REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704–0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
31–07–2005	Final Progress Report	1 August 2002 — 31 July 2005

4. TITLE AND SUBTITLE

Multi-Image Road Extraction

5a. CONTRACT NUMBER

DAAD19-02-1-0295

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

W. A. Harvey, Steven D. Cochran, and David M. McKeown

5d. PROJECT NUMBER

42993-EV

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Digital Mapping Laboratory
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213–3891

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Reserach Office
P.O. Box 12211
Research Triangle Park, NC 27709–2211

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

42993.1-EV

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official U.S. Government position, policy or decision, unless so designated by other documentation.

14. ABSTRACT Our research is focused on an investigation of automated road tracking using multiple images, toward a goal of fully automated extraction of 3D road networks with topology and attribution. The use of multiple images for road tracking makes the process more robust, due to analysis of the scene from different view points. It also supports direct extraction of 3D information along the path of the road. Determination of road elevation has significant implications for reducing cost and time in applications requiring cartographic features with full 3D attribution. These include mission planning and rehearsal, visualization in urban areas, and the automated production of digital cartographic products. Under this ARO research contract a framework for multi-image road extraction was developed and implemented (RoadMAP3D) with an interactive user interface, tailored to simplify interactions. A detailed quantitative analysis of RoadMAP3D performance is derived and presented. This includes the development of two reference data sets with 3D road geometry, metrics for error calculation with respect to automatically generated road networks, and visualizations of the extracted roads using road height and digital elevation models.

15. SUBJECT TERMS

Feature Extraction, Road Tracking, Stereo, Quantitative Analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Mr. David M. McKeown
U	U	U	UU	27	19b. TELEPHONE NUMBER *(include area code)* (412) 268–2626

Multi-Image Road Extraction
Final Progress Report: Contract DAAD19–02–1–0295

W. A. Harvey, Steven D. Cochran, and David M. McKeown

Digital Mapping Laboratory
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue,
Pittsburgh, PA 15213-3815

EMail: dmm@cs.cmu.edu
URL: http://www.maps.cs.cmu.edu/

22nd November 2005

Abstract

Our research is focused on an investigation of automated road tracking using multiple images, toward a goal of fully automated extraction of 3D road networks with topology and attribution. The use of multiple images for road tracking makes the process more robust, due to analysis of the scene from different view points. It also supports direct extraction of 3D information along the path of the road. Determination of road elevation has significant implications for reducing cost and time in applications requiring cartographic features with full 3D attribution. These include mission planning and rehearsal, visualization in urban areas, and the automated production of digital cartographic products.

Under this ARO research contract a framework for multi-image road extraction was developed and implemented (RoadMAP3D) with an interactive user interface, tailored to simplify interactions. A detailed quantitative analysis of RoadMAP3D performance is derived and presented. This includes the development of two reference data sets with 3D road geometry, metrics for error calculation with respect to automatically generated road networks, and visualizations of the extracted roads using road height and digital elevation models.

1 Introduction

Under U.S. Army Research Office (ARO) funding, the Digital Mapping Laboratory within the Computer Science Department of Carnegie Mellon University conducted a three year research project entitled "Multi-Image Road Extraction" under contract DAAD19-02-1-0295. This document is the final research report under this contract.

Our research is focused on an investigation of automated road tracking using multiple images, toward a goal of fully automated extraction of 3D road networks with topology and attribution. The use of multiple images for road tracking makes the process more robust, due to analysis of the scene from different view points, and it allows us to directly extract 3D information along the path of the road. The extraction of road elevation has significant implications for reducing cost and time in applications that require cartographic features with full 3D attribution. These applications include mission planning and rehearsal, improved geospatial visualization in urban areas, and automated production of digital cartographic products.

Under this ARO research contract a framework for multi-image road extraction was developed and implemented with an interactive user interface, tailored to simplify interactions. Finally, in order to quantitatively assess performance, a detailed analysis of how well the tracker operates in three dimensions is derived and used. This includes the development of two reference data sets with 3D road geometry, metrics for error calculation with respect to automatically generated road networks, and visualizations of the extracted roads using road height and digital elevation models. The goals of this work can be summarized as:

- Perform research into fully automated 3D road network extraction using the composition of multiple tracker methods to increase system accuracy and robustness.
- Apply rigorous quantitative performance evaluation metrics, to include the creation of reference 3D road datasets for comparison with automated extracted networks, to understand the strengths and weaknesses of various road tracking approaches.
- To design, instrument, and test and a 3D road system that would minimize operator interactions, compare well with human level performance, and provide a hundred-fold speedup over current manual collection techniques.

In Section 2 we give an overview of research accomplishments during the first two years of the contract. This work was also reported on in our Annual Research Reports provided to the ARO program manager, and is represented here for completeness and context. Section 3 describes the significant new work accomplished during the third and final year of our research contract. Section 4 provide a summary of this research project as well as conclusions and a brief description of possible future work.

2 Review of Accomplishments: Year 1 and 2

The first year of this contract was mainly concerned with porting a basic stereo road tracker that was implemented several years ago, as well as with the acquisition and/or construction of test sites and reference data. During the second year, this initial stereo road extraction system was qualitatively analyzed, and plans for an improved stereo tracker were made. Finally, the third year brought a wealth of improved reference data which allowed a detailed quantitative and qualitative analysis of the stereo tracking and a refinement of the multi-image stereo tracking model. In this Section we review our work during the first two years.

2.1 Year One Accomplishments

We began our research under this contract by collecting and summarizing extraction results generated using a previous stereo tracker implementation and based upon our RoadMAP monocular road tracking system. RoadMAP is organized as a 'composable' system, meaning that it uses object-oriented structures to allow for the interface of well defined tracker modules. Each module can post results to a higher level controller, which can make decisions regarding overall tracker behavior. This also provides a 'stateful' implementation at a level of abstraction that frees the low-level trackers from these requirements, greatly simplifying their behaviors and implementation complexity. The controlling or "master" tracker used a relatively simple control strategy to loosely manage a slave tracker for each image in the stereo pair. Tracking proceeded synchronously in all images, and the simplicity of the control did not permit much communication between slave trackers, nor did it easily allow the use of data sources other than panchromatic imagery.

This implementation was used as the starting point for our control strategy experiments and, thus, was ported from a UNIX platform to the Windows NT platform where we conducted our research. We then selected several potential test sites and began gathering reference data to be used in the evaluation of the extracted roads. These test sites are:

- Rancho Bernardo, CA
- Pittsburgh, PA
- Fort Hood, TX

2.2 Year Two Accomplishments

The second year of work focused on evaluating the performance of our current stereo tracker in order to gain insight into which control structure changes would be likely to improve our results. This detailed understanding of the current system led us to design a new, more general and flexible multi-image control structure. Inherent in this design is support for the use of multiple panchromatic images, as well as the use of multiple image modalities, including multi/hyper-spectral, LIDAR, *etc.*

The current stereo tracker implementation was built using the same composable tracker architecture used by RoadMAP, our monoscopic road tracking system (Figure 1). RoadMAP's Microsoft Windows interface is composed of an overview window (Figure 1a), which allows the operator to select regions of interest and provides context for the operator, and a tracker detail window (Figure 1b) which provides most of the user interaction and shows the tracker's operation. A number of improvements to RoadMAP have been implemented, many as a result of a detailed user-centric evaluation evaluation of a previous version of the RoadMAP interface [Harvey *et al.*, 2004]. Because RoadMAP and the stereo tracker are built using the same road extraction software architecture, any improvements or extensions to either can be utilized in both contexts.

2.2.1 The Composable Tracker

As mentioned, the stereo tracker is built on the composable tracker architecture. The current implementation uses surface correlation trackers which determine the road track by correlating potential road positions against a road intensity model derived from previous profiles. However, the control architecture (Figure 2) has been designed to be general enough to use with other types of trackers, such as road edge trackers. Our new design extends this architecture to include the use of multi-spectral/hyper-spectral and LIDAR data, as well as improving the communication between the different trackers, allowing appropriate monitoring and verification routines to be invoked.

(a) RoadMAP GUI overview window.

(b) RoadMAP GUI detail window.

Figure 1: Monoscopic RoadMAP User Interface.

Each composable tracker object is comprised of a standard set of operations that are implemented appropriately relevant to the type of tracker being constructed. For example, the `start()` operation for a master tracker would simply call the `start()` operations in each of it's slaves. A slave tracker's `start()` operation is typically more complicated, involving the initialization and/or construction of any necessary internal state. Each of the standard tracker operations are described in the following sections.

Start The controller first advances the trackers by a step, using a geometric path model. The path model incorporates the fact that roads are generally smoothly curving, both horizontally and vertically, and also any prior road location or DEM information.

Each tracker moves to the specified point and searches for points in the vicinity which best match its road model, whether that model is surface intensity, road edges, or other features. In the general stereo control model, we assume that each individual image tracker returns a score indicating how confident it is that it has found the road at each step, *i.e.*, how well the road it found fits its model. This score can be just a number, or an array of numbers indicating the relative goodness of points along a road cross-section. Our current surface correlation trackers return a normalized score indicating how well the current profile matches the road model.

Add The monitoring process first looks at the information returned by the individual trackers to see that all trackers think they are working satisfactorily, then compares the outputs to determine if the positions are

