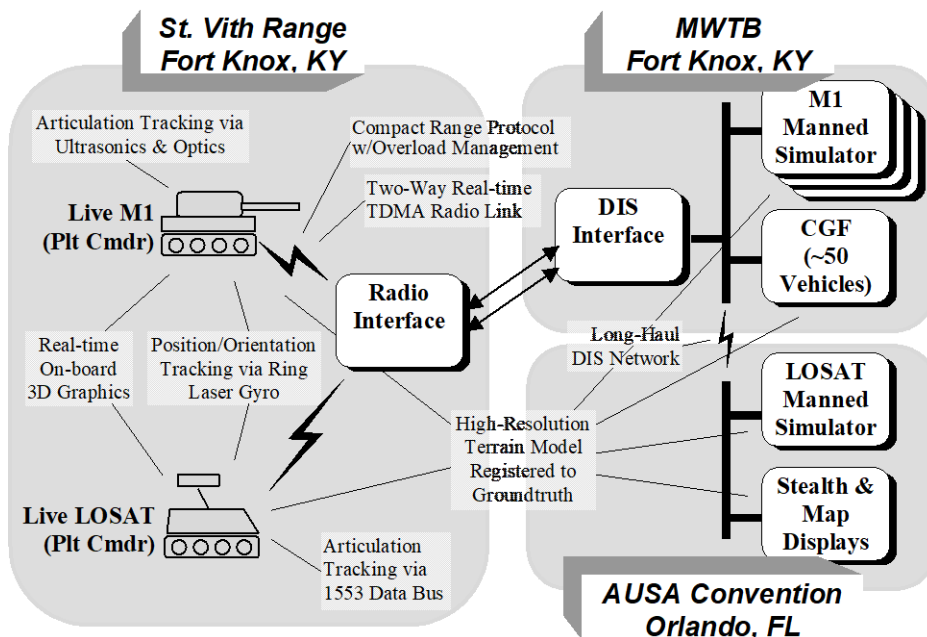


Figure 1. '93 May AUSA Demonstration. Two live vehicles in the field could see and be seen by manned simulators at Ft. Knox and Orlando, and mount a coordinate attack on a virtual enemy.

Why Do It

The motivation for the first demonstration was borne out of the simple need to do “something different” at an important trade show. At the time, manned simulators were in vogue and everyone seemed to have one, or at least the pieces for one, in the trade shows. The concept of Seamless Simulation, where live and virtual forces would somehow be combined in the same exercise, possibly on the same battlefield, and perhaps even at the same time, was just starting to circulate. We figured that putting a computer simulator on a live vehicle networked to man

simulators would certainly qualify as “something different”. The result was a compelling demonstration of the concept. Since these early demonstrations the motivations have greatly matured, and so have the operational concepts, system architectures, and benefits.

Augmented Simulation (AUGSIM)

True Seamless Simulation allows live and virtual entities to fight on the same battlefield at the same time with one another. Augmented Reality allows virtual entities to be seen and heard in the real world. In our 1995 Small Business Innovative Research (SBIR) final report on applying these concepts to training we coined the term AUGSIM — AUGmented SIMulation — to express the synergistic

combination of the two. AUGSIM allows live entities to see, hear and interact with virtual entities, and virtual entities, such as manned simulators and intelligent agents, to see, hear, and interact with live entities.

Figure 2 shows the general architecture for an AUGSIM system. Figure 3 shows a likely vehicle configuration for AUGSIM training. Because of the practical difficulties of allowing the user to run and crawl in a manned infantry simulator using VR alone, AUGSIM offers a reasonable and, in the long run, perhaps more effective alternative. Figure 4 shows a wearable AUGSIM configuration for infantry training.

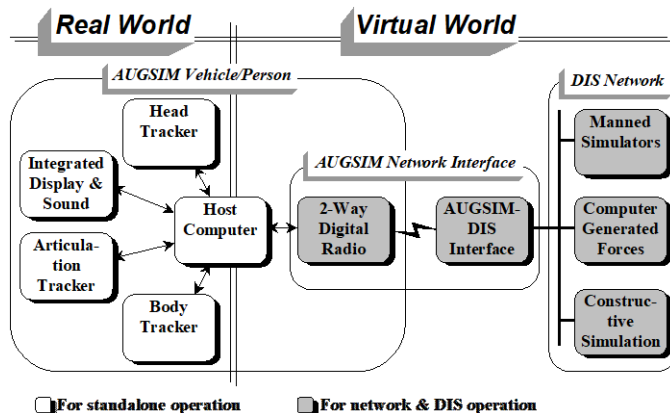


Figure 2. AUGSIM General Architecture. The AUGSIM architecture supports vehicle and wearable configurations, and allows standalone or networked operation.

