2006

Experimental Analysis of Overhead Data Processing To Support Long Range Navigation

David Silver  
*Carnegie Mellon University*

Boris Sofman  
*Carnegie Mellon University*

Nicolas Vandapel  
*Carnegie Mellon University*

J. Andrew Bagnell  
*Carnegie Mellon University*

Anthony Stentz  
*Carnegie Mellon University*

Follow this and additional works at: [http://repository.cmu.edu/robotics](http://repository.cmu.edu/robotics)  
Part of the [Robotics Commons](http://repository.cmu.edu/robotics)
Abstract—Long range navigation by unmanned ground vehicles continues to challenge the robotics community. Efficient navigation requires not only intelligent on-board perception and planning systems, but also the effective use of prior knowledge of the vehicle’s environment. This paper describes a system for supporting unmanned ground vehicle navigation through the use of heterogeneous overhead data. Semantic information is obtained through supervised classification, and vehicle mobility is predicted from available geometric data. This approach is demonstrated and validated through over 50 kilometers of autonomous traversal through complex natural environments.

I. INTRODUCTION

Due to the recent popularity of the DARPA Grand Challenge, autonomous ground vehicle navigation has received increasing attention. While on-road long range navigation has been successfully demonstrated [1], off-road navigation over long distances (multiple kilometers) without human intervention remains a challenge. This is due to the highly varying nature of off-road terrain, along with the complex surface geometry and vegetation cover that can be encountered.

The goal of long range navigation is to achieve safe and efficient traversal of an environment between two given locations. Without any prior information about the environment, safe navigation can be achieved by the vehicle’s local perception system, but path efficiency will greatly suffer. Given sufficient prior knowledge of an environment, both safety and efficiency can be achieved. Prior data can come from a variety of sources. Low resolution imagery (1 meter resolution grayscale) and topographic data (10 to 30 meter resolution) are already available for most of the world. In addition, higher resolution imagery and dense three-dimensional (3-D) data can be collected commercially upon request.

Using such data sources introduces many challenges. A system must be devised for extracting meaningful information from such data, managing and merging data from various sources, and using this data to improve autonomous navigation. This paper deals with the general problem of using heterogeneous overhead data sources (in time and resolution) to support ground vehicle navigation.

We present an approach to produce (off-line, off-board) traversal cost maps from overhead data that can be used (on-line, on-board) for global path planning for an unmanned ground vehicle. Traversal costs are computed from a combination of semantic and geometric data. Semantic information of the terrain is obtained through supervised classification using features extracted from imagery and 3-D data. Mobility analysis is performed using the ground surface recovered from 3-D data or an available elevation map. Traversal costs are assigned to reflect the capabilities of the vehicle.

This paper claims two major contributions: 1) the devise of a consistent approach to support long range vehicle navigation using heterogeneous overhead data; 2) the experimental demonstration of the validity of the same approach by integration on an autonomous vehicle, where vehicle performance was evaluated systematically by an independent observer over an accumulated distance of more than 50 kilometers of natural terrain.

The next section presents previous work in long range navigation. Section III contains an overview of our approach. Section IV details the semantic interpretation of the terrain, and the methods used for mobility analysis. Finally, Section V presents results evaluating the performance of this approach from field tests after vehicle integration.

II. PREVIOUS WORK

A. Long Range Navigation Scenarios

The challenge of long range navigation for unmanned ground vehicles exists under many scenarios, with many factors that condition them. Examples are waypoint spacing, level of prior environmental knowledge, mission context and mission constraints [2]. One of the most important factors is the vehicle’s mode of operation, which can be classified into one of three broad categories:

- Path Tracking: the vehicle follows a fixed path through the environment. If the vehicle has on-board sensing capabilities, it may be allowed to deviate from the path if it senses obstacles [3]. Provided paths may be human or machine produced, and are usually based on some form of prior knowledge of the environment.
- Full Exploration: the vehicle has no prior knowledge of its environment, and must use on-board sensing to find its way from waypoint to waypoint [4][5].
- Aided Exploration: prior knowledge is available, but (unlike during Path Tracking) the vehicle is not directed down a single, pre-planned path. Rather, the vehicle uses the combination of prior data and its own perception system to navigate between waypoints [6][7][8].

