Road Grid Extraction and Verification

Keith Price
Institute of Robotics and Intelligent Systems
University of Southern California
Los Angeles, CA 90089-0273, USA
price@usc.edu
http://iris.usc.edu/Outlines/muri/roads
Commission III, Working Group 3

Keywords: Street grid extraction, Cartography

Abstract

While maps exist for most urban areas, there are many locations where the information is not accurate, it may be out of date, or it may be incomplete or of insufficient resolution for the applications. Many difficult problems remain in automated cartography. One of them is the extraction of a street grid in an urban environment. Much of the work on road detection has concentrated on either low resolution, primarily rural roads (usually producing “spaghetti” roads with no notion of intersections) or high resolution roads without the topological information of the intersections. This paper address the problem of extracting a grid with the topological information intact.

Given an initial seed intersection, which gives the size and orientation of the regular grid, this system uses a feature-based hypothesis and verify paradigm to find the street grid. Verification uses local context, provided by an intersection model and by an extended street model, and any available sensors.

1 Introduction

Many difficult problems remain before the extracton of artificial features for the creation of maps of urban areas is completely automatic. One of the difficult problems is the extraction of the street grid. Street grids are useful in a number of applications including: urban planning, mapping, street maintenance, and traffic analysis. In this paper we describe a system for extracting regular street grids in a moderately dense urban environment (streets are not significantly obscured by buildings). Rather than just marking road pixels in the image, we generate descriptions of extended streets including the intersections connecting the streets. This complements both the work in low resolution road extraction and high resolution road following, since that work usually ignores intersections and is not applied in urban areas.

The task of extracting a street grid has difficulties caused by several factors:

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 directly from pixel values. However, multi-spectral and especially hyper-spectral data can provide some assistance.

b) The street grid is rarely present in isolation. Roads may have trees occluding them or buildings and other structures adjacent to them. Thus low level segmentation techniques such as edge detection or region segmentation, give fragmented results with many 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 multiple data sources.

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

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 (Fischler, 1981). 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. Their recent work has begun to address the problems of connections between roads, but still for rural roads (Fischler, 1998). Several other recent examples of efforts in road detection and delineation are in the proceeding of the most recent Ascona workshop (Gruen, 1997). An example of the road following approach, using detected seed roads is (Barzohar, 1996 and 1997). Rather than direct extraction, another approach is to refine approximate roads using active models such as snakes (Neuenschwander, 1997). These methods ignore the intersections and thus lose the topology of the street grid and generate spaghetti-like roads. This is the problem that is the main focus of this work: the generation of descriptions of the road grid with the topology intact.

We apply our system on regular suburban street grid patterns, the stereotypical midwestern 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.

1.2 Assumptions

We use a simple three-dimensional road segment and intersection model and have known camera parameters, which allows us to use either nadir or oblique views. We assume that roads have visible edges without significant occlusions. Since we use the geometric structure of the road and the intersection, vehicles and markings on the road are not a serious drawback. Indeed, for verification, they may be an important feature. We also assume a regular street grid but that the program must detect where the regular grid ends. Some variations in the grid are detected by using a grid model that is smaller (e.g. 1/2 or 1/3) than the actual street grid.

2 Road Grid Extraction and Verification

This experimental system requires only a few interactive steps, which could be performed by imperfect automated techniques that have been suggested in the literature. By delaying total automation, we are able to focus on the important issues of using context and grouping for street grid extraction. The only inputs from the user are three points (i.e. the center of three intersections) that give the location, direction and spacing of the street grid. This step can be replaced by automatic methods to find dominant direction and spacings, but these are less reliable and not a focus of this current work. These three points define an initial model containing four road segments and the intersection that they define. An example of this initial model is shown in Figure 1. Each intersection has four road segments and each road segment connects two intersections. The verification problem is to determine whether each of the road segments exist. Figure 2 shows a small portion of such a grid model which must be verified. Since actual road widths (in meters) vary from scene to scene, we allow the user to adjust the default width to fit the particular scene. A width refinement step, later in the extraction procedure, reduces the need for exact initial widths.

