Urban Street Grid Description and Verification

Paper Number 15

Abstract

While two-dimensional maps exist for most urban areas, the descriptions may be incomplete or out of date, or of insufficient resolution for the given application and features such a roads are not described as 3-D objects. Most of the past work on road detection has concentrated on either low resolution, primarily rural roads (usually producing "spaghetti" roads with no notion of intersections), high resolution road following without the topological information of the intersections, or pixel classification where there is no sense of the road as an object. This paper address the problem of extracting a street grid in an urban environment while maintaining the topological information of the intersections.

Starting from an initial seed intersection, which gives the size and orientation of the expected grid, this system uses a feature-based hypothesis and verify paradigm to extract a 3-D description of the street grid. The verification uses the context provided by an intersection model and by an extended street model and other available sensors.

1 Introduction

Many difficult problems remain before the extraction of artificial features in urban areas is completely automated. One of the difficult tasks is the extraction of the street grid. Descriptions of street grids are useful in a number of applications including planning, mapping, street maintenance, and traffic analysis. In this paper we describe a system for extracting approximately regular street grids in a moderately dense urban environment (streets are not completely obscured by buildings). Rather than just detecting road pixels in the image, we generate 3-D descriptions of extended streets that include the intersections connecting the streets. We apply our system on mostly regular urban street grid patterns, the stereotypical US pattern, but always keep in mind the full spectrum of street grids including combinations of regular grids, irregular street patterns, general urban streets, and those where occlusions from adjacent buildings are a major problem.

* This work was supported in part by the Army Research Office under grant No. DAAH04-96-10444.

This work complements the work in low resolution road extraction and high resolution road following, both of which ignore intersections and are not applied in urban areas.

Several factors make the extraction of a street grid difficult:

a) The intensity values depend not just on the surface material but also on illumination and viewing conditions; thus we can not infer material type solely from pixel values. However, multi- and hyper-spectral data can provide some assistance.

b) The street grid is rarely present in isolation. Trees often partially obscure roads. Buildings adjacent to roads are similar width and often similar materials plus they can obscure the roads. Thus low level extraction techniques such as edge detection or region segmentation, give fragmented results with extraneous boundaries and regions.

c) The lack of sufficient spatial or spectral resolution can compound these difficulties.

We address these difficulties by using model based extraction and grouping procedures, global and local context from extended roads and intersections, high resolution imagery, and fusion of multiple images or data sources.

1.1 Previous work

Early road extraction work dealt with low resolution rural roads. This work is best exemplified by the early SRI road work [4]. Their examples included few intersections and roads are extracted separately, so connections between them are not important. This early work also included high resolution road following using the intensity profile of the road. Recent work has begun to address the problems of connections between roads, but still for rural roads [5]. Examples of several efforts in road detection and delineation are in the proceeding of the Ascona workshop [6]. An example of the road following approach, using detected seed roads is in [1, 2]. Rather than direct extraction, another approach is to refine approximate roads using active models such as snakes [9]. These methods ignore the intersections and thus lose the topology of the street grid and generate spaghetti-like roads. This problem is the main focus of our work: the generation of descriptions of the road grid with the topology intact. Another approach [3], uses local context to reduce the problems of occlusions, shadows and intersections, but the program still applies only to mostly well defined rural roads.

2 Road Grid Extraction and Verification

This work uses a hypothesize and test paradigm guided by a strong model of the expected appearance. The model is based on several properties with extraction steps designed for each property.

The first part of the model is the street grid itself. We assume a mostly regular street grid. This means the streets fit on some regular pattern, but there is no expectation that all parts of the grid will be present and small variations in position and width are expected.

Streets meet at intersections. The intersection provides an explicit model for the topology of intersecting streets and a model for local appearance. The grid can be decomposed into a set of intersections where each intersection represents the junction of four potential street segments. The intersection and the individual street segment form the basic elements of our internal representation and ensure that connections between streets at intersections are maintained The intersection model is illustrated in Figure 1.