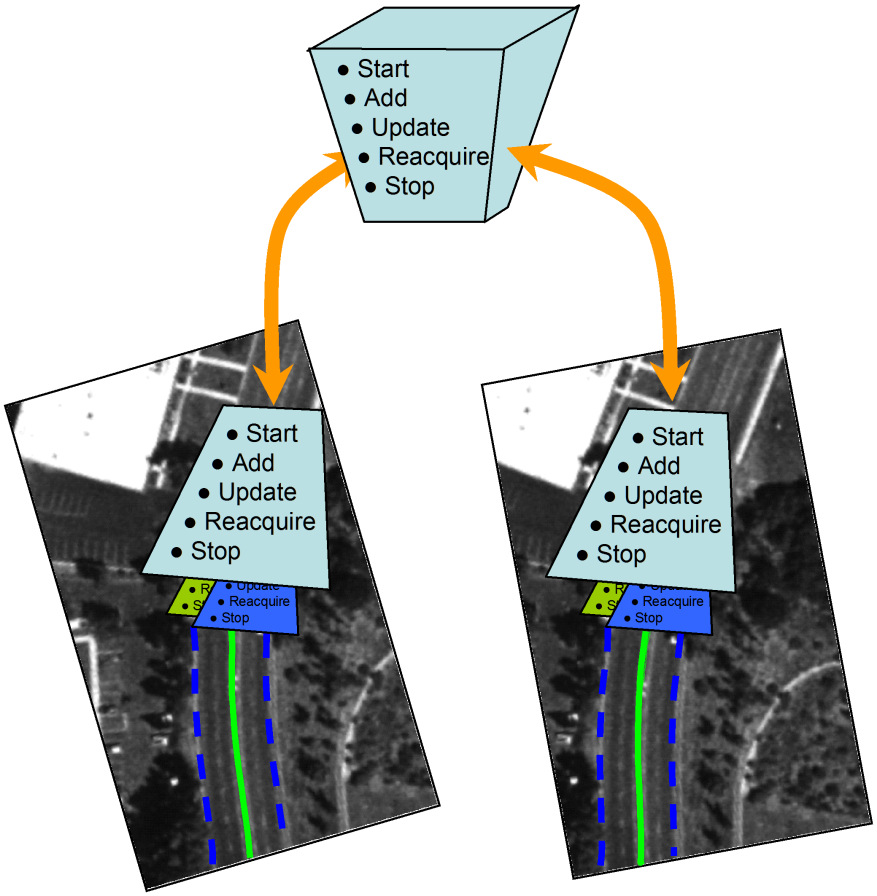


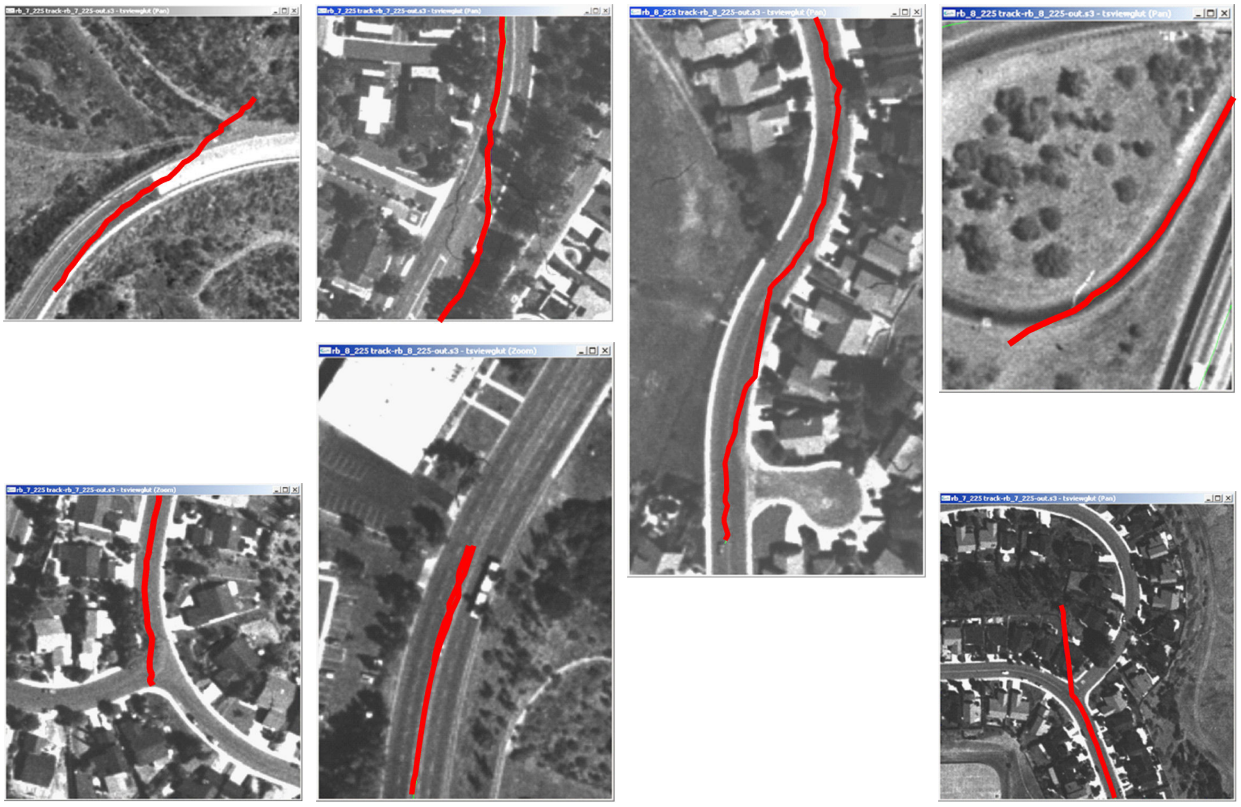
Figure 2: Multi-image Composable Road Tracker.

consistent. In its simplest form, the consistency test consists of using the camera models to intersect the image positions and looking at the residuals. Multiple monitoring methods can be applied, *e.g.*, comparing the computed 3D elevation to a DEM.

For the surface correlation trackers, the monitoring process will correlate the images in the vicinity of the positions returned by the image trackers to determine the best match. Ideally, the positions returned by the trackers should correspond to a correlation maximum. If not, more analysis is indicated.

Update If one or more trackers indicate an error or the monitoring process indicates a problem, the diagnostic process attempts to determine which trackers are correct and which are off-track, and to explain the problem(s) with incorrect trackers.

Reacquire Reacquisition of the road can occur in several ways. The geometric path can be extrapolated, thereby assuming that the road continues in the same direction, until the road can be reacquired. The position of the other image trackers can be projected into the bad image and the score re-evaluated. If the score for the projected position is acceptable, tracking can continue as before. If the score is not acceptable, possible explanations are that something is occluding or shadowing the roadway, or that the road has changed its characteristics, such as a change in surface from asphalt to concrete.



(a) Change in road albedo: Road material change.

(b) Stationary occlusions: Trees overhanging the roadway.

(c) Moving occlusions: Tree shadows in roadway.

(d) Failure to make sharp curve in roadway.

(e) Roadway local width change due to intersection.

(f) Abrupt stop for unknown reason.

(g) Conflated cause: Failure to make sharp curve and/or road width change due to intersection.

Figure 3: Errors in the Extracted Stereo Roads.

2.2.2 Failure mode analysis

To better understand the performance of the existing stereo road tracker, we generated an initial set of 3D extraction results using the Rancho Bernardo and Pittsburgh data sets, then categorized the failures that occurred. Of the 69 failures, they can be broadly grouped into three categories:

- **Road albedo changes**: Stopping errors in this category are due to both "permanent" characteristics, such as a change in road material, and temporary characteristics, such as vehicles in the road.
- **Road geometry changes**: Stopping errors occurred either because of a change in road width or a sharp change in road direction.
- **Unknown**: The stereo tracker stopped for no visually apparent reason.

These three categories represented over 75% of all of the errors observed. Several examples of stopping errors are shown in Figure 3.

2.2.3 New Multi-Image Tracker Design

Earlier experiments with the current stereo tracker highlighted the importance of the control algorithm for the individual trackers. The simple methods currently being used for controlling and combining the individual tracks from each image do not deliver satisfactory results, so we concentrated work on the design of a general control strategy with several design objectives:

1. The control system must be able to use various types of trackers, such as edge or correlation trackers for panchromatic imagery or trackers specifically designed for color, multi-spectral/hyper-spectral imagery, or with LIDAR elevation data.

2. The control system must provide a framework for monitoring and cross-verification of the different types of trackers discussed above. The control system will be able to accept interchangeable verification functions tailored to the tracker types and possibly other factors.

3. The control system must generate an accurate and consistent 3D road centerline.

Figure 4 shows the overall organization of the original stereo image tracker with the proposed improvements shown in orange. As opposed our initial implementation, we are now incorporating tighter coupling between the individual image trackers, using consistency between the trackers as well as their internal scores to monitor their tracking performance and identify which trackers are failing. The monitoring and diagnostic modules can now take advantage of whatever data is available, and should attempt to recognize common situations such as occlusion, shadowing, or overpasses.

The road path will be modeled as a 3D curve. The path will integrate information from the road as it's tracked in each image, any available prior information on road centerlines or terrain elevation, and the geometric properties of roads, such as restrictions on curvature or elevation changes. The use of a strong path model will provide guidance for extension of the road and make recognition of bad tracks more reliable. The integration of information from various sources will also make the calculation of the road path more reliable and accurate.

The road path at each point will be determined by three components:

- The 3D intersection of the tracked road points in each image.

- Smoothness and continuity from previous road points. Roads are limited in their rates of change of curvature and elevation. The maximum curvature allowed could conceivably be related to the type of road, with smaller roads or suburban streets allowed to make right-angle or hairpin turns.

- Prior information, such as road centerlines or a DEM, which will provide specific information on road location within the limitations of its accuracy and completeness.

The 3D path will be used in the control loop to predict the next road point, using the continuity constraints and any available prior information. The 3D road point will be projected into each image and each tracker will find its best estimate of the road at that position. The controller will check the tracked image points for consistency, then calculate a 3D point from the road image coordinates. By comparing the the calculated 3D point to the predicted point, we can perform an initial "sanity check" for mis-tracking. Possible problems could include bad tracking in one or more images, a change in road direction (*e.g.*, the start or end of a curve), or a problem with the prior information. The 3D path will be determined in two stages for each point: an initial blunder detection ("sanity check") step, and a final simultaneous solution using all the relevant information for fine positioning.

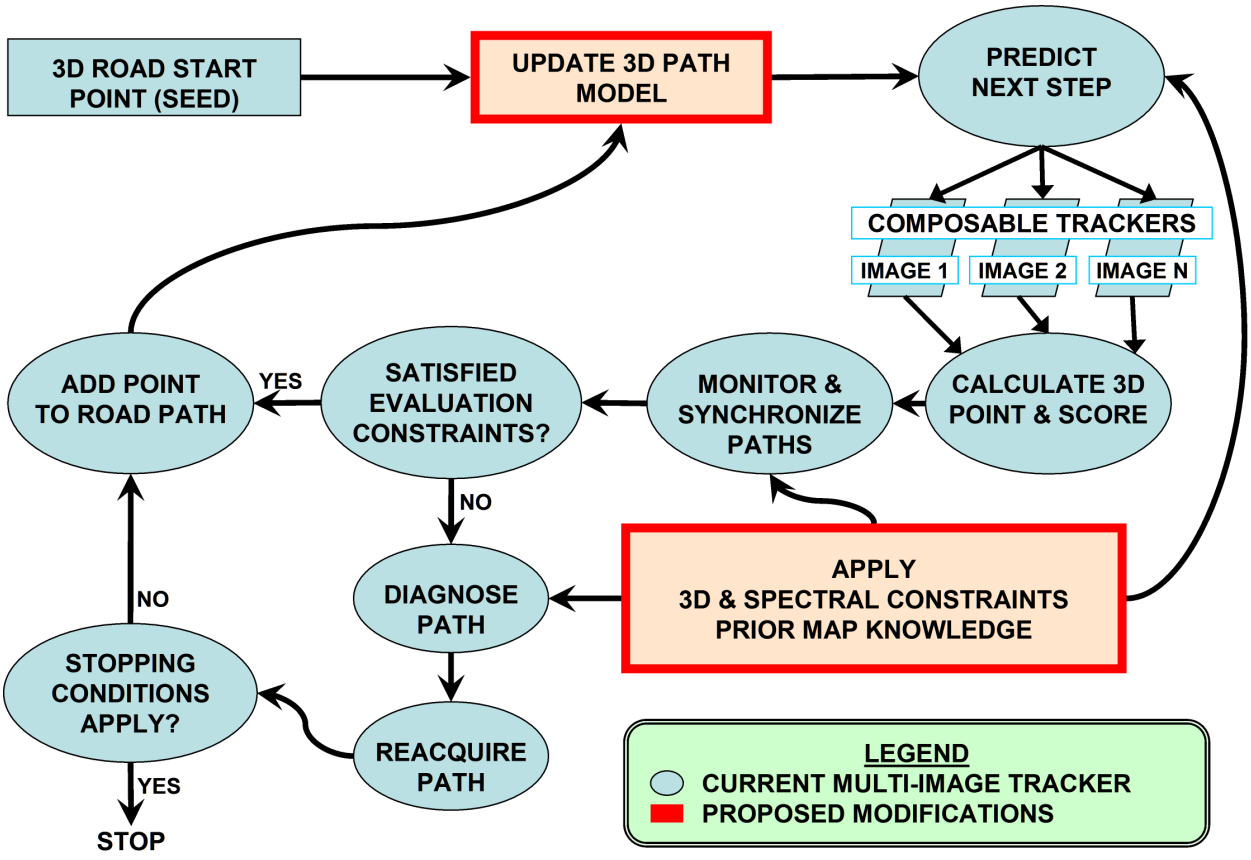


Figure 4: RoadMAP3D Organization.

