I N S T I T U T J O Ž E F S T E F A N

COMPUTER
SYSTEMS
DEPARTMENT

Technical Report

Feature Extraction from Aerial Images and Structural Stereo Matching

[1]Drago Torkar, [2]Nikola Pavešić

[1]*Jožef Stefan Institute, Computer Systems Department*
Jamova 39, 1000 Ljubljana, Slovenia
E-mail: drago.torkar@ijs.si

[2]*University of Ljubljana, Faculty of Electrical Engineering*
Tržaška 25, 1000 Ljubljana
E-mail: nikola.pavesic@fer.uni-lj.si

Technical Report **CSD-96-5** *June 1996*

Feature Extraction from Aerial Images and Structural Stereo Matching

[1]Drago Torkar, [2]Nikola Pavešić

[1]*Jožef Stefan Institute, Computer Systems Department*
Jamova 39, 61111 Ljubljana, Slovenia
E-mail: drago.torkar@ijs.si

[2]*University of Ljubljana, Faculty of Electrical Engineering*
Tržaška 25, 1000 Ljubljana
E-mail: nikola.pavesic@fer.uni-lj.si

Abstract

In this paper, a method of generating structural descriptions of stereo image pairs and their matching to recover 3-D form of objects in space, is discussed. The method consists of the following steps: firstly, straight line segments are extracted in both images and the relations among them are investigated. The relational subgraphs are constructed then, and the correspondence graph is created. Using the "stable marriage" searching algorithm the features in one image corresponding to features in the other image are found. The method was tested on a set of aerial stereo images of urban scenes.

Keywords: computer vision, feature extraction, structural stereo matching

---

[1] This work has been supported by Ministry of Science and Technology of the Republic of Slovenia under Grant Number J2-7258
This report appears in the proceedings of the International Conference on Pattern Recognition, August 1996

Technical Report **CSD-96-5** *June 1996*

1 Introduction

Feature extraction from aerial images can serve to different purposes, such as object detection, surface reconstruction, image matching, etc. Image matching is widely considered to be the most difficult to solve and clearly dependent on the choice of features [3], which are application dependent and can differ from simple pointlike ones to very complex structures.

The structural stereopsis has been proposed by Boyer and Kak [1], who have modified, as stated in [4], the exact matching approach developed by Shapiro and Haralick [8].

Horaud and Skordas [6] searched for relations among linear features and compared the relation between two features in the left image to the relation between two features in the right image.

1.1 Our approach

In this research, we focused on constructing the structured description of objects of interest, consisting of simple linear features, extracted from the left and right object image. The description of the scene was made by constructing a *relational graph* that described the relationship among image primitives. Relations among them were investigated to find groups representing particular objects. Each group was represented by connected subgraph of a relational graph. The stereo correspondence was determined by comparing these structures using a graph theoretical approach.

2 Feature extraction

Straight line segments were used as primitive features. To extract them, first Canny's edge detector [2] was applied. The width of the edge operator was determined experimentally for each image separately. It depended on image intensity and contrast as well as image contents. Next, the thinning algorithm [5] was applied to ensure the edge width of one pixel. A simple line growing algorithm was used to detect straight line segments that potentially represent object edges. Because of the noise present in almost every image, and the errors introduced by the feature detection process, the feature improvement module was developed. First, all line segments were investigated to find groups of collinear segments among them. Each such group was then replaced by a single straight line segment. Further, the distances between segment ends were compared to a certain threshold. If two ends were close enough the corresponding segments were prolonged or shortened to the point of intersection. This was done also when the line segment end was near another line segment. In the case of multiple segment ends within the threshold distance, segments were extended to the center of gravity of intersection points. Details can be found elsewhere [9].

3 Structural description

With structuring primitive features, a set of structured features for the left and right image was generated, each of them representing a symbolic description of a single object. The features structuring process was implemented in two steps. In the first step, 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 graph, representing the relationship among primitive features. In the second step, some nodes were deleted and connected subgraphs, representing single objects, were detected.