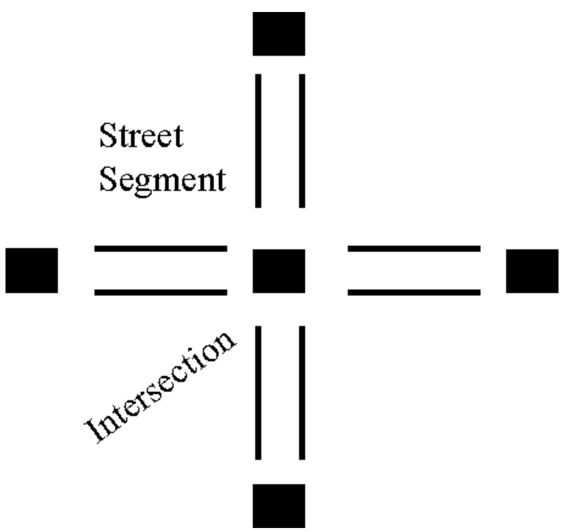

Figure 1 Basic Intersection Model

Streets are modeled as mostly straight linear features of some width. Each street may have a different width and along a given street, the width may change. The extended nature of the street is captured through a ordered list of street segments.

Streets have visible edges in the image. The first three parts of the model dealt with the representation of the street grid, this one addresses the appearance in the image. Thus, the basic detection technique will involve finding edges in the image and matching those edges to the predicted locations from the grid model.

One typical extraction technique is pixel classification, using multi- or hyperspectral data. These methods are based on the assumption that the street appearance is different from the surrounding area and can be distinguished from other features (building roofs, parking lots, etc.) [8]. While this approach is appropriate for some tasks, such as how large an area is covered by roads, it does not provide a three-dimensional model of where the street is located. However, we can use the results of a pixel classification approach as a guide in refining the locations of the road.

Finally, the automated techniques for obtaining high resolution three-dimensional information over large areas allows the use of the property that streets are lower than surrounding buildings. Automated techniques for extraction of digital elevation models included in mapping workstations produce adequate models for the use in refining street positions and eliminating detected street segments that are on buildings.

Each of these assumptions leads to a portion of the final procedure for extraction. We start from a seed intersection, which gives the location, direction and size of the street grid, then verify (and refine) the location and use the refined values to hypothesize the next intersection location. When the initial grid is completed, we use a variety of refinement procedures that adjust the position and width according to image properties, elevation models, or pixel classifications.

2.1 Initial Verification and Road Segment Matching

The initial grid generation procedure is outlined in Figure 2. This phase uses the initial model given by a seed intersection (see Figure 3 for an example), matches line features detected in the image with the predicted lines from the intersection model to find the best position for the intersection, and propagates the model to the next intersection until the full area is tested. Because the position of the intersection (and thus that of the connected road segments) may be changed with each intersection match to fit the image data, this leads to a very different grid than would occur if a uniform pattern was applied to the entire scene.

The model to image matching procedure is derived from prior work on object based change detection, and includes a number of simplifications due to the different assumptions [7]. The 3-D model for each road segment consists of the two 3-D lines corresponding to the two road sides. By assuming that both of these road sides lie on the ground surface, we can predict their location in the image from the camera models and a low resolution terrain model.

The match procedure finds the position where the model lines best overlap the linear features (i.e. extended edges) detected in the image [10]. An accumulation array technique is used for this operation. Every feature pair,

