INTEGRATION OF OBJECT AND FEATURE MATCHING FOR OBJECT SURFACE EXTRACTION

Pakorn APAPHANT*, James BETHEL**
*National Research Council Thailand
Remote Sensing Division
Pakorn@hammerhead.nrct.go.th
**Purdue University
School of Civil Engineering, USA
Bethel@ecn.purdue.edu

KEY WORDS: Matching, Digital Elevation Model, DEM, Dynamic Programming, Signal Matching, Feature Matching

ABSTRACT

A stereo matching algorithm is developed for object surface construction. The matching problems are addressed by the integration of signal and feature matching. The innovative strategy arises within the framework of global optimization of the match function. A conventional feature matching method based on dynamic programming is investigated and extended in this research. Several types of primitive features are extracted and matched. The object coordinates from the results of feature matching are used as weighted constraints during the signal matching process. In order to evaluate the performance of the algorithm, images of both urban and rural scenes have been tested. The experiments have shown promising results using this approach.

1 INTRODUCTION

The construction of 3D urban area spatial models from aerial images is a difficult problem that continues to challenge researchers in the image understanding and photogrammetry fields. The main problems are occlusion, large parallax ranges, and feature textures which are not amenable to the matching process. An urban area image by definition contains numerous cultural features. Feature based matching has been used to address these problems. However, if only matched feature information is exploited, it is not possible to reconstruct a surface model completely. Signal based matching has been used for solving such problems as well.

In order to achieve results in the presence of significant occlusions, the feature matching problem can be addressed via dynamic programming, DP. This technique permits orphan features which have no match on the other photograph. The preliminary object surface generation can then be reduced to the problem of an optimal profile search from a cost matrix generated by features from a given stereo pair. This matching technique generally consists of three parts which are (1) extracting features to be matched along with their descriptive attributes, (2) generating a correspondence cost matrix, and (3) searching for the optimal path in the cost matrix.

In this investigation, some of the problems found in stereo matching, as mentioned above, will be addressed by a strategy of combined signal and feature matching based on a global or semi-global search technique. The sequence of steps for the feature matching component would entail extracting features along epipolar lines in the images, and matching the features as suggested above by a DP-based approach. The feature coordinates in the object space are then determined. After that a signal matching technique with feature constraints is applied to generate the elevation profiles via another variant of dynamic programming. Once all profiles in a model are determined, the surface in the object space can finally be constructed.

2 FEATURE EXTRACTION

Several types of low level features are extracted from images, along with their positions in epipolar space. They are low level features which contain useful information for the matching process. Some features used in this research are already well defined from elsewhere, such as the straight line feature, and some are developed here such as end points of a plateau feature, and the spike point feature. For all of these features, either the extraction is done directly in the 1D epipolar space, or the extraction is done in the 2D image space, and subsequently intersected with the 1D epipolar lines.

2.1 Edge Feature

Edge segments are commonly used as a primitive feature in stereo matching. This feature appears in many images. In a rural area image, it will not be possible to find strong geometric features as straight lines, circles, corners, etc. These geometric features are usually parts of man-made structures such as buildings and roads. However, it is still possible to extract edges from rivers, trails, ridges, shadows, and geomorphic features in the images. There are several well developed edge detection methods. Strong edges with low noise would be ideal. Any detectors based on the first derivative can fulfill the requirements.

2.2 Line Feature

A straight line is one of the most prevalent features used for analyzing urban area images. Because of occlusions and the weakness of area based matching when abruptly changing elevations are present, features become an important factor for handling this kind of image. Some researchers have commingled straight lines and edge segments. Actually it would be more appropriate to separate straight lines from non-straight edges. Much prior work has been done on line extraction. The existing line extraction tool, *Burn's* line extraction from the *Khoros* software, has been used in this research.

Figure 1. Intersections of Epipolar Line and Extracted Lines

The final output of the extraction is given in term of the starting and ending coordinates of lines. Their endpoints can be used to compute some useful attributes for matching such as line orientation, and length of a line. Figure 1 illustrates the intersection between an epipolar line (labeled) and extracted lines overlaid on the original images.

