3D STEREO RECONSTRUCTION OF BUILDINGS FROM AERIAL IMAGES

Drago Torkar
Institut Jožef Stefan, Ljubljana
Received November 13, 1995
Revised November 21, 1995

Abstract

*In the paper our approach to the stereo reconstruction
problem is presented. First linear line segments are extracted
from left and right image. Next structures that possibly
represent buildings are detected using graph theory. Finally
stereo correspondence is determined by graph matching
algorithm.*

Keywords: *building detection, building reconstruction,
feature-based matching, feature extraction, Geodetic
workshop, graph theory, Otočec, 1995*

1 INTRODUCTION

3D stereo reconstruction from aerial images is a common application in stereo. Since it is a time consuming task increased research activity toward automatic building reconstruction (Huertas et al., 1993, Dang et al., 1994, Deren et al., 1994).

One of the problems within automatic stereo reconstruction of objects is the determination of correspondent features on the left and right image. The position of correspondent points forms the base for depth estimation in camera and the global coordinate system. The determination of correspondent image elements involves stereo matching of the left and right image and/or feature extraction from both images and feature matching. Some methods use also the digital terrain model (DTM) which can be extracted by area-based matching or by other methods (Weidner et al., 1994). The stereo correspondence determination is accompanied by several problems caused by geometry constraints.

2 AUTOMATIC METHOD

The novel method for automatic building reconstruction from aerial stereo images is based on the work reported by Horaud and Skordas (Horaud et al., 1989). The correspondent elements determination was done using the procedure from the set of feature-based disparity estimation methods. Some buildings can be quite complex so the symbolic description must be made by structured features. Each feature represents one building or a set of buildings very close together. Therefore a procedure for building detection and their representation by structured features had

to be developed. Next an efficient method for stereo matching had to be used to determine correspondent buildings. The whole procedure was done in four steps:

- □ primitive feature extraction --
- □ construction of structured features using graphs
- □ correspondent structures determination using graph matching algorithm
- □ 3D building reconstruction.

3 FEATURE EXTRACTION

Straight line segments were used as primitive features. To extract them first Canny's edge detector (Canny, 1983) was applied. The width of the edge operator was determined experimentally for each image separately. It depends on image intensity and contrast as well as image content. Since buildings are composed of straight edges the line growing algorithm was used to detect straight segments that potentially represent building edges. Because of the noise present in almost every image and the tolerances introduced by the feature detection process the feature improvement module was developed. The algorithm was implemented in two steps. First all line segments are investigated to find groups of collinear segments among them. Each such group is then replaced by a single straight line segment. Further, the distances between segment ends are compared to a certain threshold. If two ends are close enough the corresponding segments are prolonged or shortened to the point of intersection. This is done also in the case when the line segment end is near another line segment. In the case of multiple segment ends within the threshold distance, segments are extended to the center of gravity of intersection points.

4 BUILDING STRUCTURES USING GRAPHS

The goal of building structures from primitive features is a set of structured features for the left and right image, each of them representing a symbolic description of a single building. Only buildings having only straight, closed edge segments are considered. The building of structures was implemented in two steps. First the relational graph for the left and right image was constructed from primitive features i.e. straight line segments. The result of this part was an unconnected, labeled relational graph representing the relationship among primitive features. In the second step some nodes were deleted and connected subgraphs representing one building were detected.

