Using LiDAR Technology Combined with Position-Tracking Software and the Medical Training Evaluation System (MTES) in an Integrated After Action Review System for Mobile Medical Lane Training

Erin Honold, Catherine Strayhorn, Nadine Baez
Information Visualization and Innovative Research, Inc.
Sarasota, FL
ehonold@ivirinc.com, cstrayhorn@ivirinc.com,
nbaez@ivirinc.com

Brian VanVoorst
Raytheon BBN Technologies
St. Louis Park, MN
brian.vanvoorst@bbn.com

Matthew Hackett
Army Research Laboratory (ARL), Human Research and Engineering Directorate (HRED)
Orlando, FL
matthew.g.hackett.civ@mail.mil

ABSTRACT

The ability to produce a rapid After Action Review (AAR) capability provides objective data for evaluation and expands learner experience. Current video feedback does not provide a complete picture of the tactical situation during training, nor is it immediately available for evaluation. Cameras may not be positioned properly or a critical event may be blocked from view. Individual positions relative to potential enemy fire are not recorded. Light Detection and Range (LiDAR) technology combined with position-tracking software provides an effective method to produce immediate AARs. Position-tracking software and LiDAR technology when fused with digital imaging technology can capture both tactical and clinical skills to provide a more comprehensive audio/visual AAR capability with automatic scoring of performance. This capability can be used in conjunction with the Medical Training Evaluation System's (MTES) assessment checklists currently fielded at the U.S. Army Medical Simulation Training Centers (MSTC). The result is an immediate playback of performance correlated to assessment scores for training and evaluation.

Preliminary field-testing showed that LiDAR is capable of directing multiple cameras to track important actions of learners during a field training exercise. The position-tracking software accurately detected the persons of interest during a training event and cameras tracked their actions for the duration of the exercise. The AAR system successfully integrated the LiDAR data, video feeds, and skill assessment checklist to provide a fully integrated AAR system that provided all collected data in a simple user interface. Students compared the system to watching post-game footage of sports events to improve their performance. Instructors were able to pinpoint correct and incorrect actions to further educate the class and found the system to be valuable addition to training.

ABOUT THE AUTHORS

Erin Honold is the biomedical and systems engineer for IVIR Inc. Ms. Honold has a Bachelors degree in Biomedical Engineering from the University of Miami. She was the engineer for the ARL STTC Hologram Imagining Technology (HIT) Applied to Medical Modeling and Simulation program that investigated the first use of LiDAR technology for medical lane training AAR. She was also the engineer for the Mobile Medical Lane Training (MMLT) program, the follow-on to the HIT program that incorporated cameras and a MeTER AAR system with LiDAR technology, and co-authored the paper "Fusion of LiDAR and Video Cameras to Augment Medical Training and Assessment" for IEEE MFI 2015. Ms. Honold is the design manager for the TrACER software and is responsible for designing future advancements to the software. She also participated in providing engineering support for the Medical Training and Evaluation System (MTES) program, which deployed the MeTER software to MSTC sites, and the Combat Casualty Training Consortium (CCTC), which used a customized implementation of the MeTER system to collect data for the research effort.

Catherine Strayhorn is a cofounder of Information Visualization and Innovative Research (IVIR) Inc. and is company CEO and President. Ms. Strayhorn has a Bachelors degree in Aerospace Engineering Technology from Kent State University. Ms. Strayhorn is certified in ISO 9000 internal auditing, and Program Management. Besides her current duties at IVIR, Catherine has managed research programs for the Army Research Lab, Holographic and LiDAR Imaging Technology (HIT) Applied to Medical Modeling and Simulation Training that investigated the first use of LiDAR technology for medical lane training AAR, and dynamic and static holograms for advancing medical education. She also served as the Program Manager (PM) for the Mobile Medical Lane Training (MMLT) program, the follow-on to the HIT effort that incorporated cameras and AAR system with LiDAR technology, and co-authored the paper "Fusion of LiDAR and Video Cameras to Augment Medical Training and Assessment" for IEEE MFI 2015. She has over sixteen years of comprehensive experience in the medical simulation business.

Nadine Baez serves as the VP of Programs for IVIR Inc. Ms. Baez holds dual BS degrees in Micro/Molecular Biology and Forensic Biochemistry from the University of Central Florida. Ms. Baez was previously an experienced Forensic Investigator, Registry Certified by the American Board of Medicolegal Death Investigation and was also a National Registry Certified Emergency Medical Technician (NREMT-B). Ms. Baez has translated her past clinical and professional experience into simulation technology design concepts, educational and engineering requirements, and prototypes for medical training. Ms. Baez served as the PM for the U.S. Army RDECOM Medical Simulation Training Center (MSTC) Medical Training Evaluation and Review System (MeTER) Medical Pre/Post Test programs, which developed into the assessment and data collection system integrated in the MMLT development effort.