Looking beyond military applications, commercial applications for AUGSIM might include:

- Training for civil services and disaster teams.
- Visualization for business, architecture and science.
- Maintenance and manufacturing tools for industry.
- Medium for education and collaboration.
- Games for entertainment and therapy.

Benefits

Computer simulation technology offers a high degree of flexibility and cost effectiveness for training. Live exercises provide unique conditions that can not be readily achieved in computer simulation. Two-way real-time interaction between live and virtual players can bring many of the advantages of one form of

training to the other. AUGSIM augments a live training exercise with the qualities of a simulator exercise, and it augments a simulator exercise with the qualities of a live one. Many of the operational concepts and benefits for AUGSIM were identified in collaboration with military trainers at Fort Knox, KY to solve real problems.

The ability of current field systems to reproduce the complexity and “clutter” of a real battlefield is rather limited. Popup targets are for the most part stationary; indirect fire simulation is very limited; and, the introduction of air support and air threats is expensive and logistically difficult. Safety, environmental impact, range size, operating costs, and logistical factors severely limit or make impossible the introduction of supply trains, non-combatants, many types of obstacles, large opposing forces (OPFOR), and flanking friendly forces (BLUFOR). AUGSIM offers the ability to introduce more realistic and more complex battlefield environments into live exercises, with much of the flexibility, safety and cost effectiveness of simulator exercises.

Because of simulator like flexibility, AUGSIM can be tailored to the specific needs of the exercise or range. For example, if the range is small or if environmental impact is a major concern, smaller units could be trained; and, AUGSIM would fill-out the OPFOR and provide flanking BLUFOR to reduce the number of live vehicles on the range. Another scenario would place the unit leaders in live vehicles on the range; and, AUGSIM would-fill-out the units with manned

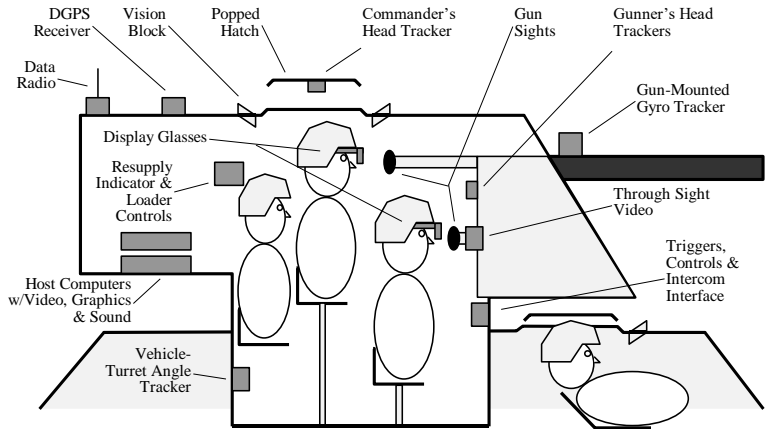


Figure 3. AUGSIM Vehicle Configuration. Advances in display and processing technologies since the original demonstrations make integrated displays feasible and the overall system more cost effective.

simulators or even intelligent agents. Estimates indicate that these approaches could reduce the number of vehicles on the range by almost a factor of four.

Because of limited fidelity, simulator exercises tend to occur more quickly and with fewer “complications” than a live exercise. Although this can sometimes result in negative training — learning to do the wrong thing — the advantages of simulator training are too great for it not to be used. AUGSIM has the potential for introducing some of the conditions, intangibles, and vagaries present in live training into simulator training. For example, with commanders in the field and the rest of their units in simulators, the pace of a predominantly simulator exercise would be controlled by the live action and field conditions instead of the artificiality of the simulator virtual environment.

Challenges Aboard

One of the biggest challenges of AR for both vehicular and wearable systems is the need for accurate spatial tracking. The position and orientation of all participants must be accurately tracked over large and often varied gaming area. In addition, the position and orientation of their articulations — head, limbs, weapon, turret — must also be tracked to varying degrees of accuracy, the head being the most critical. Other problems such as integrated displays, higher fidelity, smaller size, etc. almost pale in comparison.