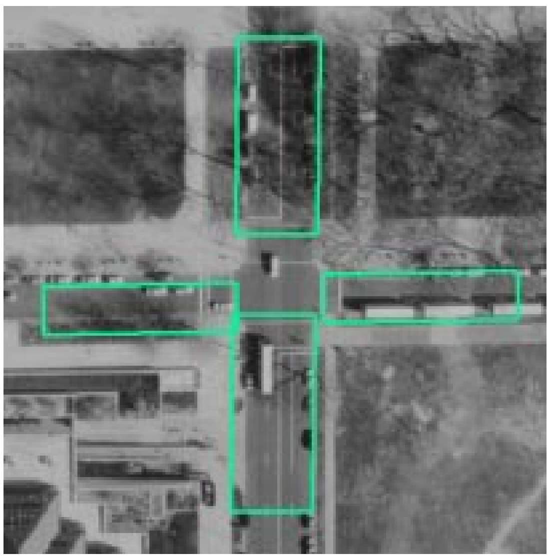

Figure 1 Initial Grid Model

The grid extraction procedure is composed of two phases, the first tests each intersection (using the four road segments for the model) to find which ones are supported by the image data. This hypothesize-and-verify phase propagates the grid across the entire scene and provides an initial geometric match for the scene. Figure 3 outlines this phase of the procedure. The second phase uses this initial match and tests triples of road segments (three consecutive road segments) to find the best location and width for each triple. Figure 4 illustrates this second phase. These results provide the input for further use of context and refinement using other data sources.

2.1 Initial Verification and Road Segment Matching

The first phase starts from the initial model and propagates the model to the next intersection until the full area is tested. During the matching step, small translations in the intersection to fit the image data are found and are reflected in the propagated model (i.e. the result is not the same as covering the images with an initial uniform grid).

The model to image matching procedure is derived from prior work within our research group on building verification, but includes a number of simplifications due to the different assumptions (Huertas, 1998). The 3-D model for each road segment consists of the two 3-D lines corresponding to the two road sides. Since both of these road sides lie on the ground surface, we can predict their location in the image. The match procedure finds the position (i.e. the optimal translation) where the model lines best overlap the linear features (i.e. extended edges) detected in the image (Nevatia, 1980). An accumulation array technique is used to find the optimal translations in two dimensions. Every feature pair, 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. Otherwise, the accumulation array is incremented for all the translations which could be valid for the two lines. This increments the array over a large area whose shape and size depend on the line lengths and angle differences. See (Huertas, 1998) for the details. After all pairs are tested, the best translation is indicated by the peak

