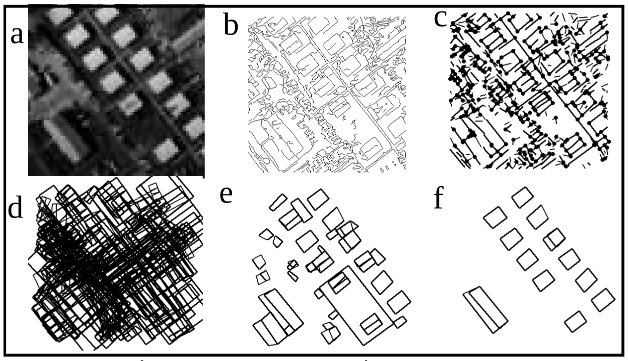

Figure 22. Fort Hood - Scene 4

structures yet. The rectangles verified however, represent a large majority of the components of the 40 or so structures in the image. Portions of the dark building on the lower right part of the image were only weakly hypothesized and thus, not selected for verification.


Figure 23. Modelboard image

## 6 Conclusion and future work

We plan to continue to extend our current system to detect the visible sides of buildings from oblique views of the scenes. This requires additional work un the use of the OTV that can be located. This will allow us to rely less on the shadow evidence, as it becomes more difficult to establish object to shadow correspondences. With oblique views, the shadows are likely to be occluded by the objects themselves or fall onto regions that belong to nearby structures. Currently we assume that the detected and verified structures lay on the ground. Some structures however are located on top of other structures. That level of refinement of the description requires an additional step in our system and is one of the subjects of our current and future work.

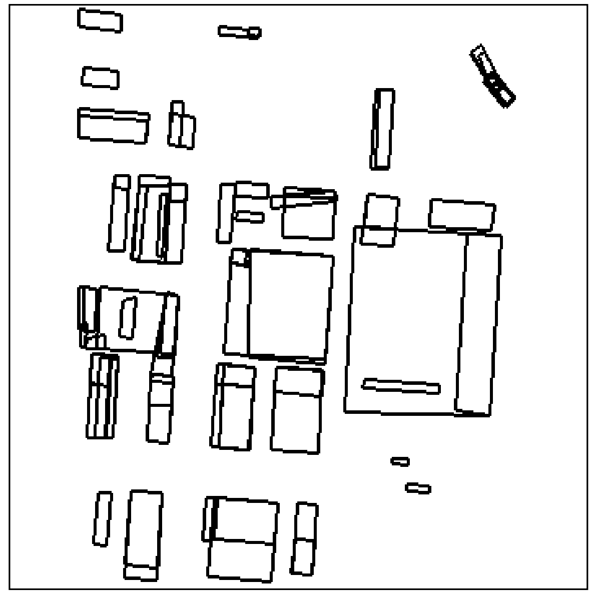
Figure 24. Verified Buildings

## References

[Canny, 1986]. J. Canny, “A Computational Approach to Edge Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679-698, November, 1986.

[Chung and Nevatia, 1992] C. Chung and R. Nevatia, “Recovering Building Structures from Stereo,” In Proceedings of the IEEE Workshop on Applications of Computer Vision, pages 64-73, Palm Springs, Calif., December 1992.

[Huertas, 1983]. A. Huertas, “Using Shadows in the Interpretation of Aerial Images,” University of Southern California USC-ISG Technical Report 104, October 1983.

[Huertas and Nevatia, 1988]. A. Huertas and R. Nevatia, “Detecting Buildings in Aerial Images,” Computer Vision, Graphics and Image Processing, 41(2):131-152, February 1988.

[Huertas etal, 1993]. A. Huertas, C. Lin, and R. Nevatia, “Detection of Buildings from Monocular Views of Aerial Scenes using Perceptual Organization and Shadows”, in Proceedings of the 1993 ARPA Image Understanding Workshop, Washington D.C., April 1993.

[Irvin and McKeown, 1989]. R. Irving and D. McKeown, “Methods for exploiting the Relationship Between Buildings and their Shadows in Aerial Imagery,” IEEE Transactions on Systems, Man and Cybernetics, SMC 19(6): 1564-1575, November/December 1989.

[Liow and Pavlidis, 1990]. Y. Liow and T. Pavlidis, “Use of Shadows for Extracting Buildings in Aerial Images,” Computer Vision, Graphics and Image Processing, 49: 242-277

[Mohan and Nevatia, 1989]. R. Mohan and R. Nevatia, “Using Perceptual Organization to Extract 3-D Structures,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(11):1121-1139, November 1989.

[Nevatia and Babu, 1980]. R. Nevatia and R. Babu, “Linear Feature Extraction and Description,” Computer Vision, Graphics and Image Processing, 13:257-269, 1980.

[Venkateswar and Chellappa, 1990]. V. Venkateswar and R. Chellappa, “A Framework for Interpretation of Aerial Images,” In Proceedings of the International Conference on Pattern Recognition, pages 204-206, Atlantic City, NJ, June 1990.