3.1 Construction of relational graph from primitive features

Every line segment was represented by a node in the graph weighted by the following values: the position of a segment, the segment size, the segment contrast and the number of pointers of each type. Two types of pointers, representing graph edges, were assigned to every node: *is_connected_to* and *is_parallel_to*. The pointers of the first type point to line segments connected to or intersect 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.

The pointers of the type *is_connected_to* (in Figure 1 labeled with 1) were determined in a way that the intersection points of a line, containing particular segment, with all the other lines were calculated. If the intersection was within both line segments the corresponding pointer was initialized.

Similary, the pointers of the type *is_parallel_to* (in Figure 1 labeled with 2) were initialized. The slope of every line segment was calculated. Among line segments, having the slope within the ± 10% and overlapping 60% or more with the segment in question, the nearest one in both normal directions was chosen. The relational graph building example is shown in Figure 1.

When the graph was built the U structures were detected. Connections of both types were used. For every edge of type *is_parallel_to* belonging to a node pair, all edges of type *is_connected_to* of both nodes were searched. If an U structure was detected it was closed, to form a quadrangular.

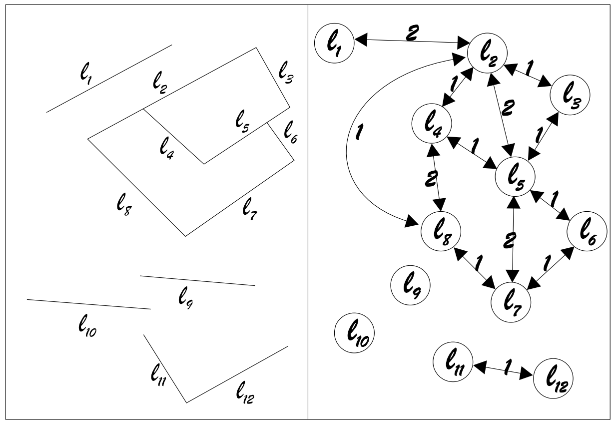

Fig. 1. Image features (left) and relational graph (right).

3.2 Deleting nodes and subgraph detection

The construction of the labeled relational graph was 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 had only one parallel line segment or were connected to only one neighbour were eliminated. It rarely happens that two features organized this way would actually represent an object.

The task was to find subgraphs, representing a single object or complex of overlapping objects on the scene. The objects should have been closed structures from the edge point of view. Therefore we searched for *cycles* among all connected subgraphs of a relational graph. Their detection was simple and efficient enough, and was done considering only graph edges, representing connections, using the well known depth-first search. The subgraphs forming the cycles were then investigated again to eliminate the ones representing line segments intersecting at a single point.

4 Stereo matching

Due to the 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 could not have been expected. Therefore the correspondence detection had to be based on similarity among subgraphs. First, the *similarity measure* between two nodes of the relational graph was defined as:

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

$$+\frac{1}{2}\frac{\min(\sum p_{1i},\sum p_{1j})}{\max(\sum p_{1i},\sum p_{1j})}+\frac{1}{2}\frac{\min(\sum p_{2i},\sum p_{2j})}{\max(\sum p_{2i},\sum p_{2j})}\Bigg)$$

The similarity measure is a number from the interval (0,1). The first term represents similarity in contrast. The second one is weighted by $\frac{1}{4}$ and checks out matching in length. The third one measures slope similarity and is also weighted by $\frac{1}{4}$. 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, only the weight of $\frac{1}{2}$ was given, because frequently, a certain feature is connected to or is parallel to the features, not belonging to the object. The slope and the length of line segments depend on the distance between the camera position during 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 normalized sum of similarity measures between all possible node pairs of the two subgraphs:

$$SM_{mn}=\frac{1}{N_mN_n}\sum_{i=1}^{N_m}\sum_{j=1}^{N_n}sm_{ij}.$$

This similarity measure reflects the relationship between every subgraph of the left image and every subgraph of the right image. A new graph, named *correspondence graph*, based on this measure was constructed. An example is depicted in Figure 2.