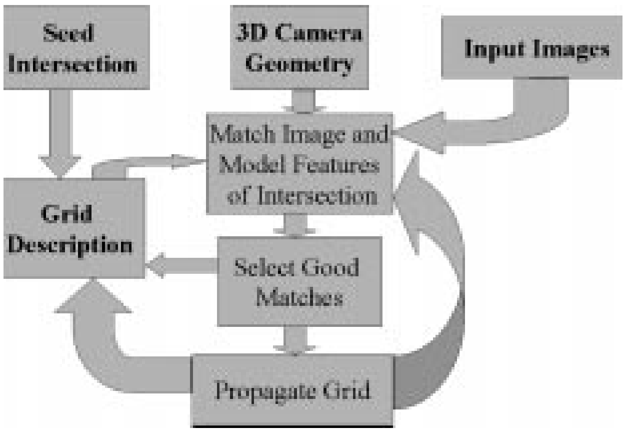

Figure 2 Hypothesize and Verify for the Initial Grid Description

Figure 3 Initial Grid Model

composed of a model line and an image linear feature, is compared and may contribute to the accumulation array. If the orientation difference is too large or the translation exceeds the allowable limits, the contribution is 0. (Note that using these limits (angle and translation) and a spatial index for the line segments not all pairs are tested.) Otherwise, the accumulation array is incremented for the range of translations that would align the two lines. This increments the array over a large area whose shape and size depend on the line lengths and angle differences. See [7] for the details. After all pairs are tested, the best translation is indicated by the peak in the accumulation array. By using the intersection rather than a single road, the translation of the road segments is better constrained.

Each individual road segment is evaluated to determine the its geometric support, defined by the degree of overlap of model features with image features (from 0 for no overlap, to 1.0 for a complete overlap). Using this geometric rating, we group road segments into four broad categories. A good match is one where most of the model line segments overlap the image lines (a threshold around 0.5 is used, with visually clear roads often having a score close to 1.0). These good matches will be used for generating extended streets. Moderate matches are the next 0.2 (i.e. down to 0.3). When the support is below 0.05 the road segment is discarded completely. A fourth category (between the moderate matches and the rejects) is maintained and is checked again in the refinement process (where the rating may change), but these are not used as validated road segments. At this point, extended streets (with intersections) may be computed from the road segments with good matches.

2.2 Road Segment Refinement

After the initial grid generation and verification, many small errors remain. For simple cases the refinement may consist of only correcting the position and width using the original images. For more urbanized areas we use classification information to adjust the positions and elevation models to adjust both the position and rating (good, moderate, etc.) of the individual road segments. Figure 4 outlines the operation of this basic refinement phase.

We compute a more precise alignment for road segments with image features using a version of the accumulation array matcher that allows translation perpendicular to the road orientation and changes in the width of the road. For this step, the allowable translation is inversely proportional to the geometric support (segments with good geometric match are not allowed to move very much). This phase uses triples of road segments (three consecutive segments) to provide context and maintain the straightness of the resulting streets. Width and position refinement accommodates for changes in the actual street widths and allows the use of approximate widths in the initial model. A similar refinement process can be applied to make adjustments in the orientation of the model road segments (i.e. less than 5°), but in practice this has not improved the results.

Width and position refinement can be applied any number of times at increasingly higher resolutions, but the gains after one application are small. Road refinement techniques using active models such as snakes [9]could also be applied to the extracted extended streets, but in our experiments using available snake algorithms, the further improvement is minimal. Given the strong edge support for the initial position of the road segments, this is not surprising. After this refinement, the individual good road segments are connected into extended streets.

This two phase procedure works very well for relatively clean images where there is little interference from adjacent buildings and occlusions from trees are not a problem. Figure 5 shows the result for an image of Yuma, AZ. Note that the grid is not completely regular over the entire image,

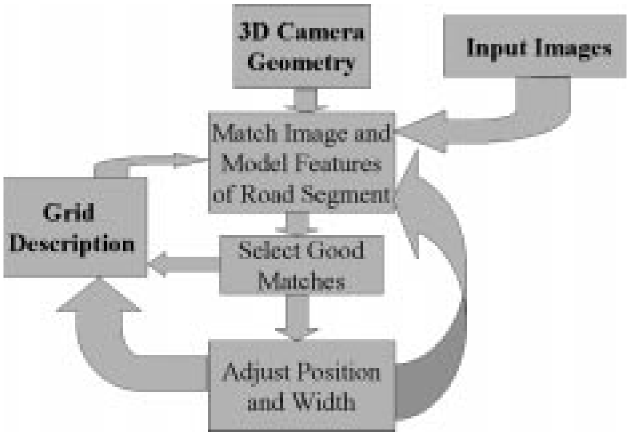

Figure 4 Road Grid Refinement

but the streets do occur at regular intervals. The initial model specified a grid size the same as the smallest blocks in the image.