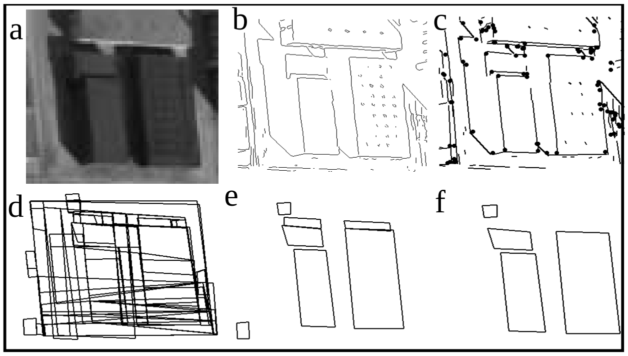

Figure 17. Modelboard - Scene 2 (oblique)

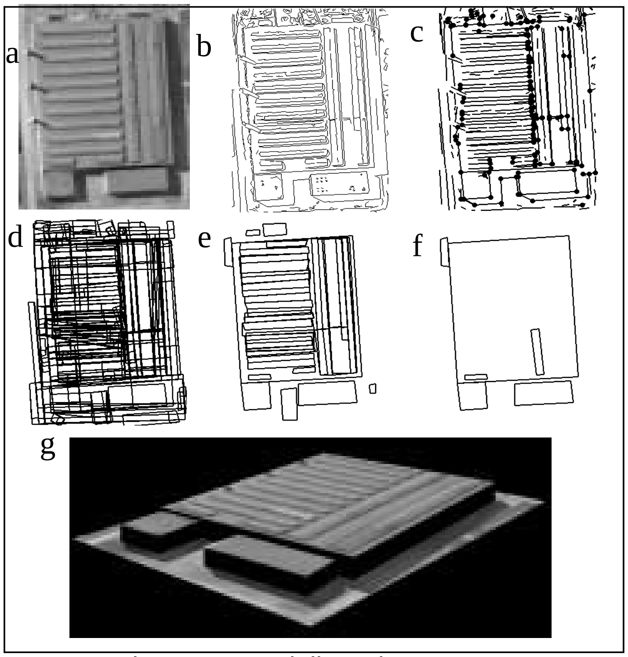

Figure 18. Modelboard - Scene 3

bitrary viewpoint.

Figure 19 shows another oblique view including some simple buildings. Note that the considerable fragmentation of the roof boundaries due to the features, such as windows, on the visible sides is tolerated well and reconstructed properly by the colinearization grouping.

Figure 20 shows a building in Ft. Hood, where some of the details of one of its sides is visible, apparently doors. These and the vehicles parked on the other side result in highly fragmented boundaries. The parallelograms verified by shadows include one that is formed from various aligned parked trailers which collectively cast a shadow. The small parallelogram on the bottom has a strong shadow junction corresponding to an actual narrow shadow cast by a vehicle. The lower wing of the building has a strong line and a corresponding medium junction. The rest of the shadow is diffused and is visible as a "dark" region adjacent to the building wings with no definite boundaries.

In Figure 21 the I-shaped building has no strong evidence of shadows. The parallelograms are weakly validated on the basis of a strong region which up a given maximum search distance remains "strongly" dark.

Figure 22 shows a group of small buildings arranged in a par-

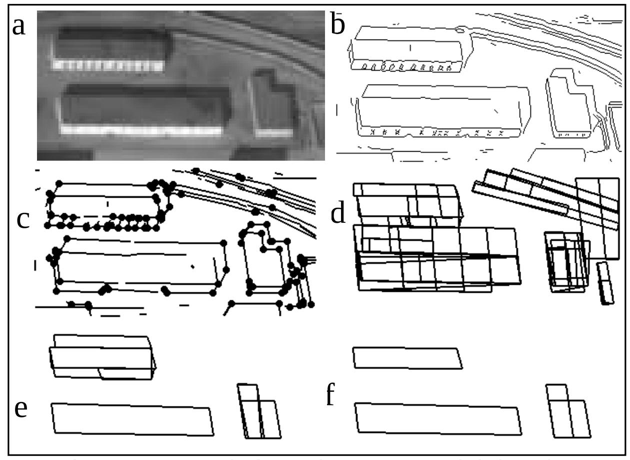

Figure 19. Modelboard - Scene 4 (oblique)

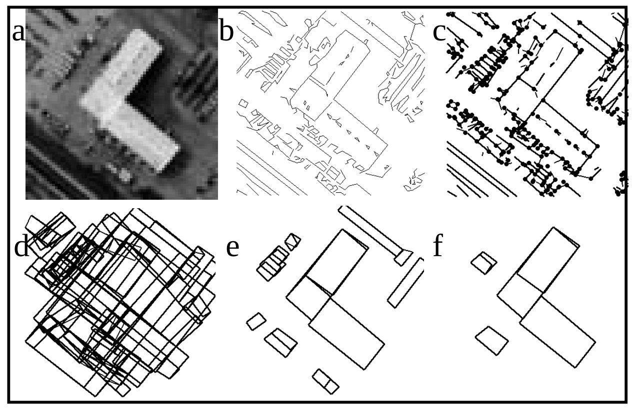

Figure 20. Fort Hood - Scene 2

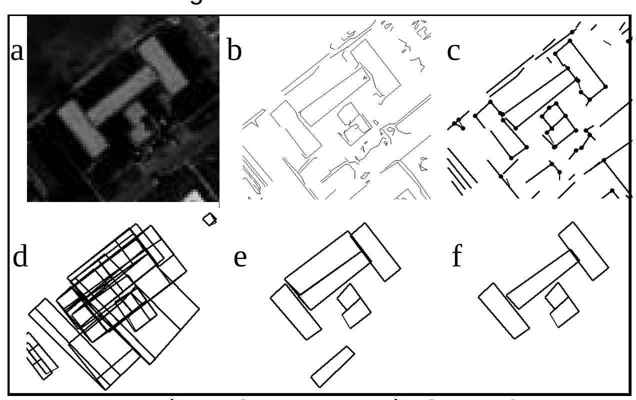

Figure 21. Fort Hood - Scene 3

allel fashion, and surrounded by other parallel structures. In spite the large number of hypotheses the system is able to select the relevant ones.