The nodes were the subgraphs of the left and right relational graphs. The edges were made from all the nodes corresponding to the left image subgraphs to all the nodes corresponding to the right image subgraphs. The edges were labeled (weighted) by similarity measures between the nodes. The task was to find such a subset of graph edges (type 1), that every node was connected only to the node, to which it had a maximal measure of similarity. Some nodes can remain unconnected. The problem of finding the correspondent structures between the left and the right image was translated to the maximum matching problem in the graph theory. This was solved using the solution of the classic stable marriage problem [7]. The algorithm assigned to every connected relational subgraph from the left image the most suitable pair among connected relational subgraphs of the right image (Figure 2).

The same algorithm was applied again within these pairs of connected subgraphs, to find out the corresponding nodes of relational graphs, i.e., correspondent features.

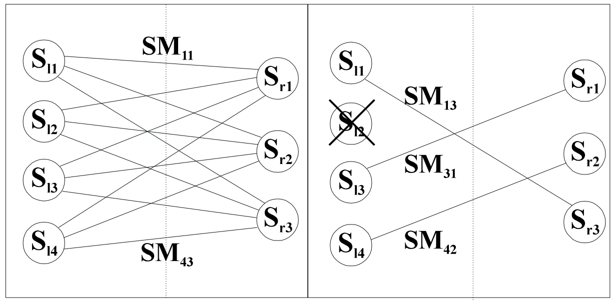

Fig. 2. Correspondence graph before (left) and after matching (right).

5 Results

The outdoor/aerial scenes are far the most unstructured of scene domains and provide complex matching problems[4]. We have tested the method on a set of aerial stereo photographs of urban scenes, taken in 1:5000 scale, overlapping 60%, and scanned at app. 1270 dpi.

Since the structural description of successfully detected and extracted objects contains only few primitive features in each image (Figure 3, 4), the number of correspondent points obtained through stereo matching is very small. As already stated in [4], structural stereopsis gives sparse disparity maps but is much less sensitive to distortions as a result in changes of viewing position.

The successfully matched features were reconstructed into 3-D world coordinate system, using the known absolute camera orientations. Reconstructed points were rendered as filled polygons (Figure 5). The heights of line segment end points were compared to manually obtained heights of the same spots (Figure 6). In Figures 3 - 6 an example with intermediate results is presented.