4.1 Construction of relational graph from primitive features

Every line segment is represented by a node in the graph weighted by the following values: the position of a segment, the segment size and the segment contrast and the number of pointers of every type. Two types of pointers, representing graph edges, are also assigned to every node: is_connected_to and is_parallel_to. The pointers of the first type point to line segments connected to or intersecting with the segment in question. The pointers of the type is_parallel_to point to nearest parallel segments. Thus every segment can have only two pointers of this type.
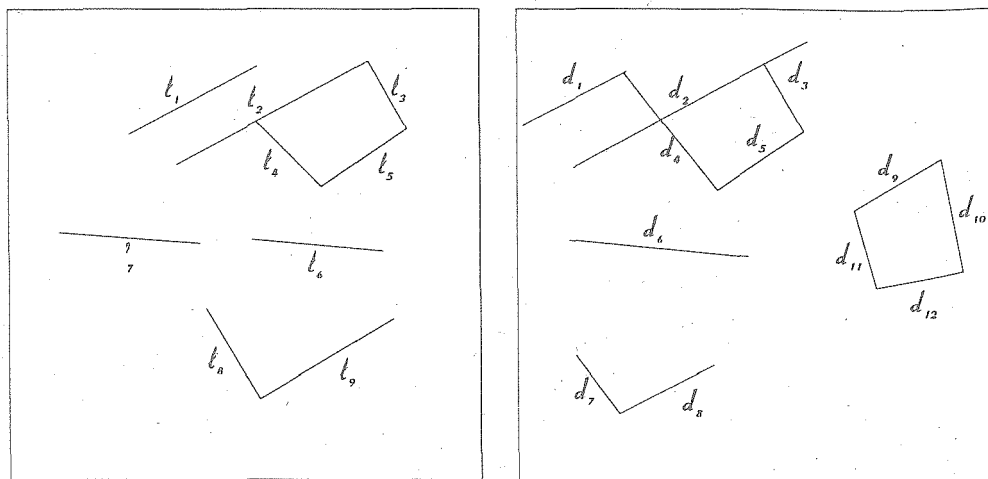
Figure 1: Left and right image features

The values of weights are defined as follows: the position of a segment is determined by its end points. Segment size is defined by the length of a segment. Contrast is calculated as a difference of the average gray value of a stripe to the left of the segment and average gray value of the stripe to the right of the same segment. Number of connected segments means the number of adjacent segments intersecting or jointing with the segment in question. Number of parallel segments means the number of segments within a certain distance from the segment in question and being parallel to that segment. The pointers of the type is_connected_to (on figure 2 labeled with 1) are determined in a way that the intersection points of a line containing particular segment with all the other lines are calculated. If the intersection is within both line segments the corresponding pointer is initialized. If no pointer of the type is_connected_to for the particular line segment is initialized the pointer number is set to zero.

Similar the pointers of the type is_parallel_to (in figure 2 labeled with 2) are initialized. The slope for every line segment is calculated. Among other line segments having the slope within the ±10% and overlapping 60% or more, the nearest one in both normal directions is chosen. The relational graph building for example in figure 1 is shown in figure 2.
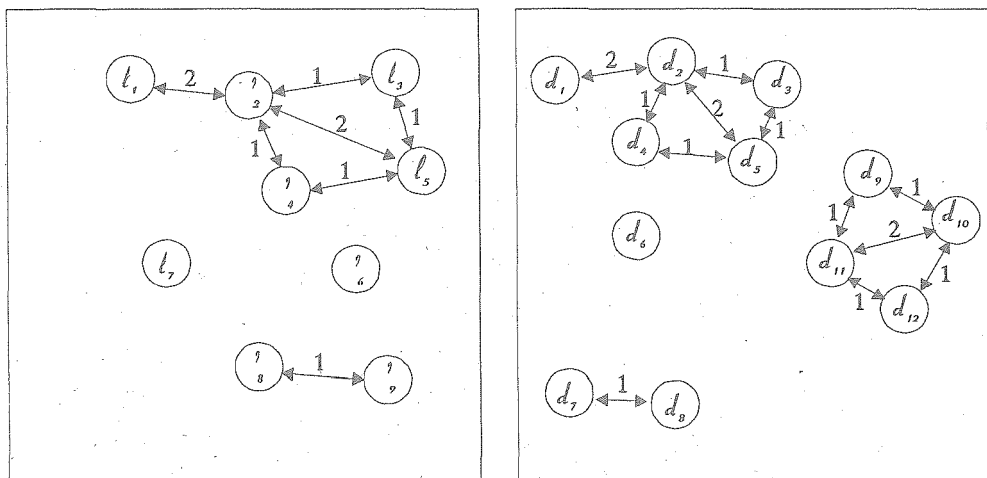
Figure 2: Relational graphs of the left and right image

4.2 Deleting nodes and subgraph detection

So far the weighted relational graph constructed is unconnected because there exists a set of features neither having intersections with neighbors nor being parallel to them. The construction of the labeled relational graph is continued by deleting the nodes in the graph having no edges of both types or having only one edge of a particular type. This way the features that have only one parallel line segment or are connected to only one neighbor are eliminated. It rarely happens that two features organized this way would actually represent a building. If they do it is the consequence of bad edge detection errors during the line growing process, which is usually due to bad choice of free parameters.