It is often difficult to clearly define the boundaries between these categories. Factors such as the density of waypoints...
and the quality of available prior data can heavily impact the behavior of the vehicle.

B. Global Planning

Global planning involves finding the path between waypoints that best achieves mission priorities based on the current knowledge of the environment. For example, it may be desirable to choose the shortest traversable path, the lowest risk path, or the path that optimizes some other metric (for example, energy acquisition and usage [9]).

Under the mission scenario where the priority is vehicle safety, a common approach is to generate a cost map of the environment. A cost map is a 2-D grid representation of the environment, where the value of each cell encodes the mobility risk inherent in traversing through that cell. The lowest traversal cost path through the grid is therefore the (perceived) safest path to the goal. Any grid search algorithm can be used in conjunction with such cost maps to plan paths.

In a path tracking scenario, global planning occurs once, before a mission even begins. In a full exploration scenario, global planning uses the perception history of the vehicle to continually replan to its goal. In an aided exploration scenario, the combination of prior data (possibly from overhead data or a previous traverse of the area) and perception history is used for continual replanning.

C. Prior Data

In an aided exploration scenario, vehicle behavior is strongly affected by the available quality of prior data. Even low resolution prior data can provide significant improvement to vehicle performance, as demonstrated by numerous simulations in [10]. In [6], low resolution (25 – 30 meter) elevation maps were used to aid long distance planning. [11] demonstrated the extraction of features (roads, trees, water, etc.) from aerial surveys, for later correlation by an autonomous vehicle.

The DARPA PerceptOR program contained an important prior data component. Aerial LiDAR data was used to predict vehicle roll and pitch over stretches of terrain, as well as to detect vegetation [7][8]. This information was used to generate prior cost maps for use in global planning.

In general, prior cost maps can be human or machine produced. However, generating cost maps manually from high resolution data over large areas is infeasible. Automated processing is therefore essential for the effective use of prior data in long range navigation.

III. OVERHEAD SYSTEM OVERVIEW

Overhead data sources fall into 3 broad categories:

- **Imagery** can be gathered by either aerial flyover or satellite and consists of gray-scale, color, or multispectral information.
- **Rasterized Elevation Data** is available at a wide range of resolutions and from multiple sources
- **3D Point Cloud** data is usually produced by aerial LiDAR flyovers

There are many inherent challenges in using such a broad set of data sources. Overhead data sets are often heterogeneous in resolution, sampling pose, and sampling time, and may not be georeferenced with sufficient accuracy. Furthermore, different data sets may cover different areas. Therefore, the first step in using such data sources is to align and sample them.

The full data flow of the overhead system is shown in Figure 1. After alignment, the data is preprocessed by sampling into a raster grid. In the case of 3-D point cloud data, the load bearing ground surface is extracted. Next, terrain-based features are generated for use in semantic classification. These features are fed to a supervised classifier (specifically a neural network) to map each raster cell into one of several classes (see Section IV-A).

Vehicle-based features are also computed by convolving a vehicle model with a Digital Elevation Map (DEM) or recovered ground surface. For each raster cell, the state of the vehicle (roll, pitch, ground clearance, etc.) is predicted (see Section IV-B).

Finally, the results of semantic classification and mobility analysis are interpreted to generate traversal costs to form a cost map. The rules for computing costs are vehicle-specific. Additional useful parameters, such as maximum safe traversal speeds for each cell, can be computed to further improve navigational performance.

IV. OVERHEAD SYSTEM INSTANTIATION

A. Terrain Classification

1) Related Work: Similar research conducted by Charaniya et al. [12] classified terrain into roads, grass, buildings, and trees using aerial LiDAR height data, height texture, and signal reflectance, achieving classification rates in the range of 66% – 84%. Cao et al. attempted to identify man-made objects from aerial imagery [13]. Knudsen and Nielson attempted to classify buildings using a previously available GIS database and RGB information for an environment [14].

2) Feature Extraction: The set of features to be extracted depends on the particular data sets that have been provided for the environment. Multiple data sets of the same type (for instance, multiple sets of color imagery) can also be used.