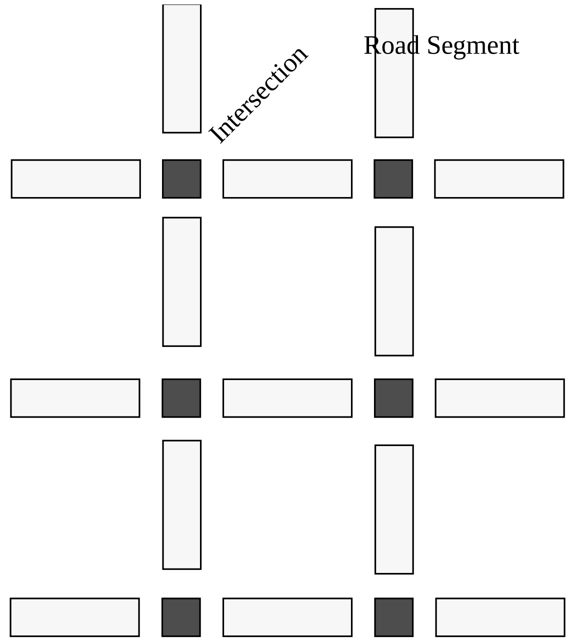

Figure 2 Small portion of road grid model

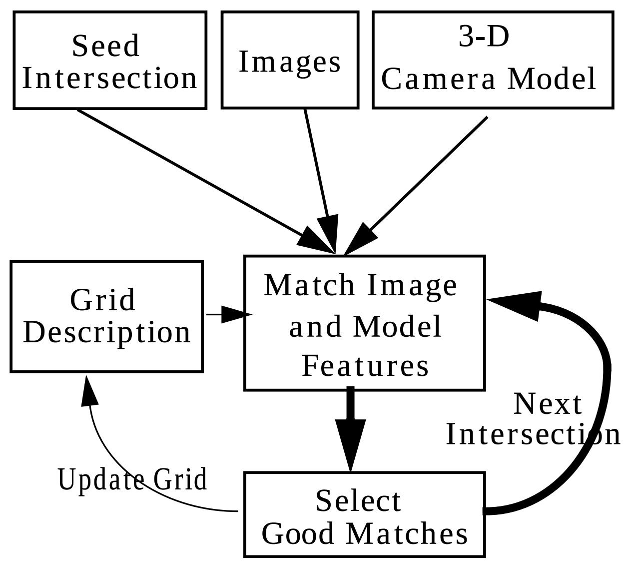

Figure 3 Road Grid Initial Verification. Bold arrows show program flow, thin arrows are data.

in the accumulation array.

Given only two parallel lines from one road segment, the optimal translation could be anywhere along the extended street. But, by using the four road segments that form the intersection (thus eight line segments in two different directions) the translation is better constrained.

We use the same matching procedure for the second phase of refinement. In this case rather than looking for a two-dimensional translation we are searching the two-dimensional space that includes the translation perpendicular to the road direction and the road width.

This initial match using a rigid model of four road segments forming an intersection finds the translation of the intersection that will best line up with the image. Using this translation, the model is shifted so that it is aligned with the image linear features, thus resulting in a change of the 3-D position of the model. Allowable changes in position from the initial predicted location are small.

Each individual road segment is evaluated to determine the its geometric support as defined by the degree of overlap 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 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). 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, but these are usually not considered validated road segments.

The processing can stop at this point, but many parts of the road will not have good matches due to variations in the position or width. To get more complete results we use the refinement procedure.

2.2 Road Segment Refinement

After the initial propagation and verification, many small errors remain. We find a more precise alignment of road segments with image features by using a version of the accumulation array matcher that only 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 rating (good matches 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. This refinement accommodates for changes in the actual street widths and allows errors in the initial model width. A similar refinement process can be applied to make small adjustments in the orientation of the model road segments (i.e. less than 5°), but in practice this has not improved the results. Figure 4 outlines the operation of this refinement phase.

The width and position refinement could be applied any number of times at higher and higher resolutions, but the gains after one application are small. Road refinement techniques using active models such as snakes (Neuenschwander, 1997) could also be applied to the extracted extended streets. In experiments using available snake algorithms, the further improvement is marginal. Given the substantial edge support for the initial position of the road segments, this is not surprising.

2.3 Refinement Using Context

When the system has completed the initial grid propagation and the width and position refinement we use local and global context for further evaluation. Primarily this is the appearance of the road segment relative to its neighbors and its extended street in a number of different ways.

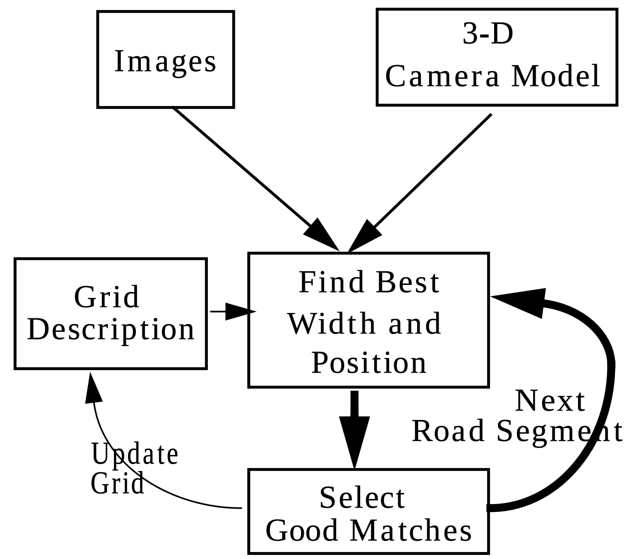

Figure 4 Road Grid Refinement

The intersection information determined which road segments are used in the first two steps described above. A more important use is to extract extended streets from the verified grid. An extended street is given by finding a sequence of good (as defined by the geometric matches) road segments. Since the extraction depends on road segments classified as good matches, we first promote those moderate matches which are bordered on both ends by good matches and remove those matches with good geometric ratings which have no support on either end. This fills in small gaps where the match is relatively weak (but where some evidence exists) and eliminates many isolated matches.

