

DETECTION OF BUILDINGS IN AERIAL IMAGES USING SHAPE AND SHADOWS*

By

Andres Huertas and Ramakant Nevatia

Intelligent Systems Group
University of Southern California
Los Angeles, California 90089-0272

ABSTRACT

Extraction of man-made 3-D objects such as buildings in aerial images is aided by analysis of shadows they cast and their restricted geometric shapes. We describe an edge based technique which exploits these clues by constructing a hierarchy of increasingly complex features composed of consistently interpreted intensity boundary segments. Feature interpretations are determined on the basis of constraints imposed on the interpretations of observed intensity boundary fragments and groups of fragments by the imaging process, the direction of illumination, and assumptions about the geometry of the objects in the scene.

I INTRODUCTION

We explore the perception of restricted shapes and shadows as important clues for obtaining 3-D information, the interpretation of image content, and building detection in aerial images. Human observers probably rely on contextual clues such as nearby roads for building detection, but the most important visual clue seems to be the restricted shapes that many of buildings have, in particular, rectangles and combination of a small number of rectangles. In some cases, buildings partially "blend" with their backgrounds preventing the detection of closed building outlines, thus the shadows cast are the only clue for successful detection. The observed shadows also allow us to differentiate a building from other objects having shapes similar to the buildings of interest such as rectangular parking lots, and give the building heights.

Our technique uses the intensity data and the line segments approximating the observed intensity edges in the image (see figures 3b, 4b and [1,2].) The method consists of constructing a hierarchy of increasingly complex features, from line segments

to building outlines, on the basis of constraints imposed on the interpretation of the segments and groups of segments by the imaging process, assumptions about building geometry, and the direction of illumination. The general idea is to extract the most obvious and strong information first and use it to guide the search for weaker evidence. Strong evidence are the corners (90-degree L-junctions) formed by pairs of line segments representing the buildings sides and, the observed correspondence between shadow casting object edges and shadow boundaries.

Previous work on object detection in aerial images by Nagao, et al. follows a region-based approach [3] to perform an elaborate structural analysis of the image. Rosenfeld and Tavakoli [4] use estimates of an ideal building gray level to evaluate edge segments as a building sides. Building segments are extracted by processing the "antiparallel" pairs extracted from the image; but in many cases, poor contrast prevents the detection of antiparallels, or result in groupings that include segments from nearby objects. Since no shadow information was used, nonbuilding groups may be extracted as well. Previous work on the interpretation of image edges and scenes with shadows are based on boundary junction constraints in the blocks world which are difficult to obtain from natural images [5,6,7,8]. Binford and Lowe [9] discuss the derivation, use and implementation of more general constraints to carry out geometric interpretations up to the volumetric level. Their technique has been tested with encouraging results on simulated image curve data derived by hand; it requires accurate curve junction information, which may be difficult to obtain from natural images. We use automatically derived edge data and our junction analysis is based on the configuration of line segments approximating the intensity edges in the image.

Since we use edge data derived automatically which explicitly implies dealing with partial information, further assumptions about image content are introduced in the form of tolerances and thresholds that may limit generality. On the other hand, a number of assumptions about object geometry and boundary continuity allow us to predict the presence of weak evidence as well as to hypothesize missing data resulting from errors in the data, poor contrast and other scene and illumination conditions that prevent detection.

*This research was sponsored by the Defense Advanced Research Projects Agency, DARPA Order No. 3119, and was monitored by the Air Force Wright Aeronautics Laboratory under contracts F-33615-80-C-1080 and F33615-82-K-1786

II ASSUMPTIONS AND MODEL

The following assumptions are made: The scene is illuminated by direct sunlight. The sun is considered to be a point source at infinity. The projected direction of illumination and the sunlight incidence angle are known a-priori; the ground is level and flat, and the top surfaces of the objects are flat and appear brighter than the shadows and at least as bright as the surrounding background. These assumptions allow us to use the directionality associated with the line segments extracted from the image. (The convention is that the bright surface is to the right of the segment. Thus, the segments representing the outline of a building will have a consistent directionality along the building outline;) the non-horizontal building edges and sides are vertical.

Our technique is based on the following model:

a) The sun incidence angle and the projected direction of illumination. The direction of the illumination rays projected on the ground surface, is given with respect to the scene. By associating coordinate systems with the scene, with an observer having an overhead view, and with the illumination field, then for a given sun incidence angle we can compute the direction of the projection of the 3-D sun rays onto the image plane, that is, the direction vector I along which an object's shadow casting point will cast its corresponding shadow point in the image.

b) Shape. We assume that objects of interest contain 90-degree corners (L-junctions) only. If we define a "box" as the boundary of a surface whose sides form 90-degree corners then the following two shape measures are used to confirm a closed building outline or, alternatively, give a rating of how complete is a partial outline.