2.3 Plateau Feature

In manual DEM collection, one of the features that human operators focus on are the boundary points of homogeneous regions. A plateau feature is defined here as the end points of a 1D region with near constant grayscale. It will be detected from an intensity profile of the points along an epipolar line. This means that it is a one-dimensional feature. The plateau is particularly suited for urban imagery. Usually it represents an entity such as a rooftop, roof panel, field, etc. If a profile is plotted, this feature is usually recognized easily by the human eye (Figure 2) i.e., it looks like a plateau.

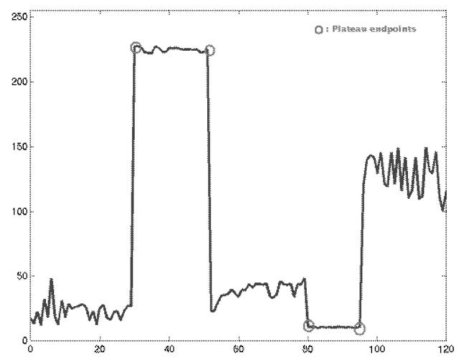

Figure 2. Ideal Plateau Features on a Synthetic Intensity Profile

Attributes of the Plateau as well as its position are retained to be compared during the matching process.

2.4 Spike Feature

In two-dimensional space, several methods have been successfully developed for point feature extraction. The spike feature is characterized by the appearance of a sharp peak or valley within a homogeneous region of an intensity profile. Like the plateau feature, ideal spike points should be easily detected from a profile by human eyes. In an image, spikes may include both point features such as manholes, and also linear features which cross the epipolar line. From this point of view, it is possible that some spike points may produce multiple detections as both spike points and line features. Actually one of the objectives in pursuing spikes was to recover missing straight line points. Although some features may be extracted twice, the matching logic precludes matching across feature types.

3 CONSTRAINED OPTIMAL SURFACE TRACKING (COST) ALGORITHM

3.1 Feature Based Matching

The features described are extracted and tagged with attribute values. After the extraction process, the features for each line are identified and ordered along each line independently from left to right. In a 2D (feature only) cost matrix, the feature points on a density profile from one image represent stages (or columns) while the ones from the other image represent the possible states (or rows) for each stage.

Correspondence metrics are defined in order to quantify the similarity of pairs of these features. Each pair formed by one feature point, *i*, on the left and one feature point, *j*, on the right yields a correspondence cost c_{ij}. The calculation of the correspondence cost c_{ij} generally depends on the descriptive attributes of the pair. A good match yields a low cost, a poor match yields a high cost. A very high non-correspondence cost, c_{nc} is assigned when the two considered feature points are determined to not correspond. A two dimensional array, the correspondence cost matrix, can then be constructed. At every transition path from an element in one column to an element in the next column, there is a path cost. This path cost is a function of the correspondence costs of the two elements in the adjacent columns.

Assume that the correspondence is known at column 1 and column m. The task is then to find the least cost path from the start point to the end point. Some stabilizing constraints are typically included to decrease the number of admissible paths. These constraints arise from the fundamental theory of stereo vision. Since one point on the left has at most one corresponding point on the right, the admissible path must proceed monotonically left to right and top to bottom. Another necessary constraint is a parallax constraint. Given a minimum and maximum allowable elevation jump, there is a corresponding minimum and maximum parallax or disparity change between successive features (columns). This constraint maintains the path within a parallax~band from upper left to lower right in the cost matrix. Among the set of all possible paths, one is optimal, i.e. it corresponds to a minimum total cost. This path determines the conjugate feature pairs, and these can be intersected via the usual photogrammetric equations to yield object space coordinates. Figure 3 shows a typical set of matched features in urban imagery.