Finally, in Figure 23 we present an image from the RADIUS modelboard set containing a large number of structures (about 40). The system forms 1,724 hypotheses and selects 177. Some rectangles are selected but not verified. These correspond to dark low buildings with a small shadow that becomes merged with the building roof, and thus harder to verify. The system verifies 112 hypotheses on the basis of shadow evidence. Those verified on weak evidence (no object-to-shadow correspondences were possible) are excluded form the set of 75 shown in Figure 24. We have not implemented a step that combines these rectangles into

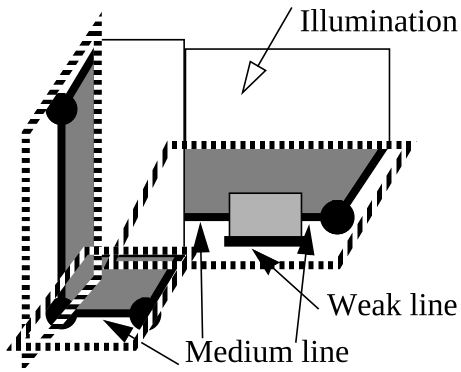


Figure 13. Windows to search for shadows evidence

quires that a minimum amount of the different kinds of evidence be present. Very high confidence requires that every kind of evidence be detected. Very low evidence is reported when no geometric correspondences can be established but the presence of a region, adjacent to and darker than, the parallelogram region itself, is found. The parallelograms selected on the basis of shadow evidence are shown in Figure 14.

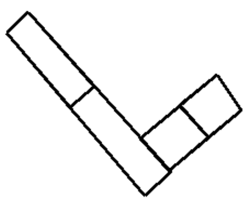

Figure 14. Verified parallelograms

### Use of shadow evidence

The parallelograms verified by shadows are used to generate an image containing the corresponding regions. The pixel values inside these regions encode the estimated height (as a function of the estimated shadow width and the sun incidence angle), thus giving an “elevation map” of the scene. This image can be viewed from an arbitrary viewpoint. The transform that projects the 3-D scene onto the 2-D screen for viewing can then be used to collect the pixel values from the input image, and use them to “paint” or render the various regions in the elevation map. Other 3-D representations such as wire frame models, can also be easily derived from the knowledge associated with the detected and verified parallelograms. A 3-D rendered arbitrary view computed from the parallelograms verified in our Ft. Hood example is shown in Figure 15.

Figure 15. 3-D view from another viewpoint

## 5 Results

Our system has been tested on a number of examples provided by the RADIUS program with very good results. We show a few to demonstrate the performance of our system and point out some of the sources of problems. As part of the RADIUS program, the system has also been ported to run on Unix workstations and transferred to two industrial sites and tested on some operational imagery. The results have been very promising and potentialy useful to the intended users. The speed of processing is a limitation however. It takes from 2 to 5 minutes to process a 512x512 image containing a few buildings. A 1320x1100 image with about 40 structures takes about one hour on a Sparc10/30. USC has currently a group working on parallel implementation of vision algorithms such as our system.

In the following figures, (a)shows an image, (b) the line segments extracted from it, (c) the linear structures and junctions computed, (d) the parallelograms hypothesized, (e) the selected hypotheses, and (f) the hypotheses verified by shadows. In particular, note figure (e), the excellent performance of the new selection technique. In the absence of shadow information, the selected parallelograms can be matched by our system if stereo views are available, thus providing verification and a 3-D model.

Figure 16 shows a set of four buildings and part of another. The difficulty here is with the building with the patterned arrangement of small objects on the roof. The shadows cast by these reach one side of the building causing it to be very fragmented. The shadow occluding the top left corner of the building and the poor boundary definition on the top right are also a source of difficulty. Accurate junction information can not be established and the systems must hypothesize a portion of the building. The strong shadow cues however help form parallelogram hypotheses for most of the building

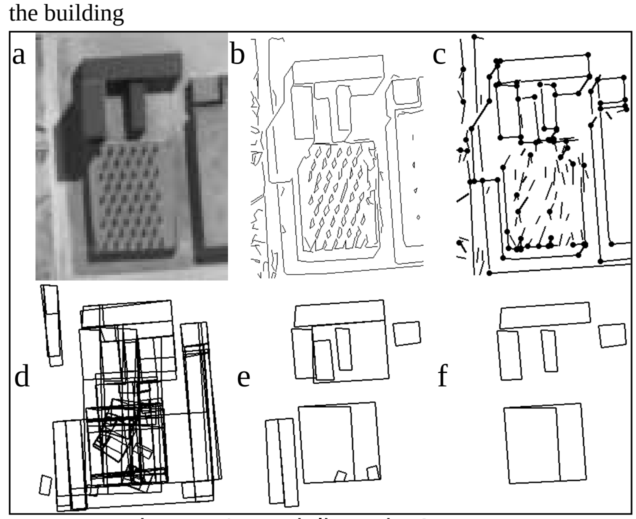


Figure 16. Modelboard - Scene 1

Figure 17 shows an oblique view of two dark buildings. The boundaries between dark buildings and shadows usually have low contrast and are difficult to detect.