- **Image Based Features:** The RGB color space is known to be sensitive to variation in illumination, while the hue, saturation, value (HSV) space is not. Therefore, HSV is used instead of RGB. To avoid discontinuities in hue space, the sine and cosine of hue, along with saturation and value, are used as features.

When Near-Infrared (NIR) imagery is available, NIR itself is used as a feature, along with the Normalized Difference Vegetation Index (NDVI) [15]. Because chlorophyll absorbs red light and reflects a larger proportion in the near-infrared channel, NDVI can be effective in detecting vegetation.

- **Elevation Based Features:** Elevation related features are computed relative to local ground levels rather than from absolute elevations in order to deal with slope. We use the minimum, maximum and average elevation for each raster cell

- **Rasterized Elevation Data:** The set of features to be extracted

- **3D Point Cloud Data:** The full data flow of the overhead system is shown
as features, along with the standard deviation of those features over a small support region.

**Point Cloud Based Features:** 3-D point clouds are first divided into ground and non-ground points using a procedure presented in Section IV-B. The ratio of ground points over the total number of points for a local support region is used as a feature to describe the density or permeability of the occluded space over the ground surface. Additional features are extracted through principal component analysis to describe the local distribution of a 3-D point cloud around a point of interest [16]. PCA based features are computed separately for ground and non-ground points.

3) **Classification:** We use a neural network with one input node for each feature, one hidden layer, and one output node for each desired classification category. The network is trained on each class by a set of labeled regions selected by a human operator. Each cell within a labeled region is used as a training example for that class. Each training example is given a desired output of 0.9 for the output node corresponding to that example’s class and a desired output of 0.1 for all other output nodes. The network is trained using back-propagation with a learning rate of 0.1 until the accuracy on a separate validation set stabilizes. The entire map is then classified, using the highest output node score to determine each cell’s class. Figure 2 shows an example of terrain classification following this approach. A more detailed description of the classification system can be found in [17].

### B. Vehicle Mobility Analysis

1) **Digital Elevation Map Processing:** Due to the rasterized nature of DEMs, it is not meaningful to predict the exact state of the vehicle at every position in the environment. Instead, mobility analysis with DEMs is reduced to the computation of an estimate of the ground slope. For each cell in the DEM, a plane is fit to all cells in a surrounding neighborhood. The size of the neighborhood depends on the size of the vehicle and the resolution of the DEM.

Slope analysis of DEMs operates under the assumption that the elevation value in each cell represents the elevation of the load bearing ground surface. However, this is not always the case as DEMs will often include the tops of tree canopy or other dense vegetation that does not get penetrated by the sensor. Our terrain classification techniques (Section IV-A) can be used to detect and eliminate many such occurrences.

2) **LiDAR Processing:** The vehicle’s state can be more effectively predicted when true 3-D point clouds are available from high resolution LiDAR.

**Ground Surface Extraction:** Before the vehicle state can be predicted from a LiDAR point cloud, individual points must be segmented into those from the supporting ground surface and those from vegetation, rocks, or other objects resting on the ground. Making this determination on a point by point basis allows better modeling not only of the ground surface, but of any non-ground objects that may come in contact with the vehicle.

Our approach is inspired by [18] and previously demon-
strated in [8]. It is based on the observation that the volume below a ground point should be free of any other LiDAR returns. For each LiDAR point, the number of other points that fall within a downward oriented cone centered at the point of interest is computed. If the cone is empty, then the point of interest is assumed to belong to the ground surface. Otherwise, it is assumed to be a non-ground point (see Figure 3). This approach is far more robust than simply assuming the minimum elevation within a window to be that of the ground surface (see Figure 4).

**Vehicle Model Convolution:** Once ground surface points have been extracted, mobility analysis can be performed. Potential positions of the vehicle control point and vehicle orientations are used in conjunction with known vehicle dimensions to determine the location of each of the vehicle’s tires. The mean elevation of ground points in the cylinder centered around each tire location is used to determine tire elevations. A plane is then fit to the tire locations to predict the roll and pitch of the vehicle. Changes in roll and pitch as the vehicle moves can also be computed. These deltas are useful for detecting both positive and negative obstacles.