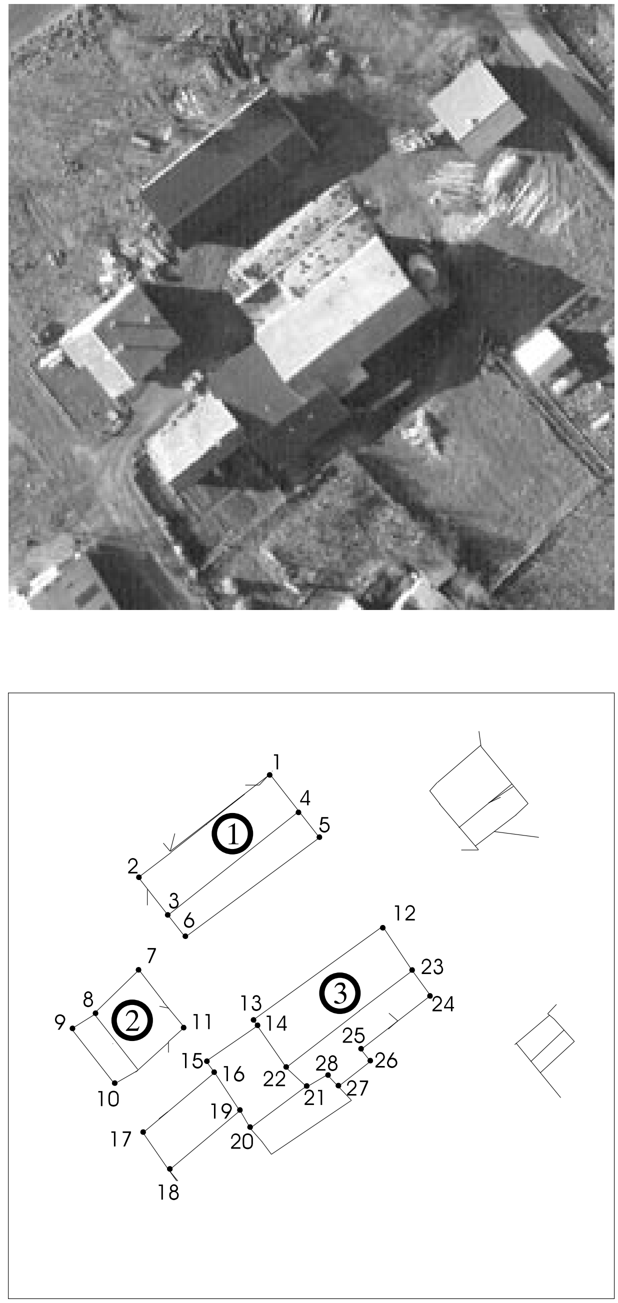

Fig. 3. Left aerial image and its successfully matched structures.

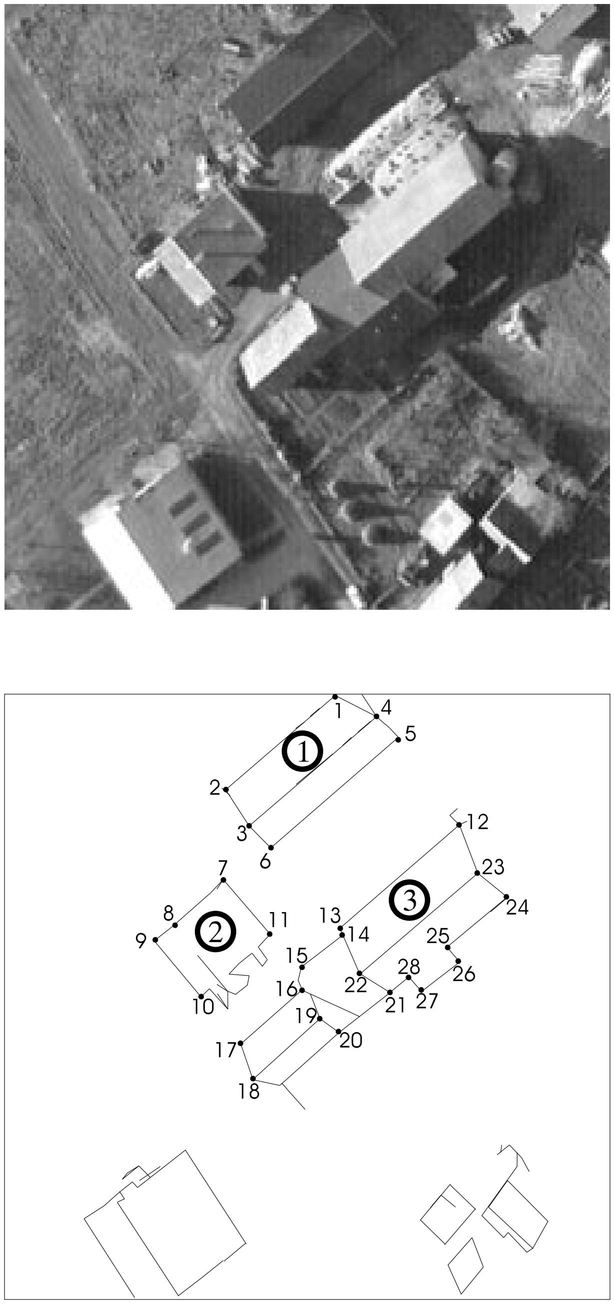

Fig. 4. Right aerial image and its successfully matched structures.