Figure 18 shows a complex building with numerous rectangular components on the roof. We are able to exploit the presence of strong shadow evidence here. It allows the system to form a hypothesis for the entire building in spite of the broken and fragmented boundaries. Note that the selection mechanism is able to select most of the rectangular components on the roof as well. Figure 18g shows a 3-D rendered view of the building, from an ar-

## 4.1 Shadow analysis

By shadow analysis we mean the establishment of correspondences between shadow casting elements and shadows cast, and the use of these correspondences to verify and model 3-D structures. We assume that the ground surface in the immediate neighborhood of the structure is fairly flat and level. The shadow casting elements are given by the sides and junctions of the selected parallelogram hypotheses. The shadow boundaries are located among the lines and junctions computed earlier from the image.

There are a number of difficulties that prevent the accurate establishment of correspondences however. Building sides are usually surrounded by a variety of objects such as loading ramps and docks, grass areas and sidewalks, trees, plants and shrubs, vehicles, light and dark areas of various materials. Nearby structures may reflect light into the shadowed areas making the objects in it more visible, and so on. To deal with these problems we have adopted the following definitions, criteria and geometric constraints to analyze the shadows adjacent to parallelograms (see Figure 11):

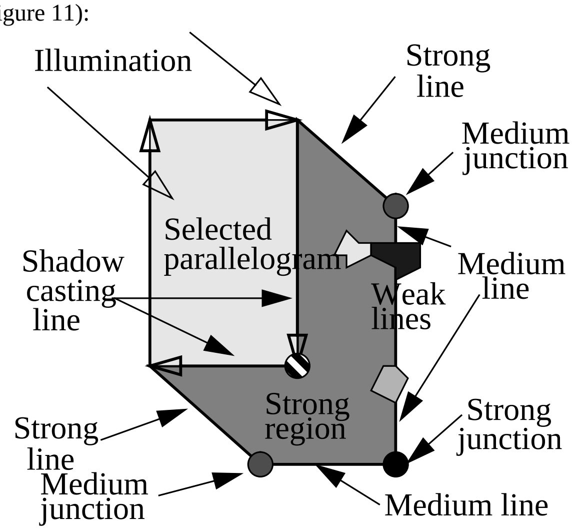


Figure 11. Shadow features

**Strong junctions:** Matching junctions along the direction of illumination, having a consistent shape and a consistent attitude. These junctions constitute the strongest monocular cue to the presence of a 3-D structure. We use knowledge of these correspondences also to help form and select parallelogram hypotheses.

**Strong lines:** Shadow boundaries cast by vertical object edges. We use this evidence also during hypotheses formation and selection.

**Medium lines:** The parallelogram sides that are supposed to cast shadows must have corresponding shadow lines.

**Medium junctions:** The junctions formed by strong and medium lines, found along the direction of the strong lines.

**Weak junctions and lines:** Junctions and breaks in the shadow boundaries between the strong and weak junctions.

**Strong regions:** Dark regions surrounded by strong and medium junctions. We require that this region be darker than the parallelogram region regardless of their gray level.

**Weak regions:** In the absence of geometric correspondences of junctions and lines, a dark region adjacent to parallelogram, consistent with the direction of illumination.

## 4.2 Shadow process

The shadow process consists of four steps:

### Extraction of potential shadow evidence

Potential shadow evidence consists of lines, junctions and intensity statistics. We extract the following:

- Lines parallel to shadow boundaries cast by vertical edges. They represent potential shadow lines cast by 3-D structures in the image.
- Lines having their dark side on the side of the illumination source are potential shadow lines.
- Junctions among the lines above.
- Pixel statistics to compare relative brightness.

The potential shadow lines and junctions extracted from the lines in our Ft. Hood example are shown solid in Figure 12. The underlying edges are shown in gray.

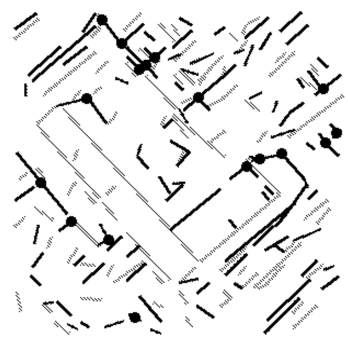

Figure 12. Potential shadow lines and junctions

### Search for shadow evidence

For each parallelogram we look in a search window (dashed lines in Figure 13) and collect all the potential shadow evidence in it. The search distance is arbitrarily chosen as a function of the maximum expected building height and the sun incidence angle. There is the possibility that lines, not relevant to the current parallelogram, be included. They however, have a reduced effect in the presence of the real evidence.

