NISTIR 7186

Terrain Characterization for TRL-6 Evaluation of an Unmanned Ground Vehicle

Michael Shneier
Tsai Hong
Tommy Chang
Harry Scott
Steve Legowik
Gerry Cheok
Chuck Giauque
David Gilsinn
Christopher Witzgall

U. S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Intelligent Systems Division
Gaithersburg, MD 20899-8230

NIST

National Institute of Standards and Technology
Technology Administration
U.S. Department of Commerce

AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

NISTIR 7186

Terrain Characterization for TRL-6 Evaluation of an Unmanned Ground Vehicle

Michael Shneier
Tsai Hong
Tommy Chang
Harry Scott
Steve Legowik
Gerry Cheok
Chuck Giauque
David Gilsinn
Christopher Witzgall

U. S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Intelligent Systems Division
Gaithersburg, MD 20899-8230

December 2004

U.S. DEPARTMENT OF COMMERCE
Donald L. Evans, Secretary

TECHNOLOGY ADMINISTRATION
Phillip J. Bond, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
Arden L. Bement, Jr., Director

Terrain Characterization for TRL-6 Evaluation of an Unmanned Ground Vehicle

Michael Shneier, Tsai Hong, Tommy Chang, Harry Scott, Steve Legowik, Gerry Cheok, Chuck Giauque, David Gilsinn, Christoph Witzgall
National Institute of Standards and Technology
Gaithersburg, MD 20899

INTRODUCTION

Mobile robotic vehicles have been under development for a number of years, but are only now becoming sophisticated enough to be useful in real-world applications. The Army has funded research into mobile robots through successive "Demo" programs, culminating in "Demo III"[1]. This program developed an experimental unmanned vehicle (XUV) intended to act as an Army scout. Following successful informal testing of the XUV, it was decided to conduct a more rigorous study to determine if the vehicle had reached Level 6 in the Technology Readiness (TRL) scale developed by NASA (Table 1)[2].

Testing was conducted at three sites with varying characteristics, including a desert site (Tooele Army Depot, UT), rolling vegetated terrain (Fort Indiantown Gap, PA), and an urban area (Fort Indiantown Gap, PA). The tests involved laying out test courses and carrying out a series of missions that required traversing all or part of the courses. Missions were evaluated using a number of factors including whether or not the mission was completed, how much human interaction was required during the mission, and how much time was required to complete the mission. The difficulty of the terrain over which the vehicle traveled was only indirectly included in the evaluation. Two test courses were set up at each site, a "black" course that was supposed to include difficult terrain, and a "gold" course intended to be less demanding. In practice, both courses contained segments that were difficult for the vehicles and those that were easier. As there are no metrics for terrain characterization for autonomous mobility, the decision as to what constituted difficult terrain was subjective and based on what was difficult for a human driver. In order to gain a better understanding of the actual difficulty of the terrain traversed by the XUVs, NIST developed methods of measuring terrain traversability and conducted terrain evaluation studies on all the test courses. This document provides an overview of the procedures and analyses that were used. The Appendix lists references that contain more details.

The goal of terrain characterization is to provide quantitative evaluation of the paths traversed by vehicles in carrying out their missions. This allows comparisons between autonomous vehicles and human drivers and between different types of vehicles. To determine if autonomous mobility has reached Technology Readiness Level 6, some measure is needed of vehicle performance over terrain of known difficulty. The Future Combat Systems program calls for vehicles capable of traversing environments including Urban, Open, Rolling-Arid, and Mixed, Open, Rolling-Vegetated, but does not specify exactly how these terrains are defined. The

selection of a particular terrain on which to test the vehicles will strongly affect the outcome of the tests. Unless there is some objective measure of terrain difficulty, the tests will have limited value. The techniques that NIST has developed and applied to measure terrain are described below.

Technology Readiness Level	Description
1. Basic principles observed and reported	Lowest level of technology readiness. Scientific research begins to be translated into applied research and development. Example might include paper studies of a technology's basic properties.
2. Technology concept and/or application formulated	Invention begins. Once basic principles are observed, practical applications can be invented. The application is speculative and there is no proof or detailed analysis to support the assumption. Examples are still limited to paper studies.
3. Analytical and experimental critical function and/or characteristic	Active research and development is initiated. This includes analytical studies and laboratory studies to physically validate analytical predictions of separate elements of the technology. Examples include components that are not yet integrated or representative.
4. Component and/or breadboard validation in laboratory environment	Basic technological components are integrated to establish that the pieces will work together. This is relatively "low fidelity" compared to the eventual system. Examples include integration of 'ad hoc' hardware in a laboratory.
5. Component and/or breadboard validation in relevant environment	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements so that the technology can be tested in a simulated environment. Examples include 'high fidelity' laboratory integration of components.
6. System/subsystem model or prototype demonstration in a relevant environment	Representative model or prototype system, which is well beyond the breadboard tested for TRL 5, is tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high fidelity laboratory environment or in simulated operational environment.
7. System prototype demonstration in a operational environment	Prototype near or at planned operational system. Represents a major step up from TRL 6, requiring the demonstration of an actual system prototype in an operational environment, such as in an aircraft, vehicle or space. Examples include testing the prototype in a test bed aircraft.
8. Actual system completed and 'flight qualified' through test and demonstration	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents the end of true system development. Examples include developmental test and evaluation of the system in its intended weapon system to determine if it meets design specifications.
9. Actual system 'flight proven' through successful mission operations	Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation. In almost all cases, this is the end of the last "bug fixing" aspects of true system development. Examples include using the system under operational mission conditions.

Table 1 The Technology Readiness Levels.

Terrain is characterized using sensors on the NIST HMMWV and on a trailer that measures soil mechanics. The NIST HMMWV is a standard military vehicle modified for the purposes of research and development in mobile robotics. The vehicle has the capability of driving autonomously or of being driven by a human. For the TRL-6 tests, it was driven manually. Mounted on the vehicle are racks to hold computers and related equipment, a generator to power the equipment, and numerous sensors (Figure 1). The sensor mounts are designed so that new sensors can easily be added.

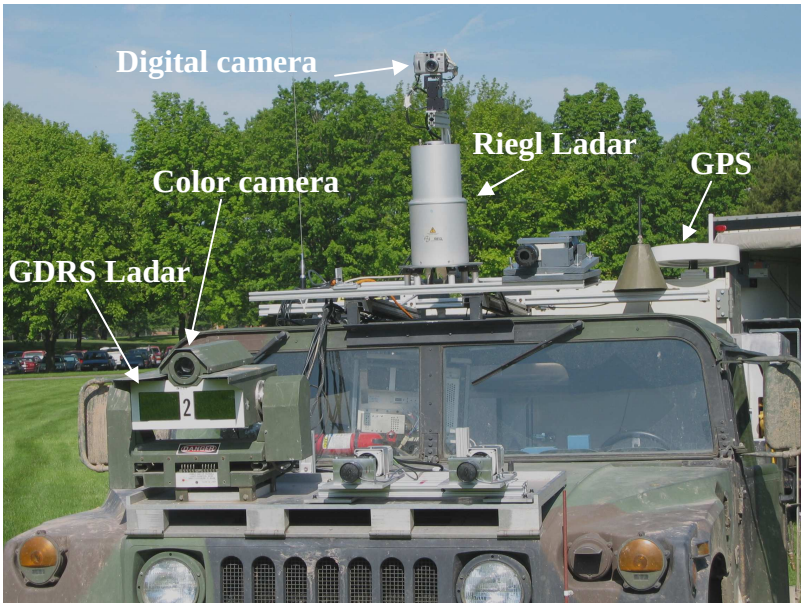


Figure 1 A view of the NIST HMMWV showing some of the sensors.

Sensors include a General Dynamics Robotics Systems (GDRS)[†] imaging ladar (Table 3) mounted on a tilt platform, a color camera mounted on top of the ladar on the tilt platform, and a highly accurate positioning system (Table 4 and Table 5)[3]. A Riegl high-resolution scanning ladar (Table 6) provides the prime source of range data for terrain characterization. In the initial phases of the program, a ring of cameras (not shown in Figure 1) was mounted concentric with the Riegl ladar to provide color data, and in later phases, a digital camera attached to a pan-tilt unit mounted above the Riegl ladar provided panoramic images of the terrain. In some of the tests, a Sick ladar was mounted on the back of the HMMWV pointing straight down at the ground to provide a measure of the topography with precision sufficient to measure the depths of ruts in the ground left by the HMMWV wheels.

[†] Certain commercial equipment, instruments, or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily best for the purpose.

COLLECTING THE DATA

In order to combine data from multiple sensors, it is first necessary to calibrate the sensors and register them to a common coordinate system. Registration also enables the spatial relationships between successive samples to be computed. Preparing for data collection includes calibrating the sensors and accurately measuring their positions and orientations on the vehicle. The cameras are calibrated using Bouguet's method[4]. The ladars are each calibrated using special-purpose methods. For example, the GDRS ladar is calibrated by mounting it on a highly accurate pan-tilt platform and determining where each laser pixel points by moving the ladar until the point is centered on a calibration target. The angle at which the laser is pointed for each pixel can be determined from the pan and tilt position of the platform.

The positions and orientations of the sensors relative to the vehicle coordinate system and to each other are determined using an external measurement system. We use an ArcSecond laser-based site measurement system (SMS) to provide these measurements. For the Riegl ladar, the approach is to park the vehicle in such a way that it faces two orthogonal walls. The Riegl ladar

Figure 2 Data from a scan of orthogonal walls using the Riegl ladar.

is then used to scan these walls and the ground plane (Figure 2). By measuring points on the walls and the ground with the ArcSecond sensor, transformations are obtained for the Riegl to the building and building to ArcSecond. Similarly, the ArcSecond sensor is used to gather points on the HMMWV to determine the ArcSecond to HMMWV transform. Finally, the Riegl to HMMWV transform can be obtained by matrix multiplication

$$Riegl \rightarrow HMMWV = Riegl \rightarrow Building * Building \rightarrow ArcSecond * ArcSecond \rightarrow HMMWW$$

Similar methods are used to locate the other sensors relative to the vehicle.

Data are collected in two primary modes. One is while the vehicle is driving normally, and the other is with the vehicle stationary. Some of the sensors do not run in real time, so can only be used when the vehicle is not moving. The trade-off between the two modes is that while data acquired in real-time approximate more closely the actual driving conditions, they are less accurate and usually of lower resolution than data from the slower sensors used when the vehicle is stopped. We expect that higher resolution data will soon become available in real time as new sensors are developed. This will enable all data to be collected with the vehicle in motion.

A critical part of data collection for mobile vehicle applications is to record the vehicle's position

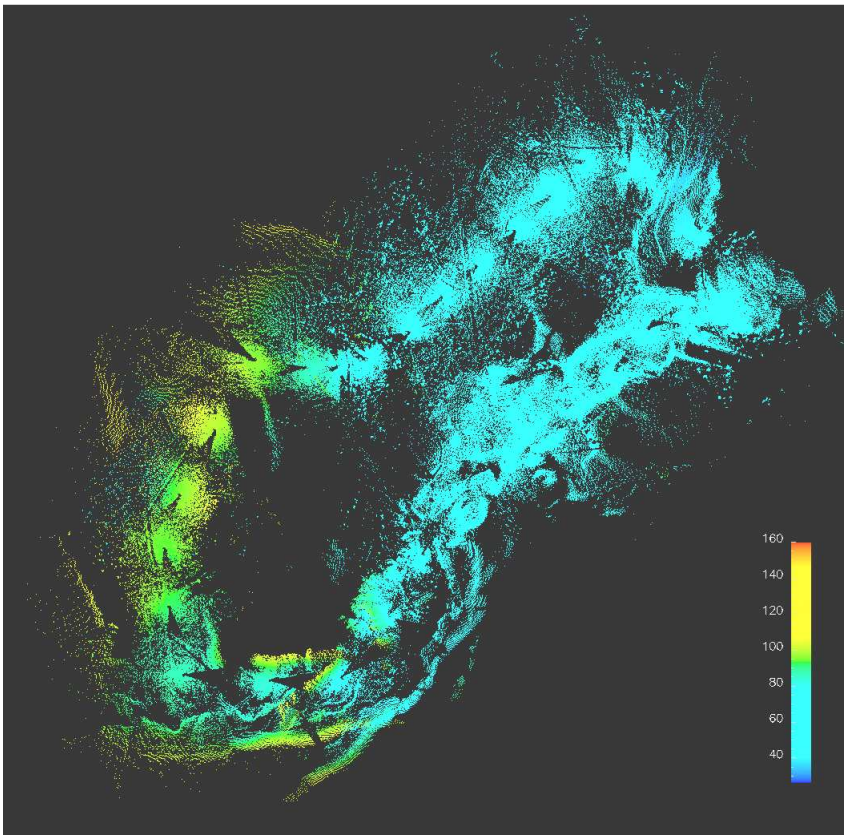


Figure 3 An overhead view of the coarse ladar scans collected on the Black course at Tooele Army Depot. The colors encode the height of the terrain.

and orientation (pose) and the time at which each sample is acquired. This enables data collected from multiple sensors to be registered, and also allows the data for a complete mission to be compiled into a reconstruction of all the terrain that was traversed (Figure 3). An Applanix navigation unit that combines an inertial component with information from the Global Positioning System (GPS) provides the vehicle pose and the time. This unit typically provides real-time data accurate to better than one meter and a few hundredths of a degree. With post-processing, the accuracy is a few centimeters in distance and angular accuracy is a few thousandths of a degree (Table 4 and Table 5).

Three sets of data are collected for each course. First, the vehicle is driven over the course collecting data with the real-time sensors (real-time Ladar, color video camera, and, in some of the runs, Sick ladar data). Next, the vehicle is moved to the first waypoint on the course. Starting from this point, and moving 10 m to 150 m between samples, scans are taken of the terrain using the Riegl ladar, GDRS Ladar, panoramic digital camera and a set of six cameras arranged in a ring around the Riegl. The scans are not taken at the highest resolution the Riegl ladar can measure, but still provide much more accurate information than the real-time sensors (Figure 3 and Figure 4). The navigation data are also stored to provide the position and attitude

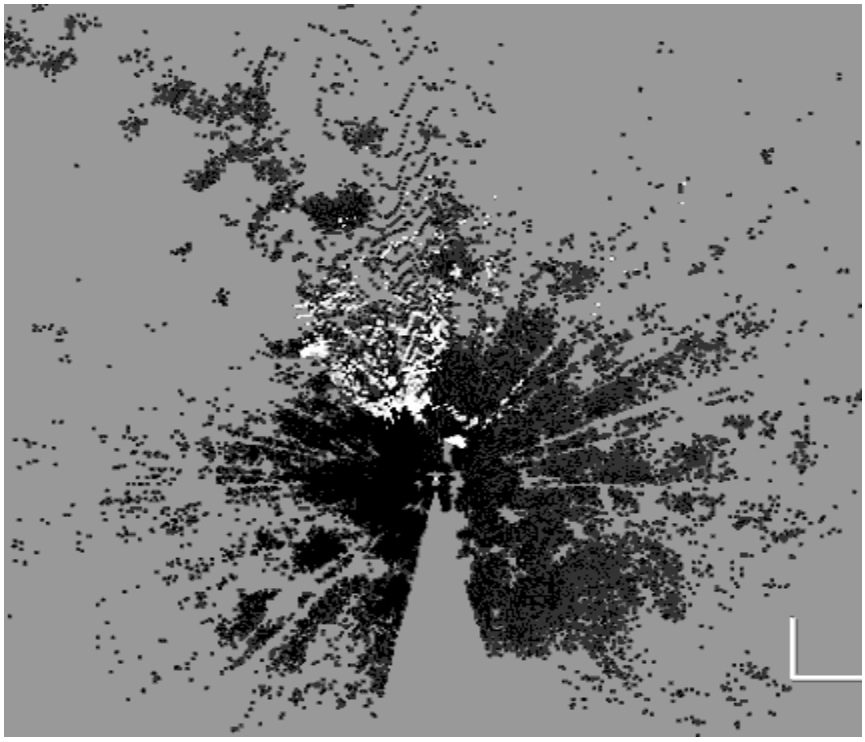


Figure 4 Top-down view of coarse resolution Riegl data overlaid with GDRS Ladar data. Riegl data are shown in dark grey and GDRS data in white. The differences in field of view and density of points are apparent.