The extended street extraction procedure is straightforward: find each sequence of road segments using the connections provided by the intersections in the model; within this sequence find the sequences of road segments classified as good matches; terminate the extended street when there are gaps caused by poor geometric matches. Extended streets can be also filtered by length to eliminate short streets.

The extended street provides 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 relatively consistent. The street should be generally straight.

Using these constraints, we could evaluate the appearance of the individual road segments relative to the overall street and eliminate any that are different. (This appearance based evaluation is part of our future current work.)

We refine the position and width of road segment to be more consistent with the position and widths of the adjacent road segments within the extended street. This alignment based on context 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 change.

2.4 Refinement Using Multiple Images

All of this processing occurs over multiple overlapping images. The initial propagation extends over multiple images with the match rating determined by the image that has the best match. This helps with the common case of buildings occluding streets near the edges of nadir images.

A second, important property 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 DEM has many problems and may not be exact, it provides a good approximation of elevation to determine when a road segment is higher or lower than the others in its extended street. Figure 5 shows a small portion of one image with the matched road segments color coded: consistent segments shown in greens, inconsistent segments (in order of decreasing consistency) in yellow, red and magenta. In this case, consistent means that the average elevation of the road segment as given by the DEM is similar to the average for the extended street, inconsistent means the elevation is much higher than the average for the extended street.

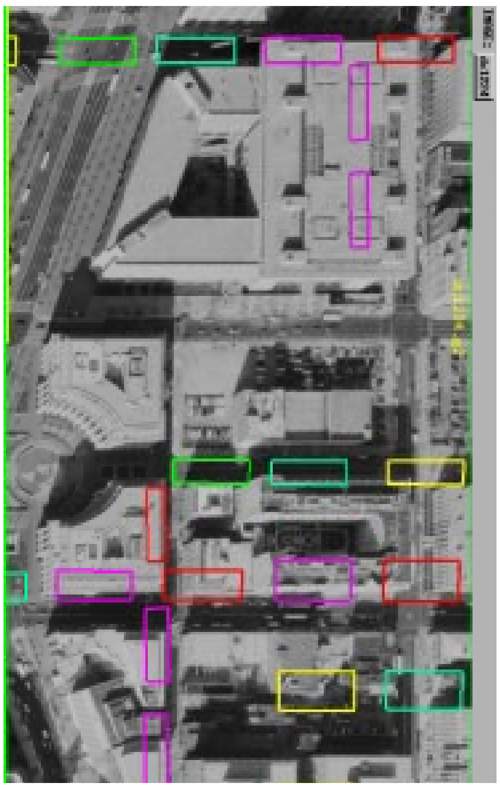

Figure 5 Initial DEM Consistency Measure

The DEM is used in two ways, first 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 6 where the consistency measure has been recomputed using the new locations of the road segments. Even with the shifts, the worst segments are still bad (i.e. magenta). We then eliminate these inconsistent road segments from the set of good matches and recompute the extended streets.

3 Results

This system has been run on three different sites including the Ft. Hood dataset used for the DARPA APGD program, images of West Lafayette, IN from Purdue, and images for Washington, DC. The only parameters that are varied are those that are site dependent, such as the allowable widths of streets and the initial seed intersection. The web site given at the top has more examples, many intermediate results, and detailed timing results. The overall computation times depend on the size (and scale) and number of images and