Brian VanVoorst is a Technical Director at BBN Technologies with over 20 years of experience working on and leading research and development programs. Most recently his work has been in the area of the automated understanding of LiDAR point clouds and sensor fusion. In addition to the MMLT work discussed here he has recently designed a LiDAR based detection and tracking system for rogue class 1 Unmanned Aerial Vehicles (UAVs) called LOTUS (LiDAR Observation and Tracking of Unmanned Systems). Once LOTUS has detected a rogue UAV it can deploy an automated intercept vehicle to chase it out of the sky. Additional work using LiDAR technology Mr. Van Voorst has participated in includes the development of two commercial automated railroad maintenance systems. Other areas of interest include computer vision, camera analytics, robotics and parallel processing. Before working at BBN, Mr. VanVoorst was a Principal Scientist at Honeywell Labs, and a Systems Analyst at NASA Ames Research Center. Mr. VanVoorst has earned bachelors and masters degrees in Computer Science from Michigan Technological University, has 5 patent awards, and has published over 20 papers in conference and journals.

Matthew Hackett is a science and technology manager for the Medical Simulation and Performance Branch of the Army Research Laboratory. He manages a variety of research projects including medical holograms, novel user interfaces, and serious games for medical training. Mr. Hackett received his Bachelor of Science in Computer Engineering from the University of Central Florida, his Masters of Science in Biomedical Engineering from the University of Florida, and is currently pursuing his Ph.D. in Modeling and Simulation at the University of Central Florida.

Using LiDAR Technology Combined with Position-Tracking Software and the Medical Training Evaluation System (MTES) in an Integrated After Action Review System for Mobile Medical Lane Training

Erin Honold, Catherine Strayhorn, Nadine Baez
Information Visualization and Innovative Research, Inc.
Sarasota, FL
ehonold@ivirinc.com, cstrayhorn@ivirinc.com,
nbaez@ivirinc.com

Brian VanVoorst
Raytheon BBN Technologies
St. Louis Park, MN
brian.vanvoorst@bbn.com

Matthew Hackett
Army Research Laboratory (ARL), Human Research and Engineering Directorate (HRED)
Orlando, FL
matthew.g.hackett.civ@mail.mil

INTRODUCTION

Military medical training is a fast-paced endeavor that includes rigorous didactic content, individual skills training, and outdoor lane training and evaluation exercises. A great deal of research has been conducted to improve military medical training, including technological advancements in medical manikins and part task trainers and new training modalities such as serious games that can utilize virtual and augmented reality technologies. These research areas have resulted in significant improvements in medical proficiency and ultimately an improvement in patient outcomes (Eastridge et al., 2012). However, there is still a need for improvement regarding the ability to assess and review these training exercises.

The culminating event in a tactical medical training course requires an individual or squad to perform actions akin to those on the battlefield. These exercises are conducted on a "training lane," an outdoor course in which Soldiers must move tactically while encountering obstacles and simulated opposing forces along the way. Additionally, simulated casualties are placed within the path of the Soldiers, requiring appropriate treatment using Tactical Combat Casualty Care (TC3) procedures. This includes movement of the patient to a safe location, performing medical treatment, and patient evacuation. During the exercise, a group of support staff may operate a variety of simulation assets to increase exercise realism, such as sound effects generators, smoke machines, and medical manikins. Instructors move throughout the lane with the group of Soldiers, assessing their performance and providing verbal instruction. At the culmination of the exercise, the instructor(s) provide an exercise debrief, or After Action Review (AAR). The AAR reviews the performance of the individuals and squad members and is one of the most important facets of training. Without understanding weaknesses, a student will be unaware of areas needing improvement. Unfortunately, the current method of conducting AARs is lacking, as instructors cannot see everything that is occurring at the same time and consequently are subjective in their evaluations without the capability to review video or other data sources to provide a thoroughly objective assessment.

During a recent effort, researchers sought to implement technology within medical field training to: 1) improve AAR capabilities by applying a system of reconfigurable sensors within a lane training environment and 2) reduce instructor workload by intelligently automating recording devices and simulation assets. The result of this research was the creation of the Mobile Medical Lane Training (MMLT) system. This system uses a Light Detection and Ranging (LiDAR) sensor as a simulation director, which tracks Soldiers' locations within the training environment. Based on their locations, the system intelligently tasks cameras to record areas of interest or send control signals to nearby simulation assets. Within minutes of exercise completion, the system processes the data from the LiDAR, video feeds, and the Medical Training Evaluation System (MTES), and creates a seamless AAR, allowing instructors to debrief students using video, an assessment checklist, and a 3D point-cloud representation of the field exercise and applied medical skills.

SYSTEM ARCHITECTURE

The Mobile Medical Lane Training (MMLT) system is a multi-sensor, rapidly deployable After Action Review (AAR) system that uses LiDAR technology combined with the Medical Training Evaluation System (MTES). The MMLT AAR system uses 3D LiDAR data, a camera array, People Tracker software, and the Medical Training Evaluation and Review (MeTER) software, which is the primary component of MTES, to produce immediately reviewable AAR data. The full MMLT AAR system consists of two subsystems: the data capture system and the AAR station (see Figures 1 and 2).
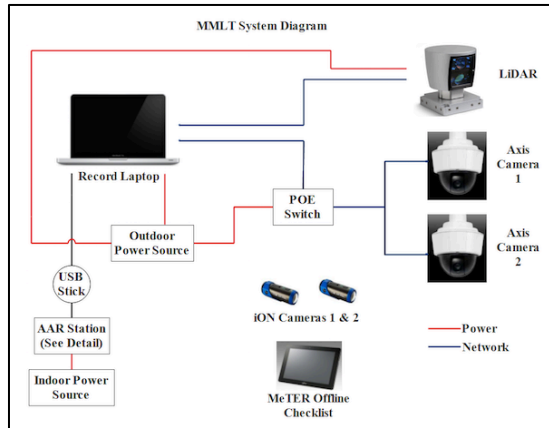