The following observations and opinions were developed by the author in the course of SBIR project and proposal work between 1994 and 1996 for AUGSIM, and more recently for AR based situation awareness. This is by no means a formal or complete treatment of the subject, and is unfortunately based

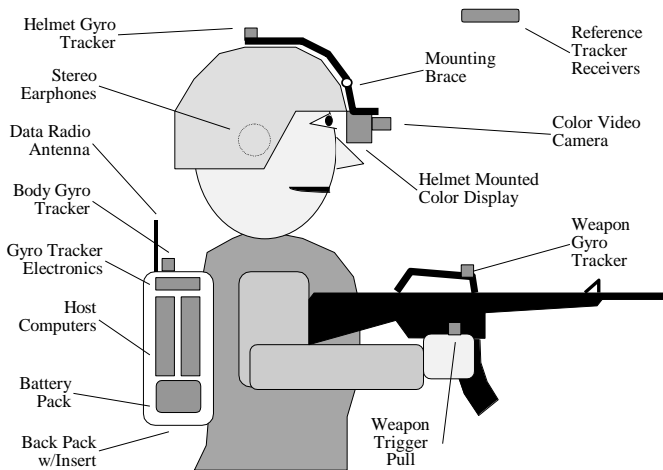


Figure 4. AUGSIM Wearable Configuration. Innovative use of available low-cost VR, video, processing, and wireless technologies put a wearable system within reach.

more on "proposed" work than it is on "actual" work. It should, however, offer a good starting point for someone beginning work in this field.

AR vs. VR Tracking

In general, commercial products developed for VR have good *resolution* but lack the absolute *accuracy* and wide area coverage necessary for AR, much less for their use in AUGSIM.

VR applications — where the user is immersed in a synthetic environment — are more concerned with relative tracking than in absolute accuracy. Since the user’s world is completely synthetic and self-consistent the fact that his/her head just turned 0.1 degrees is much more important than knowing within even 10 degrees that it is now pointing due North.

AR systems, such as AUGSIM, do not have this luxury. AR tracking must have good *resolution* so that virtual elements appear to move smoothly in the real world as the user's head turns or vehicle moves, and it must have good *accuracy* so that virtual elements correctly overlay and are obscured by objects in the real world.

Objective Accuracy

In AR the nature of positional accuracy is that its affect is the same regardless of the distance from the user to an object of interest, such as a virtual target hiding behind a live object. A 1 meter lateral error in the user’s position produces a 1 meter lateral error in that of the target. Angular error, however, is dependent on the viewing distance. A 1 degree lateral error produces a 17 meter lateral error for a target at 1000 meters, but only a 0.2 meter lateral error for a target at 10 meters. If the size of the target is on the order of 3 meters, such as a tank seen head-on, errors on the order of 1 meter (33%) may be acceptable. For a target that is only 0.5 meters across, such as a soldier, a 1 meter error would likely be unacceptable.

Making some crude but reasonable application specific assumptions, Table 1 compares the objective tracking accuracies needed for mounted (vehicle crew), dismounted outdoor, and dismounted indoor AUGSIM

AUGSIM Training Application	Typical Engagement Distance	Typical Target Size	Max Allowed Lateral Error (% of Target)	Required Positional Accuracy	Required Angular Accuracy
Mounted	1000 m	3 m	1 m (33%)	0.5 m	0.03 deg
Dismounted, Outdoor	50 m	0.3 m	0.1 m (33%)	0.05 m	0.05 deg
Dismounted, Indoor	5 m	0.03 m	0.01 m (33%)	0.005 m	0.05 deg

Table 1. Objective Tracking Accuracies. Objective accuracy depends solely on geometry and does not take into account factors of human perception. Accuracy values are halved to allow for position and angle error accumulation.

systems.

Subjective Accuracy

Looking at the overall problem of mixing live and virtual entities in the same exercise, objective accuracy is an easily understood and measured requirement, but perhaps an overly conservative one. Because of subjective and application specific factors, certain errors may not be significant or even noticeable to the user. As regards engagements — shooting at things — there are four basic situations:

- A *virtual* entity sees a *virtual* entity.
- A *virtual* entity sees a *live* entity.
- A *live* entity sees a *virtual* entity.
- A *live* entity sees a *live* entity.

Case 1, virtual-on-virtual, is simply pure VR with all objects and terrain existing in a self-consistent world. This is the situation that describes the manned simulator systems in use today. The question of live entity tracking obviously does not enter into the equation.

Case 4, live-on-live, can be easy or difficult depending on the approach dictated by the application. If a live weapon simulator, such as MILES — the military's "laser tag" system — is used in conjunction with AUGSIM, the problem of live-on-live engagement becomes self-consistent since everyone and everything, including the MILES laser beam, is operating in the real world. Performance is dictated by that of the MILES system and not AUGSIM. Without MILES the AUGSIM system is at the mercy of the combined absolute tracking accuracy of the two entities. Scoring a hit would be similar to that in Case 1. In manned simulators the size of the "hit box" around a target is generally much larger than the target to accommodate system errors in latency and tracking. This same approach could be used in AUGSIM, but for live-on-live engagements.