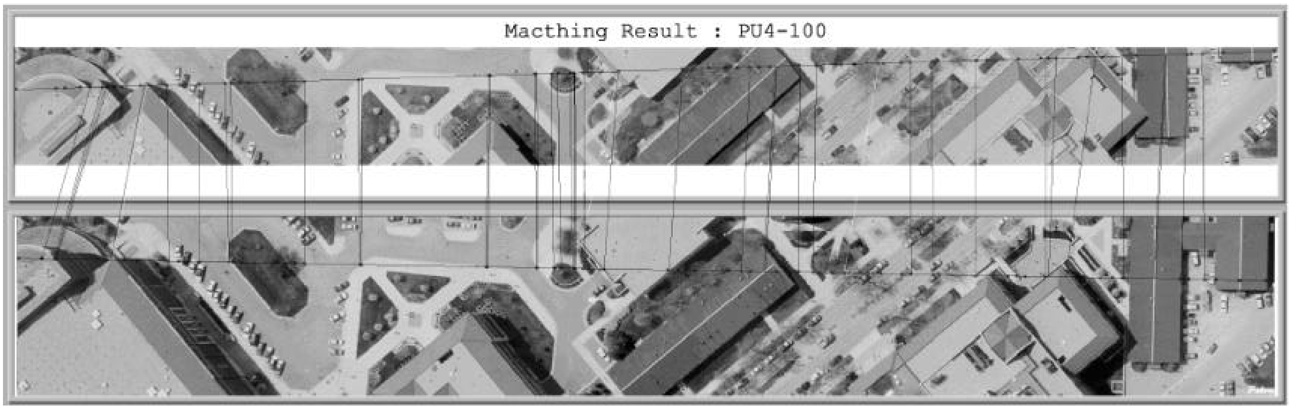

Figure 3. Typical Results from Feature Matching

3.2 Signal Based Matching

Correct correspondence at the signal level needs to be accomplished to supplement and densify the results of feature matching. The results from feature matching are only the coordinates of some points on extracted features, or all of the feature points at best. It thus represents only a sparse surface model. The proposed signal matching algorithm is driven from ground space so it is, initially, completely separate from the feature matching process. A regular grid is first established in object space. The method is designed to find the elevation profile along each gridline, for an arbitrarily selected direction. The VLL concept is adopted for use in this algorithm. Each grid line is extruded vertically to create a 2D grid in a vertical plane. Grid cells are defined by a planimetric interval and a vertical interval. Each cell in this array is projected into the left and right photos and a match coefficient is computed over a small region surrounding the point. This process is done similarly for all points in a grid profile creating a vertical plane of grid cells, each characterized by a match coefficient. If these coefficients are scaled into gray levels for display, then one can visualize the process. The picture of an approximated elevation profile would appear as a line on a noisy background, see Figure 4.

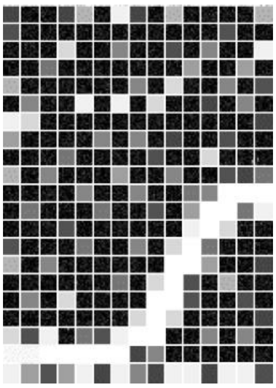

Figure 4. Cost Matrix with Costs Coded by Gray Value

3.3 Searching Procedure

A 2D dynamic programming technique for arbitrary line following is applied at this point to determine the low cost path between the endpoints of the terrain profile in each vertical plane. This variation of the technique uses path length as the stage and direction as the state. The arbitrary line following is modified in this case to restrict the search to paths that proceed from left to right, i.e. no overhangs or concavities in vertical segments. Thus at any given point in the cost matrix there are only five possible directions to move versus eight in the general line tracking problem. This path propagation mask is stepped exhaustively through the the cost matrix for each increment in path length, from 2 to the maximum specified.

3.4 Integration of Signal and Feature Matching

During the optimization process, the object coordinates from matched features are used as constraints. An elevation profile should pass through these feature points. Because the correlation method may not work well in the vicinity of some features, the signal-derived profile in the vicinity of such features may be in error. Conversely, the DP search would try to pass through the feature points if they were assigned a very low cost. To achieve this goal, we can compel or encourage the search algorithm to pass through features by providing a very small cost to any cells which represent feature points.