Finally, for each pose, the set of points that exists within the vehicle’s footprint is extracted and the distance from the tire plane is computed for each point. This distance is used to predict the ground clearance of the vehicle and to identify high-centering hazards. In the case of non-ground points, the density of such points at various heights (vehicle undercarriage, vehicle body, vehicle sensor pod, etc.) can be computed to better assess the traversability of certain areas.

**C. Producing Cost Maps**

Traversal costs are independently computed from the results of terrain classification and vehicle mobility analysis for each raster cell, and then summed to produce a final traversal cost. These costs are interpreted as relative measures of mobility risk; if the cost of one terrain type is 3 times as high as the cost of another terrain type, then the vehicle would be willing to risk traversing 3 times as much of the lower cost terrain.

Costs based on semantic classification are straightforward. Each terrain class is assigned a traversal cost by a human operator. The costs are chosen to mimic the behavior of the on-board perception system of the target vehicle on terrain of that class. Each raster cell is assigned the cost given to its category.

Traversal costs from mobility analysis are assigned more formally. A set of functions map parameters computed for each raster cell, such as roll, pitch, ground clearance, etc., to individual costs that are then summed. The specific functions used depend on the map parameter and the vehicle in question; when possible, the functions are taken directly from the vehicle’s perception system.

Finally, each cell is assigned a driving speed limit based on speed limits for each terrain class chosen by the human operator. These speed limits may then be lowered on a per cell basis by the overhead system if obstacles are detected within its vicinity. Speed limits are especially important when they correspond to negative obstacles, as the overhead vantage point allows detection of negative obstacles that a vehicle may not be able to sense until it is too late to slow sufficiently and prevent damage.

**V. Results**

The overhead processing system was validated through the DARPA UPI program. These tests involved a six wheeled, skid steer vehicle with independent suspension called Spinner (see Figure 6). The vehicle is equipped with laser range finders and cameras for on-board perception. The perception system produces traversal costs that are merged with overhead costs into a single cost map. Due to the dynamic nature of this cost map, the Field D* [19] algorithm is used for real-time global planning.

The entire system was demonstrated during two two-week field tests in the fall and winter of 2005. Each of these tests consisted of the vehicle autonomously traversing a series of
courses defined as a set of widely spaced waypoints. During these tests, nearly every important parameter of the environment (climate, season, weather, elevation, local vegetation) varied.

Each run of each course was scored according to two metrics: average speed and number of interventions per kilometer. Interventions were defined as any instance where a human operator interrupted the vehicle’s autonomous navigation for the purpose of vehicle safety (such as an emergency stop to prevent damage) or mission progress (such as an operator decision to remotely reposition the vehicle). The number of required interventions was recorded by an independent observer.

A. Test Site 1

Test Site 1 consisted of off-road desert terrain with large variations in elevation (see Figure 7). Vegetation consisted mostly of low to medium high scrub brush. Hazards included steep slopes, drop offs, and dry washes. The site was divided into three test areas. The North area consisted of a brush filled valley between two mountains, with courses running along the valley. The Center area consisted of flatter terrain, with courses leading around and over small to medium hills. The South area consisted of lowlands leading up to mountains, with both courses approaching and then ascending the mountains.

Each course consisted of waypoints spaced from 100 to 400 meters apart, with an average separation of approximately 250 meters. Each course was run with and without prior overhead data. Some courses were run multiple times under the same conditions to test the determinism of the system.

For the North area, the available overhead data consisted of 54 centimeter resolution aerial color imagery. For the Center and South areas, 80 centimeter IKONOS color and NIR imagery (pansharpened from 3.2 meters) plus 3 meter IFSAR elevation maps were available. 10 meter USGS DEMs were also available for all three test areas. This dataset therefore serves to test the performance of the system with lower resolution prior data. The classifier accuracy on a hand labeled validation set was 81% using only color imagery, and 91% using both color and NIR imagery.