The design of our general multiple-image extraction system addresses the major failure modes identified during our qualitative analysis. Integrating information from multiple images will help prevent confusion from changes in road appearance or geometry. Since the same change will be visible in all images, verification measures, such as cross-correlation scores, should be high even though one or more image trackers may have trouble with the transition.

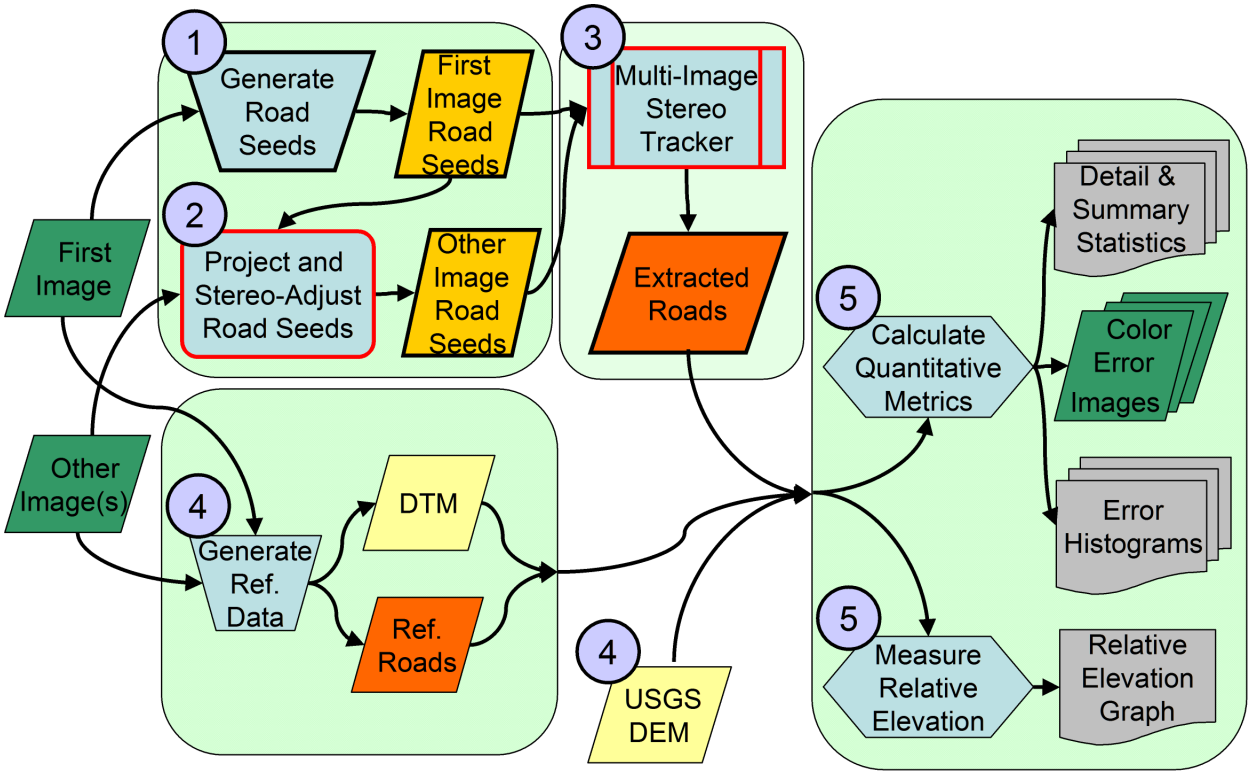


Figure 5: RoadMAP3D Experimental Methodology Process Flow.

3 Third Year Accomplishments

During the final year of this contract, we have spent the majority of our time building tools and datasets to aid us in understanding the strengths and weaknesses of our multi-image extraction process. We present the components of RoadMAP3D, the process we've built for multi-image road extraction. In addition, we provide detailed quantitative and qualitative analysis of the results of running RoadMAP3D on two complex datasets.

3.1 Description of Experimental Methodology

The method we have used for our experiments is outlined in the following five step procedure and shown in Figure 5:

1. Semi-automatically choose road starting points, "road seeds", in a single image. The operator chooses the location and the system determines direction and width.

2. Automatically generate 3D seeds by matching points in multiple images. The system calculates height and may adjust direction and width.

3. Automatically extract 3D roads by surface tracking road features in multiple images and calculating height.

4. Construct or obtain best available 3D road reference data.

5. Quantitatively evaluate the results against references, and generate qualitative displays to aid in interpretation.

The road extraction procedure we've called RoadMAP3D is comprised of steps 1–3, and our evaluation process consists of steps 4 and 5.

3.2 RoadMAP3D

In step one, an operator uses the RoadMAP (2D) interface to choose the road tracker starting parameters (initial location, width, and direction, collectively known as a road "seed") by placing the mouse pointer at the desired image location and pressing a key. The system automatically generates a road seed by estimating the local width and direction from the image. The user can then accept the seed, modify it, or reject it. In RoadMAP (2D), the system is now ready to track the road in the single image being viewed. However, for these experiments, the operator saves the seeds without tracking them.

Step two of the procedure computes 3D seeds from the semi-automatically chosen seeds generated in step one. The monocular seeds are reprojected to the other overlapping images using the image sensor model. In each image, a hierarchical stereo matching process is applied to each seed to find the best estimate for the image location of the seed points. Finally, the seeds are triangulated to generate an initial 3D point estimate.

In the third step, the 3D road seed's are used to initialize the automated 3D tracker, which tracks each road in 3D and outputs the road models in object space. The process is presented schematically in Figure 5. The 3D tracker is comprised of a master tracker, operating in object space, and a single slave image-space tracker for each input image. Each slave operates independently but synchronously, returning a 2D image point and a score at each step. The master tracker synchronizes the slave trackers at each step by maintaining the object-space distance that each slave travels. Any slaves that have advanced further than others are held in place while the others are advanced until all distances coincide. It also applies weak elevation constraints to attempt to minimize anomalous elevation swings, and it monitors the scores of all the slaves in order to determine when the tracking process should look ahead or stop. Finally, the master tracker combines the results of the slave trackers by triangulating the 2D points to generate 3D road centerline points.

3.3 Evaluation Process

To measure progress, a reference and a set of measures is required. In step four, we obtain or compile 3D data in both raster and vector formats to be used as references. Raster reference data can be DEM data, such as seamless 30 meter DEMs available from USGS, or a DTM generated by a stereo process such as ERDAS's OrthoBASE. For the 3D vector reference data we used the ERDAS Stereo Analyst to manually extract road features for the test scenes.

We used ERDAS's OrthoBASE to find the camera model, then the ERDAS Stereo Analyst provided tools for generating DTMs, as well as tools that allow the operator to extract and update the 3D road features from any pair of overlapping images. Figure 6 shows a screen dump of an operator using ERDAS Stereo Analyst to extract 3D road features as pairs of polylines. Figure 7 shows the manually extracted set of 3D reference roads in green on the automatically generated 1 meter DTM (rotated so that North is left). Looking closely, one can see that the DTM clearly represents the visible 3D features, including buildings and trees.

Given the scale of typical road features, comparisons to the USGS 30 meter DEM are not as useful for detecting 3D errors in the output of an automated road extraction process. In addition, features, such as bridges and overpasses, are absent from the DEM. A high resolution DTM yields a much better comparison, both because of the increased resolution and because it will include elevated road features. However,

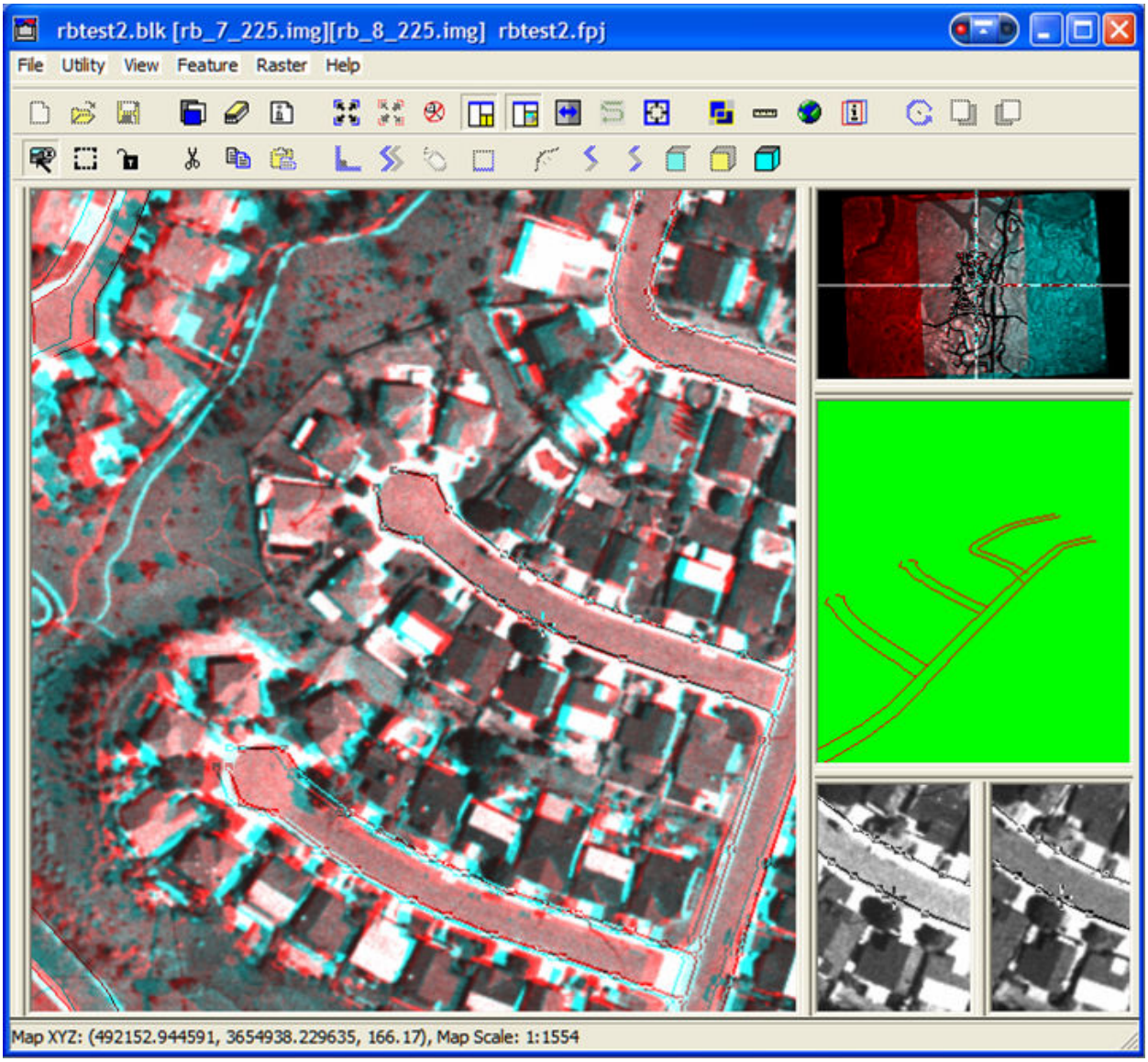


Figure 6: Using ERDAS Stereo Analyst to Extract 3D Reference Roads (Rancho Bernardo).

because the processes typically used to generate DTMs tend to smooth over depth discontinuities, we still need to analyze the differences in order to properly categorize errors in the DTM versus errors in the automated output. The most useful comparisons are performed against high-resolution, manually generated road networks, since our goal is to automatically produce extracted roads at a quality level equal to or better than those produced manually.