(roll, pitch, heading) of the data collection vehicle at the time the sample is collected. The entire course is sampled in this way. Finally, a set of high-resolution scans is taken of the difficult or interesting locations on the course.

To collect real-time data, the NIST HMMWV is driven over the course being characterized. A human driver navigates by following a visual display of the course which shows the waypoints and the HMMWV's trajectory. The sensors are started at the same instant in time and navigation and timing data are collected with the sensory data. For TRL-6, the sensors used for real-time data collection are a GDRS ladar (the primary sensor used by the XUV), a color video camera mounted on top of the GDRS ladar, and the INS/GPS position data from the Applanix unit augmented by a differential GPS base station. In some cases, Sick ladar data were also collected.

Non-real time data are collected for the entire course at a relatively coarse resolution. Higher resolution data are gathered at locations that required an emergency stop for the XUV or where the vehicle displayed "interesting" behavior, such as backing up, suddenly changing direction, or performing an unanticipated intelligent maneuver.

ORGANIZING THE DATA

An enormous amount of sensor data was collected over the course of the TRL-6 evaluations (roughly 250 GB). Cataloging it and storing it in a way that makes access easy was a non-trivial undertaking. It was important to keep all the sensor data from a particular test run together and to associate the relevant time and position data with it. The data were collected using a number of different computers, so a uniform naming scheme was necessary to keep these associations. This worked most of the time, but sometimes one of the sensors failed to take any data, or took more than one set, so numbering schemes could get out of step. Using the time the data were collected also had its problems, since the computers had their own clocks that were not always synchronized. After trying many schemes, a joint number and time label worked best, but there were still occasional errors in grouping the data that had to be manually corrected.

The goals of organizing the data were to make it easy to access for analysis and to make it available to a wide audience for algorithm development and evaluation. The approach we took was to build a database containing information about each data set. We developed a web-based query application that would enable data to be selected based on a wide range of properties[5]. We chose MySQL (http://www.mysql.com/) as the database engine, and developed a web form to enter data into the database (Figure 5). This made the data entry relatively easy, although rather slow. The user was required to provide keywords and sample images for each set of data, which meant that the entry could not be automated and could not be done while the data were being collected. Other information stored with each data record included the date and time, the location, the weather conditions, sensors used and size of each sensor's data. The sensor data are stored on a large network storage device, and the location of each file is stored in the database.

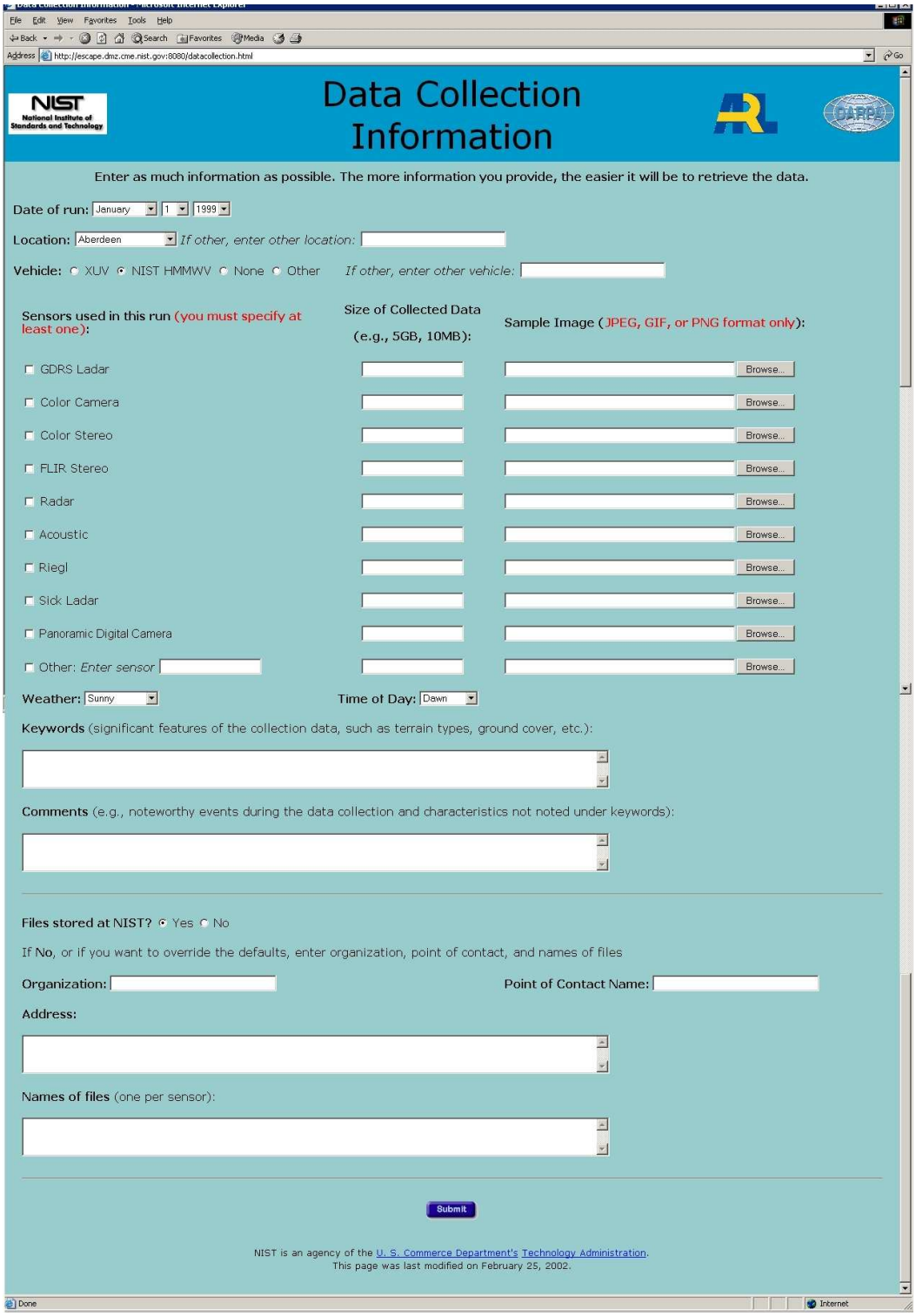


Figure 5 The web form for entering descriptions of sensor data

ACCESSING THE DATA

Searching for data sets with particular attributes is relatively easy. A web query form allows queries using keywords or any of the other attributes in the database schema (Figure 6). This

web interface will soon be made available outside the NIST firewall to enable other researchers to request data sets for research in autonomous mobile robotics. Work is also under way to add ground truth to some of the data. This involves annotating the images or image sequences with labels such as road edges and markings, pedestrians, vehicles, etc.[6]. Ground truth will allow performance evaluation of algorithms that track such features. The query interface allows

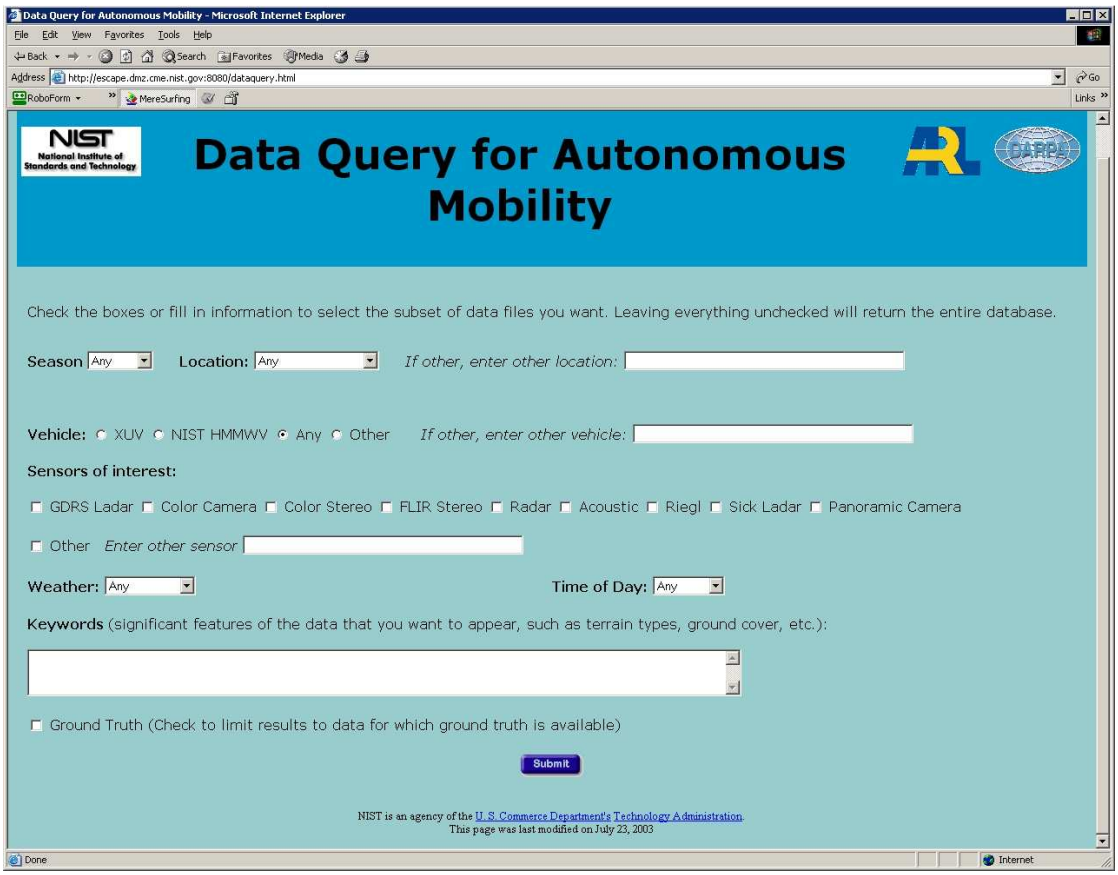


Figure 6 The web query interface for requesting data sets with specific attributes

requests that select only data for which ground truth exists. Note that some of the data are restricted for use only within the Army's Collaborative Technology Alliance.

ANALYZING THE DATA

Data analysis has not yet been automated, although efforts to develop the necessary algorithms are under way[7]. Current work makes use of visualization tools and computer-assisted measurements. First, the navigation data are post-processed to provide the highest accuracy in the position and orientation of the HMMWV when each data sample was acquired. The real-time data are then processed and inserted into a world model which represents the terrain in a grid, with elevation values in each grid cell. This enables the data to be visualized in a form close to that available to the XUV. A movie showing the terrain changing as the vehicle moves can provide valuable insight into what the vehicle knew about its surroundings at each point on the course. Watching the movie may provide an explanation of why it selected a particular path.

It must be realized, however, that the low resolution and small field of view of the real-time sensors provide such data only for the path traversed by the NIST HMMWV, and not for the range of paths selected by the XUVs on successive traversals of the course. Also, as the GDRS ladar that is on the HMMWV does not pan or tilt, it sees a smaller swath of terrain than the sensors on the XUV.

For more detailed analysis, the coarse scans taken by the Riegl ladar provide data from a much larger field of view. These data are augmented with color camera data from the ring cameras that surround the Riegl ladar and with high-resolution panoramic images from a digital still camera. The individual coarse-resolution ladar scans taken at intervals around the course are combined using the post-processed position and orientation information from the navigation system (Figure 3). They can then be used to compute terrain elevation, roughness, etc. over a region that covers several hundred meters square (Table 6). A major problem in making these computations, however, is the determination of what is ground. Is ground to be defined as bare earth or are small rocks and low vegetation such as bushes to be considered part of the ground? The inclusion of bushes and rocks as part of the ground would yield a more “bumpy” or rougher surface as compared to a bare earth surface. The next issue then is “what is roughness?” and “how is it defined?” Some possible indicators of roughness are large variability in slope, steep slopes, and high vegetation.

Support Surface Determination

It was decided that the support surface would be taken as bare earth when processing the TRL-6 data. Lacking an automated way of identifying and deleting the ground cover, heuristics are used. These include binning the data and looking at the minimum elevation value in each bin under the assumption that the vegetation is porous and some of the return signals are ground strikes or points under the vegetation. This binning procedure yields a first cut at the ground surface. These data points were further processed to remove vegetation points using a masking technique. This technique involves the creation of a TIN (triangulated irregular network) surface where vertices are adaptively selected from among the original data points. The adaptive selection results in the insertion of points in regions of high elevation variability (i.e., vegetation) and these regions can be identified by concentrations of small triangles. The data points associated with these triangles can be “masked” or removed. The remaining points are then used to create a surface against which the original data set is compared. Points above a user-specified tolerance from this mow surface are deleted. The tree and bush bins are then identified as described in the next section, and are omitted from the support surface determination. The binning, masking, mowing, and identification procedures are re-iterated. An example of a ground surface obtained using the above procedures is shown in Figure 7 and Figure 8.

Figure 7 Digital Photo of Emergency Stop 2, Fort Indiantown Gap, Gold Course.

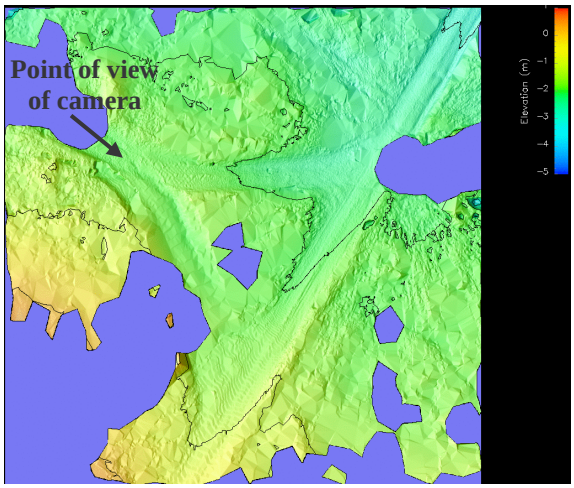


Figure 8 Estimated Ground Surface

Tree and Bush Detection

Trees and bushes are major obstacles to mobility in off-road terrain. While bushes may slow the progress of the vehicle, trees usually are obstacles that require a change in the planned path and may prevent the successful completion of the mission. Thus, detecting and labeling bushes and trees are important for determining the traversability of the terrain. Further, bushes and trees need to be identified and removed from the processing of the bare ground to determine surface roughness.

Trees and bushes are detected in the Riegl ladar point-cloud data by looking at properties of points in the binned data. The (x, y, z) points are projected onto the (x, y) plane, which is divided into uniform square bins. Each cell is then treated individually to determine if it is part of a tree or a bush (or neither). This determination is made based on the number of points in the cell and the elevation statistics of the points.

Assume that the projected data have been divided into uniformly sized bins and let n be the number of points in bin (i, j). Let $(x_{min}, y_{min}, z_{min})$ be the point with minimum in the bin. Let z_{mean} be the mean of the elevations in the bin, and z_{max} the maximum elevation. Then the criteria for trees are:

$$n > 20$$

$z_{max} - z_{min} > 3.75\ m$
$z_{mean} - z_{min} \geq 0.75\ m$

see Figure 9.