Figure 8 demonstrates the types of situations where lower resolution prior data proved most useful. The second to last waypoint of the course placed the vehicle on top of a ridge. Without prior data, the vehicle had no knowledge of how far the ridge continued, causing the vehicle to follow its border...
Table I

EXPERIMENTAL STATISTICS FOR TEST SITE 1.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Mean Net Distance (km)</th>
<th>Mean Interventions per km</th>
<th>Mean Speed (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>North 1 wo prior</td>
<td>2.39</td>
<td>6.3</td>
<td>0.80</td>
</tr>
<tr>
<td>North 1 w prior</td>
<td>2.49</td>
<td>1.2</td>
<td>1.21</td>
</tr>
<tr>
<td>North 2 wo prior</td>
<td>1.55</td>
<td>2.6</td>
<td>1.24</td>
</tr>
<tr>
<td>North 2 w prior</td>
<td>1.65</td>
<td>1.2</td>
<td>1.23</td>
</tr>
<tr>
<td>Center 4 wo prior</td>
<td>3.66</td>
<td>1.1</td>
<td>1.59</td>
</tr>
<tr>
<td>Center 4 w prior</td>
<td>3.68</td>
<td>0.8</td>
<td>1.71</td>
</tr>
<tr>
<td>Center 1 wo prior</td>
<td>4.07</td>
<td>0.5</td>
<td>1.83</td>
</tr>
<tr>
<td>Center 1 w prior</td>
<td>4.42</td>
<td>0.1</td>
<td>1.96</td>
</tr>
<tr>
<td>South 4 wo prior</td>
<td>2.75</td>
<td>4.7</td>
<td>0.92</td>
</tr>
<tr>
<td>South 4 w prior</td>
<td>2.13</td>
<td>0.9</td>
<td>1.85</td>
</tr>
<tr>
<td>South 3 wo prior</td>
<td>1.39</td>
<td>8.0</td>
<td>0.75</td>
</tr>
<tr>
<td>South 3 w prior</td>
<td>1.37</td>
<td>2.9</td>
<td>0.89</td>
</tr>
<tr>
<td>Total wo prior</td>
<td>21.6</td>
<td>2.5</td>
<td>1.26</td>
</tr>
<tr>
<td>Total w prior</td>
<td>20.2</td>
<td>0.7</td>
<td>1.54</td>
</tr>
</tbody>
</table>

in hopes of finding an easy path down. Eventually, the vehicle was forced to take a steep path down to achieve the final waypoint. With prior data, the vehicle possessed knowledge about the full extent of the ridge, and took the least dangerous path down at the beginning of the ridge.

Cumulative statistics for each course, with and without prior data, are presented in Table I. Overall, prior overhead data led to significant improvements in navigational performance. With prior data, the number of required interventions per kilometer decreased by better than a factor of three, and the average speed increased by 22%.

B. Test Site 2

Test Site 2 consisted of flatter terrain with a much greater variety of vegetation as compared to Test Site 1 (see Figure 9). Vegetation consisted of not only grass and small bushes, but trees of varying sizes. Tree canopy varied from low to the ground and dense, to sparse and high off the ground. Narrow passageways through areas of forest were often covered by tree canopy. The entire site was also heavily populated with ditches. Many of the ditches were not visible from the ground level until a vehicle (human driven or autonomous) was a few meters away. In general, Test Site 2 was considered more difficult than Test Site 1; in some cases, the vehicle chase team was unable to follow in a HMMWV, and was forced to continue on foot.

Test Site 2 was also divided into three areas. The South area consisted of flat terrain containing a mix of forest and open grass. Courses required either traversing through or around sections of forest. The Center area consisted of grassy lowlands with large hills. Courses involved maneuvering through the lowlands as well as ascending hills or plateaus. The North area consisted of ditch-filled lowlands, with courses requiring ascending and descending a large, forested plateau.

Each course consisted of waypoints spaced from 200 to 1100 meters apart, with an average separation of approximately 500 meters. This increased waypoint spacing places more responsibility for efficient traversal on overhead processing. Unlike at Test Site 1, only the South courses were run with and without prior data due to time constraints.

For the entire site, 35 centimeter aerial color and NIR imagery was available, along with 10 meter USGS DEMs. Further, LiDAR scans were available for the Center and North areas, at varying resolutions (20-40 points per square meter). This data set therefore tested the performance of the system with high resolution data. Using imagery and LiDAR, classification accuracy on a hand labeled validation set was 98%.