1- Corner consistency (CC), defined as the number of concave (bright) corners minus the number of convex (dark) corners around the outline of a region. CC is always equal to four if the surrounded region is brighter than its background, or equal to minus four if the surrounded region is darker than its background. This is due to the fact that the simplest regular shape of interest, the rectangle, has four vertices and four sides. Any additional pair of sides added to this basic shape to form more complex shapes adds one convex vertex plus one concave vertex which cancel each other in the summation.

2- Side consistency (SC), based on the fact that box sides can have only two orthogonal orientations (the line segments representing the boundary can have two possible directions for each orientation). Hence for each orientation, the sum of the lengths of the segments in one direction must be the same as the sum of the lengths of the segments along the same orientation but with opposite direction.

c) Shadow Gray Level. We assume that most of the

area in the image corresponds to smooth and continuous regions. By masking out the portions in the image where the intensity transitions occur (edges), we obtain a more accurate histogram of the distribution of the gray levels of the regions in the image. Most shadow gray levels are in the dark side of the masked histogram, and depending on the illumination conditions, histogram features can be located to provide a reliable shadow gray level range.

d) The gray level statistics in the neighborhood of the edges. The mean and standard deviation of the gray levels along several one-pixel wide stripes on both sides and, parallel to the line segments representing potential shadow casting edges and shadow boundaries are computed to look for shadow regions. The "begin" and "end" points of the stripes are computed along the illumination vector I . The mean value of the gray level in the stripes is compared to the likely shadow gray level to hypothesize the presence of a shadow, and the standard deviation is used as a measure of smoothness.

e) The edge segment contrast. The relative contrast across a portion of a boundary represented by a line segment is assumed constant along the segment. Contrast is a weak constraining parameter for most edges in an aerial image, but segments representing shadow boundaries and boundaries between objects and shadows have the highest contrast.

III FROM EDGES TO BUILDING OUTLINES

The method consists of five steps:

1. Corner Detection and Interpretation. We extract the 90-degree corners formed by pairs of line segments which meet at a vertex (see figures 3c and 4c) and, record their characteristics (position, orientation, etc.) A "bright"/"dark" corner surrounds a bright/dark region. Our goal is to assign an initial interpretation to the observed corners as object or shadow corners, or neither, on the basis of the photometric properties of the corner components and constraints imposed on their interpretation by building geometry and the direction of illumination. In general, convergence at a junction in the image corresponds to convergence in the scene thus, L-junction components are likely to have the same interpretation. We interpret the components of observed corners taken together rather than interpret the segments individually and then detect the corners, except when segments are oriented parallel to the projection of the sun rays on the ground. (These are likely to be shadow boundaries cast by vertical edges, if the dark side of the edge is likely a shadow region.)

For a given sun position some object physical edges are strongly constrained to cast shadows in the direction of illumination. Others are strongly constrained not to cast shadows at all. Figure 1

shows how the disposition of a corner with respect to the direction of illumination, shown in the upper portion of the figure, is combined with the photometric characteristics of the corner to assign an initial interpretation. The shaded/non-shaded side of a segment indicates that there must/must not be a likely shadow region to the left of the segment. The angle indicates the angle between the corner bisector and the direction of illumination. Figures 1a (1c) and 1b (1d) show respectively, bright and dark corner configurations that are interpreted as object (shadow) corners. Note that the constraints hold for angle intervals and are depicted in 45-degree steps for clarity.

Since the segment components of a shadow corner must have been cast by the components of a shadow casting object corner, located between the shadow corner and the source of illumination, then the corner vertices must lay along the same illumination vector. If correspondence matches are established between object corners and shadow corners then the interpretations based on photometry are strong and considered "very likely" to hold. Otherwise the interpretations are weak and considered only "likely" to hold.

2. Detecting Corner Compatibility. Corner compatibility is defined as the positional relationship between every pair of detected corners. Strong compatibility occurs when two corners share a component. Since both components in a corner are expected to have the same interpretation, the corners sharing a component are also expected to have the same interpretation; two corners that are in one of the four types of geometric disposition shown in figure 2 are defined to be weakly compatible. The significance of these four types of corner disposition becomes evident later in step 4

when processing partially detected building outlines.