Although the feature matching algorithm is based on an optimization technique coupled with rigorous photogrammetry and image processing concepts, incorrect matches can still occur. If we therefore integrate features into the signal matching by the requirement that an optimal path must pass through the feature points, an incorrect result may be obtained. Our approach does not force the profile to pass through the feature points, it only encourages this by a very low (or negative) cost. This integration of feature matching and signal matching can work to achieve a result such that,

- The correct path will be obtained. The search technique will benefit from the redundant information in the intensity profile as well in the features.
- A consistency check for feature matching can be accomplished. For most good matches, there will be at least an accessible path to it.

Once the optimal profiles from all grid lines are obtained, the object surface model can then be reconstructed.

4 EXPERIMENTS

A number of experiments have been carried out to verify the working of the above algorithm. A summary of one such experiment will be presented based on imagery over the Purdue University campus.

4.1 The Purdue Project

The Purdue images cover the campus area in West Lafayette, Indiana. The neat area from the corresponding images was extracted and examined as illustrated in Figure 5.

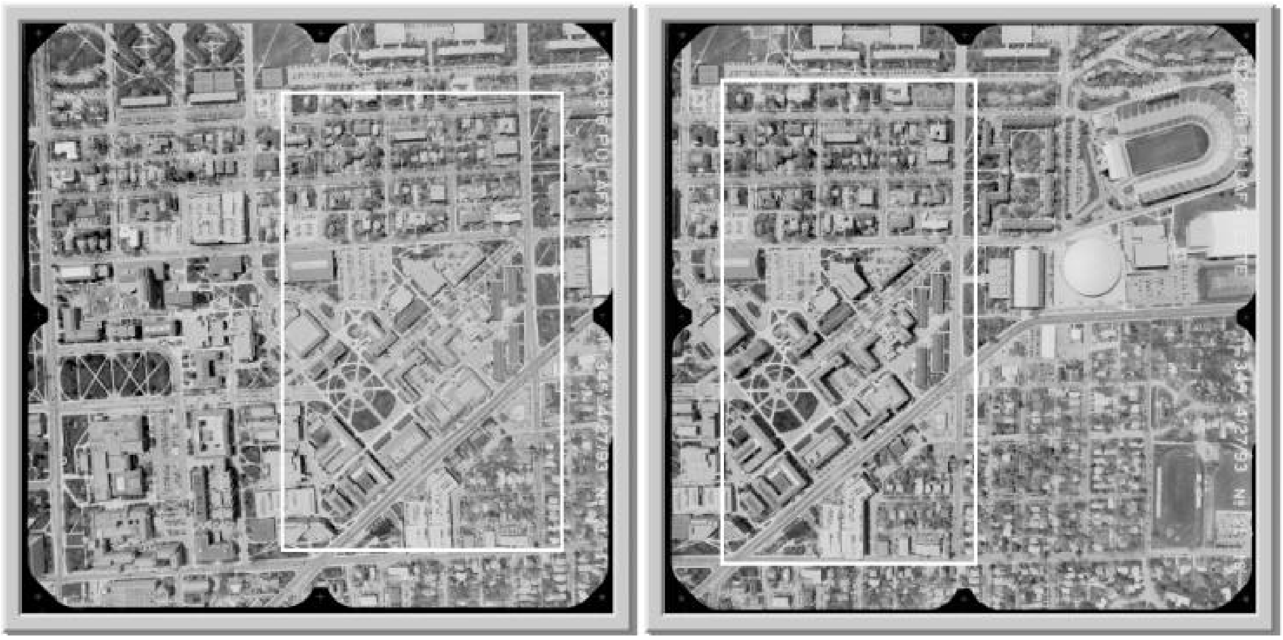

Figure 5. Left and Right Images from Purdue Pair

This area contains small and large buildings, abundant walkways, and parking lots. There are also some occlusions, shadows, and moving objects. The original photographs were nearly vertical and their nominal scale was 1:5000. They were originally scanned into digital form with a resolution of 15 micrometers. Their aggregated version at level 2 (30 micrometers) was employed in this experiment, so one pixel represents approximately 0.3 m. in the object space.