Figure 10 demonstrates the types of scenario where replanning with prior data performs differently from path tracking. The planned path from prior data chooses a more direct path between waypoints. However, the vehicle’s onboard perception system finds the first section of the planned path to be of higher cost than expected. Instead of going around this area and then coming back to the prior planned path, the vehicle replans an entirely new path, making use of a road that appears in the prior data.

Figure 11 shows the result of one traversal of course Center 1 over four kilometers of complex terrain. The course skirted...
data sources to produce traversal cost maps that can be used to generate accurate global paths through an environment. This approach was shown to significantly improve the average vehicle speed while decreasing the amount of human intervention required during navigation. This approach was evaluated through over 50 kilometers of autonomous traversal through complex natural environments.

Future research will focus on improving classification rates and more accurately predicting vehicle mobility. Future field testing will also provide the opportunity to better compare the effect of data resolution on performance, as well as determine the minimum requirements for overhead data to support efficient ground vehicle navigation.

**Acknowledgements**

This work was sponsored by the Defense Advanced Research Projects Agency (DARPA) under contract "Unmanned Ground Combat Vehicle - PerceptOR Integration" (contract number MDA972-01-9-0005). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

**References**


### Table II

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Mean Net Distance (km)</th>
<th>Mean Interv. per km</th>
<th>Mean Speed (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South 1 wo prior</td>
<td>3.41</td>
<td>1.5</td>
<td>1.74</td>
</tr>
<tr>
<td>South 1 w prior</td>
<td>3.86</td>
<td>0.3</td>
<td>1.89</td>
</tr>
<tr>
<td>South 2 wo prior</td>
<td>2.54</td>
<td>0.8</td>
<td>1.22</td>
</tr>
<tr>
<td>South 2 w prior</td>
<td>3.57</td>
<td>0.6</td>
<td>1.49</td>
</tr>
<tr>
<td>Center 1M w prior</td>
<td>4.44</td>
<td>0.1</td>
<td>1.75</td>
</tr>
<tr>
<td>Center 1N w prior</td>
<td>4.52</td>
<td>0.0</td>
<td>1.90</td>
</tr>
<tr>
<td>Center 2 w prior</td>
<td>3.19</td>
<td>0.3</td>
<td>1.83</td>
</tr>
<tr>
<td>North 1 w prior</td>
<td>6.15</td>
<td>1.44</td>
<td>0.80</td>
</tr>
<tr>
<td>Total wo prior</td>
<td>5.95</td>
<td>1.2</td>
<td>1.17</td>
</tr>
<tr>
<td>Total w prior</td>
<td>30.2</td>
<td>0.3</td>
<td>1.62</td>
</tr>
</tbody>
</table>

The average speed increased by 38%.

### VI. Conclusion

This paper addressed the problem of using prior overhead data of an environment to aid in long range navigation for an unmanned ground vehicle. The proposed approach relies on the geometric and semantic interpretation of heterogeneous the edge of a large plateau before eventually ascending it. The prior global plan for this course used all of the available prior data. When available, roads or road-like areas were used. Open paths through areas of dense trees were found. Ditches were avoided if possible; if not, the easiest place to cross them was found. Finally, the route chosen to ascend the plateau made use of the lowest grade and best maintained road. The fastest autonomous traverse of this course took 42 minutes with an average speed of 1.97 m/s (7.1 km/hr). For comparison, human navigation of the course in a HMMWV occurred at an average speed of approximately 450 meters. The fact that human driving was limited to such a low average speed conveys the difficulty of the course. It is important to note that after several days on site, the route chosen by the human driver between waypoints was nearly identical to the prior planned path.

Cumulative statistics for each course, with and without prior data, are presented in Table II. Due to system wide changes between Test Sites 1 and 2, the overall system performance improved with and without prior data, despite increased terrain difficulty. High resolution prior overhead data proved even more effective than at Test Site 1. The number of required interventions per kilometer decreased by a factor of four, and the average speed increased by 38%.
Fig. 11. Course Center 1 at Test Site 2. The global planned path from prior data and the actual executed vehicle path are shown in red and green, respectively. Select areas of interest have been expanded. The area shown is $2.4 \times 2.4$ kilometers.