2.3 Refinement Using Context

The connection of individual road segments into extended streets makes other evaluation and adjustment procedures possible. An extended street is given by finding a sequence of good (e.g. as defined by the geometric matches) road segments. We also include those moderate matches that are bordered on both ends by good matches and ignore those with good geometric ratings which have no support on either end. This connects small gaps where the support is relatively weak (but some evidence exists) and eliminates isolated segments.

The extended street extraction procedure uses the connections provided by the intersections in the model. Within a sequence or road segments, it finds the sequences of road segments classified as good matches. Streets terminate when there are gaps caused by segments with poor geometric matches (the size of the allowable gap can be specified). Extended streets can also be filtered by length to eliminate short streets.

These extended streets provide the context to analyze the individual segments in a number of different ways. The image appearance along an extended street should be more consistent than the appearance of different streets. The width and elevation should be consistent and street should be generally straight (or include a few changes).

We align the road segment and adjust its width to be more consistent with the adjacent road segments within the

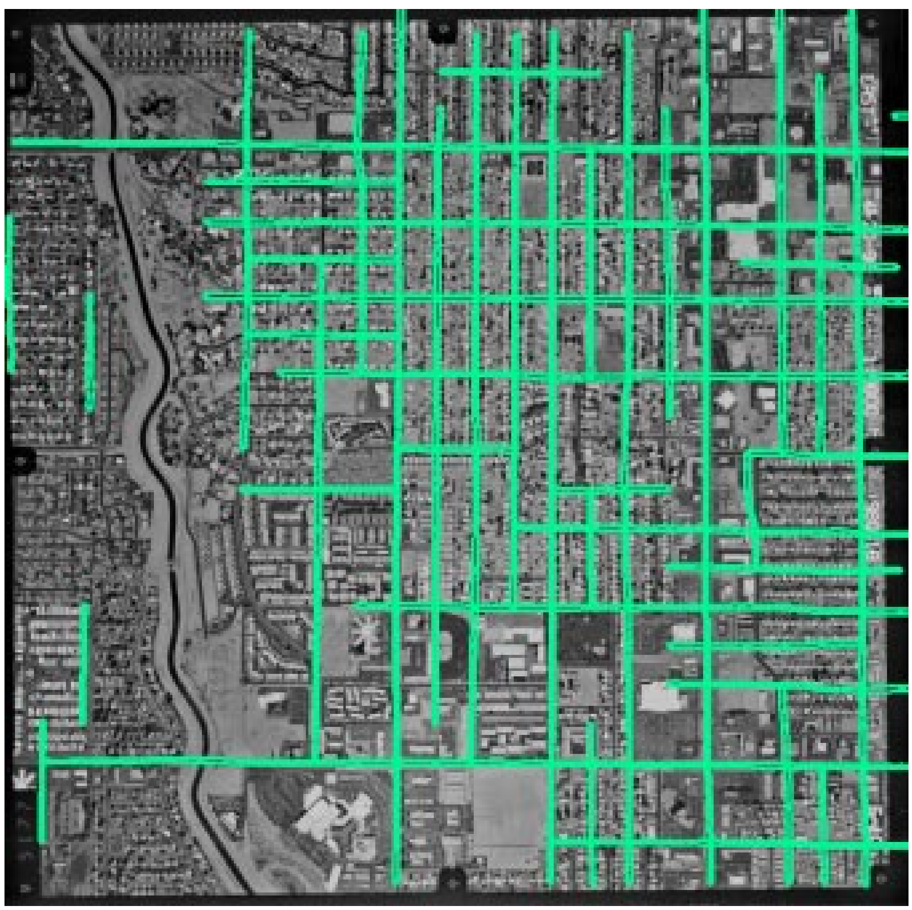

Figure 5 Street Extraction Result, Yuma, AZ.