Step 5 shows the evaluation process. We have extended our extensive set of 2D road extraction evaluation metrics [Harvey *et al.*, 2006; Harvey, 1999] to include 3D evaluation measures, as well as several other quantitative 3D evaluation methods. Once we have gathered the automatically extracted roads, the manually extracted roads and the DTM, we can compare these data sets in a number of useful ways, generating:

- Relative elevation profile graphs
- Summary statistics

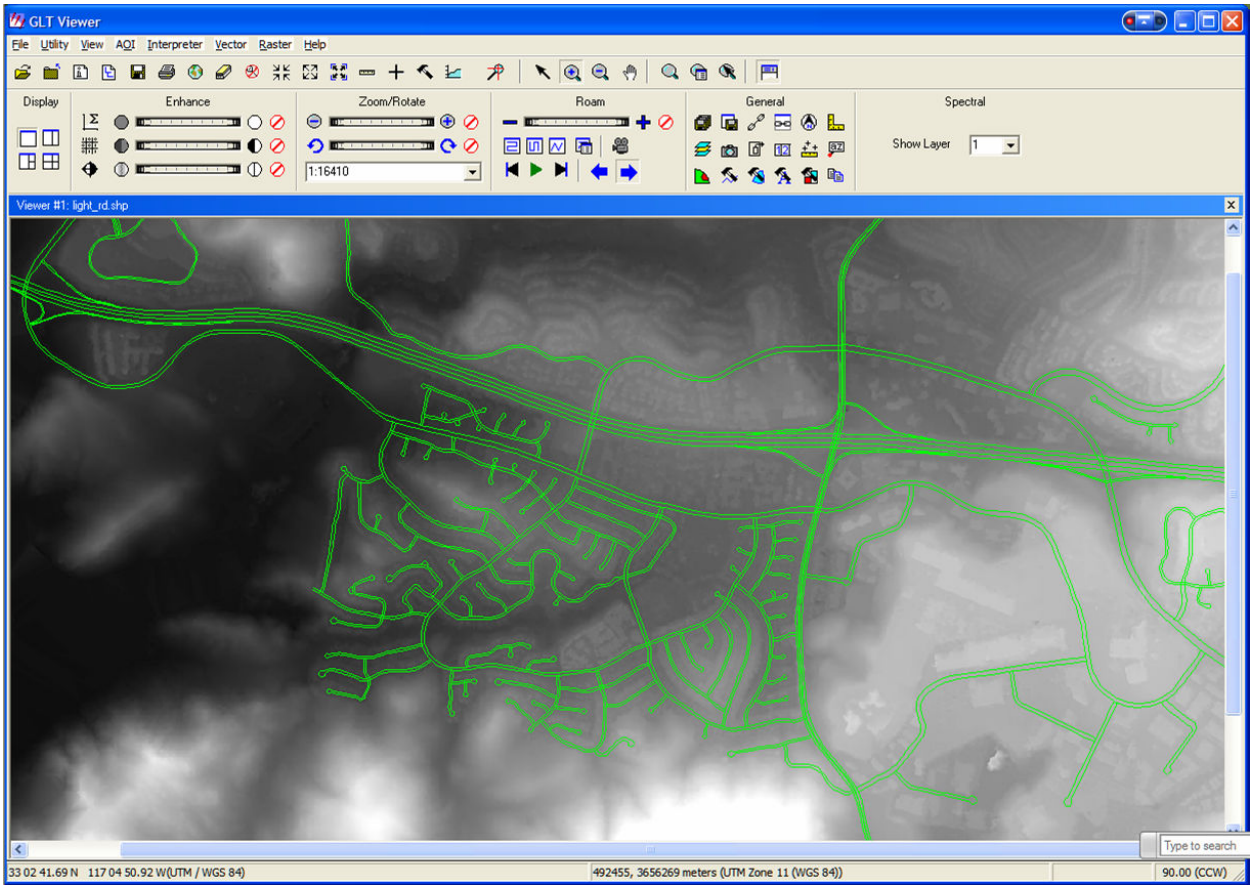


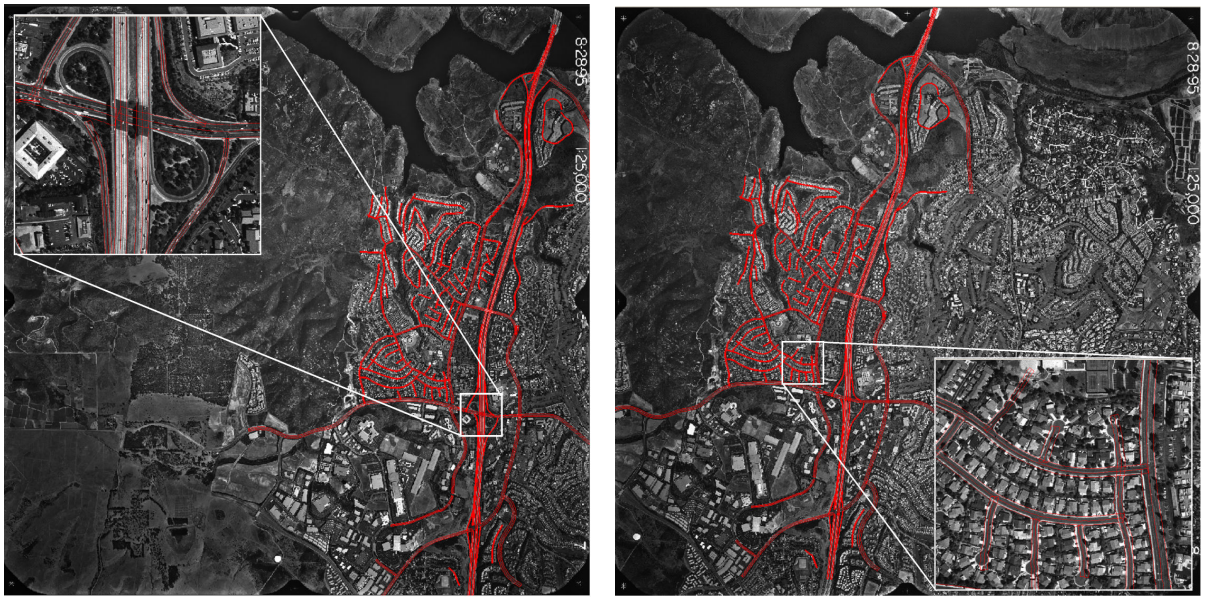
Figure 7: Manually Extracted Reference Roads for Rancho Bernardo Dataset.

- Detailed error statistics
- Colorized error images
- Error histograms

We selected two datasets for our experiments: Rancho Bernardo, CA, and Pittsburgh, PA. We present a detailed analysis of the extraction and evaluation procedure as applied first to the Rancho dataset, then to the Pittsburgh dataset. Section 3.4 described the reference data collection methodology using commercially available photogrammetric workstations, the analysis of tracker errors with respect to a digital elevation model (DEM) as well as a digital terrain model (DTM). Section 3.5 gives the analysis

3.4 Results on Rancho Bernardo, CA Dataset

The Rancho Bernardo test case is quite complex, containing rolling terrain, a mix of road types such as cloverleafs, limited access highways, intermediate feeders and residential roads including cul-de-sacs. The ground sample distance for this stereo image pair is 0.6 meters. Figure 8a shows the road model vectors extracted by RoadMAP3D overlaid in red on the source image data, and the system measurements are presented in Tables 1 and 2. Of the approximately 74 kilometers (126 roads) of roads present in the CMU reference data collected over this area, RoadMAP3D extracted almost 61 kilometers in about 11 minutes



(a) (b)

Figure 8: Manually Extracted Reference Roads (Rancho Bernardo).

Table 1: RoadMAP3D System Measurements for the Rancho Bernardo Data Set

Data Set	# of Roads	Total Length	Total Time
Reference Data Manually Compiled in ERDAS Imagine	126	74.02 Km	≈1 day
RoadMAP3D	104	60.93 Km	660 sec.

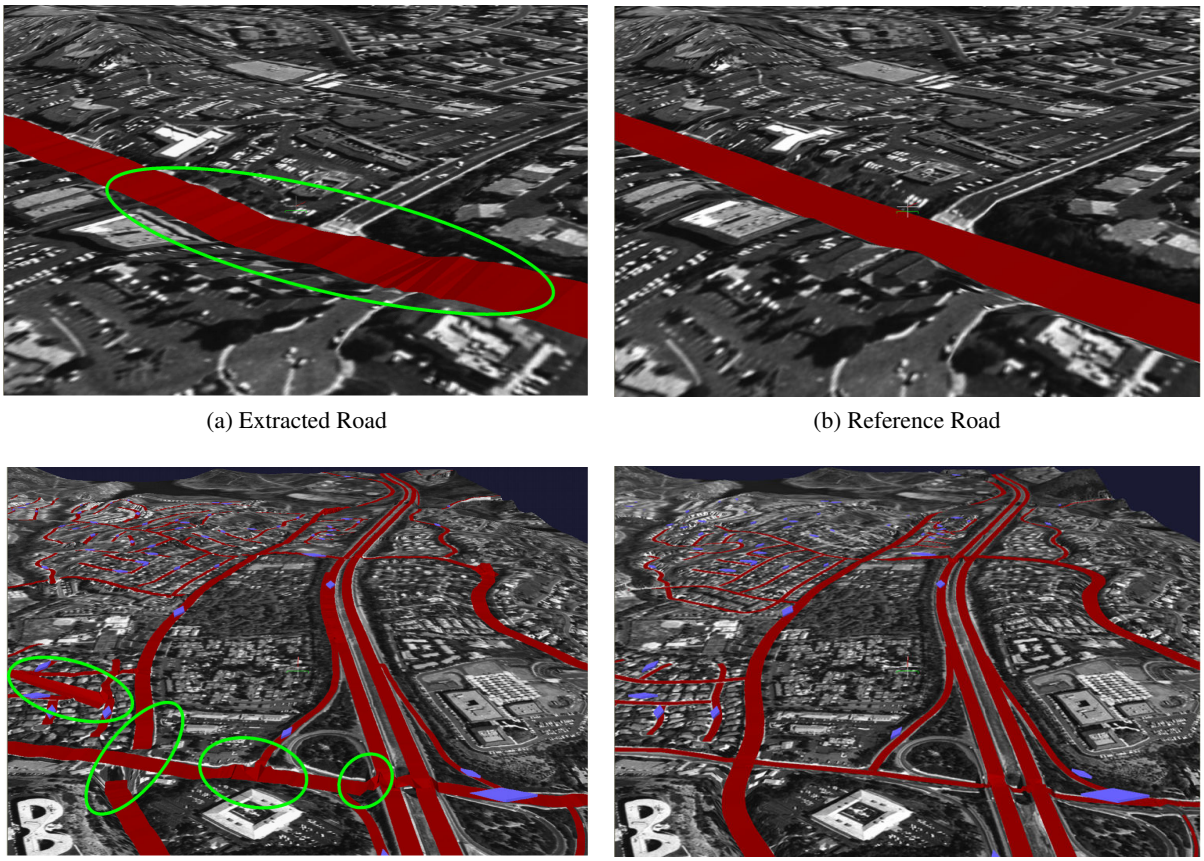
Table 2: Summary Measurements for Rancho Bernardo Data Set

Measure	Value
Match Time / Seed	1.35 sec. / Seed
Track Time / Seed	5.10 sec. / Seed
Total Time / Seed	6.45 sec. / Seed
Km / Seed	0.59 Km / Seed
Track Time / Km	8.63 sec. / Km
Total Time / Km	10.93 sec. / Km
Number Seeds / Km	1.71 Seeds / Km

(wall clock time).[1] The operator selected 104 road seeds, all of which generated road vectors in the final output. Some roads required multiple seeds for complete extraction, *e.g.*, the major highways, while other seeds tracked through intersections to cover multiple roads. The efficiency of the process, both in terms of time (11 seconds per kilometer) and seed density (1.71 seeds per kilometer), is reasonably good.