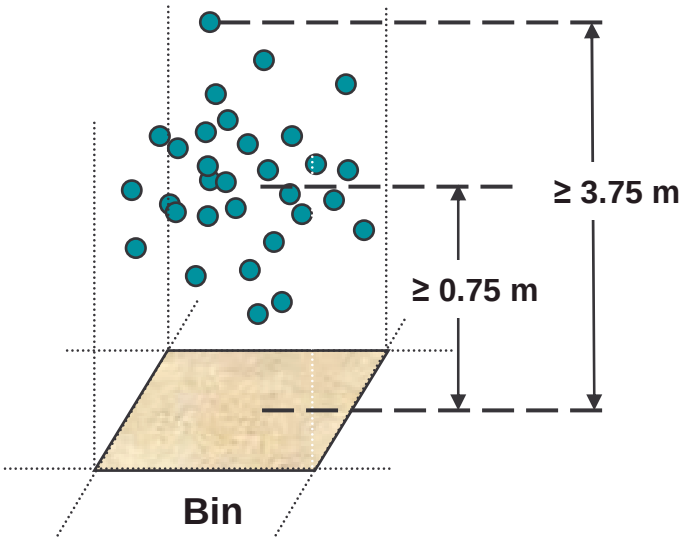


Figure 9 Criteria for a bin to be identified as containing part of a tree

The criteria for bushes (**Figure 10**) are given by:

$n > 20$
$z_{mean} - z_{min} \geq 0.25\ m$

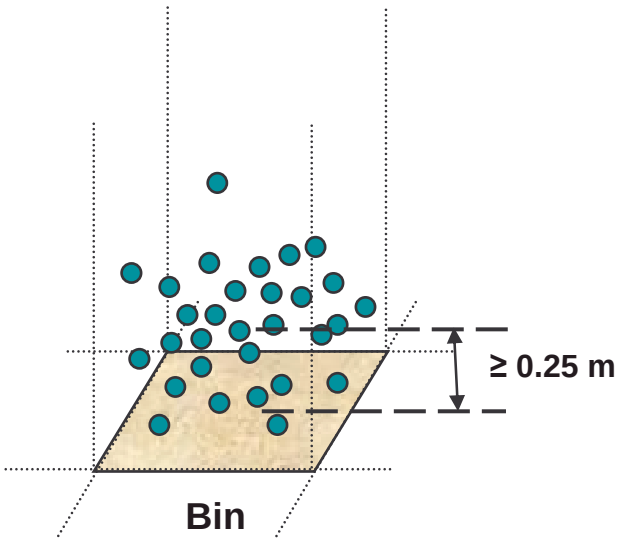


Figure 10 Criteria for a bin to be identified as containing part of a bush.

If neither set of criteria is met, the bin is not classified. Tree coverage is then defined as the ratio of the number of tree bins divided by the total number of occupied bins. Similarly, bush coverage is the number of bush bins divided by the total number of non-empty bins. Note that the criteria for trees identify the entire canopy, rather than only the trunk of the tree. This may result in some regions being labeled as not traversable when the vehicle could in fact drive under the canopy.

Surface Slope and Roughness

It is important for terrain characterization to have a measure of surface slope and roughness. Both of these measures depend on the scale at which they are computed. Small variations, such as ruts or small rocks may be acceptable to the XUV, which is a very capable platform. Larger scale roughness, however, may make a path non-traversable. The navigation unit on the vehicle can provide a measure of surface roughness given that it measures the pitch and roll of the vehicle. These values are, however, only valid for the actual path driven by the vehicle. It would be useful to be able to measure slope and roughness over a wider field of view as provided by the Riegl ladar. To this end, work was done on computing these measures using the original data points (i.e., use all the data points and not a subset of the data points) in bins classified as ground bins and from the binned point cloud data, after extraction of the ground surface [7].

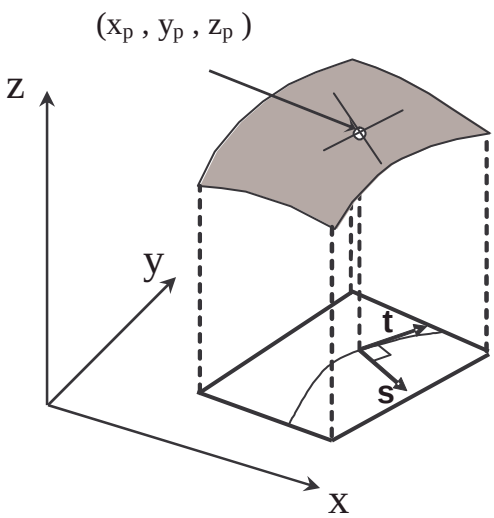


Figure 11 Path slope determination

The ground points as obtained using the procedures described above approximate a continuous surface. A path over this surface can be approximated by a sequence of points. At any point (x_p, y_p) on the path (Figure 11), let **t** be the travel direction in the base plane, $\|\mathbf{t}\| = 1$. Let **s** be perpendicular to **t**: **s** = **t** $\times$ **z**, where **z** is a unit vector in the direction of the z axis. The vector **s** is therefore also a unit vector in the base plane. The surface slopes are given by their inner

products with the surface gradient $\nabla z = (z_x, z_y)$. They are also the tangents of the pitch and roll angles, respectively. Thus pitch = $\tan^{-1}$ (<**t**, ∇z>), and roll = $\tan^{-1}$ (<**s**, ∇z>). The partial derivatives z_x and z_y and the travel direction **t** can be estimated locally from the ground surface data and the points describing the path.

Terrain roughness depends on the scale at which it is measured. Roughness includes a high-frequency component that arises from ruts, rocks, vegetation, etc., with a lower frequency component that follows the changing slope of the ground surface. Except in extreme cases, vehicle mobility is more dependent on the lower frequency component, although comfort depends strongly on the higher frequencies. The navigation sensors, which update at 200 Hz, give a good estimate of the high frequency changes and can be processed to give estimates of the lower frequencies also. Extracting this information from sensor data is also possible, but subject to substantial error because of the need for segmenting out the vegetation and other obscurants of the true ground surface.

The approach to gauge roughness followed in [7] works by binning the ground truth points into 1 m x 1 m bins. A plane is fitted to the points in each bin using least squares. This yields the slope of the plane in the *x* and *y* directions. The bins are aggregated to give an RMS slope in each direction: $xslope_{RMS}$ and $yslope_{RMS}$. Then terrain roughness is defined as:

$$\text{Roughness ID} = \sqrt{(xslope_{RMS})^2 + (yslope_{RMS})^2}$$

This measure is zero if and only if the surface is a plane, and, if it has the same value in two areas, it has the same value in the union of the areas. This roughness index is used as a relative measure to determine if one region is rougher than another.

EXAMPLES OF DATA ANALYSIS

Emergency stops were initiated when the XUV appeared about to put itself in danger of damaging itself. The locations where these stops occurred were given extra attention in an effort to understand why the vehicle got into such a situation. The region around an emergency stop was scanned multiple times from different locations using the highest resolution of the Riegl ladar. The scans were merged into a combined point cloud, either by visually lining up landmarks visible in each scan, or by using the position information from the navigation sensors to transform the data into a common coordinate system. This combined or registered point cloud formed the basis for the data analysis: surface generation, slope analysis, vegetation identification, and visualization. The data were visualized using either Data Explorer - OpenDX—http://www.opendx.org or with special-purpose software developed at NIST.

Dam Region at Tooele Army Depot

In Tooele there was a particularly difficult region at a dry dam. Figure 12 shows the dam region with the locations where high resolution ladar scans were taken. Two scans were combined to

give a high resolution topographic map of the region and a set of measurements was made to analyze the region. Figure 13 shows a combined scan, annotated as follows. The dotted line shows two traversals by the XUV. One ended in success and the other in failure. The failed path indicates that the XUV was misled by its inability to see far enough ahead to choose a good path. The slope of the ground along the path it chose is not as steep as the successful path near the top of the dam, but becomes steeper near the tree. When the XUV approaches the steeper section, it is going sideways along the slope, and an emergency stop was initiated to prevent it from tipping over.

The successful trajectory requires the vehicle to take a path that initially appears steeper than the alternatives, but the vehicle can traverse it because it does not have to travel sideways on the steeper slope. The angles shown in the figure for the slopes on either side of the dam wall were computed by fitting surfaces to the wall and measuring their slopes. The measurements match the pitch recorded by the navigation system very well, as described later in this document.

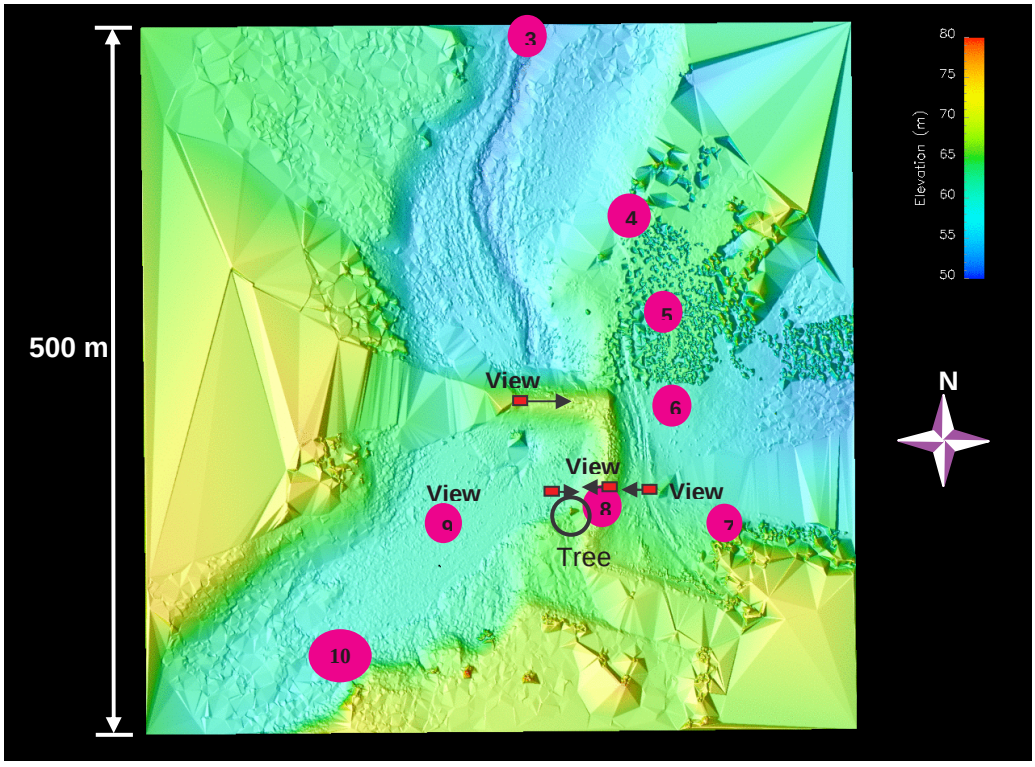


Figure 12 Dam region, Tooele. Locations where high resolution ladar scans were taken. The numbered circles are waypoints on the black course. The circle near waypoint 8 corresponds to a tree that resulted in several emergency stops.

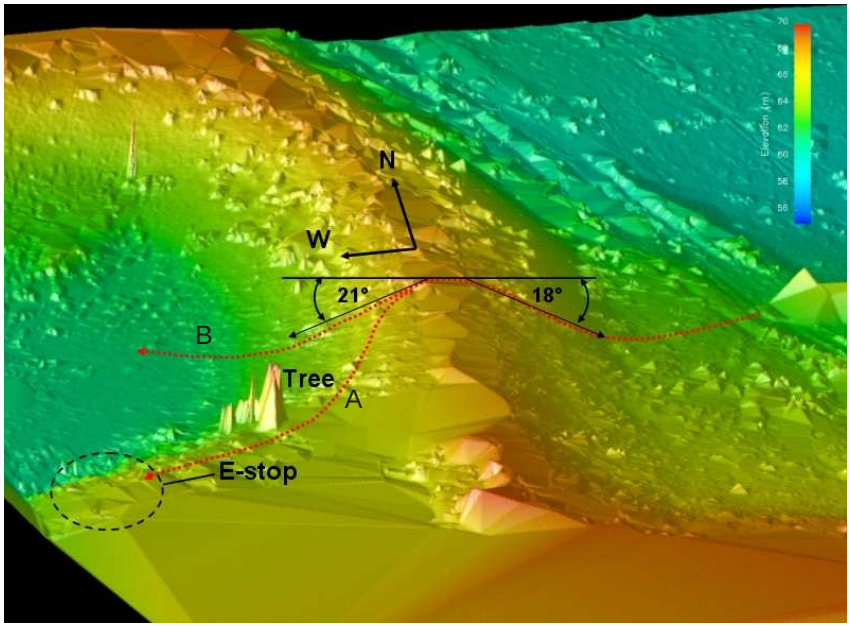


Figure 13 Dam region: Dotted lines show alternate paths taken by the XUV. One path (A) ends in an emergency stop, while the other (B) traverses the region successfully.

Urban Area Culvert

Figure 14 shows a view of a road in the urban area of Fort Indiantown Gap. There were several emergency stops at this location and other similar locations. The cause was the culvert outlined in the image. The XUV would often cut corners because it was not required to pass through all waypoints exactly. In cases where there were culverts near the corner, it would sometimes get into a dangerous situation and have to be stopped. Again, this was largely due to the inadequate field of view of the sensors. Figure 15 shows a surface created from a high resolution scan of the area around the culvert. The depth of the depression was determined to be about 0.8 m at the lowest point. If the XUV had had a sensor with similar capabilities to the Riegl sensor that could see further ahead and could identify negative obstacles, it would probably not have had any difficulty traversing these regions.

Figure 14 View of a culvert, the most frequent cause of emergency stops in the urban test area.

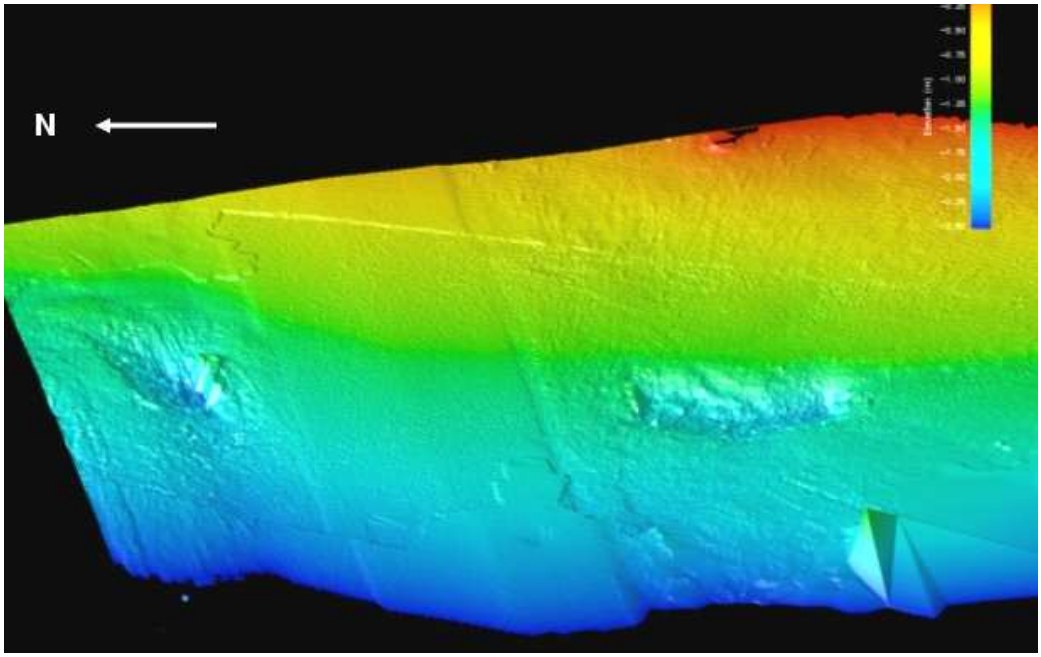


Figure 15 Surface of the culvert region. The depth of the culvert was measured at about 0.8 m.

Gravel Piles

Another trouble spot at Fort Indiantown Gap occurred on the off-road Gold Course due once more to the XUV cutting a corner in its path. At this location, there were a number of piles of gravel, as shown in Figure 16. The XUV tried to cut a corner and climbed up on the piles, leading to an emergency stop. Again, a better range sensor that could see further out would have avoided this situation (Figure 17).

Figure 16 Piles of gravel on the off-road Gold Course at Fort Indiantown Gap.