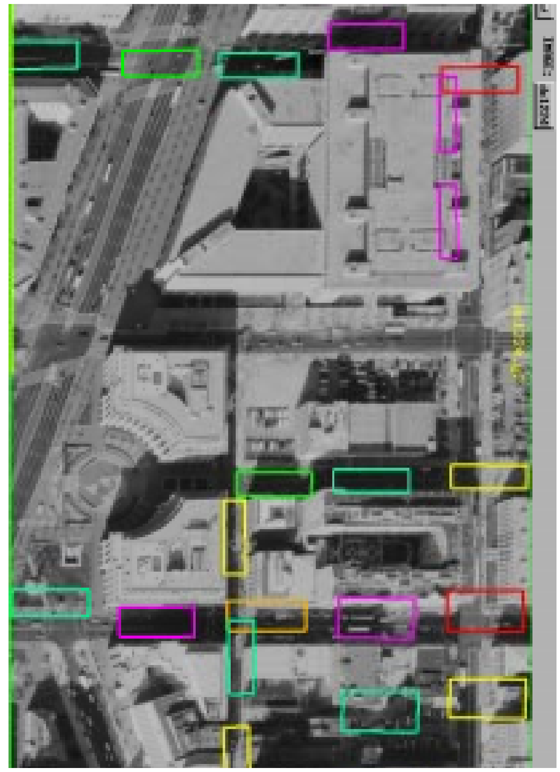

Figure 6 After using DEM

the limits on the allowable translation and width variations. For the initial intersection-based verification, three-quarters 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, three-quarters of the time is in the match accumulation procedure.

3.1 Results on a Washington DC Dataset

The final results of the extracted street grid are shown in Figure 7. projected onto an orthophoto corresponding to the extracted DEM. This orthophoto only covers most of the area of 3 of the images used for the extraction so some extracted roads are displayed off the image. The colors are used to indicate one (arbitrarily) selected street in red, with its intersecting streets in yellow and all the others in green.

The time for the initial verification (approximately 2000 intersections) was roughly 90 minutes (covering five 2000X2000 images). The refinement using the same 5 images and testing about 3400 road segment triples (some are tested in multiple images) was about 500 minutes. After all the refinement steps, approximately 46Km of streets are extracted.

A detailed analysis of these results shows one common error: road segments that are misplaced by the width of the road (i.e. the left side of the model matches the right side of the actual road and the right side of the model matches some other structure parallel to the road). These errors are caused by weak boundaries for the road itself and stronger edges in features parallel to the road.

3.2 Continuing Work

Our future work will continue the use of context and other available imagery (color, multi-spectral, multiple images, elevation maps, etc.). While the example shown here uses a perpendicular grid, there is nothing that depends on that exact arrangement.

4 Conclusions

We have presented a new approach for detecting and describing road grids from aerial imagery which maintains the topological structure of the pattern. This procedure uses multiple overlapping images to extend the area that can be covered and to provide rough elevation information for elimination of errors.

5 References

(Barzohar, 1996) 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., 42(*7), July 1996, pp. 707-721.

(Barzohar, 1997) 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.

(Fischler, 1981) 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.

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

(Gruen, 1997) 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.

(Huertas, 1998) 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.

(Neuenschwander, 1997) Neuenschwander, W.M., Fua, P., Iverson, L., Szekely, G., Kubler, O., Ziplock Snakes, *Intl. J. of Computer Vision, 25(*3), Dec. 1997, pp. 191-201.

(Huertas, 1990) Huertas, A., Cole, W., and Nevatia, R., Detecting Runways in Complex Airport Scenes, *Comp. Vision Graphics and Image Processing, 51(*2), August, 1990, pp. 107-145.

(Nevatia, 1980) Nevatia, R. and Babu, K.R., Linear Feature Extraction and Description, *Computer Graphics and Image Processing, 13*, 1980, pp. 257-269.