Strong and weak compatibility determine the initial likelihood for every pair of corners of being part of the same building outline. Strongly compatible corners are "very likely" to be in the same outline and are grouped into corner "chains", whereas weakly compatible corners are only "likely" to be in the same outline and form single corner chains. As a corollary, incompatible corners are "unlikely" to be in the same outline.

3. Tracing and Tracking of Corner Chains. The CC and SC measures discussed above roughly indicate how complete an outline is, or alternatively, how many elements, corners and sides, are needed to have a closed outline or box. We compute and record these measures by tracing the corner components in each chain. If closed chains (boxes) are detected (figures 3d and 4d) we extract their components and form building hypotheses to be verified in step 5. Otherwise we search for weaker evidence of chain continuation in the neighborhood of the partially detected outline (figures 3e and 4e.)

We search for nearby evidence by tracking object outlines in the direction of the segments in the chain, extracting those chains that become boxes. If the search does not result in a box, tracking is performed in the opposite direction too. During the tracking, the geometric constraints for corner detection are relaxed, and as a result, new corners are detected. Corners encountered while tracking the box outline which satisfy the earlier, stricter constraints must be part of another chain. In this case the chains are combined, the CC and SC ratings updated, and if necessary the tracking resumes un-

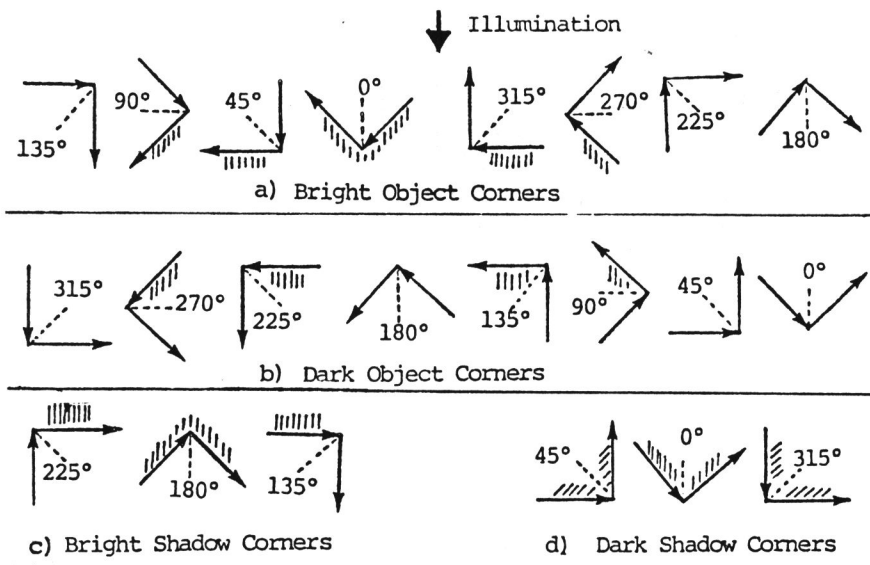


Figure 1. Corner Interpretation

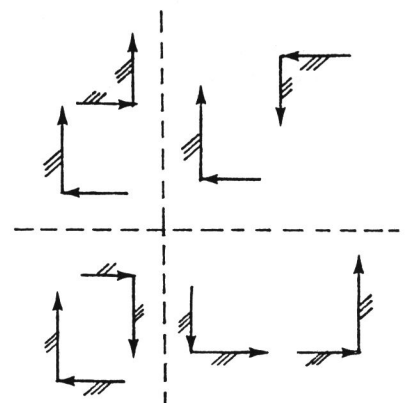
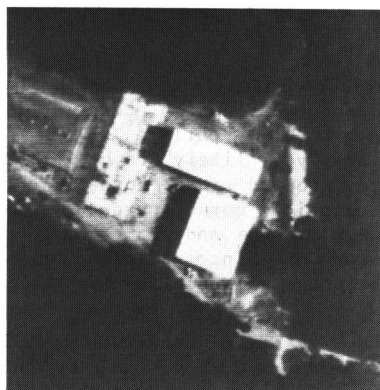
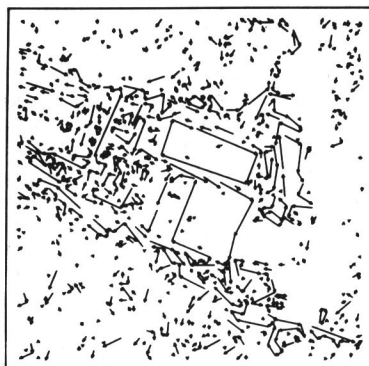