The remaining cases, 2 and 3, although not completely symmetrical, are similar enough for this discussion to be treated as one. The only thing that matters as regards a mixed live-virtual engagement is the mutual perception of the two entities in the virtual world — not the real world. This is because both entities are using the same 3D world model and each other’s reported position to view one another. Thus, if a virtual entity can see and shoot a live entity’s virtual representation, then the live entity can also see and shoot the virtual entity, regardless of live entity tracking error. Live entity tracking errors instead appear in other, more subtle ways.

From the perspective of the virtual entity, everything will be self-consistent. The virtual representation of the live entity will be perfectly aligned and occluded by a virtual object regardless of where the tracking says the live entity is relative to the corresponding live object. The reverse is not true. A virtual tank partially hidden behind a berm in the 3D virtual model will appear partially occluded to a live observer. Due to observer tracking error, however, the partial image of the virtual tank may appear to float above, below, or to the side of the corresponding live berm.

The impact of such a phenomenon is hard to judge. The target is only partially visible as it should be, but since it appears in the wrong place its detection may be artificially enhanced, by floating in the sky, or suppressed, by surrounding terrain clutter. Of course this effect is undesirable, but the degree to which it can be eliminated is directly and perhaps exponentially proportional to system cost. At long engagement ranges, such as for mounted warfare, the errors may not be very noticeable due to the perceived small size of the target, the clutter of its surroundings, and the blur and haze of intervening atmospherics. For closer engagement ranges, such as for dismounted urban warfare, these errors would be more predominant.

Active Tracking

A number of active tracking schemes have been developed using a variety of means, including ultrasonic, magnetic fields, scanning lasers, and encoded radio. They require a transmitter and a receiver unit, with one in an accurately known position and the other on the entity being tracked, and some means to get tracking data from one to the other. Some provide orientation tracking. Most provide position tracking. All have serious drawbacks in one or more areas as concerns AR — accuracy, resolution, responsiveness, interference, clear line-of-sight — and AUGSIM — scalability for multiple entities, wide coverage, and real gaming areas such as stair wells and nooks in buildings, streets and alleys around buildings, and forested hills and valleys.

GPS is often proposed as a cheap means for accurate tracking. Differential GPS (DGPS) is getting close to the necessary accuracy, at least for body tracking, but it generally lacks the necessary responsiveness and orientation tracking needed for head tracking. Also, GPS only works reliably outdoors, in the open, away from building and not under tree cover, which does not bode well for the training warfighter.

Inertial Tracking

Inertial systems offer the greatest promise for achieving the necessary responsiveness, resolution, accuracy in position and especially orientation

tracking. The use of a self contained vehicle-based Inertial Measurement Unit (IMU) was proven in the live demonstrations using an off the shelf Ring Laser Gyro (RLG) system. Although costly and heavy, it provided the necessary performance for vehicle body and weapon tracking over very large gaming areas. More recent systems combine GPS with an IMU to lower cost and to simplify operation.

For wearable configurations, such as for infantry and vehicle crew head tracking, a head mounted IMU with comparable performance is a ways off. In the interim, wearable inertial systems are being supplemented with active “reference” trackers to provide accurate position data, with all their attendant problems. A brute force solution to ultimately solving the tracking problem is to wait for inertial systems to evolve to the necessary accuracy, size, and power. There is nothing in physics to prevent it, but it might take a while. An alternative is to use passive vision tracking as the reference tracker in a hybrid inertial system.

Vision Tracking

Vision tracking is a passive technique that relies on computer vision techniques to determine the position and orientation of an entity. One approach uses area video cameras to track the players in the gaming area. Players must remain relatively close to and in the line of sight of a camera. Tracking data is transmitted to each player in real-time.

Another approach is more self-contained and potentially more accurate. Given a 3D model of the gaming area, or at least a 3D model of key landmarks, the system computes absolute orientation and position from video cameras mounted on the entity being tracked. The system avoids having to transmit high-bandwidth tracking data to each player and offers benefits in scalability, accuracy, and robustness. It is also a lot less mature.

The need for a priori 3D models is a concern, but nothing that money and lots of surveyors and photogrammetrists can't solve. Besides, even though most of an AR presentation is real, to provide convincing occlusion of virtual objects by those in the real world an accurate 3D model of objects and the terrain is needed anyway. Also of note are programs underway in the military to allow remote capture and rapid 3D modeling of urban and non-urban areas for VR and someday, presumably, AR based training.