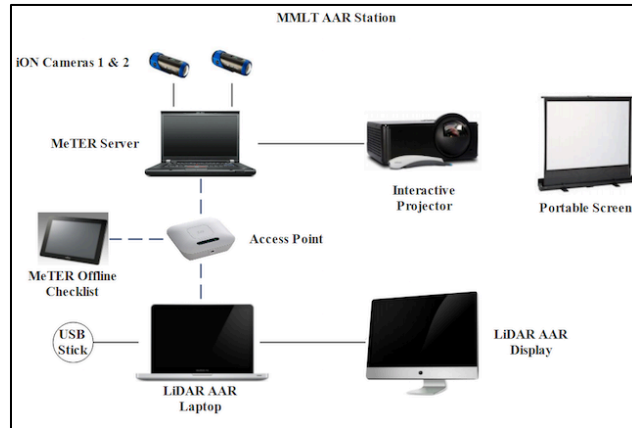Figure 1. MMLT System Diagram   Figure 2. MMLT AAR Station

The data capture system consists of one Velodyne HDL-64 LiDAR unit mounted on a tripod, two Axis Pan-Tilt-Zoom cameras on tripods, two helmet-mounted cameras (iON cameras), a data collection computer for the LiDAR People Tracker data and video data, a tablet for the MeTER offline checklist, and networking equipment. The system is expandable to allow for multiple LiDAR units to cover a wider area, up to 8 Axis cameras, and multiple helmet-mounted cameras. The AAR station consists of the MeTER server, an AAR computer for LiDAR and video data, a router, and two displays for the AAR data. One of the displays investigated was an interactive projector display that allows the instructors to use a small pen-like device to control the system from the projected visual (on a wall or a screen) rather than having to operate a computer.

MTES Assessment System

The system uses MeTER software, which was fielded as part of the MTES. MTES includes all of the hardware and networking equipment required for a site to use MeTER, including a server and student and instructor workstations, while MeTER is the software installed on the server. MeTER was developed by IVIR Inc. with funding through the Army Research Lab (ARL) Human Research and Engineering Directorate (HRED) and the Program Executive Office for Simulation, Training and Instrumentation (PEO STRI). MeTER is a computer-based educational assessment tool kit for medical and tactical knowledge and skills designed for the United States Army and currently installed at 23 Medical Simulation Training Centers (MSTC). The software runs on a server and is accessed through a browser. MeTER software allows for customized and editable content, and provides automatic cognitive tests and skill assessment checklists.

The MeTER checklist provides a list of tasks and steps decomposed to the terminal level, with GO and NO GO scores for each step. Each score and step on the checklist are timestamped when completed, and instructors can take notes for each step if necessary. The checklist can be used both online and offline, allowing the instructors to conduct training away from the server. The offline checklist is a separate application that is installed on a Windows tablet and can be synchronized with the server to download and upload the student data. See Figure 3 for an example of the checklist.
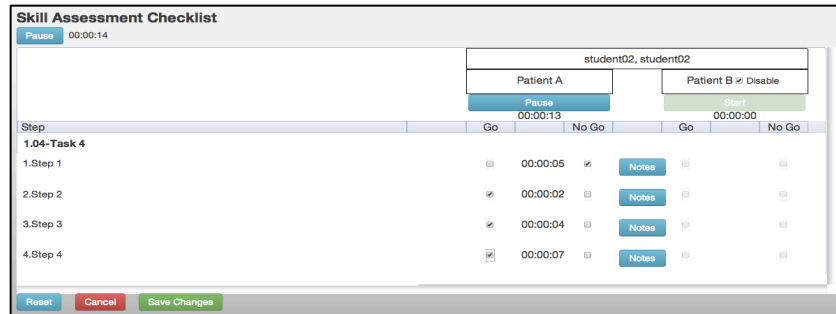
Figure 3. MeTER Skill Assessment Checklist

In the MMLT system, the MeTER system captures the student/team performance and allows for the data to be directly correlated to trainee actions being recorded by the LiDAR and cameras. When an instructor scores the completion of a step, the timestamp can be used later to index into the video and LiDAR data, which have been time synchronized prior to training. The MeTER system creates an interactive display that allows the instructor to choose specific events from the checklist to highlight during the AAR. Patient simulation sensor data may also be captured and transmitted into the MeTER system providing objective measurements of student and team performance, fulfilling a requirement of Army Learning Concept (ALC) 2015 (TRADOC, 2011).

LiDAR People Tracker

The LiDAR People Tracker is composed of two parts: a *data capture system* (LiDAR, set of cameras, laptop, and networking equipment) that is rapidly deployed in the outdoor lane, and a *review station* (laptop) that is typically positioned at the end of the lane or back in a classroom. Data is passed from the capture system to the review station by use of a USB flash drive (using a DVD is optional).