The extraction results and the manually produced reference data were both exported as sets of ESRI Shapefiles. These vectors, along with the USGS 30 meter DEM and a 0.6 meter orthophoto, were used

[1]The procedure was executed on a dual-processor 2.4GHz Xeon PC with 1GB of memory running Microsoft WindowsXP Pro.



(a) Extracted Road

(b) Reference Road

(c) Extracted Roads

(d) Reference Roads

Figure 9: Errors in the Extracted Stereo Roads for Rancho Bernardo.

to construct several visualization databases from which the screen dumps in Figure 9 were made. The RoadMAP3D extracted data is shown on the left and the corresponding areas with the reference data are shown on the right. The blue lozenge-shaped markers scattered throughout visualizations indicate the 3D position and direction of the manually created road seed points that were used to initiate the 3D road tracker.

The figures in Figure 9 illustrate some typical tracking errors observed in the road vectors extracted by RoadMAP3D. The top left figure shows ripples appearing in the road surface. These correspond to locations where the object-space master tracker is having trouble maintaining synchronization between each of the independent image trackers, reflecting the lack of a smoothness constraint applied by the tracker when performing matching. An alternative to embedding such a constraint in the tracker would be to apply road construction smoothness constraints as a post-processing step. Of course, under some road construction conditions, these undulations could be possible, *e.g.*, avoid cut-and-fill engineering costs in rural roads. The figure in the bottom left highlights several other problems, including road width errors, gaps along the road, and road intersection errors including overshoots, undershoots, and height discontinuities. The green ovals point out the locations of these errors as it can be difficult to find them by visual inspection.

More quantitative statements about 3D accuracy and road network completeness can be made by comparing the automatically extracted vectors against various types of reference data. From these comparisons, we can identify the strengths and weaknesses of our approach, as well as measure the effects of system modifications. Table 3 shows the standard statistical measures we compute (mean, standard deviation, min-

Table 3: Quantitative Summary: 3D Metrics for Rancho Bernardo Dataset

Reference	Percentage		Elevation Error (meters)				
	TP	FP	RMS	Mean	s.d.	Max	Min
Manual Roads	63.68	36.32	3.69	-0.11	3.69	34.42	-28.15
1 meter DTM	99.23	0.00	4.57	-1.10	4.44	39.70	-29.52
30 meter DEM	93.27	0.00	6.19	1.42	6.02	23.64	-28.53

imum/maximum error, and RMS error). Additionally, we classify as "true positives" (TP) those areas of the extracted data that overlap the reference data, and "false positives" (FP) otherwise. When using a raster reference, such as a DTM, areas that lay outside the bounds of the reference are classified as "ignored". This classification is most useful when performed against a vector reference, such as the manually generated roads, as it provides a measure of the 2D accuracy of the automated extraction. In this example, one can see that almost 64% of the manually extracted reference roads are covered by the automatically extracted data. Other 2D correctness/quality measures are also computed [Harvey, 1999].

The 3D accuracy of these results can be better understood by analyzing the elevation error distribution histograms. Figure 10 shows a set of three histograms comparing the RoadMAP3D results to the reference road vectors (blue), the 1 meter DTM (orange), and the 30 meter DEM (white). One can see from these histograms that there is very good correlation between the elevations computed by RoadMAP3D and both the reference and DTM elevations. This is verified by the statistics, where the mean and standard deviation for both the reference roads ($\mu = -0.11$m; $\sigma = 3.69$m) and the 1 meter DTM ($\mu = -1.10$m; $\sigma = 4.44$m) are displayed on the right side of the plot. The comparison to the 30 meter DEM yields good agreement ($\mu = 1.46$m), but shows a wider error distribution ($\sigma = 6.02$m) due largely to the coarse pixel size. Though the elevation errors in the automatically extracted 3D roads are evident in the spread of the histograms, these initial results agree very well with the 1 meter DTM and the road reference data.

To compliment this analysis, we use the reference comparison to generate colorized road elevation difference maps such as those presented in Figures 11a–c. The maps are color coded to depict road height variations while also showing spatial locality for road error analysis. A labeled color palette in the upper left corner of Figure 11b presents the range of error values ($-20 \ldots +20$ meters) with their corresponding colors.

Figure 11 shows three colorized road elevation difference maps that were generated by plotting all of the 3D roads extracted by RoadMAP3D and comparing their elevation estimates against the 30 meter DEM, the 1 meter DTM, and the reference roads. baseline reference elevation. For a perfect agreement we would expect to see mostly green road segments. This provides a quantitative depiction of extraction accuracy given a high spatial resolution comparison.

Figures 11a, 11b and 11c: This set of error images shows that the automatically extracted roads are a good match to the 1 meter DTM and the hand-generated reference image, but not to the 30 meter DEM. This shows both that the automatically extracted roads are reasonably good and that the 1 meter DTM (which is easier to produce) serves as a good first approximation for the hand-generated reference roads. As the automatically generated roads improve in quality, multiple hand-generated reference sets will be needed to establish the accuracy of human-generated roads, and more extensive ground reference will be needed to guide the improvements to the automatic tracking.

We can use these evaluations to aid in diagnosing the causes of various observed errors. For example, some errors can be detected by generating elevation profile plots along the centerline of individual extracted roads. An example of such a plot is shown in Figure 12. This plot was generated from an extracted road

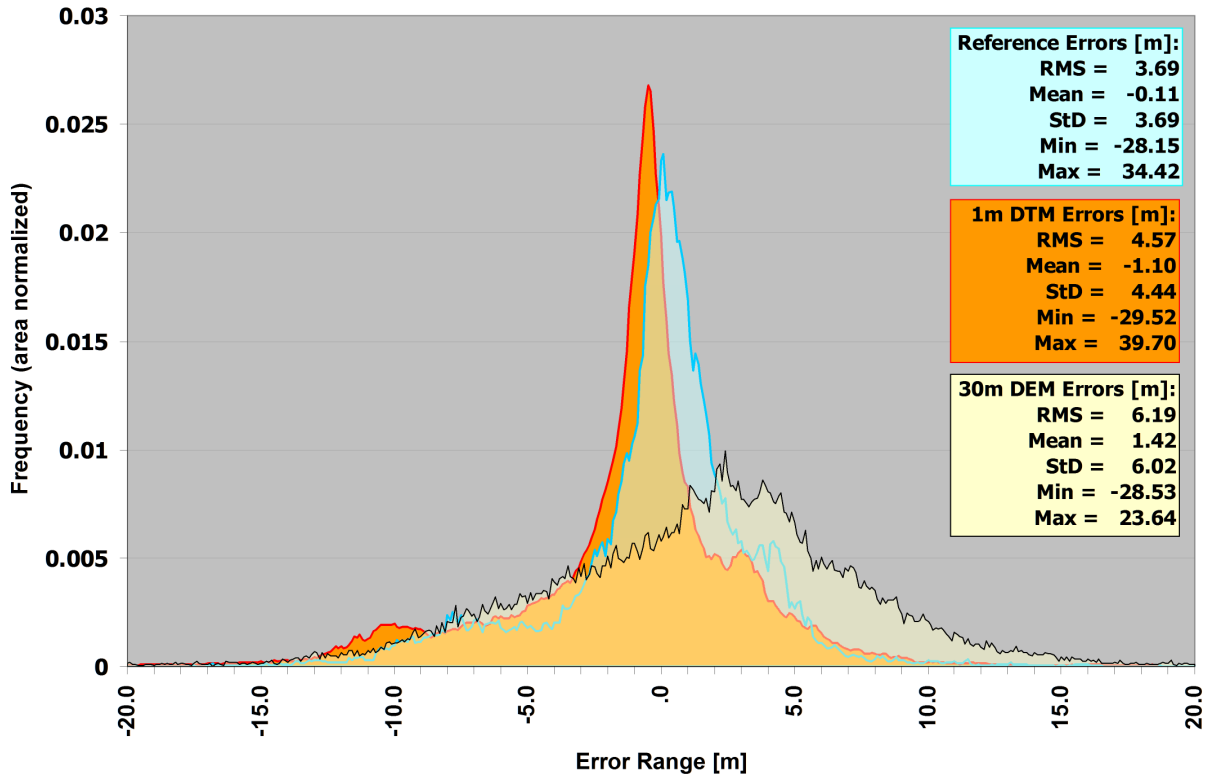
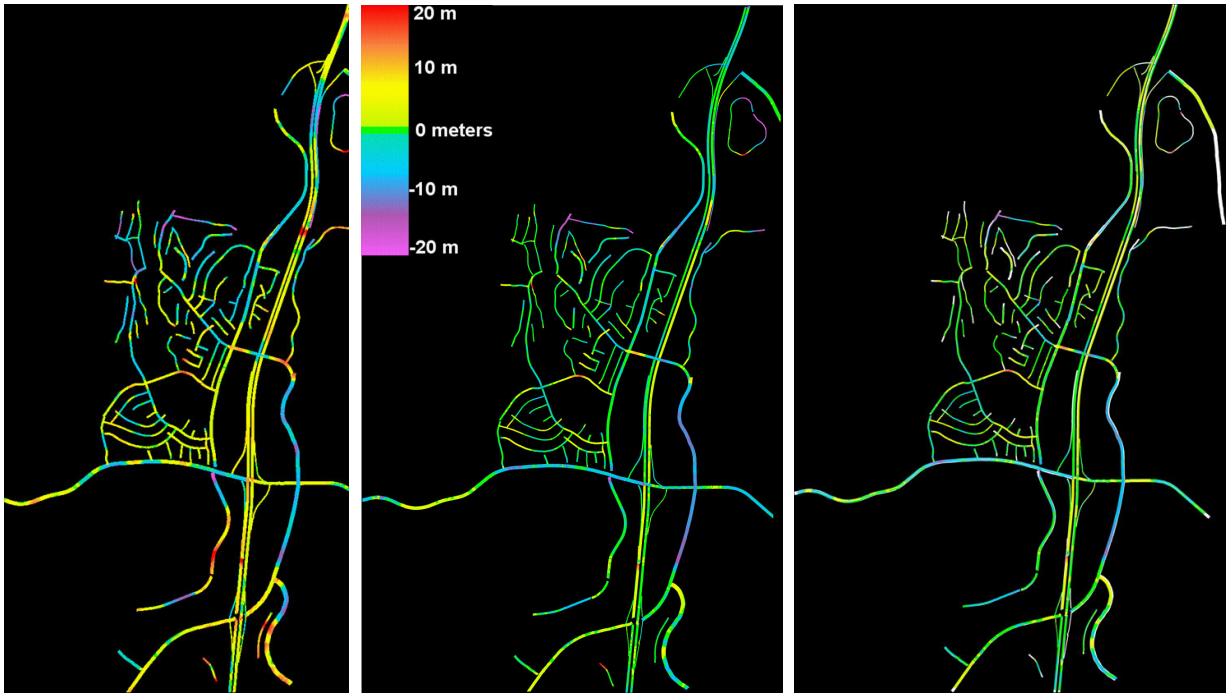


Figure 10: Road Tracker Elevation Errors for Rancho Bernardo Dataset.

segment over the right side of the main highway traveling north-south through the images. Only the northern portion of the segment is plotted so that some of the detail can be seen. The four datasets being plotted are the values from a 30 meter USGS seamless DEM, a 1 meter and a 5 meter DTM generated using ERDAS Stereo Analyst, and the height value computed by RoadMAP3D. The horizontal scale is the length along the road feature (here, approximately 1.5 kilometers) and the vertical scale is the elevation value (both in meters).