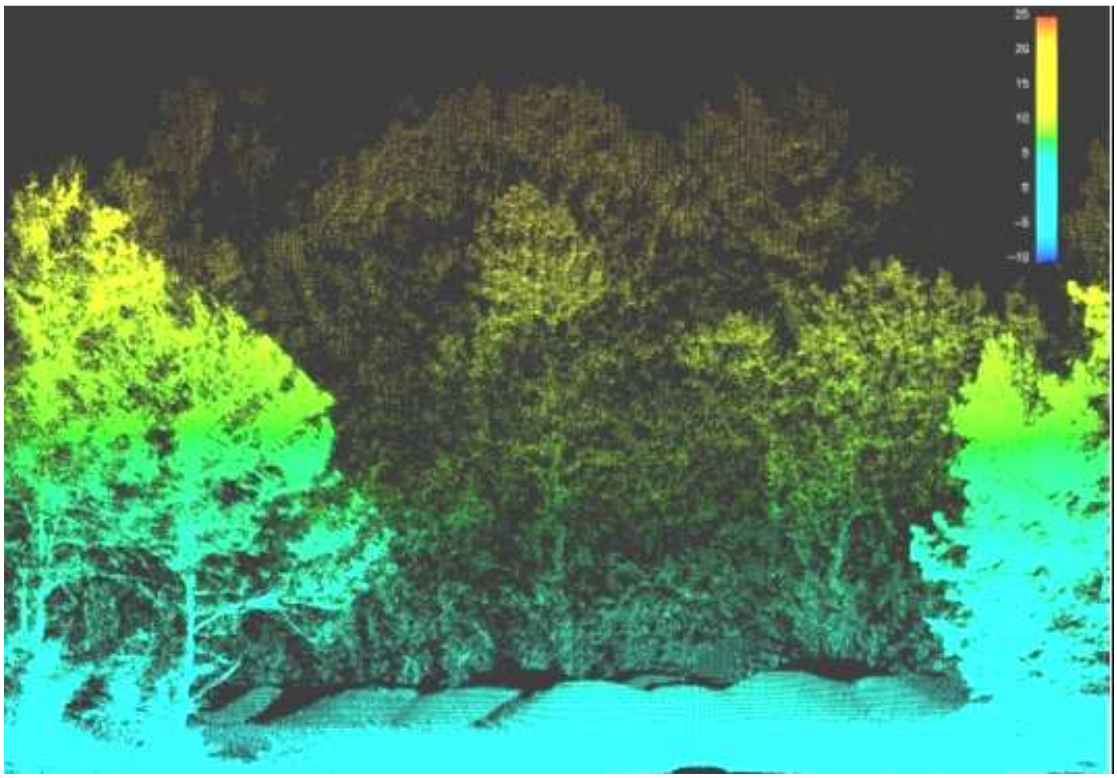


Figure 17 A view of the gravel piles as seen by the Riegl ladar.

Logs

An interesting case is a pile of logs on the Black course at Fort Indiantown Gap that was always seen by the XUV when traveling counter-clockwise, but frequently missed when traveling clockwise. As can be seen in Figure 18, the logs are obvious when the vehicle is traveling in a counter-clockwise direction on the course, but very hard to see when traveling clockwise Figure 19. In this case, the high-resolution sensor is not much help, as is shown in Figure 20.

Figure 18 Logs seen from anti-clockwise direction of travel.

Figure 19 Logs seen from clockwise direction of travel.

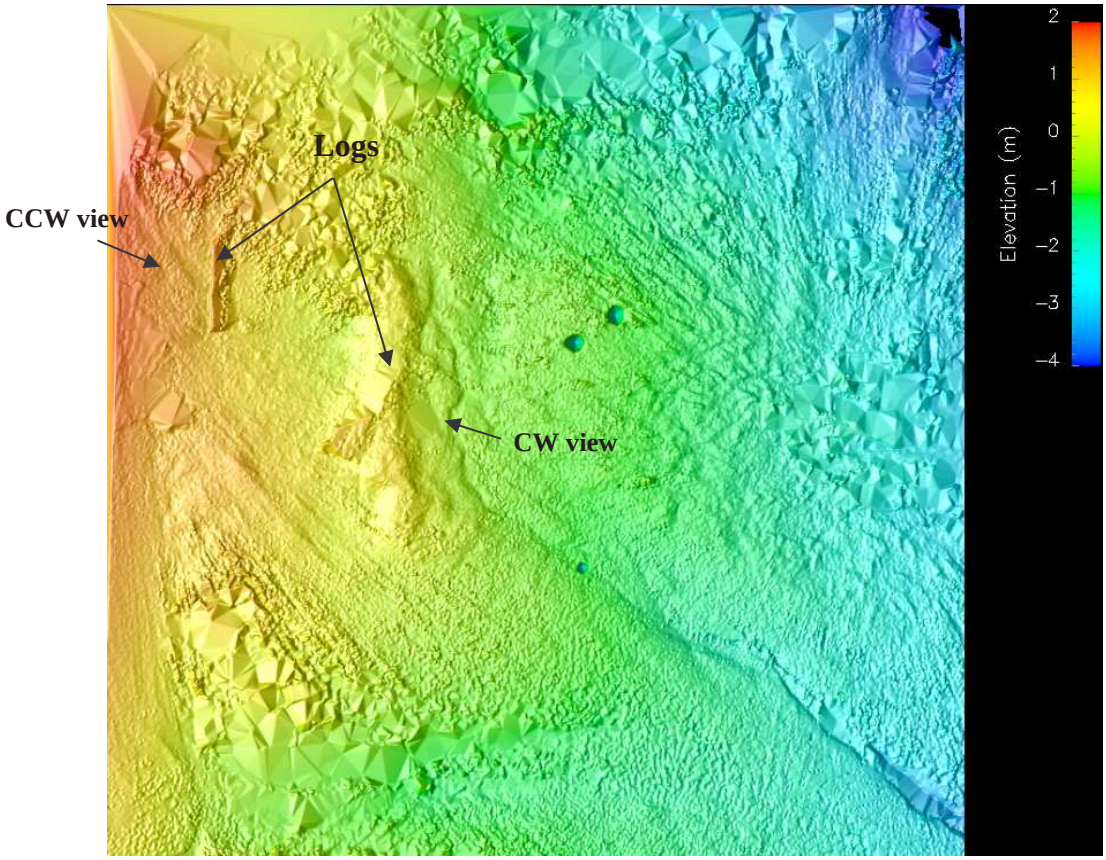


Figure 20 Merged clockwise and anti-clockwise Riegl scans. Some of the logs are very difficult to detect even in this high resolution data.

Air/Ground Registration

During the Urban part of the TRL-6 exercise, the opportunity arose to work with ladar data taken from a small helicopter that scanned the urban terrain. NIST worked on registering the overhead ladar data with ground-based Riegl data[8]. There were two main goals for this work. One was to build more complete models of the terrain, and the other was to determine if registration could help in maintaining good position estimates if the GPS on one of the vehicles failed.

The approach used was a hybrid of a feature-based and a point cloud-based Iterative Closest Point (ICP) algorithm for registering two different sets of LADAR range images. The feature-based method detected corners in each of the sets of data. By matching corresponding corners, the two data sets could be brought closer together. At this point, an ICP algorithm was used to improve the correspondence. As can be seen from Figure 21, the registration succeeds in bringing what are initially unregistered data sets into close registration.

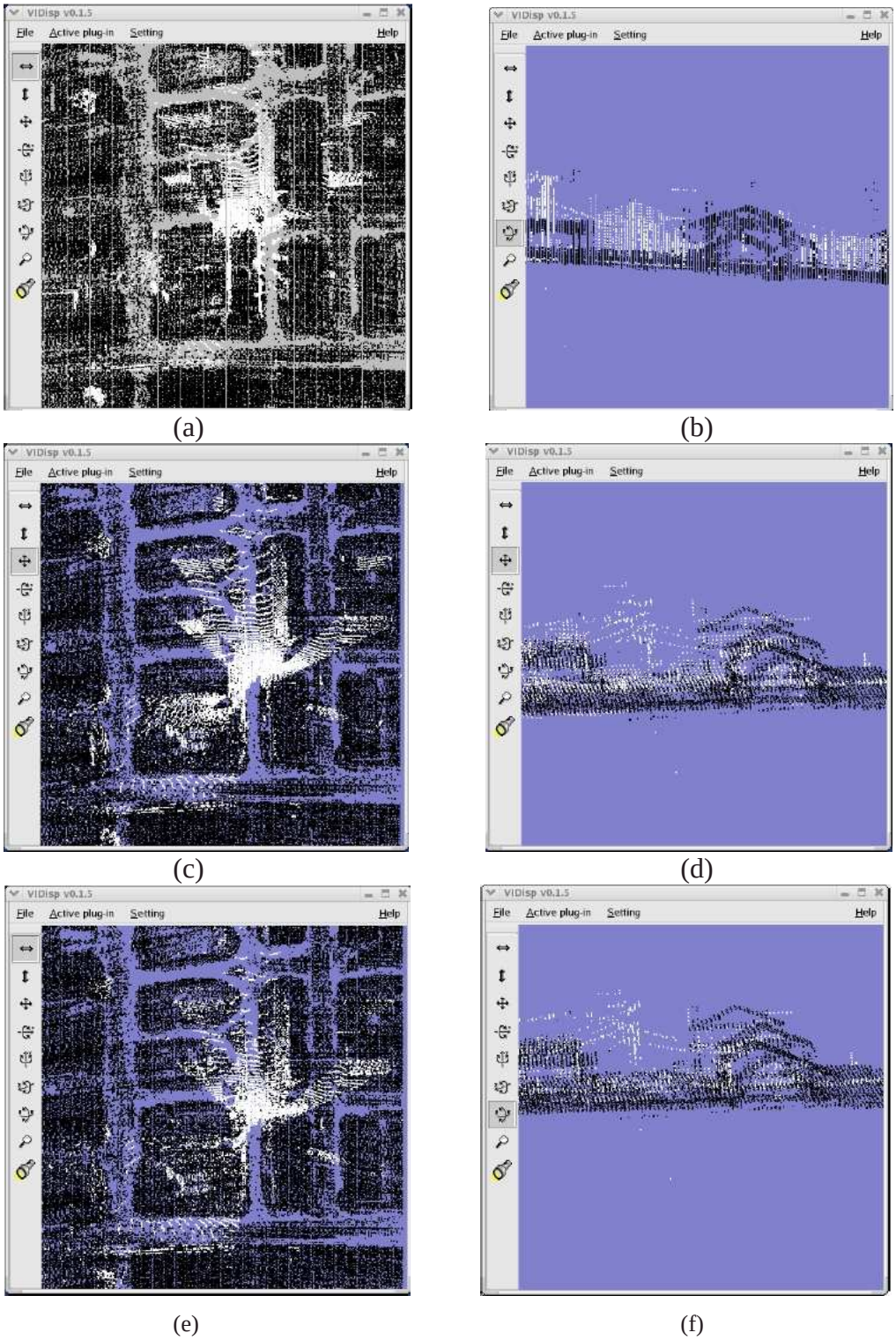
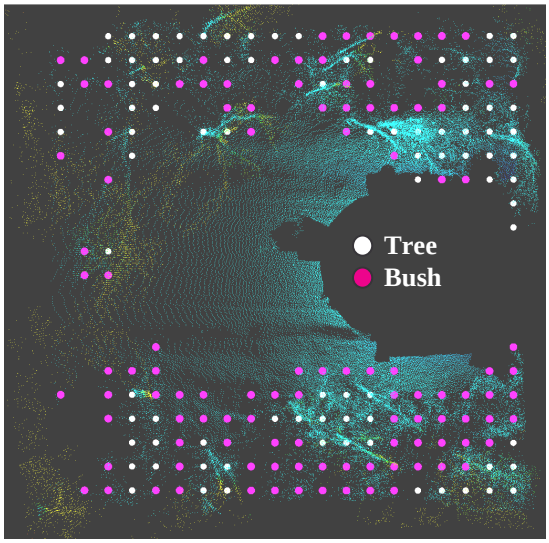


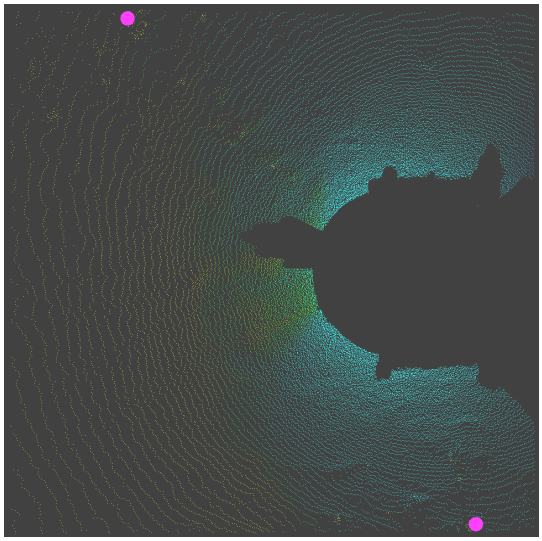
Figure 21 (a) A top view of unregistered range images of UGV and UAV Ladars, (c) the feature-based translation obtained using the extracted corners, and (e) the registered UAV and UGV Ladar range images obtained by utilizing the feature-based translation results. (b), (d) and (f), respectively, show magnified side views of their counterparts in the left column.

Tree and Bush Detection

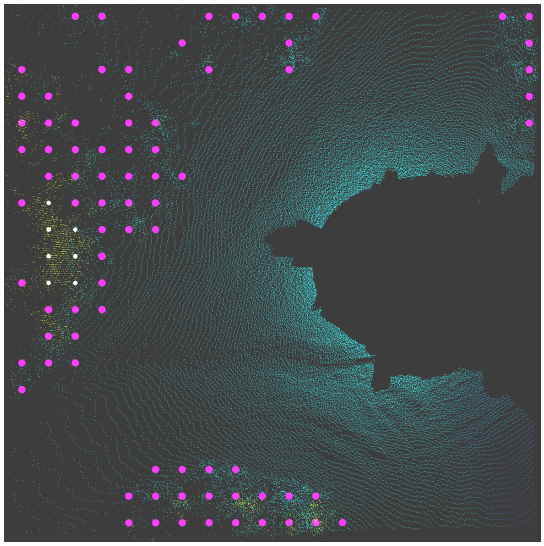
When the vegetation detection algorithm described above was applied to the 90 coarse scans from the black course at Fort Indiantown Gap, the mean vegetation density was computed as 27 %, with a standard deviation of 19 %. For the 50 coarse scans of the gold course, the average vegetation density across all the scans was 37 %, with a standard deviation of 19 %. The high standard deviation implies, correctly, that there are some regions with a lot of vegetation and others with only a little. Figure 22 illustrates the range of vegetation densities encountered in the black course and Figure 23 illustrates the range encountered in the gold course.



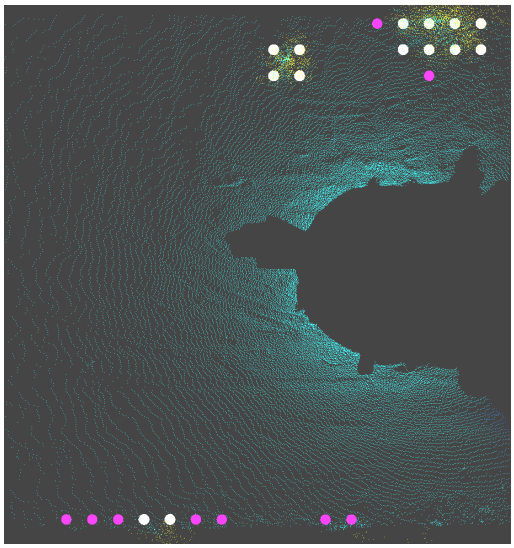
a. Stop 7, Vegetation Density = 54 %



b. Stop 35, Vegetation density = 0.6 %

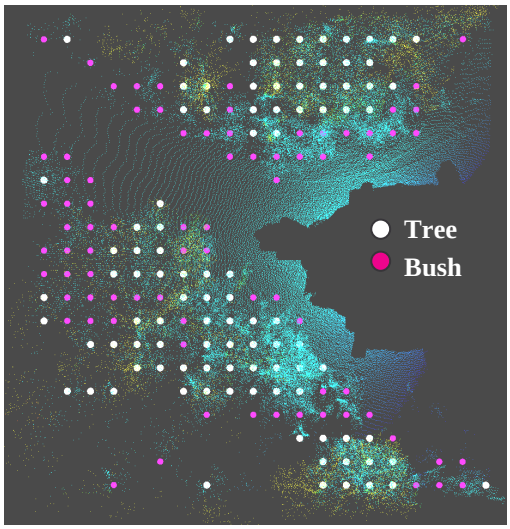


c. Stop 32, Vegetation density = 25 %

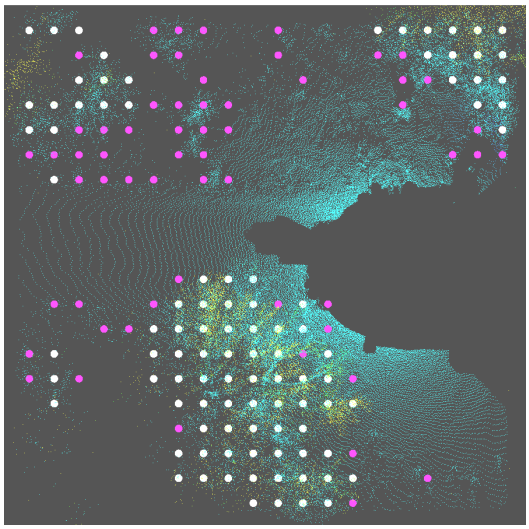


d. Stop 63, Vegetation density, 7 %

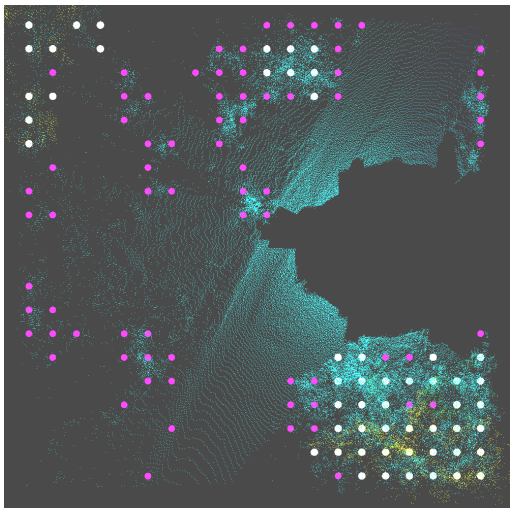
Figure 22 Fort Indiantown Gap, January 2003, Black course coarse scans. Vegetation density computed for a 20 m x 20 m region around HMMWV.