We favor medium and weak lines that are parallel to the parallelogram side. In some cases there may be various sets of lines, all parallel to the building side but at various distances from the parallelogram side. This is actually a common occurrence since many side walks, grass areas, streets, vehicles and so on, will be found to be arranged or located parallel to building sides. In this case we choose those shadow lines at the distance from the parallelogram side such that the sum of their lengths is greater, but not exceeding the length of the parallelogram. We determine the width of the shadow by averaging the distance to the lines selected. The selected evidence is then considered to surround the shadow region. We compute the mean intensity of this region and compare it to the parallelogram region.The evidence collected for both sides is combined to give the evidence for the parallelogram.

### Evaluation of Shadow Evidence

We evaluate the shadow evidence and give a confidence value as a weighted sum of the evidence of strong junctions, medium junctions, strong line, weak lines, strong and weak regions. We designated five levels of confidence. Each level of confidence relines crossing any side of a parallelogram, existence of L-junctions or T-junctions in any side of a parallelogram, existence of overlapping gaps on opposite sides of a parallelogram, and displacement between four sides of a parallelogram and its corresponding edges support. See Figure 9. Negative evidence is as important as positive evidence, because it helps us to remove those parallelograms which are less likely to be part of buildings.

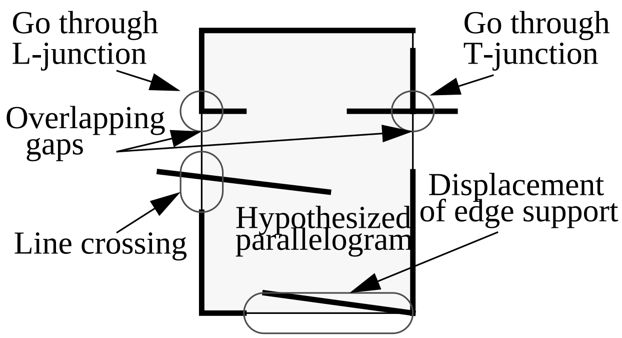


Figure 8. Negative evidence

Each kind of evidence of support is formulated into an evaluation criterion. There are no formal definition of goodness of a parallelogram thus, our evaluation criteria formulated from evidence of support are based on analysis of likely and unlikely events. For example, four junctions are very unlikely to fall on the four corners of a parallelogram accidentally. So the existence of four corners on a parallelogram strongly suggests that the parallelogram is good. Also, from the Gestalt Laws of Perceptual Grouping, the Law of Closure suggests that the existence of L-junctions or T-junctions on a side of a parallelogram will make a closure on part of the parallelogram and it means that the hypothesized parallelogram is not good. Some evidence of support are not always available such as the shadow evidence and the OTV corner evidence, but they are very important because it is very unlikely that some shadow features will appear around the hypothesized parallelogram by chance and the probability for three lines to form an OTV corner by chance is very small. We can emphasize the importance of an evaluation criterion by assigning a higher weight to it.

Positive weights are assigned to those evaluation criteria formulated from positive evidence of support while negative weights are assigned to those evaluation criteria formulated from negative evidence of support. A weight should be assigned to each evaluation criterion according to the probability of existence of buildings under the condition of presence of the evidence of support from which the evaluation criterion is derived. However, we do not have the probabilistic analysis of goodness of a parallelogram, but the problem of optimal weights assignment for a given set of examples could be formulated into a search problem.

Good parallelograms surviving local selection may compete with each other. For example, some parallelograms could share the same edges or corners support and some parallelograms might overlap with each other. The goal of global selection criteria is to select a minimum set of parallelograms which best describe the rectangular composition of the scene.

Global selection criteria examine overlapping parallelograms and choose one if appropriate. The selection is based on relative properties of each parallelogram, the amount and kind of overlap, and whether they share support or not. Note that a parallelogram fully contained in another is not necessarily removed. If a parallelogram does not overlap with any other parallelogram then it is not in competition, and it remains. There are four global selection criteria in our system. They are the criterion for duplicated parallelograms, the criterion for mutually contained parallelograms, the criterion for fully contained parallelograms, and the criterion for overlapping parallelograms.

It is very easy to extend and improve the criteria-based selection process. If a new kind of evidence of support is found to be crucial for the goodness of a parallelogram, we can formulate an evaluation criterion from the evidence of support and merge the evaluation criterion to the original set of evaluation criteria by assigning appropriate weight to it and adding the weighted value to the goodness value. On the other hand, if a new global relationship between parallelograms is found to be important, we can also implement a new filter to enforce the relationship and add the new filter in appropriate position to the original pipeline of filters.

The parallelograms selected in our Ft. Hood example after both the local and global selection criteria have been applied are shown in Figure 9.

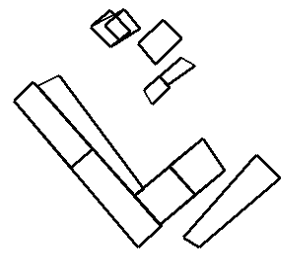

Figure 9. Selected parallelograms

## 4 Verification of hypotheses

The purpose of verification is to validate the selected hypotheses to correspond to buildings. Our validation step segments the objects, generates a description of the shape of the structures and derives a 3-D model. The use of shadow evidence, discussed below, uses methods described in [Huertas,1983, Huertas and Nevatia, 1988, Huertas etal, 1993] with the appropriate extensions to handle oblique views. Oblique views require at least two sun angles (see Figure 10), the direction of illumination and the sun incidence angle. For testing we have gotten these angles from image measurements.