All four datasets agree at a gross level, though many differences, some significant, can be seen. The 30 meter DEM is much smoother and travels above and below the other values. This is consistent with our expectations, since the pixels are larger, thus generating a smoothed profile, and the fact that it is supposed to represent elevations on the ground implies that it will not represent 3D road features such as bridges. That the DEM values stay mostly below the others is also expected because of the smoothing inherent in generating a DEM at a 30 meter scale. The 1 meter and 5 meter DTMs track each other closely, with the 1 meter DTM expectedly exhibiting more noise. Since the DTMs are high-resolution, we expect them to depict 3D road features, though because they are are automatically generated, we also expect them to contain some errors.

We've highlighted three portions of this plot in order to emphasize several comparisons. The annotated images are provided so we can correspond the positions in the plot to the locations in the images. At comparison point 1, the RoadMAP3D result corresponds very well (within a meter or so) with both of the DTMs, whereas the DEM dips about 20 meters below the others. Looking at the corresponding image location, it is clear that RoadMAP3D (and the DTMs) are correctly detecting the elevation of the bridge deck along the highway. At comparison point 2, the RoadMAP3D result rises almost 10 meters above the



(a) 30 meter DEM (b) 1 meter DTM (c) Reference Roads

Figure 11: Elevation Errors for Rancho Bernardo Dataset.

others. This is due to one of the 2D image trackers incorrectly repositioning itself after guessing ahead, introducing a false rise in elevation. Though this looks extreme, this elevation rise and the correction occurs over a distance of about 200 meters. Finally, at comparison point 3, RoadMAP3D and the DTMs again agree, with all falling below the corresponding DEM values. In this case, the road is traveling through a cut in the hillside where we've manually verified that the RoadMAP3D and DTM values more accurately represent the road height.

An alternative method for diagnosing error sources is to assign colors to "interesting" portions of the error histogram, then generate a colorized elevation error map that assigns those colors to errors within the selected error ranges. For example, comparing the automatically extracted roads to the manual reference roads in the Rancho dataset yields the blue histogram in Figure 10. The statistics suggest that these two sets of roads compare very well. However, at least two anomalies in the histogram stand out, namely the height of the tail on the left and the small peak on the right. In Figure 13, we've divided this histogram into four ranges and assigned a color to each of them:

- The left tail of the histogram ($<$ -3.7 meters or 1 standard deviation) is colored BLUE and shows areas much lower than the reference roads. Approximately 16% of the errors are found in this range.

- The central peak ($-3.7 \ldots +3.6$ meters, errors within 1 standard deviation of the mean) is colored GREEN and shows areas where the extracted roads and the reference roads agree. Approximately 66% of the errors are found in this range.

- The local peak to the right (3.6 ... 5.0 meters above the reference roads) is colored YELLOW. Approximately 11% of the errors are found in this range.

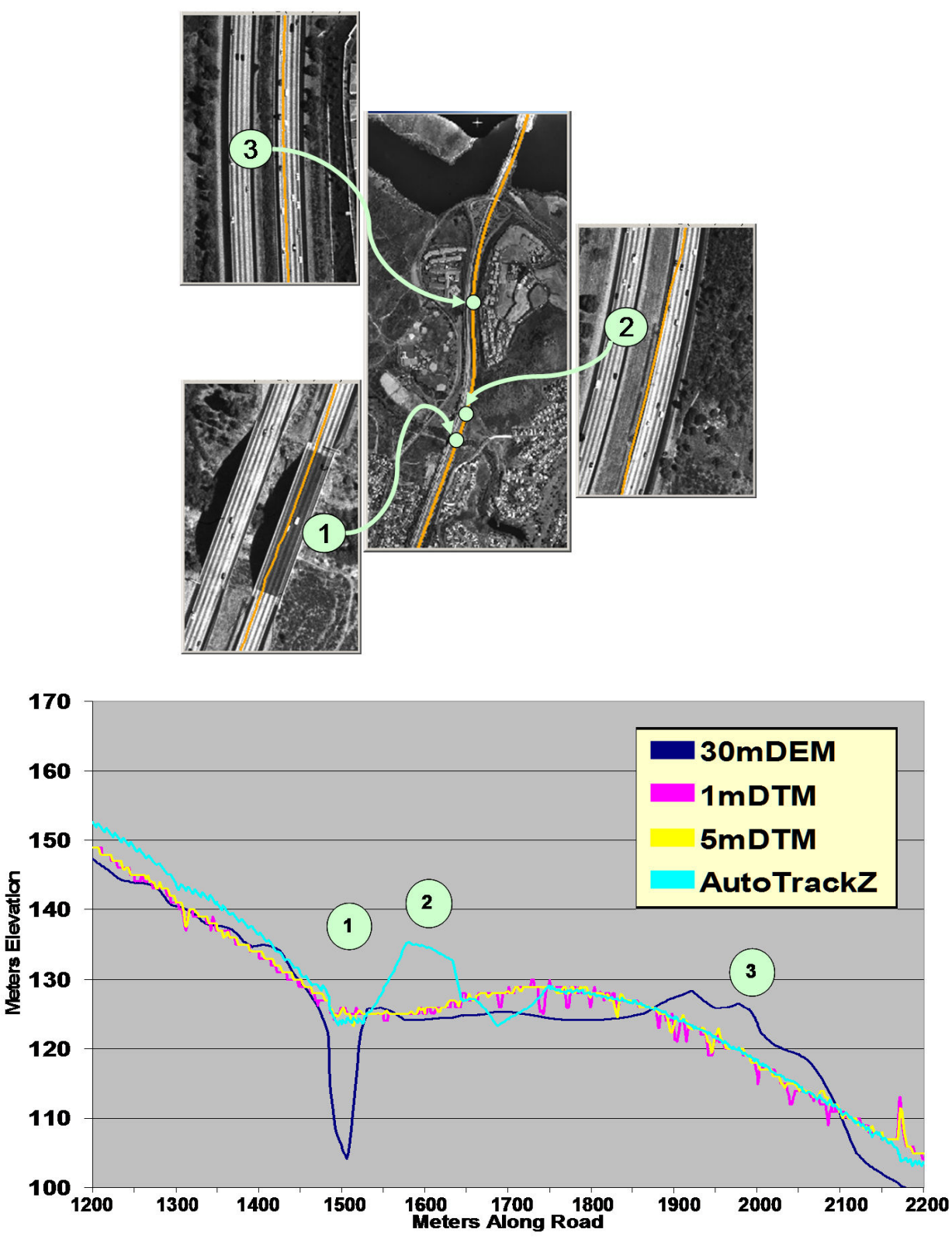


Figure 12: Elevation Profile Comparison for a Single Road.

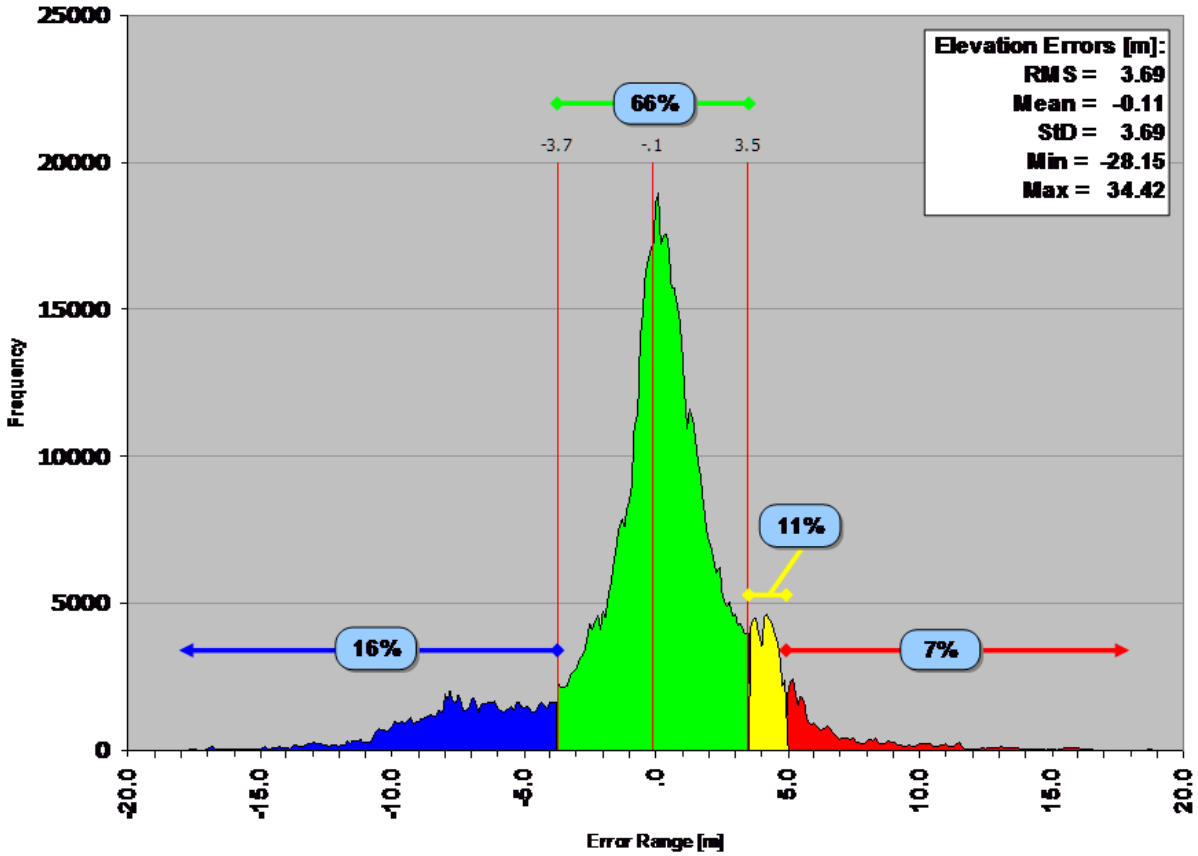


Figure 13: Analysis of Elevation Errors: Histogram Sections.

- The right tail of the histogram (> 5.0 meters) is colored RED and shows areas much higher than the reference roads. Approximately 7% of the errors are found in this range.

It is interesting to note that this analysis verifies that the errors between these two datasets are very nearly normally distributed about the mean.