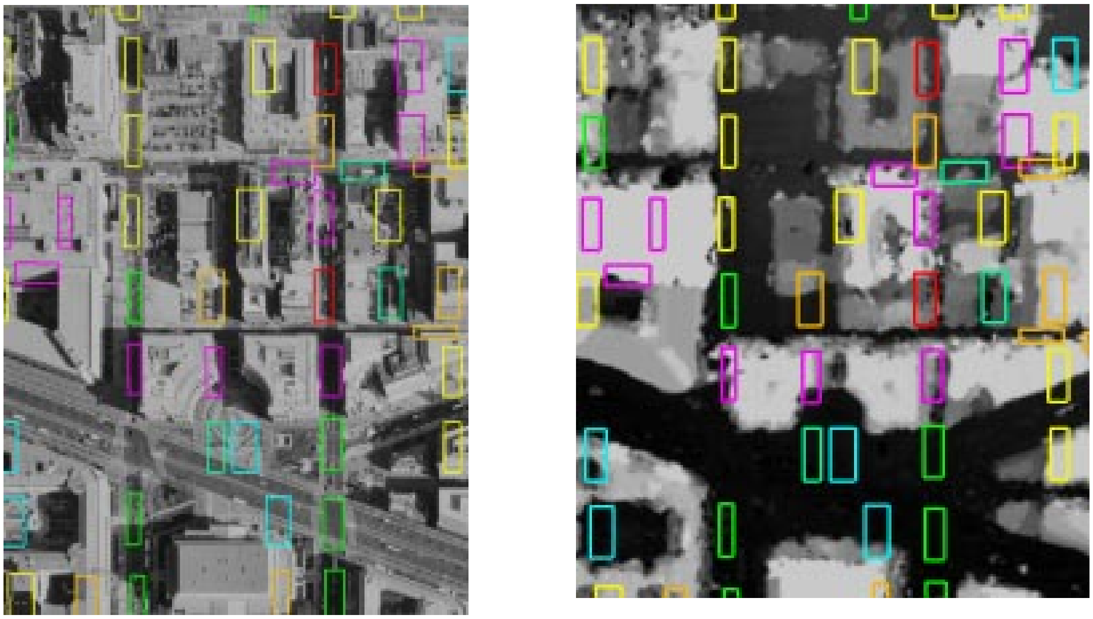

Figure 6 Before Using DEM for positioning. Road segments on image and on computed DEM. Green/Blue indicate relatively low elevation, red/purple relatively high.

extended street. This often corrects for the errors caused by features parallel to the streets (e.g. sidewalks). Long term shifts in the alignment of the street are still allowed since this refinement is applied only if segments on both sides indicate a similar change.

2.4 Refinement Using Multiple Images

Both the initial grid generation and refinement are computed over multiple overlapping images. By choosing the best match among the (usually two) possible matches, the common case where buildings occlude streets in oblique views or near the edges of nadir images can be handled.

A second, important capability of multiple images is the possibility of extracting 3-D information using standard stereo analysis techniques. We use automatically extracted digital elevation models (DEMs) to aid in refinement of the road segments. While a raw computed DEM has many problems, especially near buildings, it provides a good approximation of the elevation of the area of the road segment and indicates when a road segment is higher or lower than the others in its extended street. Also, it can indicate local elevation minima that usually correspond to the gaps between buildings where the street is located. Figure 6 shows a small portion of an image with the selected road segments color coded: segments with consistent elevation are shown in greens, those with inconsistent elevation (in order of increasing elevation) in yellow, red and magenta. In this case, consistent means that the elevation under the road segment in the DEM is similar to that for the extended street, inconsistent means the elevation is much higher than the average for the extended street.

We use the DEM in two ways to refine the results. The road segment is shifted (perpendicular to its primary direction) to a minimum elevation location. Rather than allowing arbitrary shifts, the distance is limited according to the quality of the geometric match (with a perfect match the segment will not be shifted). The results of this refinement step are shown in Figure 7 where the consistency measure has been recomputed using the new locations of the road segments. Even after shifting, some segments may still fall on buildings. In the second use of the DEM, we eliminate these segments from the set of good matches and recompute the extended streets.

