Detecting Changes in Aerial Views of Man-Made Structures

Andres Huertas and Ramakant Nevatia*
Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, California 90089-0273
{nevatia|huertas}@iris.usc.edu
www: iris.usc.edu

Abstract

Many applications require detecting structural changes in a scene over a period of time. Comparing intensity values of successive images is not effective as such changes don't necessarily reflect actual changes at a site but might be caused by changes in the view point, illumination and seasons. We take the approach of comparing a 3-D model of the site, prepared from previous images, with new images to infer significant changes. This task is difficult as the images and the models have very different levels of abstract representations. Our approach consists of several steps: registering a site model to a new image, model validation to confirm the presence of model objects in the image; structural change detection seeks to resolve matching problems and indicate possibly changed structures; and finally updating models to reflect the changes. Our system is able to detect missing (or mis-modeled) buildings, changes in model dimensions, and new buildings under some conditions.

1 Introduction

One of the key applications for aerial and space image analysis is for detecting significant changes on the ground. This could be for various purposes such as urban planning, agricultural analysis, environmental monitoring and military intelligence. We focus on changes in man-made structures, particularly buildings, rather than changes in the vegetation. Note that while changes in the desired structures should be reflected in some changes in the image, not all changes in the images may be caused by 3-D structural changes. Image intensities can vary due to a number of factors such as changes in illumination, viewpoint, atmospheric conditions. Further, seasonal variations may cause changes in vegetation and ground cover; while these represent "real" changes on ground, they may not correspond to structural changes.

There is little previous work in structural change detection. Early work in change detection was based on determining pixel intensity changes [Lillestrand,1972], We believe that the solution to finding structural changes lies in not comparing images taken at different times directly but rather in comparing new images (or descriptions derived from them) to an abstract model derived from previous observations; such models have come to be called *site models* [Gerson & Wood, 1996]. This approach also allows an updating of the site models which may be one of the prime goals of the change detection analysis.

A new image can not be directly corresponded to an abstract model. Instead, we must compute descriptions from the image that can be corresponded with the model or descriptions that can be derived from the model. This is a common problem in object detection in computer vision and various techniques such as *alignment* have been developed to solve it [Huttenlocher & Ullman, 1990]. The change detection problem is simpler to the extent that some parameters of the *pose* of the objects may be known *a priori*. However, the objects themselves may have changed and not fit a prior model exactly. Also, aerial images typically contain a large number of man-made and natural objects, not all of which may have been modeled (for example, we do not assume models for trees in a scene). The objects of interests may be partially (or totally) occluded by other objects and shadows cast by them may cause confusion. The images also contain significant amount of texture, thus leading to a large number of features at the lower-levels (such as edges) that prohibit use of combinatorial techniques to search for desired objects. Finally, we need to verify that the suspected changes actually correspond to some 3-D structures and to derive a description of the changes where possible.

The system described in this paper is designed for detecting changes in 3-D building structures. We further assume that the buildings are rectilinear and have either flat or simple gable roofs. Composite shapes (such as "L" or "T") are allowed. Each part is represented by its 3-D wireframe (consisting of any number of vertices and edges). We assume that the camera geometry and approximate viewpoint from which images are taken are known. Spe-

* This research was supported by Contract No. DACA-76-93-C-0014 from the Defense Advanced Research Projects Agency (DARPA) and monitored by the Topographic Engineering Research Center of the U.S. Army.

cifically, we assume that the errors are such that a *projected* model can be corresponded with the image by *translation* only; this is a reasonable assumption for many imaging situations; errors in other parameters would only affect the registration stage of our system as described below.

The change detection process consists of the following five major steps:

- *Site Model to Image Registration*: This step is to register the new image(s) to the stored site model(s); our system uses line feature matching.
- *Site Model Validation*: This step verifies whether the objects in the site model are present in the new image by comparing predicted features with observed features. A confidence value is computed for each object in the model; low confidence values are likely to represent possible changes to the objects.
- *Structural Change Detection:* In this step, we analyze in more detail possible change in the site indicated in the previous step, and determine if the missing correspondences can be explained by the imaging and viewing conditions or whether evidence exists for actual changes. Our system is able to detect missing (or mis-placed) buildings, buildings with dimensional changes, and new buildings under certain conditions.
- *Site Model Updating*: In this step, 3-D models of changes are constructed where possible. These can be reported to a human analyst and reflected in an updated site model (which can then be used to process new images at the next cycle).

The following sections describe the processing at each step and illustrate with an example. More results and evaluations are described in section 6. Our system has been tested primarily on data available for the Ft. Hood, Texas, site. The site models for our tests were constructed by using tools provided in the Radius Common Development Environment (RCDE) [Strat et al 1992; Fua, 1996]. The kinds of changes we are looking for occur over relatively long periods of time; unfortunately, we were not able to acquire data reflecting such changes for this site. Instead, we have modified the site models which should have the same effect as our system only compares images with site models rather than previous images. This method also provides a check on the accuracy and validity of previous site models. Our system has been ported to an industrial laboratory for possible use in current applications.