From the corresponding error map in Figure 14, one can immediately see the locations of the the various errors ranges. The majority of the roads fall into the green range, demonstrating that most of the automatically extracted road points agree with the manually extracted reference roads. The blue range occurs primarily on some road ends and two long road segments. The long segments were both verified as having been started with poor initial height estimates. The yellow range is localized primarily on the East side of the main divided highway and appears to be due to the trackers falling slightly out of sync after attempting to look past a surface material change. The red range is seen to be occurring primarily on road ends, bridges, epipolar-aligned roads. These errors occur because the independent image-space trackers can fall out of sync shortly before stopping, or when having reacquired the road just after anomalies and surface material changes.

We need to perform further analysis to determine, and eliminate, these and other error sources. Even so, we consider these results very encouraging, though they illustrate that there will always be differences due to the interactive processes used when producing DTMs and 3D reference roads. We understand that a robust 3D feature extraction system must be able to cope with such errors in a predictable manner.

3.5 Results on Pittsburgh, PA Dataset

The extraction and evaluation methodology demonstrated with the Rancho Bernardo dataset was also applied to a dataset over Pittsburgh, PA. As with the Rancho dataset, the Pittsburgh imagery also contains a rich set of road types, including a bridge and several overpasses. Extensive tree cover and shadowing provides additional challenges.

Figure 14: Location of Histogram Section Errors.

Two sets of overlapping stereo image pairs (0.6 meter GSD) were selected and block-adjusted using ERDAS OrthoBASE. ERDAS Stereo Analyst was used to extract a set of 12 reference roads and to produce a 1 meter DTM. Figure 15 presents the set of extracted reference roads overlaid on the generated DTM. We also acquired a 30 meter DEM from the USGS Seamless web site. Finally, we manually selected a set of 26 road seed points and ran the 3D extraction process.

Tables4 and 5 summarize the extraction results, and Table 6 shows the statistical results of comparing the extracted data to the three reference datasets. In all three cases, the mean error value is within 2 meters of the references. It is encouraging that the standard deviation is smallest ($\sigma = 4.85$m) when we compare against the reference roads. In addition to modifications to reduce the standard deviation of the 3D error, we see that 2D accuracy with respect to the reference roads (**TP** $= 59.79\%$) is high, but could be improved.

From the error histograms presented in Figure 16, one can see that the automated results compare reasonably well to both the reference roads and the DTM, though the errors versus the DTM are more widely spread about the mean and have longer tails. As with the Rancho data, the large standard deviation reported when comparing to the USGS 30 meter DEM is primarily due to the coarse pixel size. However, there are several peaks in the distribution that we intend to investigate further.

Figures 17 and 18 present color-coded error maps superimposed on the one meter DTM (top) and an orthoimage (bottom) with north at the top. The error map in Figure 17 was generated with respect to the reference roads and the map used in Figure 18 was generated with respect to the one meter DTM. In both sets of figures, the striped segment was manually added to show the location of the tunnel connecting the highway that runs east-west through the image.

The white regions in Figure 17 are portions of the extracted roads that do not overlap the reference roads. In some cases, this is because the 3D tracker extracted a road that was not present in the reference data, *e.g.*, the regions in the extreme east and west. In the central portion of the scene, it appears that the 3D tracker extracted the road shoulder. This is due, in part to tree shadows on the road surface.

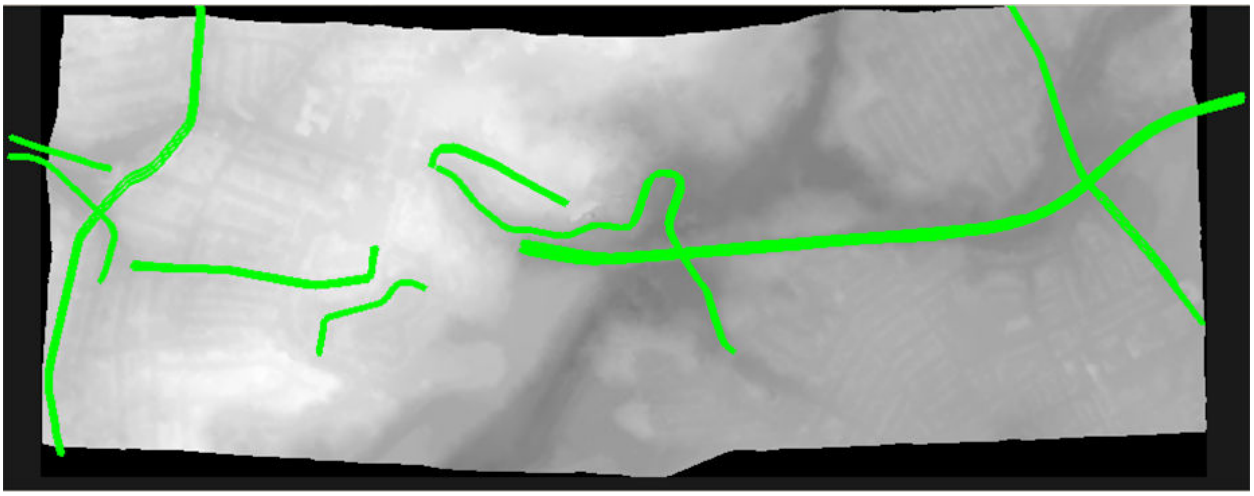


Figure 15: Manually Extracted Reference Roads and 1 meter DTM over Pittsburgh.

Table 4: RoadMAP3D System Measurements for the Pittsburgh Dataset

Data Set	# of Roads	Total Length	Total Time
Reference Data Manually Compiled in ERDAS Imagine	12	13.00 Km	≈3 hours
RoadMAP3D	26	10.74 Km	131 sec.

Table 5: Summary Measurements for the Pittsburgh Dataset

Measure	Value
Match Time / Seed	1.21 sec. / Seed
Track Time / Seed	3.82 sec. / Seed
Total Time / Seed	5.03 sec. / Seed
Km / Seed	0.41 Km / Seed
Track Time / Km	9.24 sec. / Km
Total Time / Km	12.17 sec. / Km
Number Seeds / Km	2.42 Seeds / Km

The green (correct) and yellowish (slightly higher than reference) portions of the maps are well distributed throughout the extracted roads, though there are several notable exceptions. The red road portion on the right in Figure 18 is a false elevation rise near a road end, and the orange portions at the left and right ends are overpasses. A bluish section of the central portion of the main highway in Figure 18 is due to the ERDAS stereo process extracting the tree canopies that overhang the road.

We consider this a good result, though significant improvements can be made to the 2D performance. A more thorough classification of the errors needs to be performed, but it appears that many of the premature stoppages are due to trees and shadows.

Table 6: Quantitative Summary: 3D Metrics for Pittsburgh Dataset

Reference	Percentage		Elevation Error (meters)				
	TP	FP	RMS	Mean	s.d.	Max	Min
Manual Roads	59.79	40.21	5.13	1.69	4.85	50.41	-32.45
1 meter DTM	98.21	0.00	8.45	1.63	8.29	58.13	-33.21
30 meter DEM	99.96	0.00	10.94	-1.76	10.80	59.41	-38.02

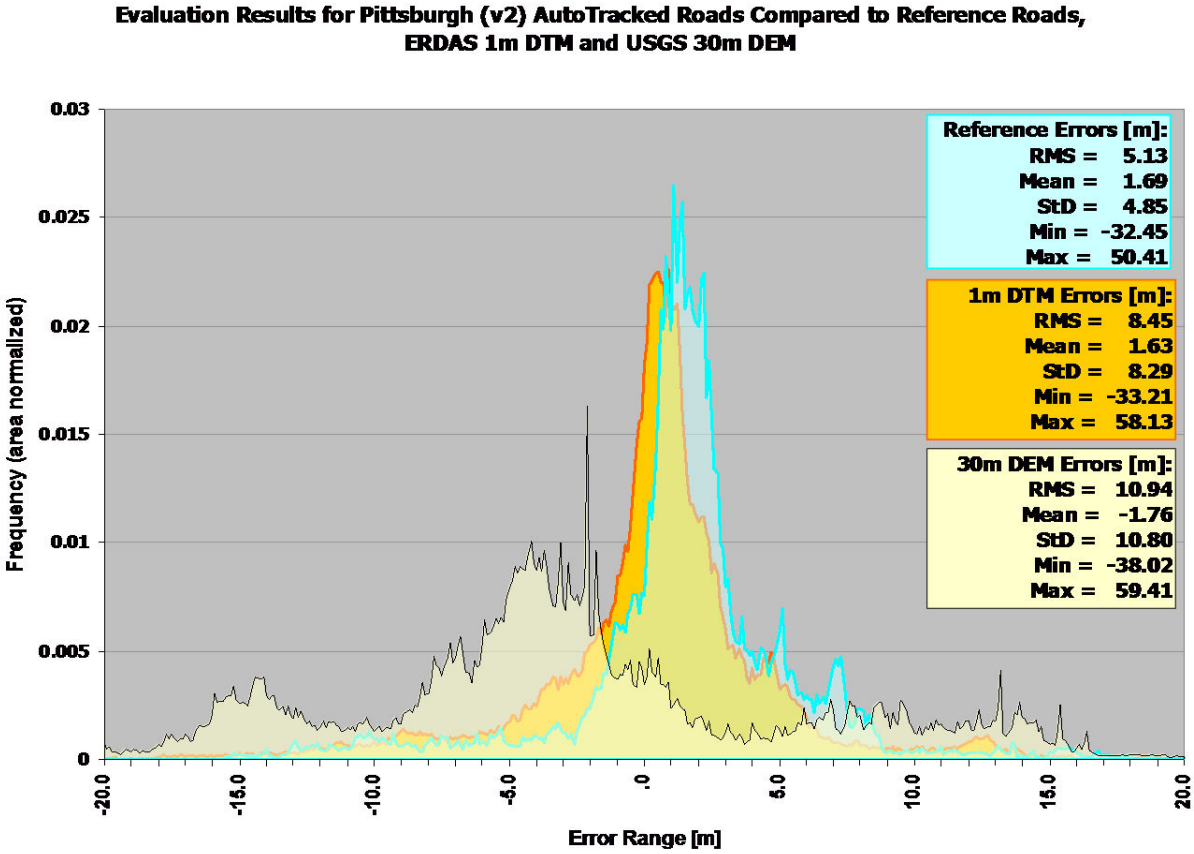
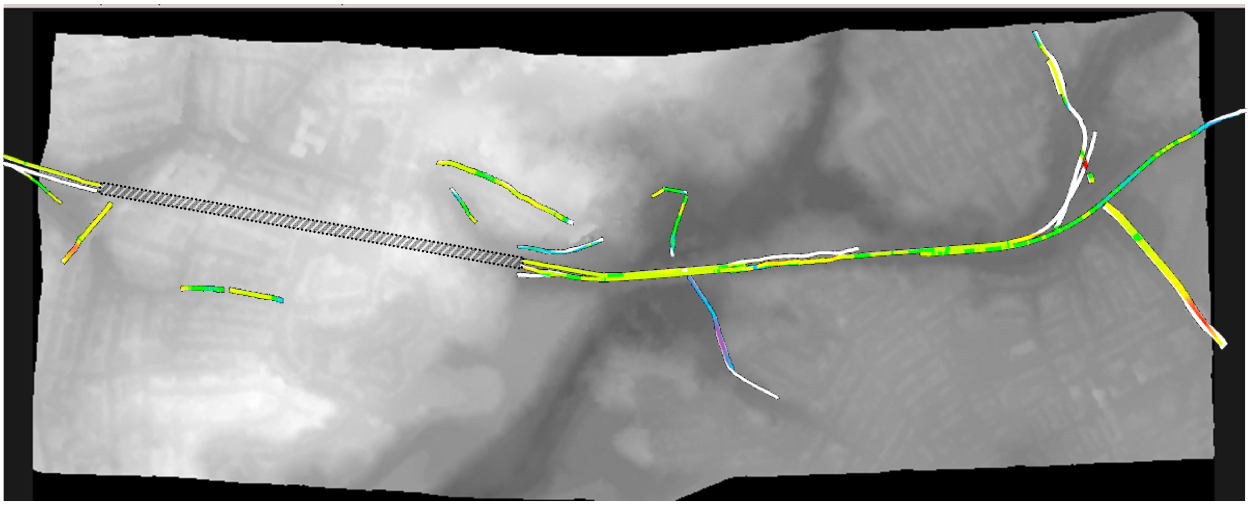


Figure 16: Road Tracker Elevation Errors (Pittsburgh).