The task is to find subgraphs representing a single building or complex of buildings on the site. The buildings should be closed structures from the edge point of view. Therefore we search for cycles among all connected subgraphs of a relational graph. Their detection is simple and efficient enough. We do this only on graph edges representing connections using the well-known depth-first search. The subgraphs forming the cycles are then investigated again to eliminate the ones representing line segments intersecting at a single point. The remaining connected subgraphs are then used to determine correspondence between left and right image features.

5 DETECTING CORRESPONDENT FEATURES USING GRAPH MATCHING ALGORITHM

Due to several reasons, such as photometric differences between left and right image, occlusions, shadows, errors during primitive feature extraction etc., double subgraph isomorphism among connected subgraphs of relational graphs can not be expected. Therefore the correspondence had to be detected based on similarity among subgraphs. First the similarity measure between two nodes of the relational graph was defined.

$$mp_{ij} = \frac{2}{5} \left(\frac{min(K_i,\ K_j)}{max(K_i,\ K_j)} + \frac{1}{4} \frac{min(l_i,\ l_j)}{max(l_i,\ l_j)} + \frac{1}{4} \frac{\pi - 2\phi_{ij}}{\pi} + \right.$$

$$\left. + \frac{1}{2} \frac{min(\Sigma p_{1i},\ \Sigma p_{1j})}{max(\Sigma p_{1i},\ \Sigma p_{1j})} + \frac{1}{2} \frac{min(\Sigma p_{2i},\ \Sigma p_{2j})}{max(\Sigma p_{2i},\ \Sigma p_{2j})} \right)$$

The similarity measure is a number from the interval (0,1). The first term represents similarity in contrast. The second one is weighted by and checks out matching in length. The third one measures slope similarity and is also weighted by. The fourth term determines the difference in the number of pointers of the type is_connected_to and is weighted by $\frac{1}{2}$. The fifth one finds out the difference in the number of pointers is_parallel_to and is also weighted by $\frac{1}{2}$. The greatest weight was given to the similarity in contrast. To the pointers of both types was given only the weight of $\frac{1}{2}$ because frequently a certain feature is connected to or parallel to the features not belonging to the building. The slope and the length of line segments depend on the distance between the camera position taking the left and right image.

Next the similarity measure between two connected subgraphs of relational graphs of the left and right image was defined as the sum of similarity measures between all possible node pairs of the two subgraphs and normalized to the interval (0,1).

$$M_{mn} = \frac{1}{N_m\ N_n} \sum_{i=1}^{N_m} \sum_{j=1}^{N_n} mp_{ij}$$

The similarity measure that was defined reflects the relationship between every subgraph of the left image and every subgraph of the right image. Now a new graph, called a correspondence graph, can be constructed. The nodes are the subgraphs of the left and right relational graphs. The edges are made from all the nodes corresponding to the left image subgraphs to all the nodes corresponding to the right image subgraphs. The edges are labeled (weighted) by similarity measures between the nodes. The task is to find such a subset of connected nodes in which every node appears only once. In other words every node can be connected to only the node to which it has a maximal measure of similarity. Some nodes can remain unconnected. The problem of finding the correspondent structures between the left and the right image is so translated to the maximum matching problem in the graph theory. This is solved using the solution of the classic stable marriage problem (Sedgewick, 1988). The algorithm assigns to every connected relational subgraph from the left image the most suitable pair among connected relational subgraphs of the right image. The same algorithm is applied again within these pairs of connected subgraphs to find out the correspondent nodes of relational graphs, i.e. correspondent features. Using the known positions of line segment ends in the left and right image, its orientation and camera orientations in the moments of taking the left and right image, the 3D position of feature points can be calculated in the left or right camera coordinate system. If the camera perspective center is known in absolute coordinates the point position can be expressed in any of absolute coordinate systems.