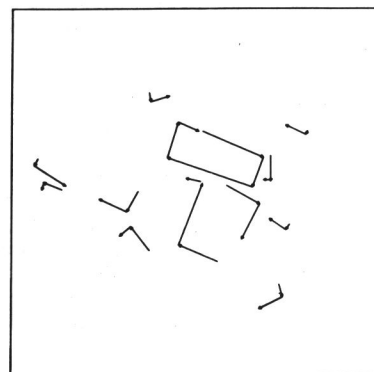
Figure 2.
Weakly Compatible Corners



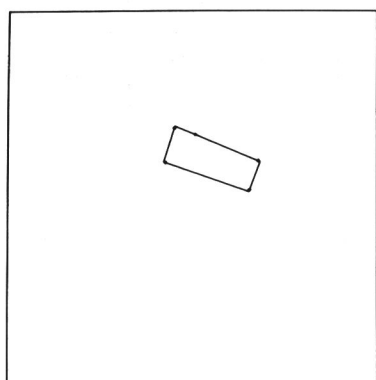
a) Original Image



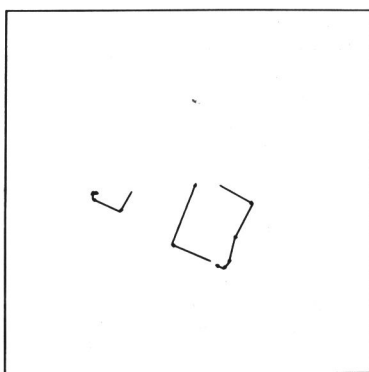
b) Line Segments



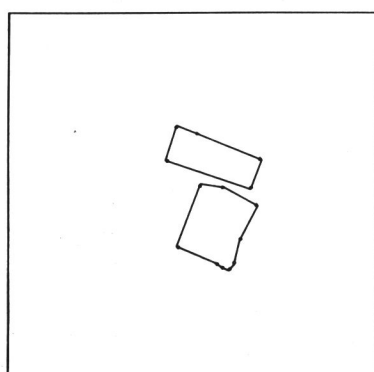
c) Corner Segments



d) Closed Chains

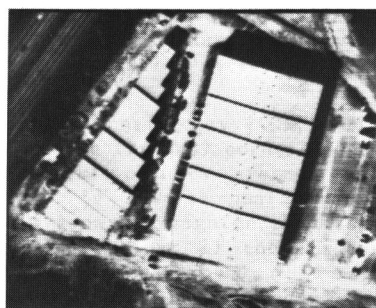


e) Open Chains and Nearby Evidence



f) Closed Boxes

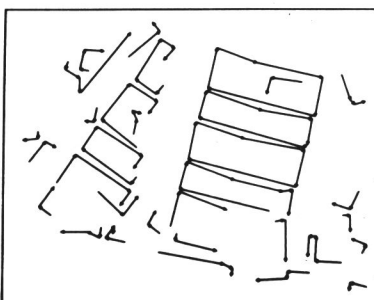
Figure 3. Image 1



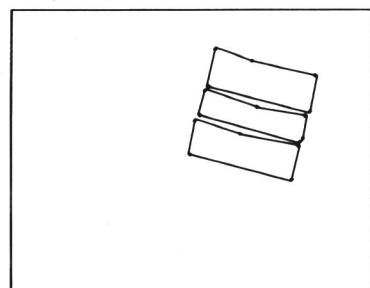
a) Original Image



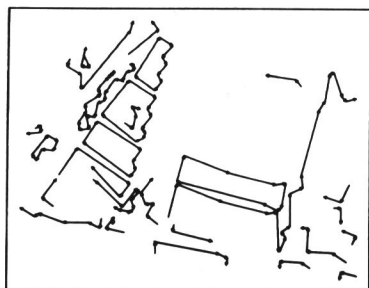
b) Line Segments



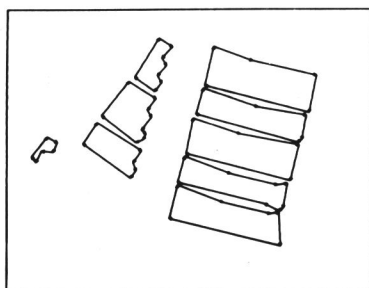
c) Corner Segments



d) Closed Chains



e) Open Chains and Nearby Evidence



f) Closed Boxes

Figure 4. Image 2

til the outline becomes a box or no new evidence is found.