4 Conclusions

In this report we have described the development of a semi-automated system for 3D multi-image road network extraction called RoadMAP3D. This work was based upon previous research in composable road trackers working with a single image. The flexibility of the composable design is demonstrated in that we were able to include the tracking of multiple images as an instance of the tracker architecture. In this report we present an extensive performance analysis including strengths and weaknesses of the current research system. RoadMAP3D requires limited user interaction in a single image, demonstrates reliable extraction of 3D road network, has been evaluated in a variety of complex urban scenes, and generates accurate road height estimates without a DEM.



(a) Superimposed on 1 meter DTM

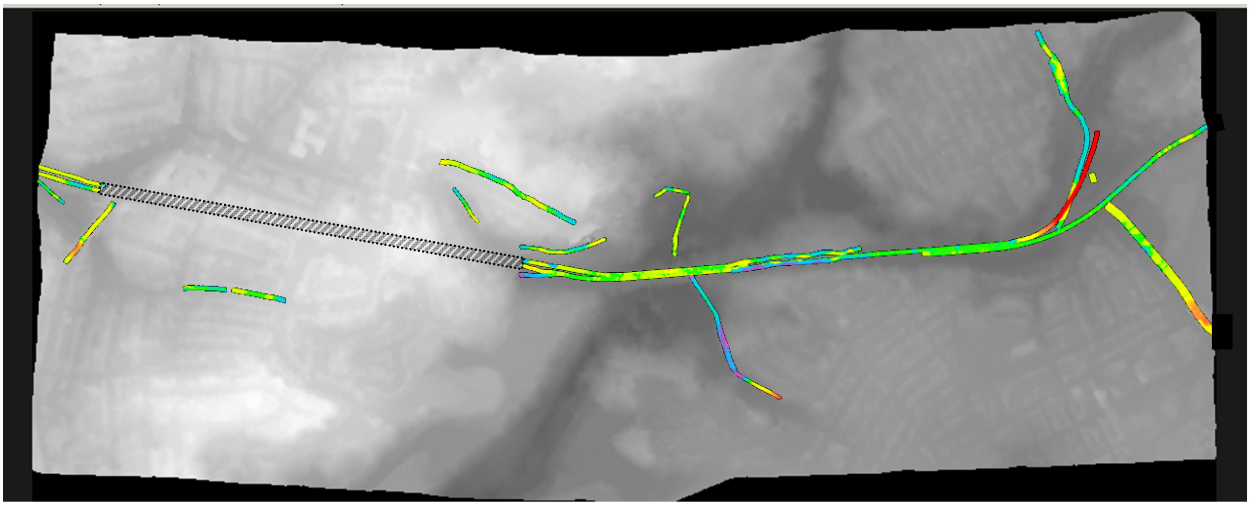
(b) Superimposed on Orthoimage

Figure 17: Pittsburgh Road Error With Respect to Reference Roads.

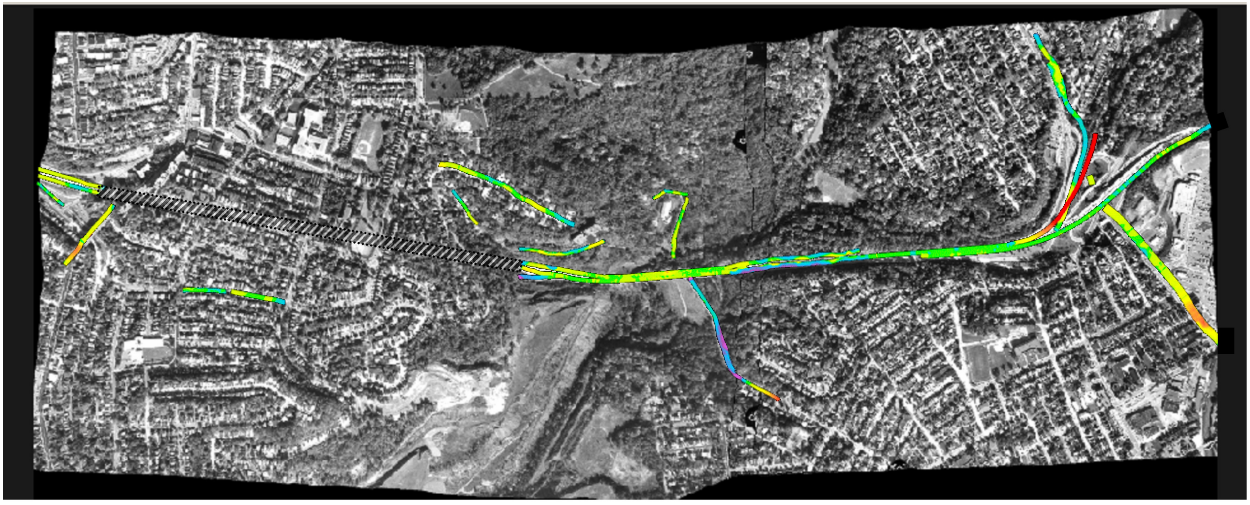
A detailed quantitative analysis of RoadMAP3D has focused our research on real problems and allows for a true measure of the effects of system changes on the extracted roads. From this evaluation, we have determined that RoadMAP3D produces generally good results on major roads, including good 3D accuracy, regardless of orientation with respect to epipolars. It also provides an efficient use of user inputs by automating as much of the process as possible, and providing effective editing and annotation tools to finish the final output.

4.1 Identified Issues With the Current RoadMAP3D System

Our observations of the feature extraction behavior of RoadMAP3D, as well as our quantitative analysis process discussed in Sections 3.4 and 3.5, highlight several problems with our current 3D road extraction process. These are:



(a) Superimposed on 1 meter DTM



(b) Superimposed on Orthoimage

Figure 18: Pittsburgh Road Error With Respect to 1 meter DTM.

- Errors at road ends, particularly overshoots and stopping conditions.
- Road intersection errors including overshoots and undershoots.
- Synchronization (ripples), especially after the **reacquire** stage.
- Gaps in coverage along the extracted road.
- Height discontinuities that could be addressed by post processing.
- Improve the ability for RoadMAP to deal with poor initial match point height estimates.

With this said, RoadMAP3D represents the only fully automated 3D road network extraction system that the authors are aware of. It provides a rapid estimate of the 3D location of each road centerline, generates

an estimate of road width, automatically computes topology, abeit currently 2D topology, and generates its output in an standard commercial format, with user specific attribution.

4.2 Future Work

The issues found during the evaluation of the RoadMAP3D system, as well as discussions with field operators processing road data, have identified the following areas that need to be investigated to move research forward to improve the fully automated production of road networks.

- Integrate a full 3D path model into the current RoadMAP3D system. This would solve synchronization problems and better handle road ends.
- Complete support for using more than two images, which would improve synchronization by using multiple, weighted epipolar constraints.
- Integrate and evaluate the 3D/MSS constraints, which would aid in coping with problems due to shadows and overhanging vegetation.
- Use the multi-image framework to integrate other data sources, including LIDAR and Multi- & Hyperspectral sensor data.
- Investigate 3D topology derived from the extracted road network to improve the consistency and coherence of attributed road network.
- Explore the integration with prior vector data sources. Particularly, to support automated update of previously extracted datasets, in light of more recent imagery with potentially higher spatial resolution.
- Evaluate the use of commercial graphic processing units (GPU) hardware to greatly improve system performance by executing key image analysis and matching for RoadMAP3D on the graphics card.

5 Acknowledgments

We would like to acknowledge the support of Russell S. Harmon, PhD, who is the Senior Program Manager for Terrestrial Sciences at the Environmental Sciences Division of the U.S. Army Research Office (AMSRD-ARL-RO-EV) in Research Triangle Park, NC 27709-2211. Russ provided guidance as to the importance of finding research topics and directions of interest to the U.S. Army, particularly the RDECOM and ERDC centers.

J. Chris McGlone, Senior Systems Scientist, was the Principal Investigator for this project until his departure from the university in April, 2005. His work on 3D photogrammetric road point matching was largely integrated into the RoadMAP system by Mr. Wilson Harvey , Project Scientist and chief architect of the road tracking system.

Mr. Harvey provided overall supervision of the software development for the RoadMAP3d system. He also contributed to automated performance evaluation software development, 3D tracker enhancements, particularly in the areas of 3D reasoning, path maintenance, and data export. Wilson organized several variants of RoadMAP3D, using commercial source control, so that we could perform a variety of experiments while maintaining version control.

Steven D. Cochran, Senior Systems Scientist, joined the project during the third year and was instrumental in preparing the reference datasets for both the Rancho Bernardo and Pittsburgh test areas. Steve generated Digital Terrain Models (DTM's) and well as compiled all of the 3D reference road networks using the ERDAS photogrammetric system as a baseline manual system. Without his work the quantitative comparisons of automatically extracted results, a significant research result, would not have been possible. He also greatly contributed to the performance analysis software development, and the joint analysis of the data with the entire group.

During the first and second year of this contract, three part-time Carnegie Mellon undergraduate programmers were supported under this contract. One independent study by David Camillus was supported in the area of "Failure Analysis in Automated Road Tracking". Funding shortfalls did not allow us to continue undergraduate employment during the third contract year.

Upon the departure of Chris McGlone, David M. McKeown, Research Professor of Computer Science, took over the overall management of this research contract. Mr. McKeown was supported by internal Computer Science Department funds during his participation in this research project.

6 References

[Harvey *et al.*, 2004] Wilson Harvey, J. Chris McGlone, David McKeown, and John Irvine. User-centric evaluation of automated road network extraction. *Photogrammetric Engineering and Remote Sensing*, 70(12):1353–1364, December 2004. 4

[Harvey *et al.*, 2006] Wilson A. Harvey, Steven D. Cochran, and David M. McKeown. Performance analysis of autonomous multi-image road tracking. In *ASPRS Annual Conference*, Reno, Nevada, 1–5 May 2006. American Society for Photogrammetry and Remote Sensing. Accepted. 12

[Harvey, 1999] W.A. Harvey. Performance evaluation for road extraction. *Bulletin de la Société Française de Photogrammétrie et Télédétection*, n. 153(1999–1):79–87, 1999. Colloque « Production de Données Géographiques 3D : vers le Respect des Contraintes Applicatives ». 12, 16