For some sites we have multi- or hyperspectral data that can be used for the classification of image pixels as road or other material. This data is usually lower resolution than the image data and will contain classification errors due to occlusions from trees or cars, but it aids in better locating the road features. We have a classification map generated for the DC data set [8] with one class for road (and a second for gravel paths). Using this map we can, in a manner similar to the shifting of road segments to local minima on the DEM, force the alignment of the road segments to more completely cover the points classified as road. Using the classification in this manner means occlusions or missing pixels are not a problem since the road segment covers a large area and perfect classifications are not required. Figure 8 shows some of the changes when this refinement is applied to the DC dataset. The primary improvement is in the alignment of some streets in areas where the DEM would provide no assistance. The classification-assisted results include more shorter segments since the better alignment with streets allowed this limit to be relaxed.

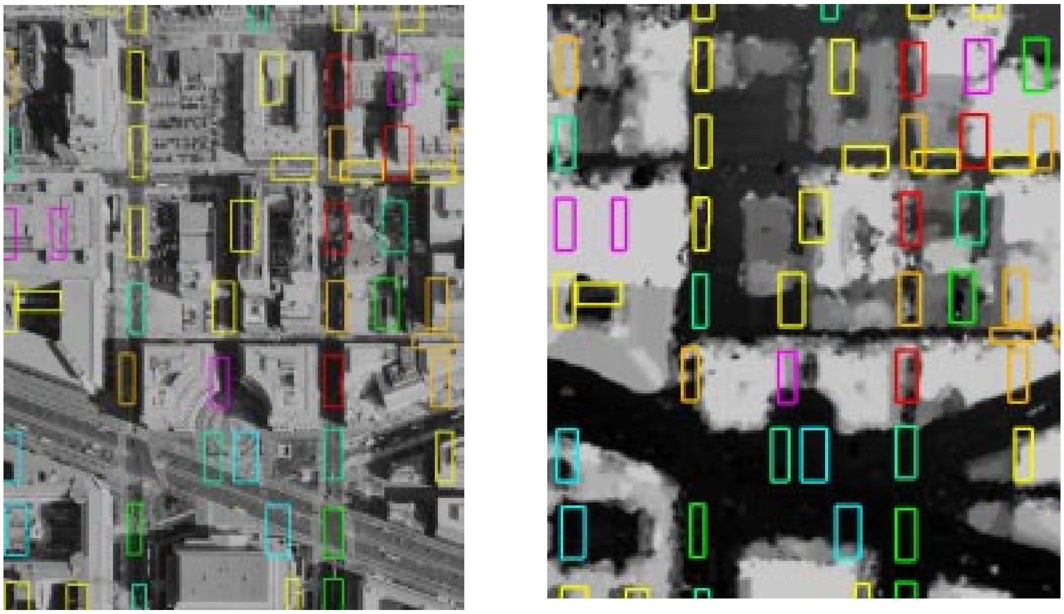

Figure 7 The same area after DEM adjustments.

3 Results

This system has been run on a variety of sites including the Ft. Hood dataset used for the DARPA APGD program; Yuma, AZ; images of West Lafayette, IN; and images of Washington, DC. The only parameters that are varied are those which are site dependent, such as the allowable widths of streets and the initial seed intersection. More examples of results with intermediate results and detailed timings are available from the web site. The overall computation times depend on the size and number of images and the limits on the allowable translation and width variations. For the initial intersection-based verification, 60% of the time is in edge and line extraction from the image windows with the rest dominated by the accumulation array computation. In the width refinement, two-thirds of the time is in the match accumulation procedure (since the width test involves multiple applications of this procedure). Total execution time is roughly 60 minutes for the intersection verification (using 3 2000X2000 images and 3000 road segments) and 220 minutes for the refinement. After the