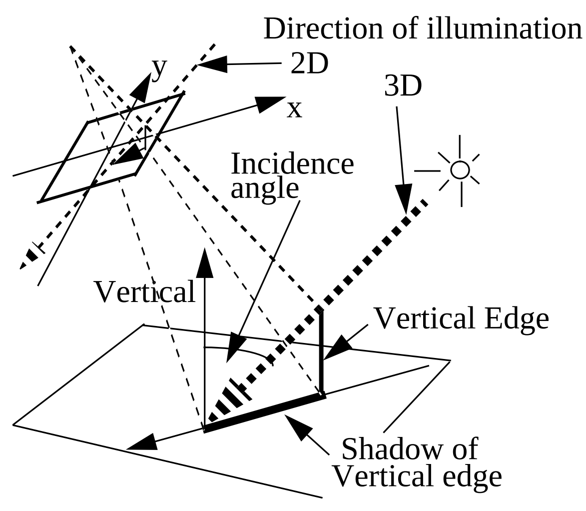


Figure 10. Sun angles & oblique shadow geometry

**Lines and junctions**

A group of close parallel lines represent a linear structure at a higher granularity level than the edges (see the common boundary between the building wings in Figure 2..) The resulting lines have a length and an orientation derived from the contributing elements. Figure 5 shows the lines obtained from grouping the segments in Figure 2. We use these lines to detect L-junctions and T-junctions also shown in Figure 5. For oblique views we also look for evidence of vertical edges in the immediate neighborhood of the L and T-junctions, thus allowing us to detect potential OTV's. Vertical edges are detected by looking for line segments that are parallel to the image's principal line.

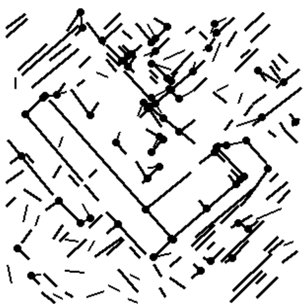

Figure 5. Linear structures and junctions

**Parallels and skewed U-structures**

Structures in urban scenes like buildings, roads and parking lots are often organized in regular grid-like patterns. These structures are composed of parallel sides. As a consequence, for each significant line-structure detected in the scene, there is not one but many lines parallel to it. For each line, we find lines that are parallel and satisfy a number of reasonable constraints. Note that the formation of a parallel structure also aids in the formation of new lines, as they suggest extension and contraction of the parallels to achieve full skewed overlap.

When the two lines in a parallel structure have their ends aligned as a function of the viewing angles, they strongly suggest the presence of a line with which the parallel structure would form a skewed U-structure. Even if the third line does not exist in the set of lines, we hypothesize it and generate the U-structure.

**Skewed rectangles or parallelograms**

Skewed rectangle or parallelogram structures are generated from the U-structures. The parallelograms formed in our example are shown in Figure 6. In practical applications this number can be reduced by restricting the formation of parallelograms on the basis of size, as a function of image resolution, for example. Parallelograms are also generated from matching junctions along the direction of illumination (see strong junctions in section 5.) We hypothesize the missing portions of a parallelogram having a corner with a matching shadow corner or evidence of an OTV.

## 3 Selection of hypotheses

After the formation of all reasonable parallelograms, a selection process is applied to choose parallelograms having strong evidence of support and having minimum conflict among them. Earlier versions of our system used a Constraint Satisfaction Network (CSN) [Mohan and Nevatia, 1989]. In the current system, we use a criteria-based method which seems to give much more predictable results. Next we only summarize our current method due to lack of space

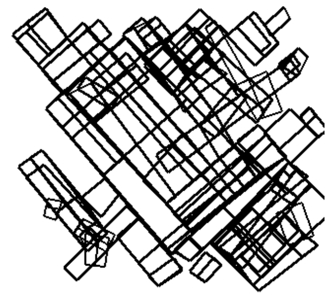

Figure 6. Parallelogram hypotheses generated

Our new system uses two kinds of criteria: local selection criteria and global selection criteria. Local selection criteria determine whether or not a parallelogram is "good" based on the local supporting evidence. Only good parallelograms are retained for global selection. It is possible that some of the good parallelograms retained after the local selection are mutually contained or duplicated or overlapped with some other good parallelograms. Global selection criteria select the best consistent parallelograms from good parallelograms.

We apply local selection criteria and global selection criteria differently. Local selection criteria (also called evaluation criteria) work together to evaluate the goodness of a parallelogram, while global selection criteria work separately. Each global selection criterion acts like a filter. The set of retained parallelograms pass through all filters and the set of parallelograms coming out from the last filter will be the set of parallelograms selected by the selection process.

The local selection criteria are used to remove parallelograms formed using weak evidence. For each parallelogram the evaluation criteria compute a goodness value. If this value exceeds a given threshold, the parallelogram is selected, otherwise the parallelogram is removed.

Every evaluation criterion is weighted according to its importance. The goodness of a parallelogram is then measured by the sum of the weighted values calculated by the evaluation criteria. The problem of measuring the goodness of a parallelogram now becomes a problem of finding and formulating good evaluation criteria, and assigning appropriate weights.