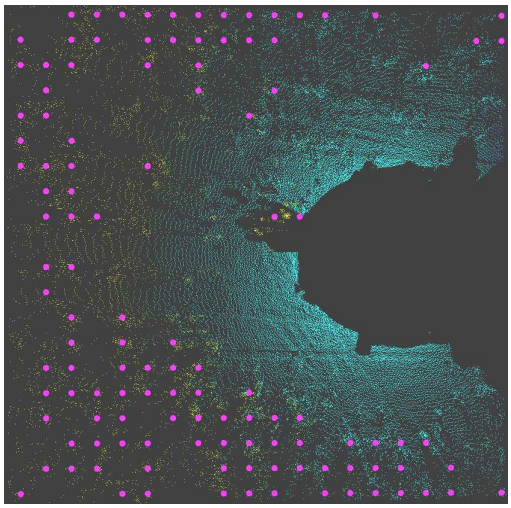
a. Stop 1, Vegetation density = 53 %



b. Stop 2, Vegetation density = 46 %



c. Stop 3, Vegetation density = 38 %



d. Stop 50, Vegetation density = 35 %

Figure 23 FORT INDIANTOWN GAP, January 2003, Gold Course coarse scans. Vegetation density computed for a 20 m x 20 m region around the HMMWV.

Slope and Roughness Determination

For selected regions on the course a more detailed analysis is required. These regions include those that either proved difficult for the robotic vehicles or were considered interesting for some reason. The analysis starts with a series of fine scans from the Riegl ladar, registered to a set of images from the ring of cameras around the ladar, and with the position and orientation readings from the navigation system. In the experiments conducted at Tooele Army Depot, three regions, the Ravine, the Dam, and the Wash, were identified as being particularly interesting (Figure 24). In these regions, a number of high-resolution ladar scans were made from different viewpoints.

The point clouds for each region were registered visually to provide detailed range and color information. Two locations were used for the Ravine, four for the Dam, and fourteen for the (much larger) Wash area.

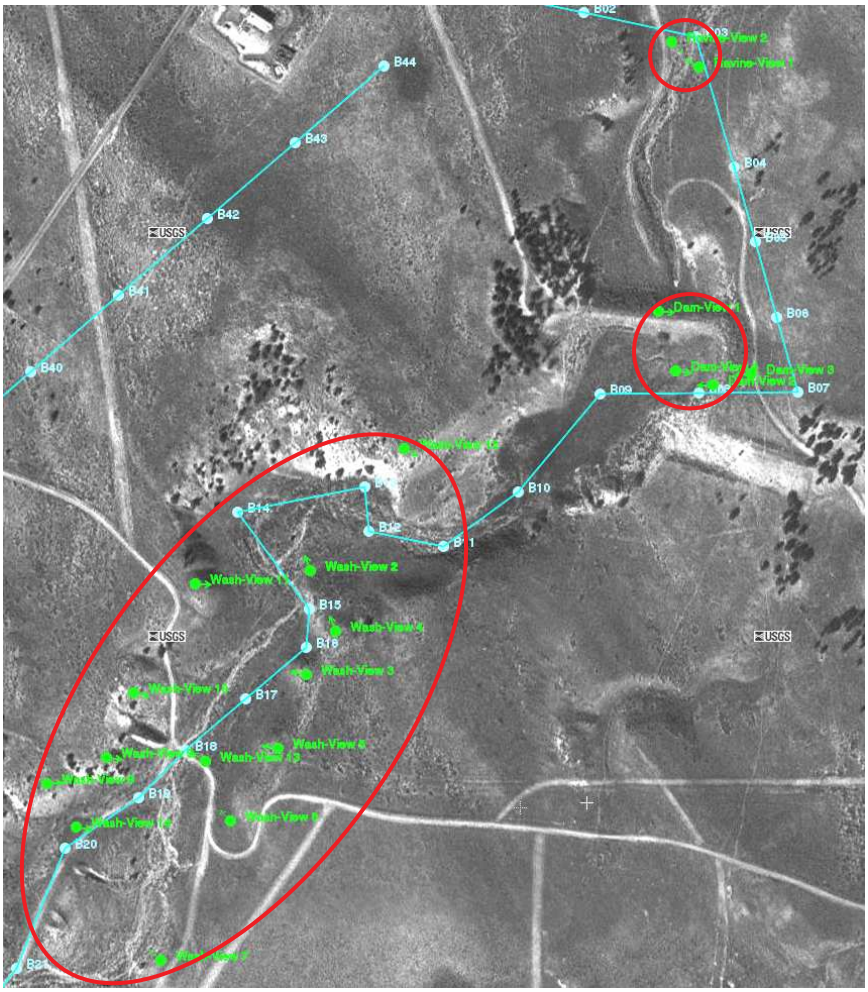


Figure 24 Regions of interest, Tooele UT: Dam (middle circle), Ravine (top circle) and Wash (bottom oval) regions. The green dots represent locations where high resolution scans were obtained.

The paths traversed by the robotic vehicles were then mapped onto the ground surface (e.g., Figure 13), and various measurements were computed. Of particular interest are terrain roughness and surface slopes. There are four methods available for measuring surface roughness. They are largely independent, so that consistency between them leads to greater confidence in the results.

The first method is to use the elevation data provided by the navigation system. For the path actually followed by the vehicle, the changes in elevation give a good measure of how bumpy the terrain is. A way of measuring the slope is to look at the pitch angle of the vehicle as it traversed the slope. The pitch can be obtained from the navigation sensor, and is shown in

Figure 25. The navigation data measures the Western slope of the dam as about 21º, and finds that the Eastern slope is about 13º. For more analysis of the navigation data, see Appendix B.

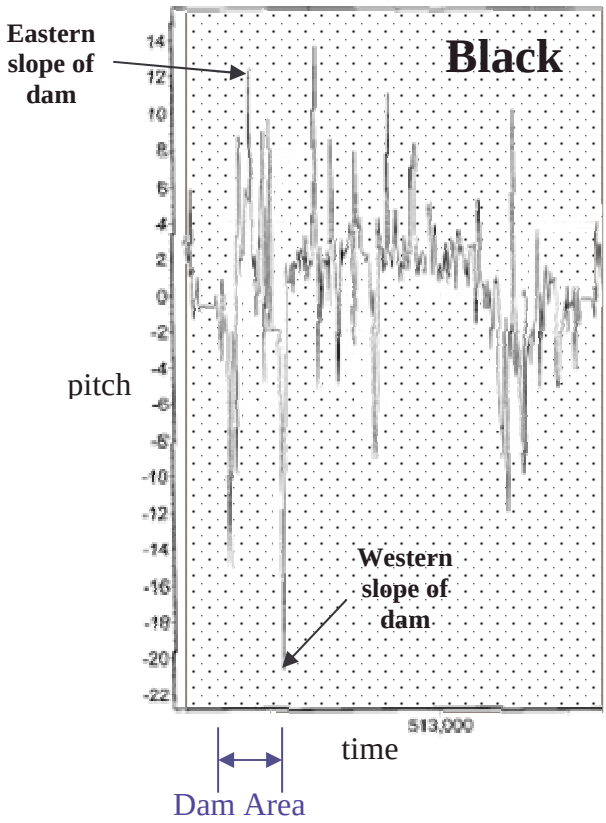


Figure 25 Pitch angle of the NIST HMMWV as it traversed the Black course in Tooele Army Depot.

A second measurement comes from the soil mechanics trailer (Figure 26) that also computes elevation changes as well as how soft the ground is and how much friction there is between the ground and the tires. This measure is also local to the path traversed, and should correlate well with the navigation system's measurement. Analysis of these measurements is still under way.

Figure 26 The NIST HMMWV coupled to the TARDEC trailer during a data collection run at Tooele Army Depot.

The third approach is to use the Riegl range sensor. While this suffers from the problem mentioned above of having to remove the vegetation to measure the support surface, it provides a roughness measure for a larger region compared to a single pathway. As described earlier in this paper, a method was developed in an attempt to gauge the roughness of the terrain—a roughness index—and this value provided a relative measure as a means of comparison. The average roughness index for Fort Indiantown Gap, gold course (n = number of runs = 50) is 0.1076 with a standard deviation of 0.0352 and the average roughness index for the black course (n = 90) is 0.1183 with a standard deviation of 0.0418. Statistical analyses show that there is no statistical difference between the average roughness indices for the black and gold courses.

The fourth approach also uses the data from the range sensor and is similar to the first method. However, instead of using navigation data for slopes, a ground surface, generated from the range data, is used to obtain elevations and slopes for a given region (Figure 27). By overlaying the paths traversed by the XUVs on individual missions, it is then possible to determine how rough each path is, and perhaps also provide a quantitative measure of how well the vehicle chose its path.

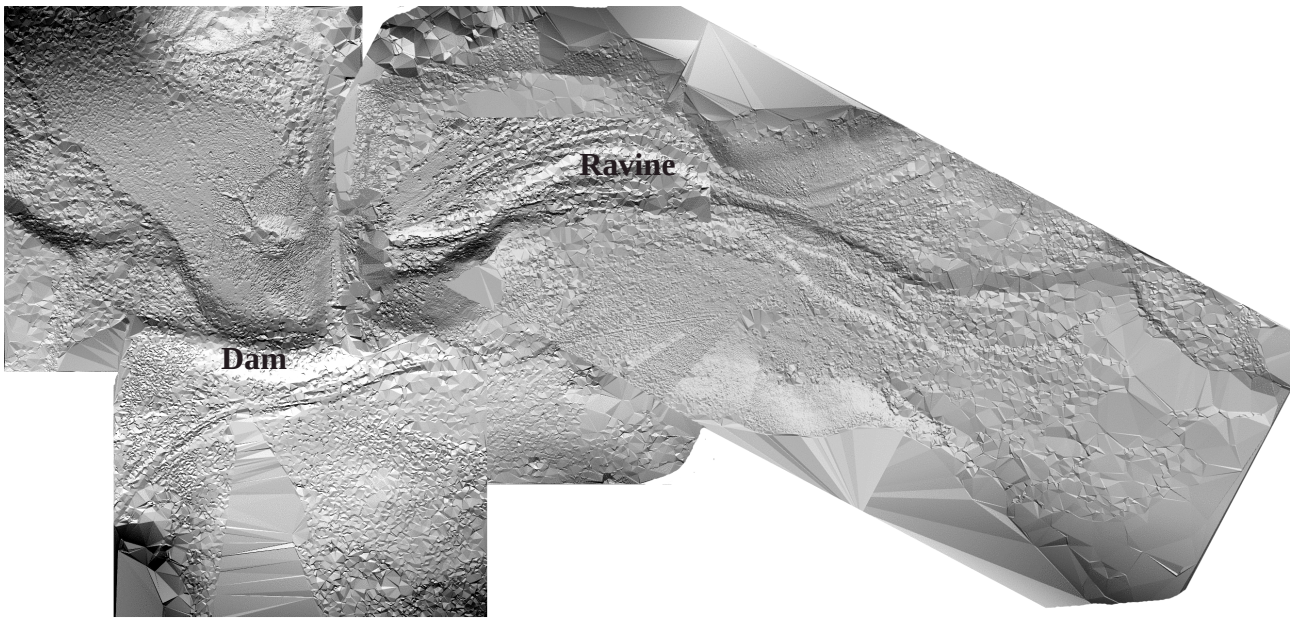


Figure 27 Surface of the Ravine and Dam region. Actual or simulated XUV paths may be overlaid on such a surface to determine the "roughness" of a selected path.

Figure 28 shows a portion of the Dam region and a path over the dam and the computed pitch and roll angles of a vehicle traversing this path. The calculated pitch along the path is about 13º on both the eastern and western slopes of the dam.

Direction of travel

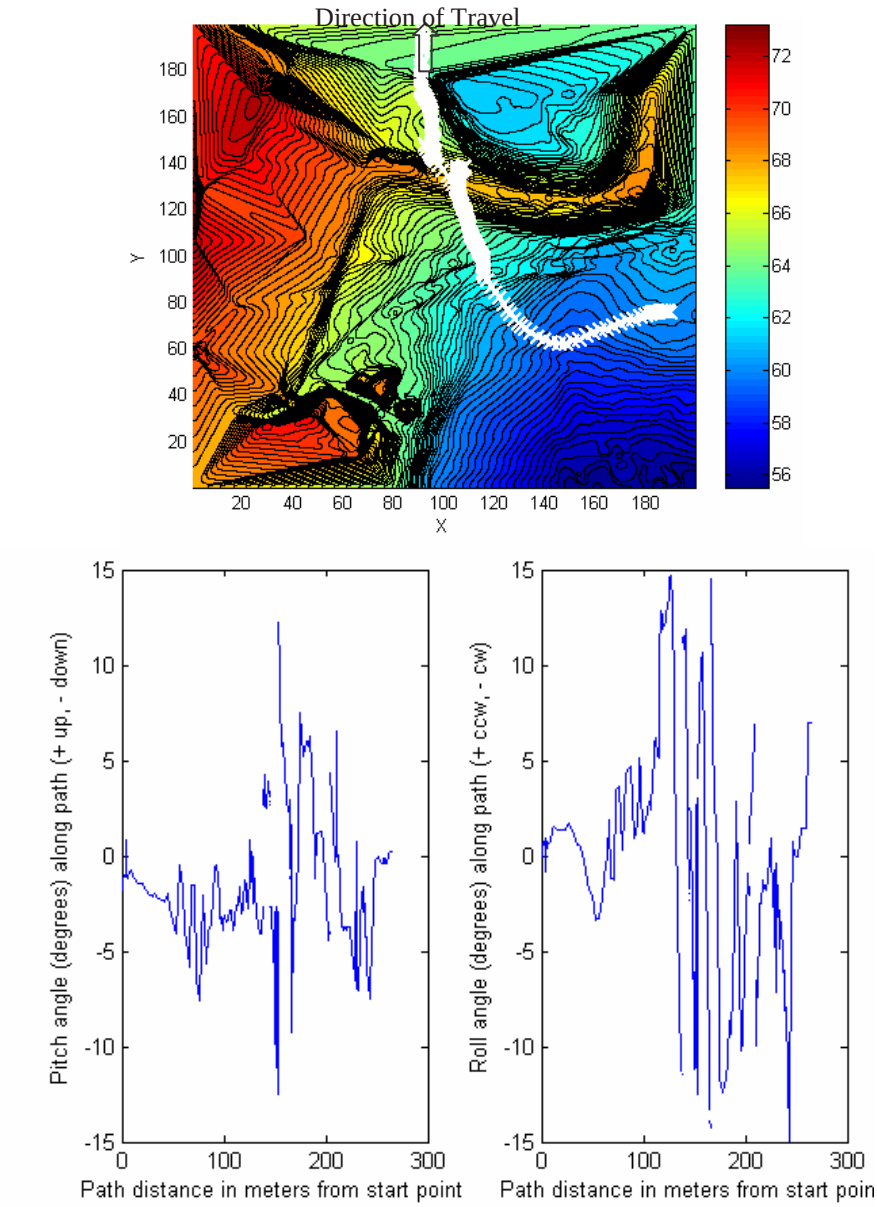


Figure 28 (Top) Surface of the Dam region with a vehicle path overlaid. The path begins on the lower right. (Bottom) The computed pitch and roll angle of a vehicle along the path.