The straight line, plateau, and spike features were extracted from both images. About 140 epipolar lines covering the whole area were examined. There were approximately 50-70 extracted feature points from each epipolar line. However, only 35-40 correspondences were found from each set. Some false alarms and missed detections of match pairs still occurred. These errors are caused by the imperfect results from the feature extraction process, occlusions and other factors. The 3D information computed from feature matching was used as constraint information in the signal matching process.

Searching in the signal matching process was performed independently in two directions, north-south and east-west as shown in Figure 6.

Figure 6. Orientation of the Vertical Planes Defining the Cost Matrices

The inclusion of the feature constraints within the signal matching showed favorable results in many parts of this experiment. Figure 7 shows a cost matrix with the extracted surface profile without and with the feature constraints.

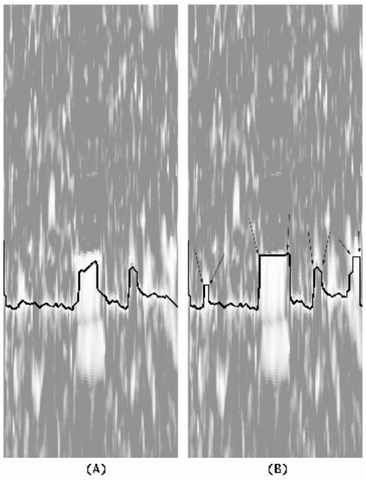

Figure 7. Cost Matrices and Extracted Profiles Without and With Features

After integrating the feature information the building shapes were substantially improved and several structures that had been completely missed were found. This example indicates the strength and the benefits of using both feature and signal matching. This process was carried out over the entire overlap area. The object surface reconstruction is shown in Figure 8.