Besides simple corners, the tracking process is likely to encounter many segment configurations at which a decision has to be made as to which path to follow. These include straight junctions, L-junctions with non-converging components, and various types of T-junctions. The CC and SC ratings are updated according to the path followed and evidence added to the chains. In some cases the process results in nearly closed outlines. A simple analysis of the CC and SC ratings show that only small portions of the outline are missing and, can easily be added (see the top building in figure 3c.) This is detected by noticing that either CC or SC is satisfied, with the unsatisfied constraint being "nearly" satisfied: If CC is satisfied (i.e. CC=4) then SC is "nearly" satisfied when the segments at the end of the chain are colinear and within a short distance. If SC is satisfied (within a small tolerance) then CC is "nearly" satisfied when only one corner vertex is missing (i.e. CC=3).

4. Combining Open Corner Chains. In this step we take each remaining open corner chain and search for evidence among other chains which allows us to create longer corner chains by pairwise joining of corner chains. We take each corner in an open chain and use its characteristics to predict the likely position and attributes of a matching compatible corner, according to the four types of geometric dispositions shown earlier in figure 2. Candidate corners in the remaining open chains are matched against the predicted corner. Suitable candidates for matching are determined by corner color, corner interpretation and, geometric constraints among the angles between the segment components of the two corners. Also, the region between the candidates is tested for uniformity. A successful match results in the combination of the two chains which include the matching corners (see the bottom building in figure 3e) and, the CC and SC measures updated. Since it is unlikely that chain combinations result in closed outlines, outlines with their CC and SC constraints nearly satisfied have their missing portions added (see figure 3f). Missing portions of sides facing the source of illumination are also added (see the bottom box in figures 4e and 4f).

5. Verification for Building Extraction and Description. We use the shadows to extract the boxes that can be said to correspond to buildings. If the segment components in a box have been interpreted as "very likely" to be the outline of a building, there is a record of a matching shadow. The recorded shadow width is used to estimate the height of the building and, the box attributes (perimeter, area, centroid, etc) are used to describe the building. In the absence of shadow matches (i.e. the corner components in the box are labeled "likely"), the width of a shadow is computed by analyzing the region adjacent to box segments that are constrained to cast a shadow if the box was a building. The method consists in growing several one-pixel wide slices over the shadow, starting at the edge of the box, in the direction of illumination until the shadow boundary is found.

The description obtained for the buildings found in the image shown in figure 3a is:

Building	Perim	Area	Orient	Centroid	Seg	Height
1	134.5	847.6	250.8	(81,110)	5	30.77
2	139.9	1166.3	158.5	(117,113)	9	54.26

For the image shown in figure 4a, we obtain a similar description for each individual box extracted.

IV CONCLUSION

We have described a building detector using shadows that is based on many simplifying assumptions about the scene. However the examples shown are from real aerial images and represent fairly complex scenes. We believe that basic methods can be extended for less constrained environments.

REFERENCES

- [1] R. Nevatia and R. Babu, "Linear Feature Extraction and Description," Computer Graphics and Image Processing, 13:257-269, 1980.
- [2] A. Huertas and R. Nevatia, "Edge Detection in Aerial Images Using Laplacian-Gaussian Filters," USCIP Technical Report 1010, pp. 16-26, September 1981.
- [3] M. Nagao and T. Matsuyama, "A Structural Analysis of Complex Aerial Photographs," New York, Plenum, 1980.
- [4] M. Takavoli and A. Rosenfeld, "Building and Road Extraction from Aerial Photographs," IEEE trans SMC, 12:1 (1982) 84-91.
- [5] D. Waltz, "Understanding Line Drawings of Scenes with Shadows," in Winston, P. H. (editor), The Psychology of Computer Vision, chapter 2, pp. 19-91, McGraw-Hill, 1975.
- [6] M.B. Clowes, "On Seeing Things," Artificial Intelligence, 2:2 (1971) 79-116.
- [7] D.A. Huffman, "Impossible Objects as Nonsense Sentences," Machine Intelligence, 6:7 (1971) 295-323.
- [8] S. Shafer and T. Kenade, "Using Shadows in Finding Surface Orientations," Carnegie-Mellon Technical Report CMU-CS-82-100, 1982.
- [9] D. Lowe and T. Binford, "The Interpretation of Geometric Structure from Image Boundaries," Proc. Image Understanding Workshop, pp. 39-46, 1981.