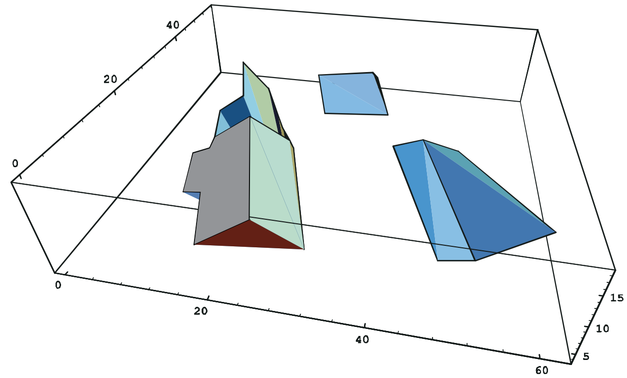

Fig. 5. 3D reconstruction of successfully matched features.

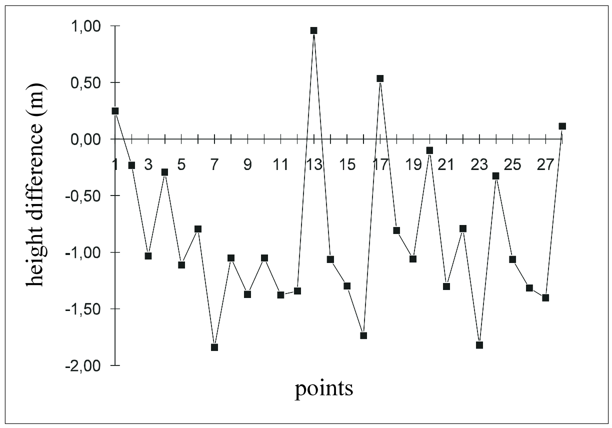

Fig. 6. Height errors.

6 Conclusion

A method for generating a structural description of image primitives and determination of stereo correspondence using such description was developed. It is more suitable for scenes containing objects with straight edge segments of the same size range. This type of objects can usually be described with small number of primitive features. As a consequence, the disparity map obtained through stereo matching is very sparse.

Using the proposed structural description of scene, object detection in images can be performed. We searched only for a few types of organization among connected primitives such as closure, connectedness and parallelism. Taking in account other types of organization, such as proximity, similarity and symmetry, the set of detectable objects would be enlarged.

References

[1] K. Boyer and A. Kak. Structural stereopsis for 3-d vision. *IEEE Trans. Pattern Analy. Machine Intell.*, PAMI-10(2):144–166, Mar. 1988.

[2] J. Canny. Finding edges and lines in images. Technical Report AI-TR 720, M.I.T. Artificial Inell. Lab., Cambridge, MA, 1983.

[3] S. Cochran and G. Medioni. 3d surface description from binocular stereo. *IEEE Trans. Pattern Analy. Machine Inetell.*, 14(10):981–994, Okt. 1992.

[4] U. Dhond and J. Aggarwal. Structure from stereo - a review. *IEEE Trans. Syst. Man Cybern.*, 19(6):1489–1510, Nov.-Dec. 1989.

[5] C. Holt, A. Stewart, M. Clint, and R. Perrott. An improved parallel thinning algorithm. *Comm. of the ACM*, 30(2):156–160, Feb. 1987.

[6] R. Horaud and T. Skordas. Stereo corresponcence through feature grouping and maximal cliques. *IEEE Trans. Pattern Analy. Machine Inetell.*, 11(11):1168–1180, Nov. 1989.

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

[8] L. Shapiro and R. Haralick. Structural description and inexact matching. *IEEE Trans. Pattern Analy. Machine Intell.*, PAMI-3(5):504–519, Sep. 1981.

[9] D. Torkar. 3-D building reconstruction from aerial stereo images (in Slovene). Master's thesis, University of ljubljana, Faculty of Electrical Engineering and Computer Science, 1995.