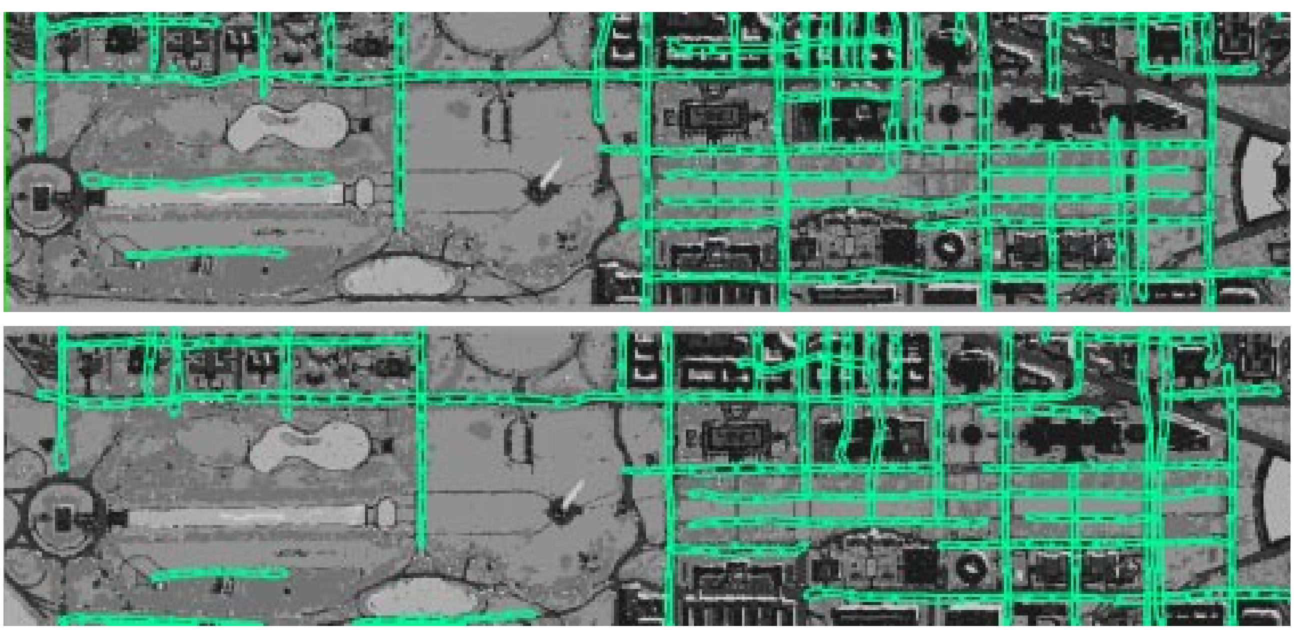

Figure 8 Extracted streets with no classification information (Top) and with classification information (Bottom)

refinement, roughly 36Km of streets are extracted.

By running the procedure with several different initial seeds over a number of images, a more complete set of streets is possible. Figure 9 shows the accumulated results for the DC data displayed on an orthophoto of part of the site (3 additional images, north of the mall area, were used to generate the results but were not included in the orthophoto generation).

A detailed analysis of the results shows one common error: road segments widths are incorrect due to matching other strong edge features (sidewalks, stripes, etc.). Without the use of the DEM or classification map another error is that streets are misplaced by the width of the street (i.e. the left side of the model matches the right side of the actual street and the right side of the model matches some other structure parallel to it).

4 Conclusions

We have presented a new approach for detecting and describing urban road grids from aerial imagery which maintains the topological structure of the pattern. This procedure uses a model-based approach that easily incorporates multiple overlapping images to extend the area covered and can incorporate other forms of image data to improve the results. Further analysis is needed to provide detailed analysis of completeness and accuracy for the extraction results.

5 References

[1] Barzohar, M., Cooper, D.B., Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation, *IEEE Trans. Pat. Anal. and Mach. Intel., Vol. 42.* No. 7, July 1996, pp. 707-721.

[2] Barzohar, M., Cohen, M., Ziskind, I., and Cooper, D.B., Fast Robust Tracking of Curvy Partially Occluded Roads in Clutter in Aerial Images, in (Gruen, 1997), pp. 277-286.

[3] Baumgartner, A., Steger, C.T., Mayer, H., Eckstein, W., Ebner, H., “Automatic Road Extraction Based on Multi-Scale, Grouping, and Context,” P*hotogrammetric Engineering and Remote Sensing, Vol. 65*, No. 7, July 1999, pp. 777.

[4] Fischler, M.A., Tenenbaum, J.M., and Wolf, H.C., “Detection of Roads and Linear Structures in Low Resolution Aerial Images Using Multi-Source Knowledge Integration Techniques,” *Comp. Graphics and Image Processing, 15(*3), March 1981, pp. 201-223.