6 3D STEREO RECONSTRUCTION OF BUILDINGS

The described method for 3D building reconstruction from aerial stereo images was tested on aerial photographs during the systematic periodic aerial survey of Slovenian territory (CAS). The absolute camera orientations were known. Photographs at a scale of 1:5 000 were used. The scanning of photographs was made with pixel size of 20 m which means a resolution of app. 1 270 dpi. The size of the image from the photography of 23x23 cm was therefore about 11 500 x 11 500 pixels which means about 130 MB of data. Because of this enormous size overlapping squares of only 2x2 cm were used. The overlapping of images was 60%. Below the example of 3D reconstruction of buildings with an intermediate result is shown.

7 CONCLUSION

A method for almost automatic stereo reconstruction of buildings was developed. The user must specify some parameters that define the behavior of the edge detection module and tolerances for the feature improvement module. The appropriate values can be obtained by trial and error or through experience. The extracted features are usually part of building roofs because of vertical camera position. This and errors during feature extraction and structure building cause only partial reconstruction of buildings. The number of correctly reconstructed points is too small to make this method a serious competitor to semi-automatic or non-automatic methods. The described procedure thus represents only one of the possible research directions in the evolution of automatic methods for stereo reconstruction of buildings from aerial images.
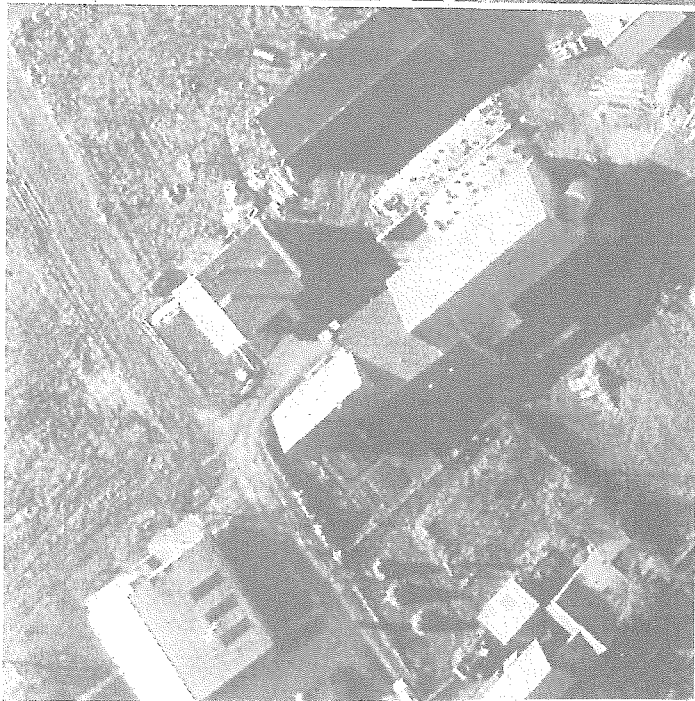
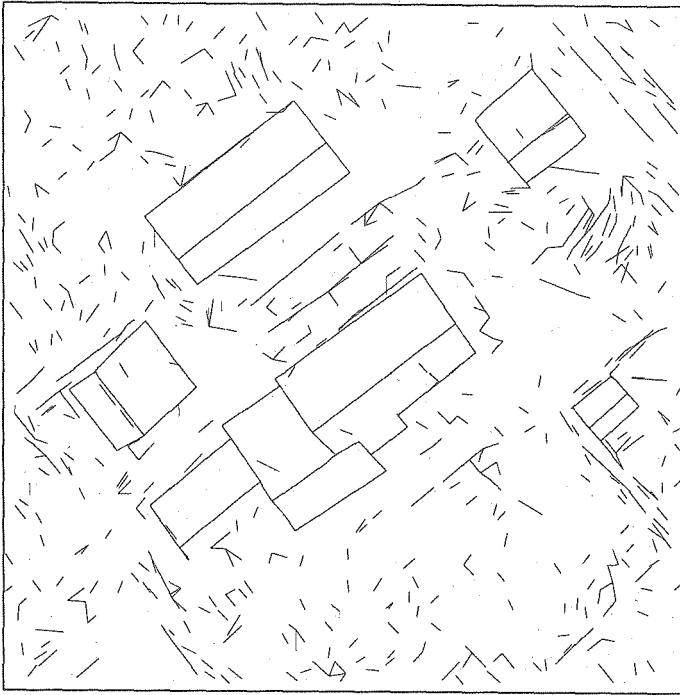*Figure 3: Left (upper) and right view (lower) of stereo pair*