Figure 13 shows a reconstruction from range data of the Dam area. The path of one of the XUV missions is shown in red. By fitting surfaces to the western and eastern slopes of the dam, the slopes were measured as 21º and 18º, respectively. A comparison of the eastern and western slope values derived for the various methods is given in Table 2.

Method	Eastern Slope	Western Slope
(a) Navigation system	13	21
(b) Compute pitch from surface along selected path	13	13
(c) Fit planes to dam surface	18	21

Table 2 Computed Slope Values for the Dam Using Various Methods.

The discrepancy between the slope values for the different methods could be due to the fact that the slopes were obtained at different locations. For example, the path traversed by the HMMWV (method a in Table 2) is likely different from the path chosen in Figure 28 (method b in Table 2).

DISCUSSION

Collecting and analyzing data for terrain characterization was a learning experience for the whole team. The methods were constantly evolving during the course of the project. This was reflected in the time taken to collect each set of data, the way the collected data were organized, and the amount of user interaction needed. By the end of the project, the time taken for a set of data was less than half what was needed at the beginning, and there was substantially less post-processing required to get the data into a useful format.

Experience showed that collecting data in stop-start mode required very small spacing between the scan locations – about 20 m rather than the initial large spacing of 50 m to 100 m. Because the Riegl ladar is mounted looking forward, the grazing angle on the ground puts most sample points in close proximity to the sensor, and causes the spacing of subsequent ground points to increase until the data are no longer useful. If there are vertical surfaces in the far field, however, they will be represented with good resolution out to 350 m.

Another lesson learned was that registering successive data sets with a common coordinate system was not straightforward even with good position information. It had been expected that successive scans would be easily transformed into a common reference frame because the Applanix navigation system would provide highly accurate position estimates. In practice, we found that a human could manually register the data better than an automated registration program using the navigation data. This may be the result of one of several factors or a combination of some or all the following factors: (*i*) errors in the transformations of the Riegl ladar to the HMMWV and world coordinate frames, (*ii*) ladar position shifted from initial alignment position – shift may have occurred due to vibration of vehicle or because the sensor had to be mounted and dismounted from the HMMWV many times in the course of the data collection, (*iii*) vehicle frame is off from initial alignment pose because of a difference in tire pressure or vehicle load, and (*iv*) drift in sensors.

As there is no standard definition as to what constitutes terrain roughness, a trial and error method had to be used to determine what approach worked or didn't work. As an example, one of the approaches to determining terrain roughness was to examine the triangles in the TIN. It was felt that a rougher terrain would have more triangles in the mesh because of the adaptive

insertion method used to create the TIN. Also, it was felt that larger triangles should indicate a flatter terrain. The slopes of individual triangles, or the change in slope from one triangle to another could be used as an indicator of terrain roughness – a higher variation in the slopes could indicate a rougher terrain. Another approach was to examine the slopes of all the 1 m bins in the 20 m x 20 m region. We had hoped that a trend may be evident when the results were analyzed; however, on examination of the number of triangles, triangle areas, triangle slopes, and bin slopes, no conclusive results could be made. Also, given that there was no "truth" or a reference surface for comparison, it was difficult to validate these approaches and to determine if the parameters examined could be used as indicators of terrain roughness.

Visually, the algorithm for bush and tree identification appears to work well. Inspection of the 140 point clouds with the tree and bush points overlaid showed there were some mis-identifications, but on the whole, the tree and bush locations identified by the algorithm and the visual locations of the bushes and trees showed very good correspondence (Figures 20 and 21). Again, lacking a reference or performance metric and more experience with the algorithm for other types of terrain, we are currently not able to quantify how well the algorithm works.

Very little of the post-processing has been automated. The topic of terrain characterization is still very much in the realms of research, and the procedures and algorithms are in the developmental stages and are constantly evolving and being refined. However, some of the procedures developed can be automated. It should also be noted that roughness and vegetation coverage are just two possible factors that contribute towards the characterization of terrain for mobility. Other factors include the physical vehicle dimensions, soil conditions/mechanics, and the incidences of steep slopes, ditches, and potential obstacles.

The current status of terrain analysis and classification is that small, but carefully chosen, parts of the data have been analyzed. Indications are that useful information can be extracted from the sensor data that will allow quantitative measurement of terrain difficulty.

Property	Specification
8 laser beams, 1 rotating mirror	With 8 facets
Scan resolution	32 lines × 180 pixels
Scan coverage	20 ° × 90°
Angular resolution	0.658° × 0.5°
Maximum frame rate	60 scans/s but 30 scans/s
Range	5 m to 70 m (vertical surface)
Range resolution/accuracy	±7.6 cm / 15cm
Data measurement rate	Range: 345,600 measurements/s.
Day/Night Operation	Range Independent of ambient light

Table 3 GDRS Area-Scan Ladar Specifications

POS LV 420-RT (Using DGPS)	GPS Outage Duration (minutes)					
	0 min	1 min	3 min	5 min	10 min	20 min
X, Y Position (m)	1.0	1.5	1.75	2.0	2.5	3.5
Z Vertical Position (m)	1.5-2.0	2.0	2.0	2.0	2.5	3.0
Roll & Pitch (deg)	0.02	0.02	0.02	0.02	0.02	0.02
True Heading (deg)	0.02	0.02	0.04	0.06	0.10	0.20

Table 4 Real-Time Performance of Applanix POS LV 420 Intertial Navigation Unit.

POS LV 420 (Post processed)	GPS Outage Duration (minutes)					
	0 min	1 min	3 min	5 min	10 min	20 min
X, Y Position (m)	0.02	0.12	0.40	0.75	1.5	2.5
Z Vertical Position (m)	0.03	0.15	0.50	0.65	1.0	2.0
Roll & Pitch (deg)	0.005	0.005	0.005	0.007	0.007	0.09
True Heading (deg)	0.02	0.02	0.02	0.03	0.035	0.035

Table 5 Performance of Applanix LV 420 with post-processing.

PROPERTY	SPECIFICATION
Scan coverage	80°x 330°
Angular stepwidth	0.072°to 0.36°
Angular readout accuracy	0.036°
Frame scan rate	1 ℅s to 15 ℅s
Minimum Range	2 m
Maximum Range	350 m (25 mm resolution, natural target)
Range resolution	25 mm or 50 mm, selectable
Standard uncertainty	Resolution + Distance error of ≤ ±20 ppm

Table 6 Manufacturer's Specifications - Riegl LMS Z210 Ladar.

Bibliography

[1] Shoemaker, C. M. And Bornstein, J. A. The Demo3 UGV Program: A Testbed For Autonomous Navigation Research. Proceedings Of The IEEE International Symposium On Intelligent Control. Sept.,1998. Gaithersburg, Md.

[2] Camden, R., Bodt, B., Schipani, S., Bornstein, J., Phelps, R., Runyon, T., French, F., And Shoemaker, C. Autonomous Mobility Technology Assessment Interim Report. Army Research Laboratory (ARL-MR-565). 2003.

[3] Scott, H. And Szabo, S. Evaluating The Performance Of A Vehicle Pose Measurement System. Proceedings Of The Performance Metrics For Intelligent Systems (PerMIS) Workshop. Aug.,2002. Gaithersburg, Md.

[4] Bouguet, J.-Y. Camera Calibration Toolbox For Matlab. Http://Www.Vision.Caltech.Edu/Bouguetj/Calib_Doc/ . 10-2-2002.

[5] Shneier, M. O, Chang, T., Hong, T., And Cheok, G. Scott H. A Repository Of Sensor Data For Autonomous Driving Research. SPIE Aerosense Symposium. 2003.

[6] Hong, T., Chang, T., Takeuchi, A., Cheok, G., Scott, H., And Shneier, M. Performance Evaluation Of Sensors On Mobile Vehicles Using A Large Data Repository And Ground Truth. Proceedings PerMIS 2003. 2003.

[7] Witzgall, C., Cheok, G., And Gilsinn, D. E. Terrain Characterization From Ground-Based Ladar. 2003.

[8] Downs, A., Madhavan, R., And Hong, T. Registration Of Range Data From Unmanned Aerial And Ground Vehicles. Proceedings Of The Applied Imagery Pattern Recognition Workshop. 2003.

APPENDIX A. Papers Resulting From This Research in FY2003

1. W. F. Oberle and G. A. Haas, "Three-Dimensional Stereo Reconstruction and Sensor Registration With Application to the Development of a Multi-Sensor Database", ARL-TR-2878, December 2002.

2. Aya Takeuchi, Michael Shneier, Tsai Hong, Tommy Chang, Chris Scrapper, Gerry Cheok, "Ground Truth and Benchmarks for Performance Evaluation", SPIE Aerosense Symposium, Orlando, FL, April, 2003.

3. Michael Shneier, Tommy Chang, Tsai Hong, Gerry Cheok, Harry Scott, Steve Legowik, Alan Lytle, "A Repository Of Sensor Data For Autonomous Driving Research", SPIE Aerosense Symposium, Orlando, FL, April, 2003.

4. C. Scrapper, A. Takeuchi, T. Chang, M. Shneier, "Using A Priori Data for Prediction and Object Recognition in an Autonomous Mobile Vehicle", SPIE Aerosense Symposium, Orlando, FL, April, 2003.

5. R. Madhavan and E. Messina, "Iterative Registration of 3D LADAR Data for Autonomous Navigation," in Proceedings of the IEEE intelligent Vehicles Symposium, June 2003.

6. C. Witzgall, G. S. Cheok, and D. E. Gilsinn, "Terrain Characterization From Ground Based Ladars", Proceedings of PerMIS03, Sept. 2003.

7. A. Downs, R. Madhavan, and T. Hong, "Registration of Range Data from Unmanned Aerial and Ground Vehicles", in Proceedings of the Applied Imagery Pattern Workshop, Oct. 2003.

8. P. Conrad and Mike Foedisch, "Performance Evaluation of Color Based Road Detection Using Neural Nets and Support Vector Machines", in Proceedings of the Applied Imagery Pattern Workshop, Oct. 2003.

9. G. A. Haas and W. F. Oberle, "Juxtaposition of Inertial Navigation Sensor and Camera Egomotion Estimates of Ground Vehicle Trajectory: Results and Implementation Details", ARL-Technical Report.

APPENDIX B Position and Orientation System of NIST HMMWV

System Description

The NIST HMMWV is equipped with a commercial, high precision position and orientation system that integrates inertial measurements from accelerometers and gyros with Global Positioning System (GPS) and a wheel encoder aiding data to produce a full navigation solution. The system produces a complete solution in real time at rates up to 200 Hz and can log required data at that rate as well to permit post-processing for a more precise navigation solution.

The GPS information may itself be raw GPS data or can be differentially corrected in real-time (RTK) for better real-time accuracy. For the Tooele Army Depot (TAD) and Fort Indiantown Gap (FTIG) TRL6 data collections, a GPS base station was put in place to enhance the real-time navigation solution quality. Because of shortcomings discovered in the existing GPS base station hardware, the best navigation solutions were provided through post processing. This is discussed below. New GPS base station hardware has since been acquired to address weaknesses found in the existing, nearly obsolete, base station equipment.

Figure 29 The NIST-erected GPS base station placed on a local berm at TOOELE ARMY DEPOT and powered by generators on the NIST support van.

Post processing of the GPS data was accomplished using RINEX (Receiver Independent Exchange Format**)** data files from CORS (Continuously Operating Reference Station) differential correction base stations. CORS stations are located close enough to both Tooele Army Depot (Midvale CORS data) and Fort Indiantown Gap (York CORS data) to make this approach acceptable, providing post-processed position errors in the few centimeter range. Post-processed navigation data for the Tooele Army Depot runs was developed and post processing of Fort

Indiantown Gap is performed as needed. Note that the ability of the fielded base station to improve the real-time solution is not related to the quality of the post processed solution, since post-processing employs raw data from the vehicle system and a public reference base station.

Data

A number of variables can affect the quality of the final navigation solution, including such things as satellite geometry and tree cover, which of course are changing continuously. Thus, it is important to have knowledge about the quality of the solution at any point in time as well as the solution itself. The position and orientation system used provides estimated RMS error values for each of the components of the navigation solution for all points in time. Earlier testing at NIST has shown that these RMS error values are accurate, though to date only position errors were evaluated. The quality of the post processed navigation data solutions ranged across data sets from a few centimeters to meter level estimated errors.

Examples of the data collection capability and descriptions of some of the data follow later under *Tooele Army Depot Navigation Data Description.*

The data collected has initially been used for two purposes. One use has been to provide position and orientation information for the vehicle which, with the appropriate transformations, can be used to register ladar images to each other. Such navigation data has been used to piece together scans of the course at 150 m intervals into a “movie” of the courses. This is an example of how to ultimately build world models of the environment from sensors, with resolutions that should be available for real-time use in the future. In addition, the ladar data itself has been studied to determine methods to evaluate terrain characteristics.

Secondly, the position and orientation data from the Tooele Army Depot Gold, Black, and practice area runs has been provided to support analysis of the TARDEC Terrain Characterization Trailer data collections. That trailer, with its own vertical accelerometers and hitch load cell, collected additional data while simulating the weight of the XUV over the Tooele Army Depot courses. The information below, along with a CD of the data, was provided to TARDEC. This data was aligned by TARDEC to the Terrain Characterization Trailer data by using the start and end times for each run along with information on vehicle speed shown plotted below.

Navigation Information for TARDEC Trailer Runs at Tooele

Shown below are plots and Start/End times based on the navigation data collected from the NIST HMMWV navigation system and post-processed. Data is for the run performed with the TARDEC trailer on 12/12/02 at Tooele Army Depot. Times are GPS time.

****
The entire run, including traverses between courses and to/from the base station covers the following period:

Thursday Tardec Data

Data/Time:

12/12/02
GPS week 172

401800.0082	to	411641.0121	(GPS seconds)
15:36:40.008	to	18:20:41.012	(UTC)
8:36:40.008	to	11:20:41.012	(Mountain Standard)

Description:

Nav data corresponding to runs pulling the TARDEC Terrain characterization trailer.

Clockwise traverse of Black
then CCW traverse of Black
then CW traverse of Gold.

Differential postprocessing of the GPS data was performed using files from the Midvale, UT CORS reference station.
Integrated Inertial Navigator and Smoother were used to post-process all data.