Whether a parallelogram is good or not depends on the evidence of support. We distinguish between positive evidence and negative evidence of support for a parallelogram. The positive evidence of support includes the presence of edges, corners, parallels, OTV's and shadows (see Figure 7).

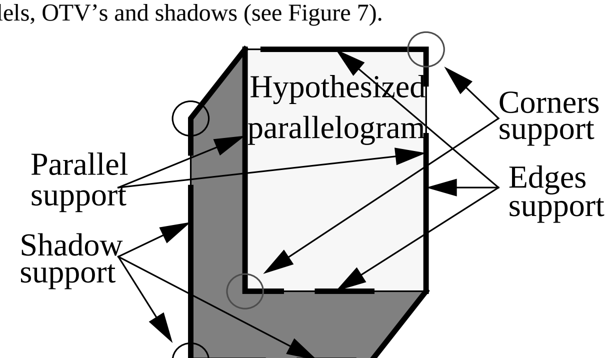


Figure 7. Positive evidence

The negative evidence of support includes the presence of tia, 1988, Venkateswar and Chellappa, 1990]. These are essentially local techniques that must make a decision of which path to trace at each local junction. Of course, all paths could be traced using backtracking but then the search space may become prohibitively large. Region based techniques construct closed curves that often do not correspond to the objects of interest.

Model based techniques can deal with fragmentation but require a-priori shape models. For example, it is not sufficient to say that the building is a rectangular parallelepiped; you must also supply the relative dimensions of the sides. In summary, these systems have shown interesting performance but on limited examples. None of these systems can generate a description of the buildings at the level of shape descriptions of the different wings.

We have proposed, instead, to use a perceptual grouping approach. Cultural features such as buildings represent structures that are not random but have specific geometric properties. In this we restrict the shapes of buildings to be a single or a composition of rectangular parallelepipeds (thus allowing L, T and I shapes for example).

Previous systems have assumed that the viewpoint is more or less overhead. The system described here uses the viewpoint angles (swing and tilt) needed to deal with images acquired from an oblique viewpoint. The geometric constraints relevant to shape take into consideration, as a function of the viewpoint angles, the expected skewness of the rectangular surfaces that most buildings are expected to have. This property is used to organize the detected line segments into roof hypotheses. While the visible building sides (walls) can be hypothesized similarly, we do not handle them now. We believe that this approach leads to many fewer hypotheses than would be generated by a complete contour tracing scheme.

Our approach combines several of the techniques from previous work. Our perceptual grouping approach comes from the work described in [Mohan and Nevatia 1989], however, we use a very different hypotheses selection technique. Mohan and Nevatia, in fact, used perceptual grouping for stereo analysis, here we apply it to monocular analysis. Our shadow analysis method is an extension of the approach first described in [Huertas, 1983, Huertas and Nevatia, 1988].

The diagram in Figure 3 shows the main components in our system. The system uses the line segments approximating the intensity boundaries to compute linear structures and relevant junctions among them. A hierarchy of features including parallel relationships and portions of skewed rectangles or parallelograms leads to the formation of building hypotheses. These consist of instances of rectangular shapes that potentially correspond to building roofs and parts of building roofs (see section 2). Next, promising parallelograms are selected and verified to correspond to roofs of building structures. Shadow information, if available, is used to help form, select and verify hypotheses. It also, as a function of the sun angles, provides estimates of the height of the structures, leading to a 3-D description of the scene.

Our philosophy in the design of this system has been to make only those decisions that can be made confidently at each level. Thus, we choose to generate as many hypotheses as seem feasible at the first level. Our selection process too is conservative and favors keeping hypotheses that may be viable. The verification process has the most global information and can make stronger decisions. Even here, if our system is to be embedded in a larger system, some of the decisions would be deferred to that system where more context is available for decision making.

The technique we describe in this paper, we believe, significantly extends the range of scenes that can be analyzed though

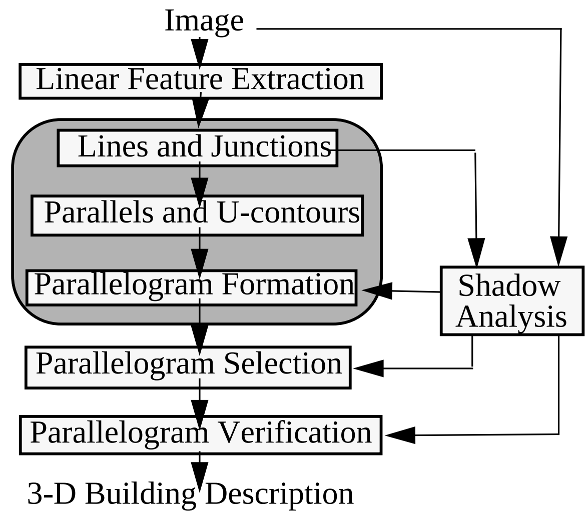


Figure 3. Block Diagram of the System

many problems remain. We show several examples taken from the images provided by the U.S. Government sponsored RADIUS program to demonstrate the effectiveness of our technique. Also in the context of the RADIUS program we have transferred our software to two industrial sites and continue to support their testing tasks. The results in general have been very good, and the experience of technology transfer, albeit a difficult one, has been a successful one.