The capture system used in this project is comprised of a single Velodyne HDL-64 LiDAR, two Axis 6034-E cameras, a MacBook Pro laptop running the People Tracker software, a power-over-Ethernet switch, and tripods on which to mount each sensor. The system is capable of working with more cameras if needed, and functions with any model of Velodyne LiDAR. The LiDAR unit uses Class 1 (eye safe) lasers to detect objects in the environment by emitting beams in a 360° area and capturing the reflections of the beams off objects (see Figure 4).
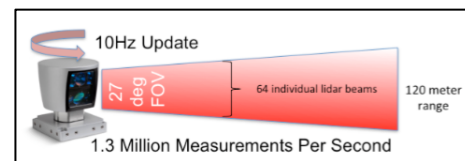

Figure 4. Velodyne HDL-64 LiDAR

The People Tracker software takes the raw data stream from the LiDAR, applies calibration corrections, segments the point cloud into objects, classifies the objects to discern people and tracks the people as they move; all in real time. Cameras are assigned to the tracks based on algorithms for priority and optimization discussed in the next section. Based on these assignments, the cameras are commanded moment to moment to focus on the tracks of interest.

LiDAR AND CAMERA MULTISENSOR FUSION APPROACH

To better understand the multisensor fusion approach, some background in the tracking software is needed. At the highest level, the LiDAR People Tracker takes the LiDAR data as input, segments the LiDAR pointcloud data stream into objects, classifies objects that are likely to be people, tracks the people with a Kalman filter (Kalman, 1960), assigns cameras to tracks or groups of tracks based on a priority scheme, and maintains a real-time data display.

A few key data structures are kept for the LiDAR data. A *background mask* of voxelized cells is populated by the first eight seconds of data as an initial approximation of the static objects present in the scene (e.g. ground and trees). This background mask is aged out over a period by one thread, and re-enforced by incoming evidence of new

data. In this way, if an object that was initially believed to be static leaves the scene, the notion that is was background is discarded. If however, an object persists, it remains part of the background.

A second key data structure is a *vacant voxel map*. This data structure consists of occupied voxels. Periodically, the system traces rays between the sensor origin and each sensed return from the LiDAR. Any unoccupied points along that ray then become observed vacant voxels. Due to the beam spacing of the LiDAR there are large regions of space that are never observed to be vacant. Since the background mask ages out, it is possible for a person to cast a shadow on a surface that no longer belongs to the background. Therefore, the points used to spawn a new track cannot come from within a background mask voxel, and must originate from a vacant voxel.

Calibration of LiDAR and Cameras

The goal of calibration is to be able to translate LiDAR coordinates to camera pan, tilt and zoom (PTZ) commands to follow tracks of people (see Figure 5). To begin, each tripod mounted sensor is manually leveled using a tribrach with a circular bubble level installed to
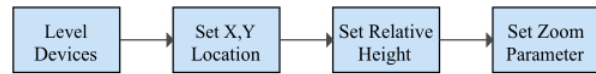
Figure 5. Calibration Steps

each plate. The PTZ cameras are manually located by the user in a display showing a live view of the LiDAR's pointcloud, defining cameras' X,Y location. Next, the height of the camera relative to the LiDAR is determined by setting the camera's tilt to 0 and focusing the camera on a graduated pole placed in the scene. Crosshairs in the scene mark image center. The height of the camera (crosshairs on pole) is marked for the LiDAR by placing an IR retro-reflective dot on the pole. Using the calibration software, the user can find the IR dot in the coordinate frame of the LiDAR. Once complete, the height of the camera relative to the LiDAR is known and recorded. Once each camera has been localized in the LiDAR frame, that localization can be used to convert a point in the LiDAR's frame to the corresponding pan and tilt values for the PTZ camera. The next step is to determine the zoom parameter. From a theoretical standpoint, if the target's size and distance are known, the required field of view angle can be computed as shown in equation 1:

$$\frac{fov}{2} = atan(\frac{width_t}{2 \times dist_t}) \qquad (1)$$

where $width_t$ is the width of the target and $dist_t$ is the distance of the target to the camera. That field of view can then be used to determine the desired focal length as shown in equation 2:

$$\frac{fov}{2} = atan(\frac{width_i}{2} \times fl) \qquad (2)$$

$$fl = \frac{(width_i \times dist)}{width_t}$$

where $width_i$ is the width of the image in pixels, $width_t$ is the width of the calibration target in meters, and fl is the focal length in pixels. These computations hold similarly for height. These equations calibrate the zoom parameters of the cameras, ensuring that the resulting field of view is adequate to capture the training event. However, the Axis PTZ cameras that were used did not allow the focal length to be set directly, but instead they use an abstracted zoom level. In order to map focal length to zoom level, the camera matrix was computed at several zoom levels (Hartley and Zisserman, 2003) and fit to a linear function that maps focal length to zoom level.

Prioritization for Aiming Cameras

Determining which objects to target falls into three hierarchical categories: tracks, groups, and zones. A track is an individual person whom the People Tracker is following. Tracks have a centroid, an extent, a heading, behavior properties such as crouching or prone, and a priority. Priority is an integer value used to give relative weighting of importance between tracks. A group is a set of tracks that have a similar velocity and heading, and sufficient proximity to each other. Zones are polygon areas on the ground plane of the pointcloud that are defined *a priori* by the user through a point and click interface. Based on necessary parameters, a user creating a zone will select from the following types: *ignore zones, observer zones, action zones, priority zones and priority pool zones*. Ignore zones are areas in which the LiDAR does not process the data (used for regions in the range of the LiDAR not to be

considered) such as a nearby street. Tracks that originate in *observer zones* will never have a camera resource assigned to them. Ideally, instructors will stand in an observer zone at the beginning of a data collection cycle.