****

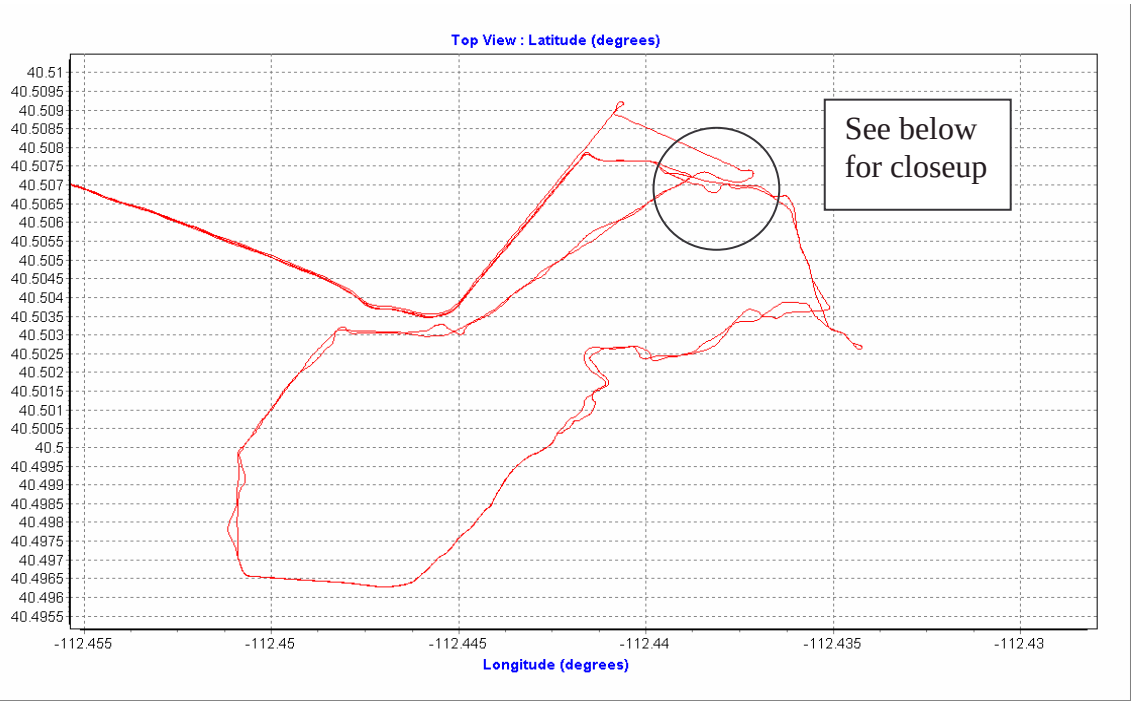


Figure 30 Plot of black course at Tooele.

Below, Start and End times are shown for each course loop. Start and End shown were selected from the data because they appear to be locations near the beginning and end of course loops where the data indicate the vehicle was stopped.

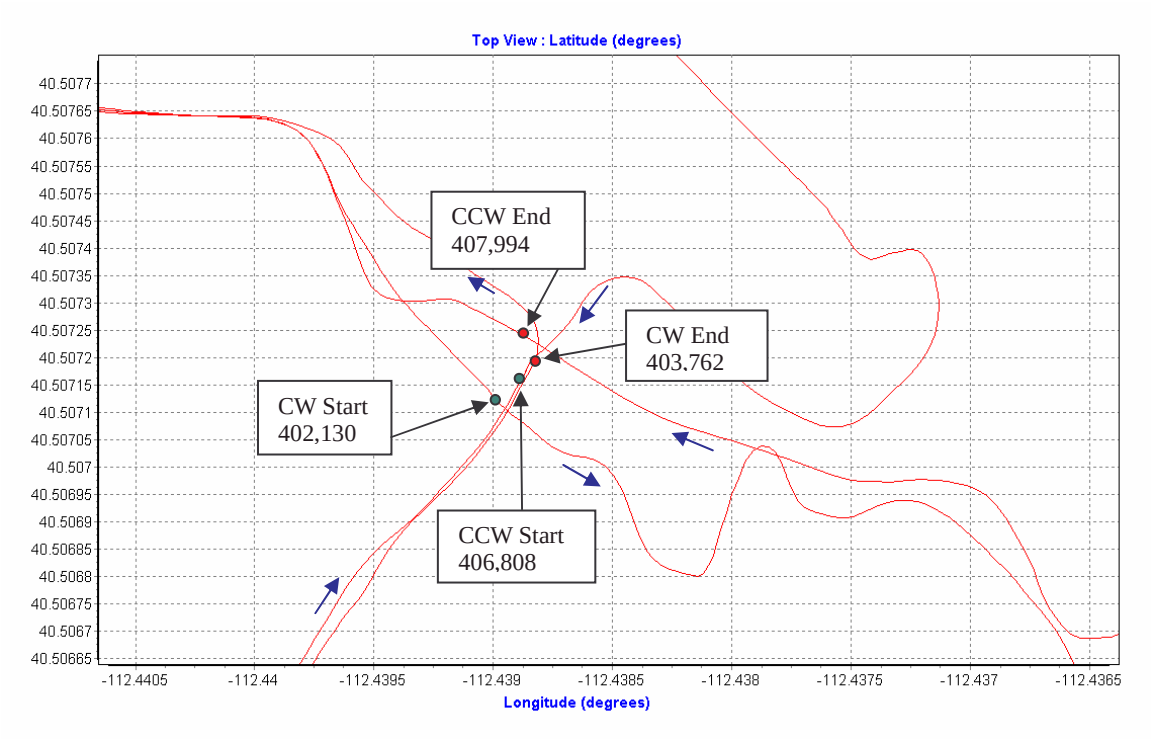


Figure 31 Close-up of Black Start/End points (area circled in Figure 30).

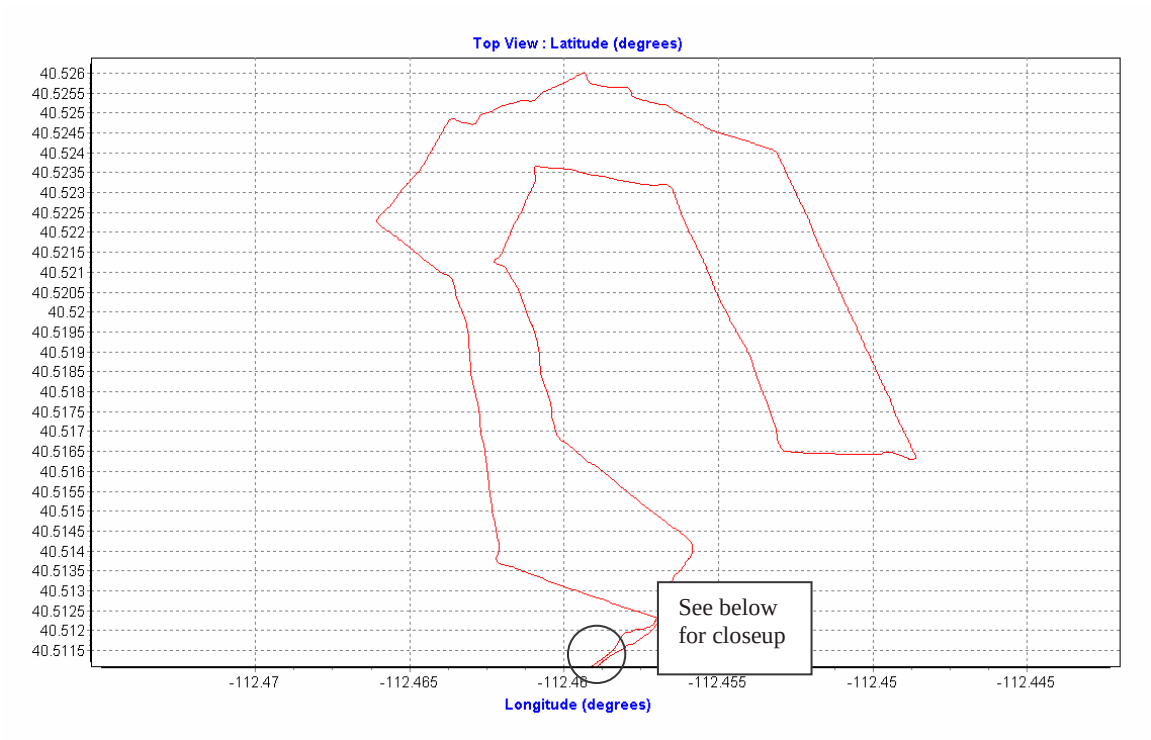


Figure 32 Plot of Gold Course at Tooele.

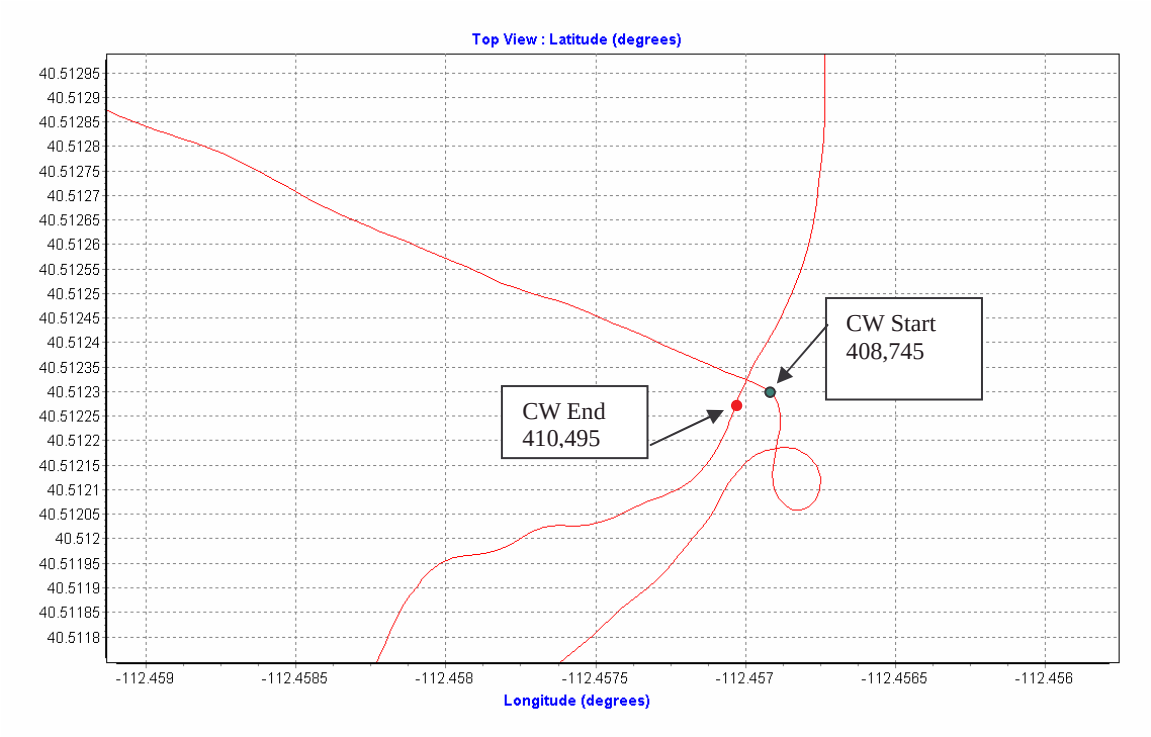


Figure 33 Close-up of start and end point on gold course (area circled in Figure 32)

Speed plot (entire data run):

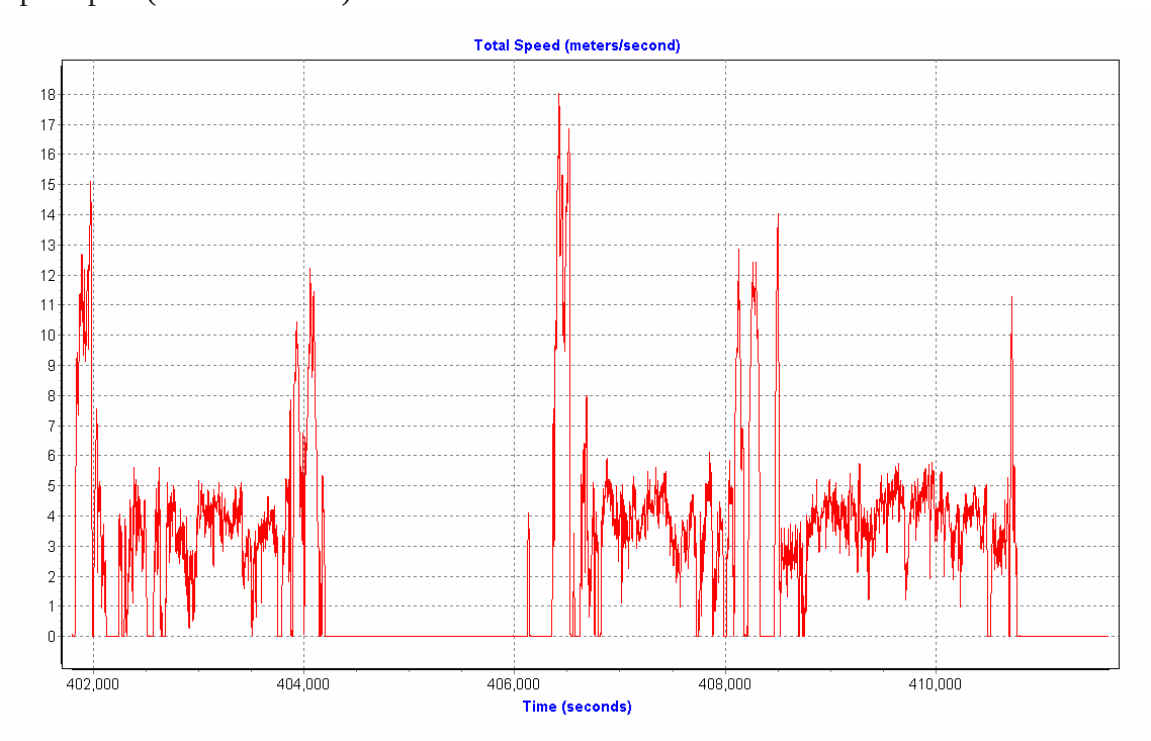


Figure 34 Plot of vehicle speed over both courses. Black is on the left, gold on the right).

Tooele Army Depot Navigation Data Description

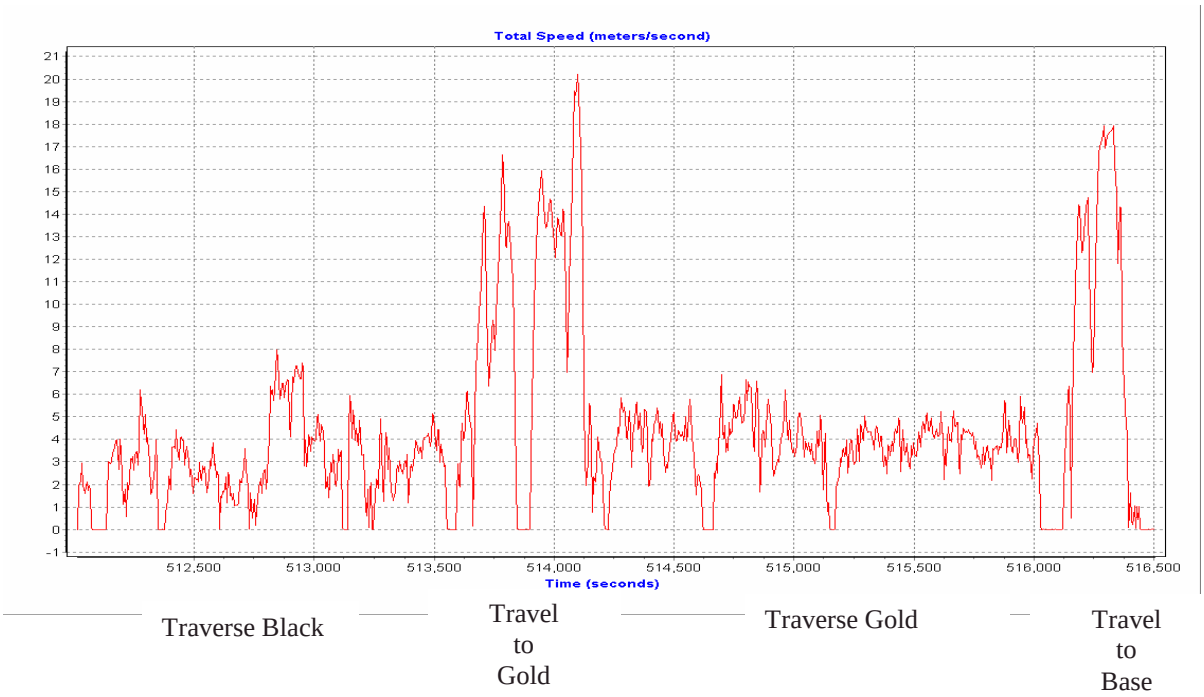


Figure 35 Vehicle speed on the black and gold courses.