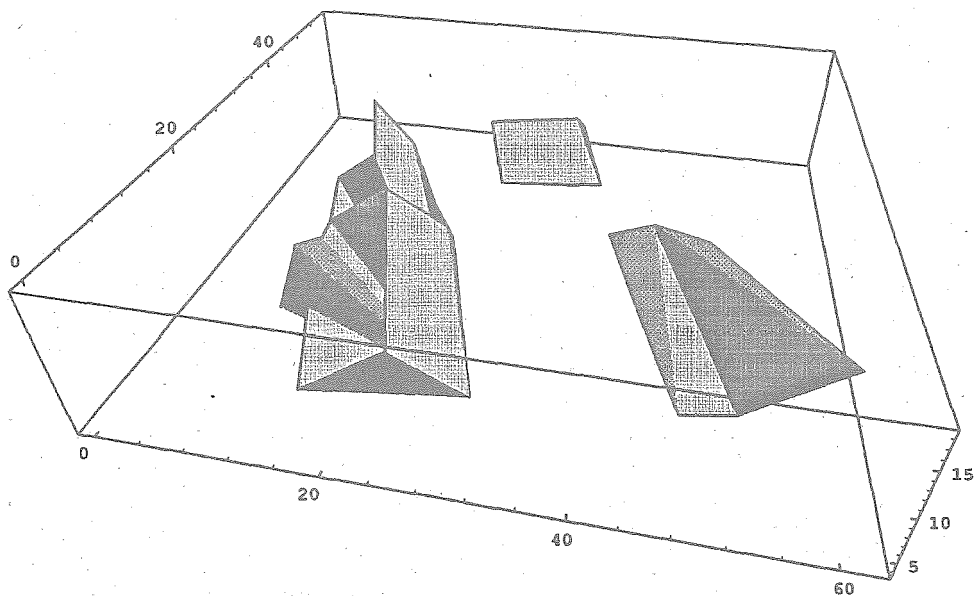*Figure 4: Detected features on the left (upper) and right (lower) image*

Figure 5: 3D rendering of successfully reconstructed points

References

Canny, J.F., *Finding edges and lines in images. Technical report AI-TR 720, M.I.T. Artificial Intell. Lab., Cambridge, MA, 1983*

Dang, T. et al., *Applying perceptual grouping of surface models to the detection and stereo reconstruction of buildings in aerial imagery. In: Proc. ISPRS 1994 – Spatial Information from Digital Photogrammetry and Computer Vision, SPIE, DGPF, IEEE, 1994*

Deren, L., Juliang S., *House extraction with multiresolution analysis and informational fusion. In Comm. III Symposium 94, WG/2, 1994*

Horaud, R., Skordas, T., *Stereo correspondence through feature grouping and maximal cliques. IEEE Trans. Pattern Analy. Machine Intell., 1989, vol. 11, no. 11, p. 1168-1180*

Huertas, A. et al., *Detecting of buildings from monocular views of aerial scenes using perceptual grouping of shadows. In: Proc. DARPA Image Understanding Workshop, 1993*

Sedgewick, R., *Algorithms. Addison-Wesley Publishing Company, 1988*

Weidner, U., Foerstner, W., *Toward automatic building extraction from high resolution digital elevation models. In: Digital Photogrammetry, Muenchen, 1994*

Review: Vasja Bric, M.Sc.
          Radovan Dalibor, M.Sc.