Action zones are regions wherein if tracks enter, all cameras should be assigned to the tracks in that zone. *Priority zones* transfer a pre-determined priority onto tracks when entering. Lastly *priority pool zones*, contain a quantity (pool) of priority to transfer onto the tracks that are in the zone. Priority drains from the pool to the individual tracks at a set rate until it is exhausted. In

Figure 6. Aiming Prioritization

this way, the first tracks to reach a priority zone may accumulate priority that later tracks will not receive.
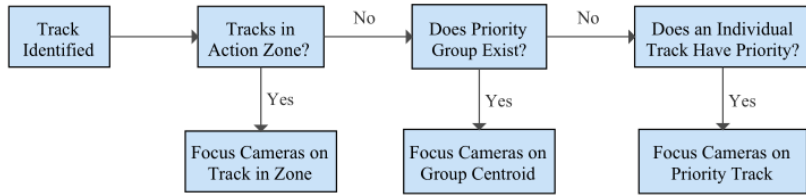
Camera assignment is prioritized first to tracks in action zones, then to groups exhibiting crouching behavior, and then to individual tracks based on priority (see Figure 6). Groups are defined by creating a dissimilarity matrix based on the following function in equation 3:

$$d_g(t_1,t_2) = \alpha d_s(t_1,t_2) + \beta d_h(t_1,t_2) + \gamma d_v(t_1,t_2) \qquad (3)$$

$$d_s(t_1,t_2) = \sqrt{(X_{t_1} - X_{t_2})^2 + (Y_{t_1} - Y_{t_2})^2 + (Z_{t_1} - Z_{t_2})^2}$$

$$d_h(t_1,t_2) = \cos^{-1}\left(\frac{(V_{t_1} \cdot V_{t_2})}{(|V_{t_1}| \cdot |V_{t_2}|)}\right)$$

$$d_v(t_1,t_2) = \text{abs}(|V_{t_1}| - |V_{t_2}|)$$

where t_1 and t_2 are two tracks, α, β, and γ are weighting factors that determine the importance of each factor, and $d_s(t_1,t_2)$ is the spatial distance component, $d_h(t_1,t_2)$ is the heading component, and $d_v(t_1,t_2)$ is the velocity component. The density-based clustering algorithm (Ester, Kriegel, Sander, & Xu, 1996) is then used to cluster the tracks for a particular frame. This equation is used by the software to identify when multiple individual tracks are near each other and when they are moving in a similar direction at a similar speed, signaling a cohesive group of tracks rather than a random cluster of people.

Once groups of tracks have been defined, priority groups are created. A priority group is defined to be one where some member of the group has a non-zero priority, and the group is determined to be crouching and stationary. These conditions correspond to expected behaviors of medical personnel during casualty treatment. Any such priority group gets the attention of all available cameras. The cameras zoom out to encompass the entire group and aim at the group centroid.

In the case of targeting zones and groups, it is undesirable to have the camera constantly making small adjustments due to the shifting centroid of the group. To mitigate this, a set of "lock-on" behaviors has been defined. If the average velocity magnitude for a targeted group of tracks is observed to be below a threshold, the camera target position is only updated every 5 seconds, and similar behavior is used for zones after 3 seconds.

The final category, tracks, is reached only when, after having considered the previous categories, there are untasked cameras left. Here an affinity matrix is computed between each untasked camera and unmatched track, where the affinity is defined as shown in equation 4:

$$d(t_j,c_j) = \begin{cases} 0 \text{ if } t_i \text{ is an observer,} & (4) \\ 0 \text{ if } t_i \text{ is occluded} \\ w_p P(t_i) + w_f F(t_i,c_j) - w_o O(t_i,c_j) + C \end{cases}$$

where t_i is the ith track and c_j is the jth. The w_p, w_f and w_o parameters are all weights that define the importance of the corresponding functions. The P(t_i) is the priority for track t_i, $F(t_i,c_j)$ is a function that returns 1 if camera c_j is already following track t_i, and $O(t_i,c_j)$ is a function that returns a value between 0 and 1 that indicates the level of occlusion between camera c_j and track t_i. This equation allows the software to assign the correct camera to the priority track, ensuring that a camera with a blocked view of the track is not the camera assigned to the track. This status is continuously updated while following the track to allow for seamless handoffs between multiple cameras if the original camera becomes occluded due to track movement. To determine whether a camera/track pair is occluded by a static background object (e.g. bush), nine rays are traced through the vacant voxel map between the camera center to various target points on the track. Multiple rays are used to account for sparse occlusions like vegetation. A majority vote is made across the nine rays to decide whether the track is occluded from the camera. These scores are then accumulated across several past frames and the occlusion level is determined by the percentage of occluded frames. Once the camera/track affinity matrix has been created, cameras are associated to tracks using the Munkres algorithm (Munkres, 1957), which creates an optimal matching. By incorporating the 3D modeling of occlusions (e.g. walls, trees), the result is a seamless handoff of a track from one camera to another for tracks of high priority.

AFTER ACTION REVIEW

While the video and LiDAR data is being collected, an instructor uses a tablet with the MeTER checklist to score trainee performance during the scenario. The checklist is a list of steps, each with a GO or NO GO score for objective scoring. Each score is timestamped, so when the checklist scores are loaded into the MeTER server, the timestamps can be used to index the video and LiDAR data during the AAR.