The speed data shown in Figure 35 above, as with all the collected data, is available on any 5 ms interval (200 Hz rate). Note the larger changes in speed, but lower average speeds, were attained on the Black course. This is consistent with the nature of the differences in the TAD Black and Gold course terrain characteristics, with Gold being bumpier, but more open.

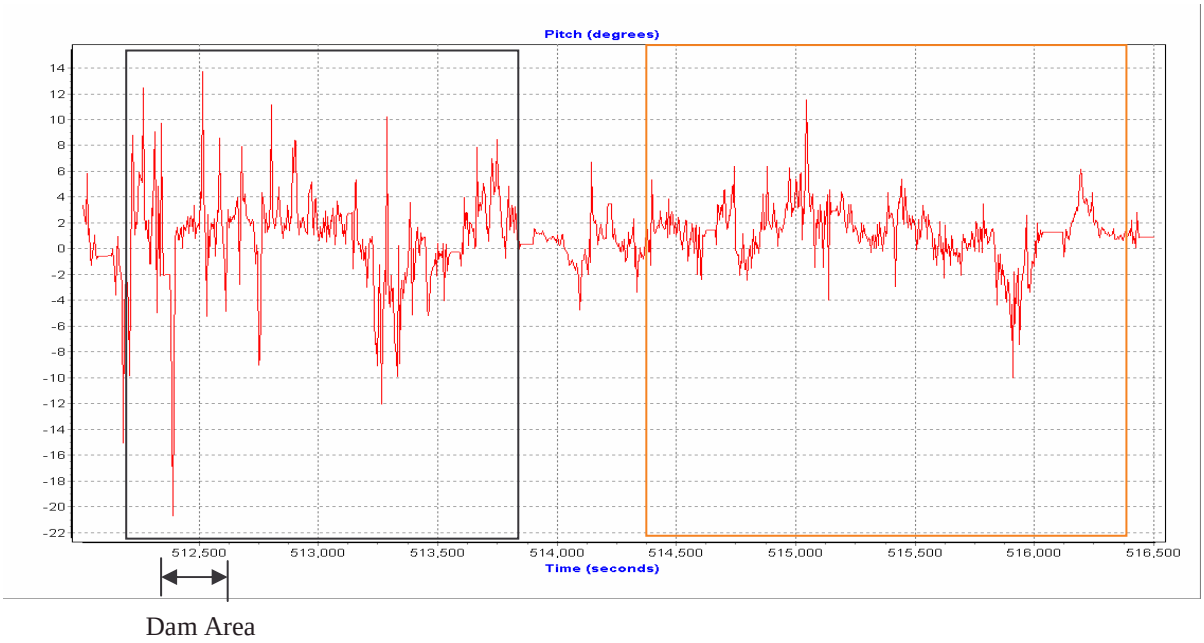


Figure 36 Pitch data on black and gold courses

The pitch data shown in Figure 36 above clearly expresses the hilliness of Black course vs. the Gold course. Note that the maximum pitch value of about 20° corresponds well to the terrain slope of the Dam derived from the collected ladar data.

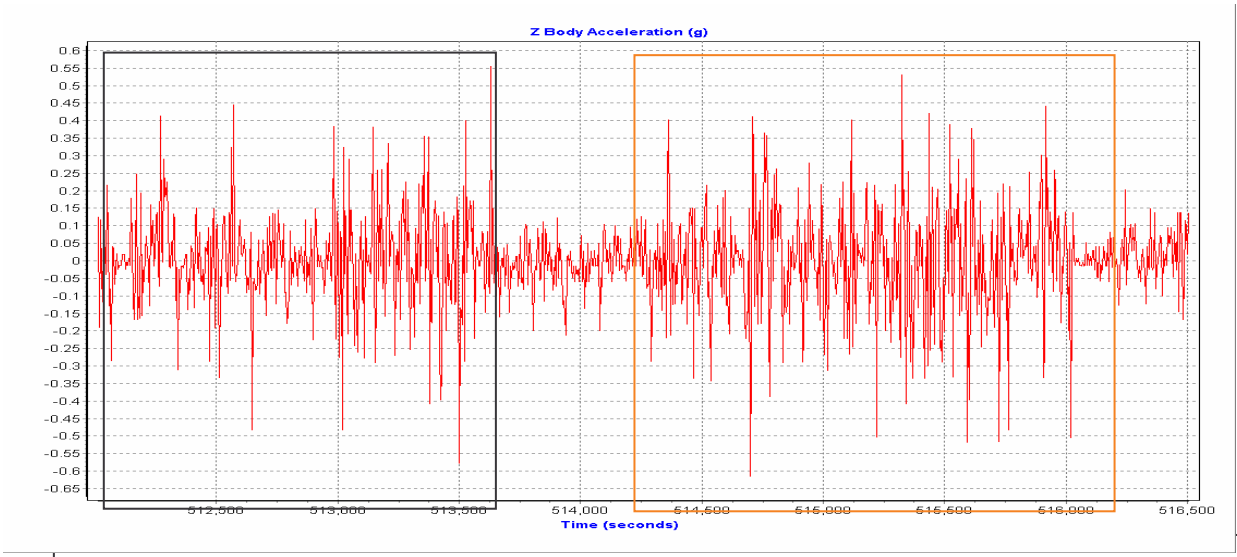


Figure 37 Vertical acceleration on the two courses

The z (vertical) acceleration data shown in Figure 37 is consistent with both the roughness and higher speeds found on the Gold TAD course.

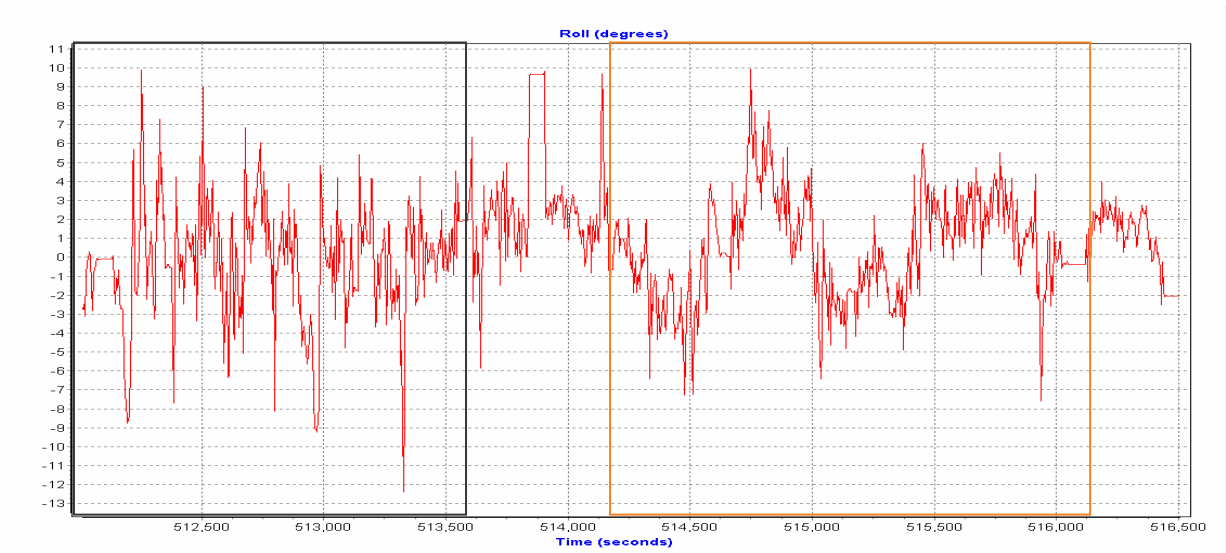


Figure 38 Roll data for the black and gold courses.

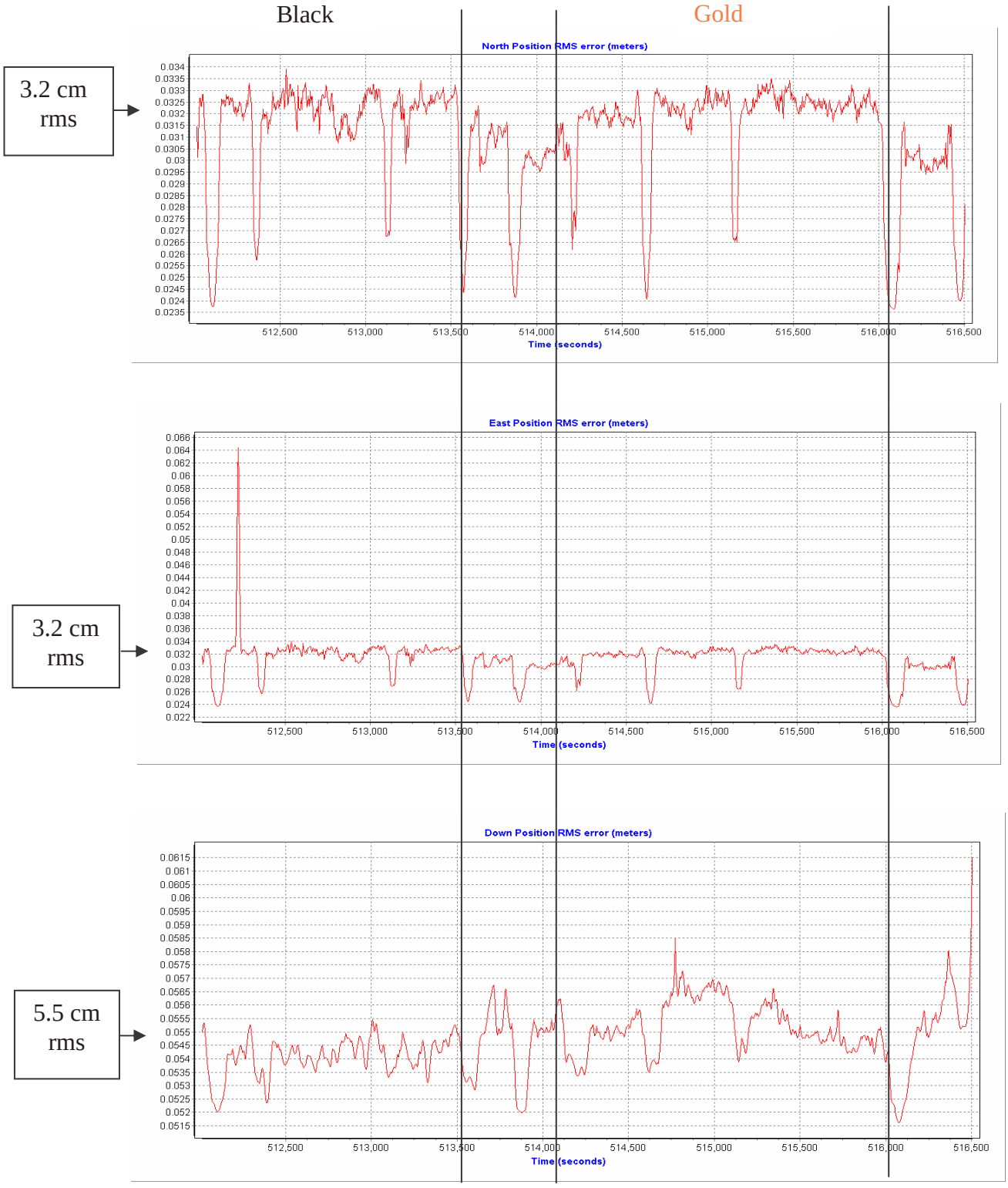
The roll data, shown above in Figure 38, also attests to the more challenging terrain of the TAD Black course.

Data Quality

As mentioned earlier, it is important to understand the uncertainty in the data being used. The plots that follow summarize the uncertainty in several of the important position and orientation measurements.

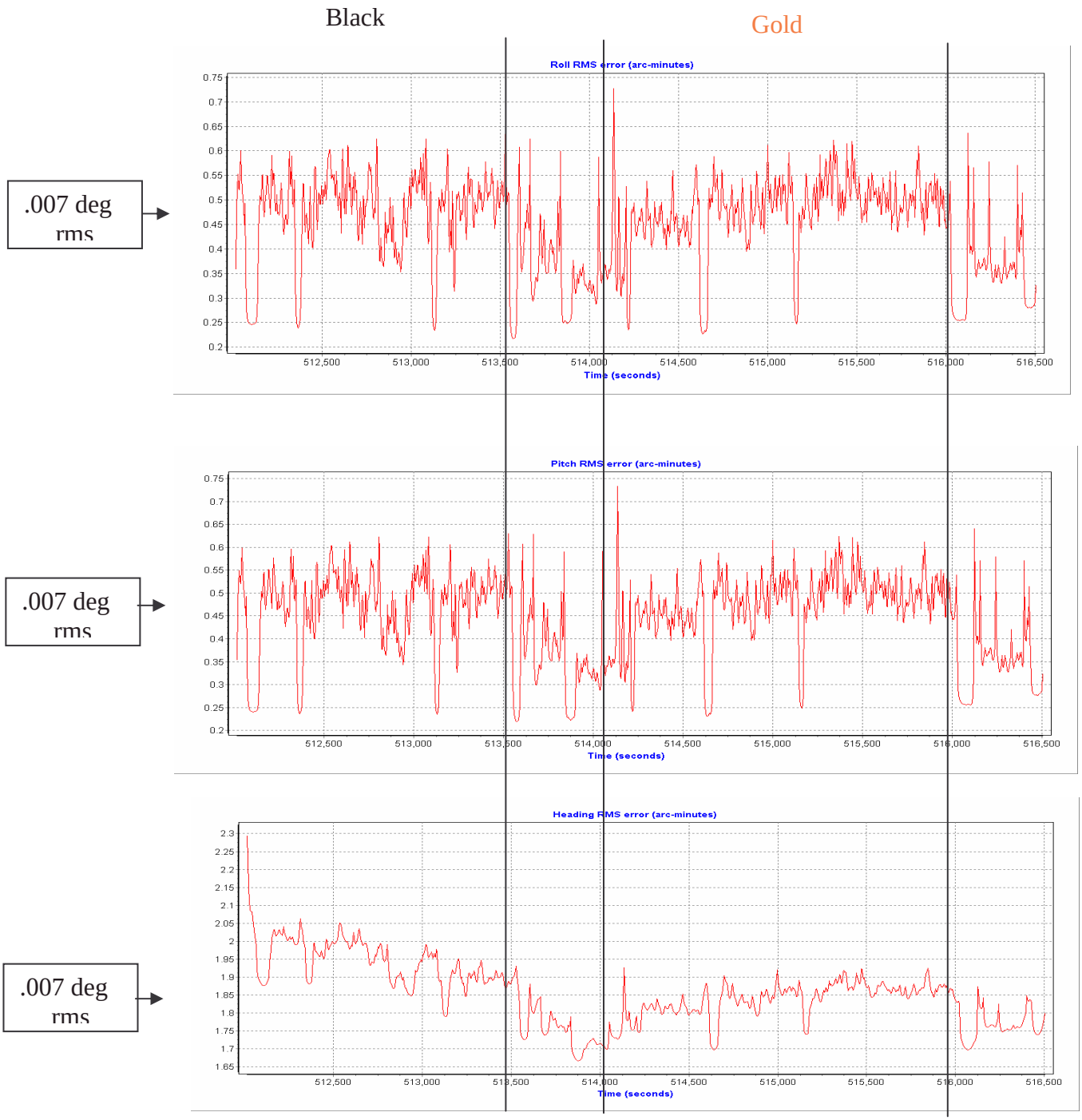
The first three plots below indicate uncertainty in the position solution data. Note the rms uncertainty is about 3.2 cm in Northing and Easting and 5.5 cm Vertical for the post-processed data.

Measurement performance metrics



The next three plots characterize errors in the orientation data. The plots, in units of arc-minutes, correspond to rms errors in units of degrees of about:

Roll: .007 degrees
Pitch: .007 degrees
Heading (yaw): .027 degrees.



Other available data includes:

- x y z Accelerations
- N E D (Northing Easting Down) Velocities
- Heading
- x y z Angular Rates
- Distance vs. time
- N E D Velocity errors
- Expected rms errors for all above
- Low level sensor data, such as IMU data
- Raw and Processed GPS data
- Kalman Filter Errors and Residuals
- The real-time (non post-processed) solution and expected errors.