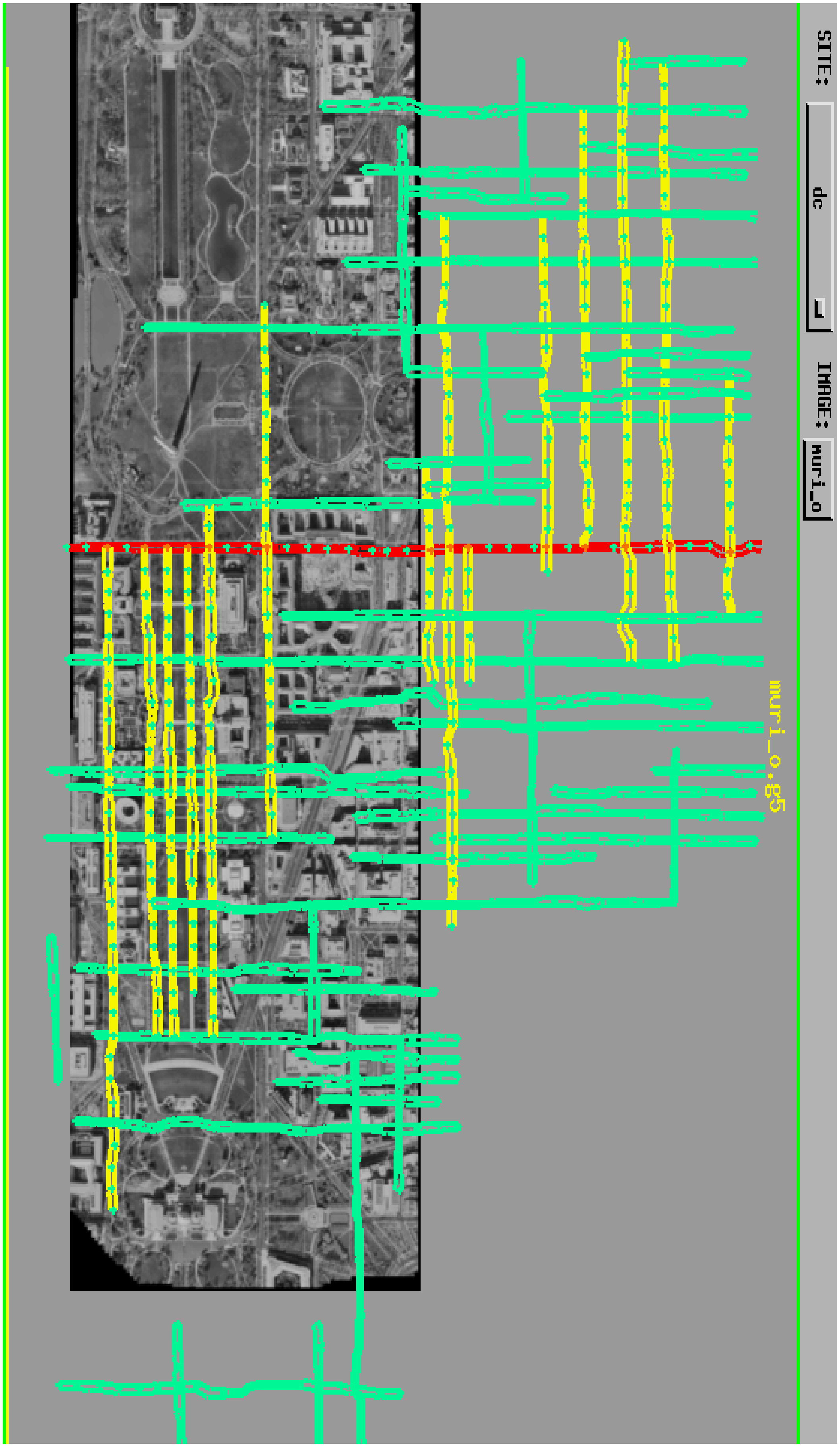

Figure 7 Extended streets combined on one oblique view.

Road Grid Extraction and Verification

Keith Price
Institute of Robotics and Intelligent Systems
University of Southern California
Los Angeles, CA 90089-0273, USA
price@usc.edu
http://iris.usc.edu/Outlines/muri/roads
Commission III, Working Group 3

Keywords:

Abstract

While maps exist for most urban areas, there are many locations where the information is not accurate, it may be out of date, or it may be incomplete or of insufficient resolution for the applications. Many difficult problems remain in automated cartography. One of them is the extraction of a street grid in an urban environment. Much of the work on road detection has concentrated on either low resolution, primarily rural roads (usually producing "spaghetti" roads with no notion of intersections) or high resolution roads without the topological information of the intersections. This paper address the problem of extracting a grid with the topological information intact.

Given an initial seed intersection, which gives the size and orientation of the regular grid, this system uses a feature-based hypothesis and verify paradigm to find the street grid. Verification uses local context, provided by an intersection model and by an extended street model, and any available sensors.