When the training scenario is complete, all the data is transferred to the AAR station via USB drive. The AAR station provides instructors with a comprehensive set of AAR information, including checklist scores, all video and LiDAR data, and a calculated shooter score. The shooter score calculation is based on the LiDAR data. Using the LiDAR data, the system can determine when any given trainee track is exposed to an enemy position (virtual or actor), which can be set by the instructor prior to training. The MeTER system calculates that every 3 continuous seconds of exposure equates to one shot on the exposed track. This information is displayed on the AAR page in MeTER as a graph, with one block for each "shot" on each track, each with a timestamp.

The system provides two AAR displays, one containing the LiDAR and Axis camera data (see Figure 7) and one containing the MeTER data, including checklist scores, helmet-mounted camera videos, and shooter score graph (see Figure 8).

The instructor can use the MeTER AAR display to select timestamps from either the checklist scores or shooter score graph to automatically advance all video and LiDAR data feeds to the same point in time, allowing the instructor and

Figure 7. LiDAR and Axis Camera AAR Display

trainees to review all actions at that point in time to determine what was done correctly, or what could be improved during the next scenario. In addition to the AAR page, the MeTER system provides individual and summary reports to indicate strengths and weaknesses of a trainee or a team based on both cognitive tests and the skill assessment checklist.

The resulting AAR system provides multiple streams of data, automatically correlated, that can be used by the instructors immediately after the training exercise. The data provided in the AAR stresses both medical training performance and tactical training performance through the shooter score and LiDAR data.
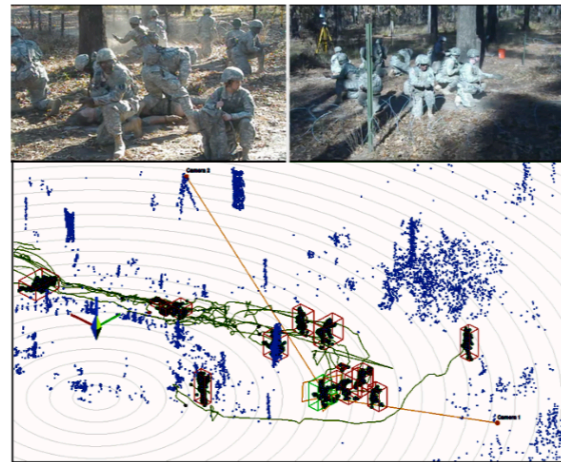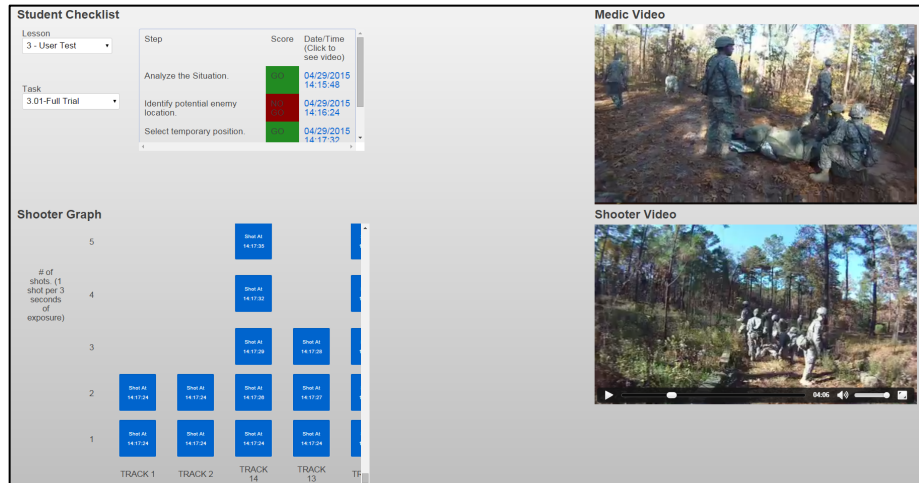
Figure 8. MeTER AAR Display

SYSTEM PERFORMANCE DURING USER TESTS

The MMLT system was used for field evaluation of Combat Medics (68W) and Combat Lifesavers (CLS) at the Ft. Bragg MSTC. The 68W class consisted of twenty-four students, divided into eight groups of three. The CLS class had forty students, split into four teams of ten. The first day consisted of a set of collections during which pairs of combat medics performed simulated medical procedures, including patient relocation, on a third soldier who played the role of the wounded. The second day was spent with squads of combat lifesavers performing lane training in a wooded environment.

First Day Results

In the first day's exercise, two combat medics would start from a location of cover and proceed to recover a casualty hidden behind a tree. They would drag the casualty to safety, examine and/or treat them, place them on a stretcher, and prepare them for evacuation. This activity took place behind the training center in a pine forest, with discarded vehicles and other cover (natural and placed objects) available. One or more instructors walked with the students, offering real-time instruction and feedback during the drill.

Prior to starting the drills, the system was deployed and calibrated in less than an hour. During the drills, data was recorded from seventeen exercises. Table 1 shows the performance of the system on these exercises. Performance is computed by sampling every ten seconds to determine whether the camera is aimed at an appropriate subject. The overall camera score is the percentage of these samples that target an appropriate subject. The result yielded ample video footage with which to evaluate student performance.