Integrated Displays

To be an effective form of training most if not all visual operating modes on a vehicle need to be supported, such as looking out of an open hatch, looking through vision blocks (small periscopes), and of course looking through a gun sight. One obvious

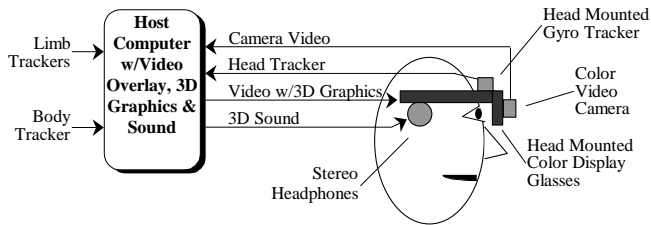


Figure 5. AR Video Overlay. Video overlay is inexpensive and avoids the many problems of see-through overlay. A head mounted gyro provides accurate head tracking.

approach is to mount displays in front of the vehicle's vision ports. Upon closer study, however, the number and types of ports that must be instrumented, and the infrastructure needed to support this array on a moving vehicle is sobering. For example, on an M1A2 tank the number and types are roughly: (2) popped hatches with 360 deg. visibility, (>16) vision blocks, (2) gun sights, (1) commander's thermal viewer, and (1) situation/map display. There is also the question of how to capture and integrate into the displays the corresponding live view from each unique vantage point.

The best overall approach is head mounted displays (HMDs), which offers a single unified solution to all of these display needs. The key to effective HMDs, unfortunately, is accurate head tracking under some of the most challenging conditions — inside of a cramped, cluttered, and metallic environment such as a tank turret, or on open remote terrain.

Short of a direct tap into the optic nerve, the ultimate HMD technology appears to be Virtual Retinal Display (VRD), where the image is directly scanned onto the retina. Although VRD is in development at this time, it may be a while before it becomes commercially available. More traditional technologies using lightweight low profile Liquid Crystal Display (LCD) technology have been commercially available for several years, and their performance is continuing to improve.

Direct View Display

A direct view AR display system allows the user to see the real world directly, such as with see-through optics. A common approach is to use a partially silvered mirror to combine the real view with a graphics display of the virtual entities. Another approach is for the user to look at the world directly through the graphics display device, such as a transparent LCD. A third approach is to directly scan the virtual world graphics onto the user's retina using VRD. The big advantage of direct view is that it allows reasonably normal vision; but, the

disadvantages are significant:

- Very precise alignment is required between the display and the eye.
- The focal distance of the graphics must match that for objects in the live world.
- The position of the graphics must allow for eye convergence.
- The graphics must not wash out when the real world is bright or dominate when it is dim.
- The graphics must effectively block out the portion of the real world it is occluding.

Note that outdoor scenarios where objects of interest are at or near "infinite" eye focus mitigate problems regarding eye accommodation and convergence. They also tend to aggravate lighting problems.

Indirect View Display

An indirect view AR system typically uses video to capture the real world view, and video processing to overlay the virtual world graphics onto of the live video. Figure 5 shows such a configuration. To the user it is similar to watching a television screen. The advantages are significant, and it may be the best choice while waiting for VRD to mature:

- The position of the eye relative to the device can be non-critical.
- The focal distance is the same for graphics and live video — the display screen.
- There are no eye convergence problems since live and virtual are on the same display screen.
- Graphics and video use the same display so graphics will not wash out.
- The live video signal can be monitored to automatically adjust the graphics intensity.

The biggest drawback, at least for the real part of the presentation, is low resolution in comparison to that of a direct view system — the real world. As with VR, the potential for "simulator sickness" in both display approaches and AR in general is a concern.

Conclusions

Significant first steps have been made regarding the demonstration of AR in live training. Operational concepts combining the best of live and virtual simulation and training, in the form of AUGSIM, have been developed and their potential benefits identified. The number one challenge for AR is accurate and robust tracking. As the technology for VR improves AR will no doubt reap the benefits; however, there are fundamental differences in the needs of AR versus VR. Work in AR is ongoing in both the military and academic sectors. One effort of note is the DARPA sponsored Warfighting Visualization program, which

involves leading academic and institution research groups. Although the emphasis of this work is on battlefield information visualization, it will likely provide answers and technology applicable to more general AR applications, such as AUGSIM. If anything is certain it is that many interesting questions lie ahead that will need to be answered before significant headway can be realized in effectively applying AR to live training.