## 2 Generation of hypotheses

The process of hypotheses formation is similar to the one described in [Mohan and Nevatia, 1989] with the appropriate extensions to oblique views and the use of strong shadow cues (if available). In this process we construct a feature hierarchy which encodes the structural relationships specific to oblique views of rectangular shapes, presumably corresponding to the visible roof surfaces: Lines, skewed parallels, skewed U-contours, and skewed rectangles or parallelograms. The degree of skewness is computed as a function of the swing and tilt angles denoting the viewpoint. Figure 4 show the angles involved. We expect that images from aerial scenes have a camera model associated with them from which these angles can be derived.

Next, we describe the hierarchy of features in our system:

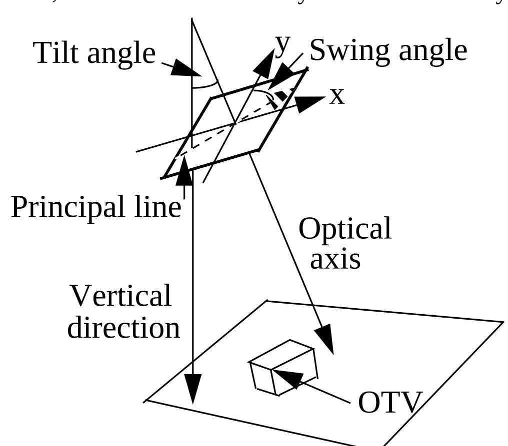


Figure 4. 3-D Viewpoint angles

# Detection of Buildings Using Perceptual Grouping and Shadows

C. Lin, A. Huertas and R. Nevatia*
Institute for Robotics and Intelligent Systems

University of Southern California
Los Angeles, California 90089-0273

## Abstract

*We describe a system for detection and description of buildings in aerial scenes. This is a difficult task as the aerial images contain a variety of objects. Low-level segmentation processes give highly fragmented segments due to a number of reasons. We use a perceptual grouping approach to collect these fragments and discard those that come from other sources. We use shape properties of the buildings for this. We use shadows to help form and verify the hypotheses generated by the grouping process. This latter step also provides 3-D descriptions of the buildings. Our system has been tested on a number of examples and is able to work with overhead or oblique views.*

## 1 Introduction

The goal of this work is to detect and describe buildings from monocular views of arbitrary aerial scenes. This is a difficult but important task for many applications such as photo-interpretation, cartography and surveillance. Building detection is difficult for several reasons. The contrast between the roof of a building and surrounding structures such as curbs, parking lots, and walkways can be low. The contrast between the roofs of various wings, typically made of the same material, may be even lower. Low contrast alone is likely to cause low-level segmentation to be fragmented. In addition, small structures on the roof and objects, such as cars and trees adjacent to the building will cause further fragmentation and give rise to "noise" boundaries. Roofs may also have markings on them caused by dirt or variations in material. Shadows and other surface markings on the roof cause similar problems.

There are other characteristics of these images which may cause problems. Roofs have raised borders which sometimes cast shadows on the roof. This results in multiple close parallel edges along the roof boundaries and often these edges are broken and disjoint. At roof corners and at junctions of two roofs, multiple lines meet leading to a number of corners making it difficult to choose a corner for tracking. A roof cast a shadow along its side and often there are objects on the ground such as grass, trees, trucks, pavement, etc., which lead to changes in the contrast along the roof sides.

* This research was supported primarily by a subcontract from Hughes Aircraft Co. on the RADIUS program, and in part by the Defense Advanced Research Projects Agency under contract F49620-90-C-0078, monitored by the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

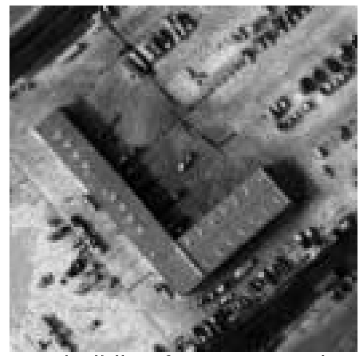

Figure 1. A building from Ft. Hood, Texas

Consider the building in Figure 1 (from a scene of Ft. Hood in Texas.) We use an overhead view as a running example, for simplicity. The building is easy for humans to see and describe, but it is difficult for computer vision systems. Figure 2 shows the line segments detected in the image, using LINEAR, our linear feature extraction software [Nevatia and Babu, 1980, Canny, 1986]. We are still able to see the roof structures of the buildings readily and easily, but the complexity of the task now becomes more apparent. The building boundary is fragmented, there are gaps and missing segments. There are also many extraneous boundaries caused by other structures in the scene.

Figure 2. Line segments extracted from the image

There have been many previous attempts to solve this problem [Huertas, 1983, Huertas and Nevatia, 1988, Mohan and Nevatia, 1989, Huertas etal 1993, FuaHan87, Irvin and McKeown 1989, Liow and Pavlidis, 1990, Venkateswar and Chellappa, 1990]. Building detection requires robust segmentation techniques and methods to infer the 3-D structure. These methods rely on edges or regions extracted from the image. Simple edge-based methods attempt to collect linked edge curves into the desired object boundaries, and succeed only for relatively simple scenes. Some edge-based methods have used some form of a contour tracing technique, see for example [Huertas, 1983, Huertas and Neva-