Table 1. First Day Experiment Camera Pair Utilization

Group	Group Statistics		
	Total Time	*Camera 1*	*Camera 2*
1	9 min	48.1%	53.7%
2	6 min, 30 sec.	71.8%	53.8%
3	3 min	77.8%	50%
4	6 min, 50 sec.	100%	95.1%
5	8 min	95.6%	100%
6	7 min	87.6%	100%
7	7 min, 40 sec.	91.3%	100%
8	9 min, 20 sec.	100%	100%
9	5 min, 20 sec.	87.5%	96.9%
10	6 min, 20 sec.	100%	100%
11	9 min, 50 sec.	100%	100%
12	9 min, 50 sec.	100%	100%
13	9 min, 40 sec.	86.2%	94.8%
14	9 min, 50 sec.	96.6%	98.3%
15	9 min, 30 sec.	82.5%	98.2%
16	7 min, 40 sec.	80.4%	100%
17	5 min, 20 sec.	90.6%	100%

The first day's arrangement stressed the system in four distinct ways. First, due to the location where the drills were being practiced, the cameras were spread far apart with a shipping container between them. This resulted in considerable time in which one of the two cameras could always see the action, but the other camera was obscured.

Second, the Axis camera mpeg codec seemed particularly challenged by the "digital camo" pattern of the soldiers' uniforms in certain lighting conditions during moments of high motion. This resulted in video artifacts that were unpleasant to watch. By reducing the video resolution from 1080p to 720p these video effects were eliminated. This was not reflected in the results shown in Table 1.

Third, the training area was covered in a deep carpet of pine needles. The training exercise involved dragging wounded soldiers across the ground, which in some cases accumulated large piles of pine needles that in turn caused enough change in the environment that the tracker considered them new objects in the scene. If a soldier's track was later lost, and a camera was assigned to these "disturbed earth" tracks, the camera could persist in watching the disturbed earth track for the duration of the exercise. The first collection exhibited this problem more so than others.

Fourth, in nearly all cases the instructors did not start the exercise from the agreed *observer zone* location. In these cases the system could not differentiate the instructor from the soldiers, introducing the possibility that a camera resource would track the instructor rather than the soldiers.

Second Day Results

On the second day, the MMLT system was used to collect video for four groups of combat lifesavers performing lane training. The location selected was an area in which the soldiers would encounter a simulated casualty underneath barbed wire. The first two soldiers through the course retrieve the wounded, pull them to a safe location, and begin treatment. Additional soldiers arrive and assume tactical positions to protect the medical providers. This portion of the exercise could take five to fifteen minutes, during which one or both cameras should be aimed at the soldiers performing the medical intervention.

The soldiers would approach from a distance and be picked up by the sensor at a range of approximately 100 meters. For this collection, a *priority pool* zone was placed around the casualty simulator, so that the soldiers who first reached the simulator would become the highest priority tracks. Table 2 reports the performance for the second day using the same methodology as in Table 1.

Table 2. Second Day Experiment Camera Pair Utilization

Group	Group Statistics		
	Total Time	*Camera 1*	*Camera 2*
1	1 min, 50 sec.	90%	60%
2	2 min, 50 sec.	84.6%	69.2%
3	14 min, 10 sec.	97.6%	97.6%
4	3 min, 20 sec.	80%	73.3%

Despite the rather good camera utilization the system did experience challenges. First, the group sizes became rather large. When a group is found, the cameras lock on and treats them as one unit. When criteria for a group are no longer met, the system re-targets to individuals. It was observed that the walking instructor could sometimes cause a group to momentarily dissolve (either because of added velocity, or by being upright) thereby causing the system to re-target the cameras to individuals. Even if the resulting camera positions continue to record the medical procedure, the change in camera focus is undesirable to the user.

Second, the latency to complete a zoom command is much longer than pan and tilt commands. In some cases, while the Axis camera is in the process of executing a zoom command, a subsequent zoom command is ignored. Since the MMLT system sends zoom commands without checking whether they are actually executed, the system may enter "bad zoom state." Consider a sequence that begins with the camera focused on a group in a wide-angle view. The group is then momentarily dissolved, reassigning the camera to a single individual in a close-up view. While the camera is busy zooming in, the group re-forms and the camera is re-assigned; however, the "zoom out" command needed to view the group is never executed. This results in the camera aiming zoomed in closely at some arbitrary portion of the whole group. Third, disturbed earth continued to be a challenge on the second day, particularly under the barbed wire. This is especially problematic as it occurred in areas where the undesired objects could draw from the priority pool.

EVALUATION OF THE USE OF MMLT IN TRAINING

The Day 1 and Day 2 exercises had different characteristics. The notable differences were the length of the exercises (Day 1 median time was 7 minutes 40 seconds, while Day 2 median time was 3 minutes 20 seconds), and the

number of participants per group (Day 1 was 3, while Day 2 was 10). Quantitatively, both days had relatively high automated camera utilization. Day 1 averaged 88% on camera 1, and 90.6% on camera 2. Day one's first three data sets are markedly lower, which may be due to environmental changes (early students moving all the loose ground). Day two's exercise had slightly lower camera utilization (88% and 75% respectively), again with the first data set being lower in utilization. Table 3 provides an overview of the issues and solutions identified during testing.