[5] Fischler, M.A., Heller, A.J., Automated Techniques for road Network Modeling, Proc. DARPA Image Understanding Workshop, Nov. 1998, pp. 501-516.

[6] Gruen, A., Baltsavias, E., Henricsson, O., **Extraction of Man-Made Objects from Aerial and Space Images, II**, Birkhauser Verlag, 1997, Proceedings of meeting held in Ascona, Switzerland, May 5-9, 1997.

[7] Huertas, A., and Nevatia, R., Detecting Changes in Aerial Views of Man-Made Structures, Proc. Intl. Conf. on Computer Vision, January 1998, pp. 73-79.

[8] Jimenez, L.O., Landgrebe, D.A., “Hyperspectral Data Analysis and Supervised Feature Reduction Via Projection Pursuit,” IEEE Trans. Geoscience and Remote Sensing, Vol 7, No. 6, November 1999, pp. 2653.

[9] Neuenschwander, W.M., Fua, P., Iverson, L., Szekely, G., Kubler, O., Ziplock Snakes, *Intl. J. of Computer Vision, Vol. 25,* No. 3, Dec. 1997, pp. 191-201.

[10] Nevatia, R. and Babu, K.R., Linear Feature Extraction and Description, *Computer Graphics and Image Processing, Vol. 13*, 1980, pp. 257-269.

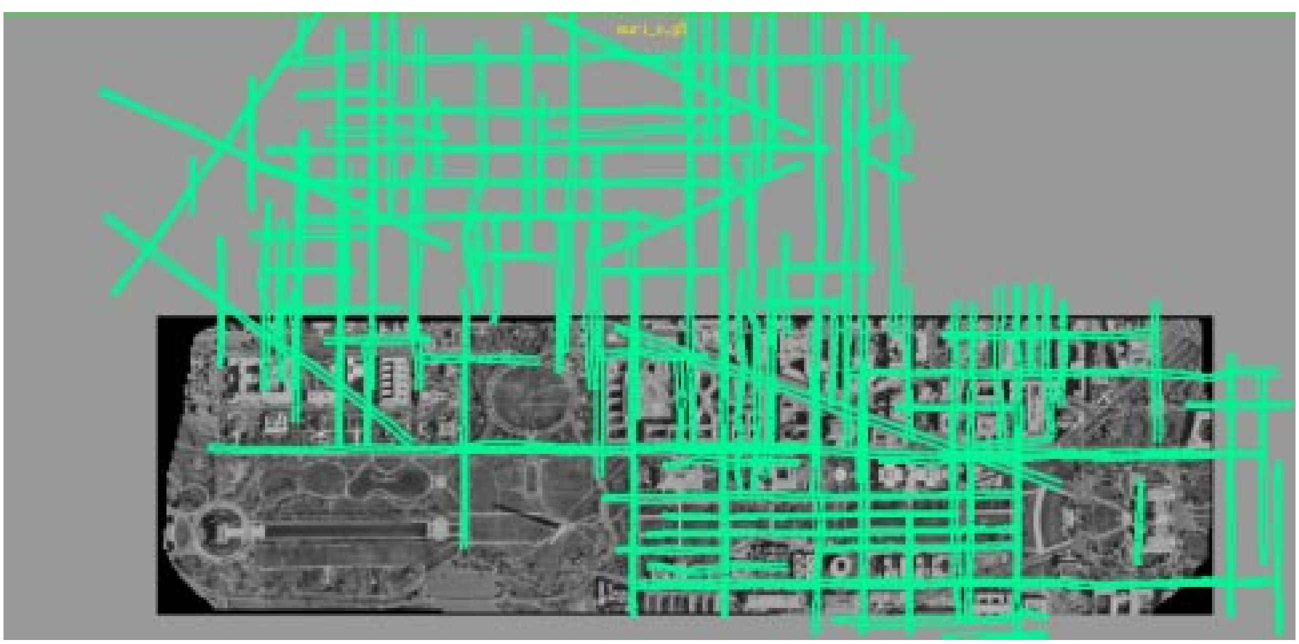

Figure 9 Accumulated results for the Washington, DC dataset.