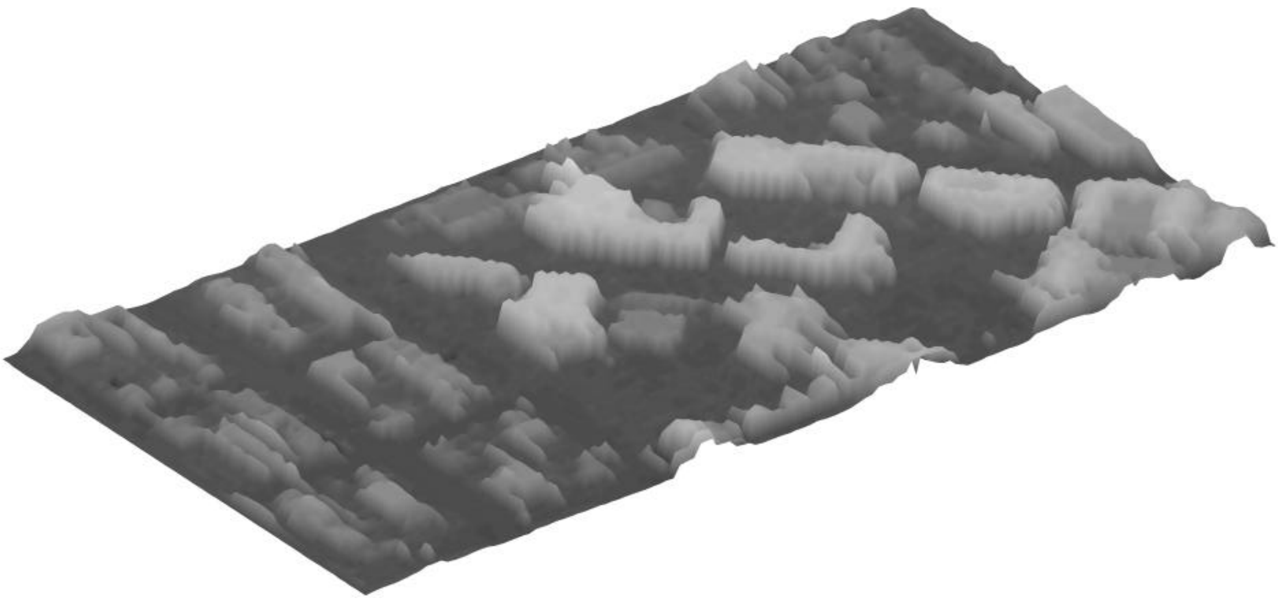

Figure 8. Perspective Rendering of Extracted Surface of the Purdue Model

A preliminary evaluation of the results was carried out by comparing the surface model to that produced by a commercial system using an *automated terrain extraction*. The error histograms from the two results are shown in Figure 9.

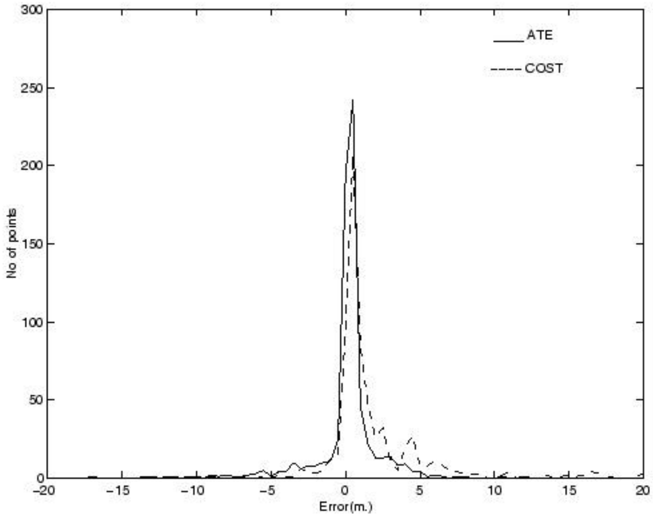

Figure 9. Error Distributions from the the COST and ATE Systems

5 CONCLUSIONS

This paper has described an integrated stereo matching technique for object surface reconstruction. The matching problem is addressed by the integration of both signal and feature based approaches. To address stereo matching in urban area images, we developed two innovative features, i.e., the plateau and spike. The classic dynamic programming for feature matching is modified and enhanced to integrate signal and feature matching into a simultaneous profile/surface determination. The initial global error statistics indicate that no improvement has been achieved over strictly correlation based methods. However examination of some individual area (see Figure 7) indicate that the methods offers promising potential in exactly the circumstances where area correlation is weak. This approach can be further developed with other features, other match criteria, and improved search methods to yield even better results in the future.

ACKNOWLEDGEMENTS

The work described in this paper was partially supported by the U.S. Army Research Office under Grant Number DAAH04-96-1-0444. The authors would like to acknowledge the support of Dr. Edward Mikhail, the project director. They would also like to thank Ahmed Elaksher for his help in data analysis.

REFERENCES

1. Bellman, R., 1957, *Dynamic Programming*, Princeton University Press, Princeton NJ
2. Bernard, M., Boutaleb, A., Kolbl, O., Penny, C., 1986, *Automatic Stereophotogrammetry: Implementation and Comparison of Classical Correlation Methods and Dynamic Programming Based Techniques*
3. Bethel, J., Mikhail, E., Kim, K., Marshall, J., 1998, *Intelligent Map Understanding – Automated Feature Recognition and Delineation*, Technical Report, School of Civil Engineering, Purdue University
4. Burns, J., Hanson, A., Riseman, E., 1986, "Extracting Straight Lines", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8:429-455
5. Canny, J., 1986, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679-698
6. Ohta, Y., and Kanade, T., 1985, "Stereo by Intra- and Interscanline Search Using Dynamic Programming", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7:139-154