Table 3. Testing Issues and Resolutions

Identified Issues	Potential Resolution
Software may track human-like objects (e.g. disturbed earth) rather than the tracks of interest	Add more specific filters for identifying a moving, human track
The grouping algorithm was recalculated each frame, potentially dissolving some groups temporarily and refocusing cameras away from the action	The grouping algorithm persists between frames, so smaller changes in the group dynamic will not dissolve the group
Difficult setup and calibration for non-technical users	Automate aspects of the setup process and provide a detailed walkthrough in the software
Cameras may get stuck in a "bad zoom" position	Add a check to verify the zoom command was successful
Camera focus and movement may appear jerky, the resulting video is difficult to watch	Ignore smaller changes in group and individual position, so the camera only refocuses during large position shifts

Qualitatively, both groups felt that the MMLT AAR visualization of their performance in the field was valuable. All students felt it would enhance the training experience to have access to the visual data performance after training and prior to final validation. Due to the site training schedule during the user test, the test team was unable to formally collect student and instructor surveys; however, student and instructor comments were noted. Students found that the MMLT system can assist in self-performance evaluation, similar to a "pre and post-game review" in sports. The instructors found the MMLT system to be a valuable tool for after action reviews of individual and team performance in a field environment.

CONCLUSIONS AND FUTURE WORK

MMLT embodies a novel multi-sensor approach to documenting training exercises. From the end user's point of view, the two data feeds (video and LiDAR) give complementary information, and by recognizing events (reaching wounded or instructors' time stamped grades), key moments can be indexed for easy retrieval. From an implementation point of view, tracking in the LiDAR data feed is easy and gives accurate coordinates for aiming the camera. The 3D model of the scene allows for anticipation of occlusions of tracks for camera handoff. The inclusion of several different data sources (checklist, Axis cameras, helmet cameras, and LiDAR) was found to increase the validity of the AAR. In several cases, one or more of the data sources could not accurately detect all of the actions: an object or person may have occluded an Axis camera or the LiDAR, or the instructor may have been distracted. In those cases, data from the other sources served to "fill in the blanks" to provide an accurate account of what occurred during those moments.

A limitation of the original approach was that groups were (re)discovered with each frame of LiDAR data. The MMLT system now allows groups to persist from frame to frame, which allows for greater reasoning about how to prioritize groups relative to other tracks, reducing the effect of instructors walking through a group. The challenges with disturbed earth have been addressed but still need to be tested. A step was added to compute statistics of height, number of points, and velocity for a long window of time, creating filters that can ignore disturbed earth tracks.

The MMLT system shows great potential as an AAR system for students and instructors in a field-training environment. The students greatly appreciated having video feedback and were interested in the LiDAR feed, especially for tactical movement review. The students also found value in how the LiDAR view showed the movement of the entire team, not just the individual, so they could see how well they worked together.

Future Development

With new technology becoming available, and to make MMLT more mobile and affordable, there should be further investigation into alternative LiDAR systems, including smaller and less expensive units. With less expensive units, more units can be combined into one system, allowing the coverage of more area. The system can also be made to operate on a local wireless network for data transfer. Future development will also include simplifying the calibration and setup process, allowing non-technical personnel to quickly set up the system prior to training.

The MMLT system could be used to fully automate the lanes during training. Using the zones and LiDAR data, virtual triggers can be implemented for certain effects like smoke and simulated artillery fire as the students move through the lane, removing the need to place instructors in the lane waiting for students to pass by. Further automation can be included, such as automatic medical skill checklist scoring using sensor data from manikins, and automatic tactical skill checklist scoring using LiDAR data and individual identifiers on each student (the LiDAR data only sees track 1 and track 2 currently, it does not know that track 1 is John Smith).

Future Applications

While the current MMLT system was designed for and tested with MSTC Combat Medic and Combat Lifesaver training, the system can be expanded for use in other applications, particularly training scenarios that currently do not have robust or objective data collection and assessment methods and those that focus on team training exercises. The MMLT system can be used to capture team training events to conduct comparisons of best practices/approaches for refining procedures specifically where lifesaving events can be improved upon. The system can track individual and team movements of experienced providers and compare those movements to a less experienced team, highlighting differences in efficiency during a scenario (e.g. better communication among team members for positioning). Training for disaster preparedness and incident response such as active shooter situations would also benefit from the MMLT system allowing providers to measure the effectiveness of patient triage and the timeliness of medical intervention. The 3D data and virtual shooter capability can highlight the importance of situational awareness during an active shooter scenario, ensuring the provider is aware of both the patient and of their position in relation to additional danger.

ACKNOWLEDGEMENTS

REFERENCES

Eastridge, B. J., Mabry, R. L., Seguin, P., Cantrell, J., Tops, T., Uribe, P., … Blackbourne, L. H. (2012). Death on the battlefield (2001–2011): implications for the future of combat casualty care. *Journal of trauma and acute care surgery*, *73*(6), S431-S437.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*, *96*(34), 226-231.

Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge University Press New York, NY. ISBN:0521540518.

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the society for industrial and applied mathematics*, *5*(1), 32-38.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, 82*(1), 35-45.

TRADOC. (2011, January 20). The U.S. Army Learning Concept for 2015. TRADOC Pamphlet 525-8-2.