2 Site Model to Image Registration

The first step is to register a site model to an image. In our task, it is reasonable to assume that the imaging parameters are known to some accuracy and that the errors are such that if a site model is *projected* by the known parameters, its features will correspond with those from the image except for translational errors (which may be quite large, in the order of tens of pixels). The registration problem is then that of determining this translation, which we model as being uniform across the image. We need to decide what features of the models and image should be matched to determine the translation. The models are abstract, 3-D wire frame structures, the image is a 2-D array of intensity values. We have chosen to match lines extracted from the image with the lines projected from the site models by the known (approximately) camera geometry. The line matching technique is adapted from an earlier method [Medioni et al., 1990] and has been described in [Huertas, et. al, 1995].

Figure 1 shows the line segments extracted from portion of an image of the site. The model registered with the image is shown in Figure 2. We find this process to be quite robust, whether applied to small windows containing just a few buildings and to very large windows containing many tens of buildings (and other structures which may not be in the site model).

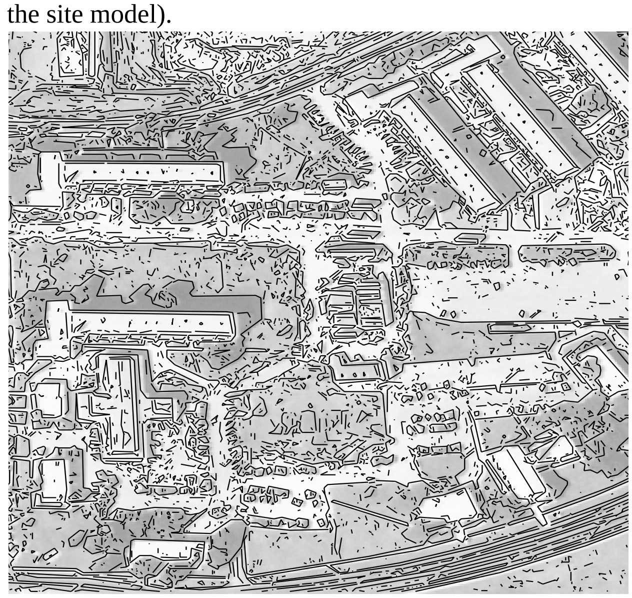


Figure 1. Line segments extracted from a portion of an image from Fort Hood, Texas

3 Site Model Validation

The purpose of model validation is to verify whether the objects in the site model are present in the new image in the same form or whether they should be examined in more detail for evidence of significant changes. The previous step (registration) provides a correspondence between model and image segments. For model validation, we combine evidence from a variety of object features such as lines and junctions. We also look for 3-D evidence by evaluating expected shadows cast on the ground; in multiple images, this could come from feature correspondence information. Note that not all the features of the model may be visible in the image, some will be missing due to self and mutual occlusion. These occlusions can, however, be predicted from the viewing geometry and accounted for. There will also be missing evidence due to difficulties of feature extraction in images: low contrast edges may not be detected and line segments fragmented due to surface and ground texture.

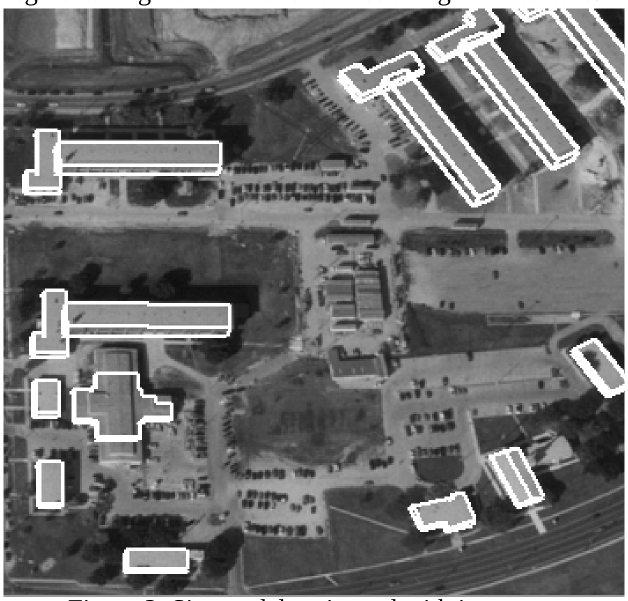


Figure 2. Site model registered with image

Before confidence values are calculated (as explained in section 3.3), the system deals with a number of problems and ambiguities inherent to any matching process. We discuss these next.

3.1 Missing Features

To validate a model accurately we need to study the source of missing model-to-image correspondences. Some missing image features will be due to viewing conditions such as self-occlusion, occlusion by other objects, self shadows and shadows cast by nearby objects. These, however, can be predicted and explained from the site model itself. Missing correspondences may be due to over- or under-modeling of objects (Figures 3 and 4) and are more difficult to predict from the model. The confidence associated with over- or undermodeled objects may thus be underestimated or difficult to calculate. Over-modeling is due to the use of modeling primitives that introduce elements that do not correspond to actual physical elements or boundaries. Figure 3 shows a building that has been modeled by two rectangle parallelepipeds. The thick lines represent portions of the elements on the building model that do not correspond to physical boundaries. These can not be matched and the missing correspondences result in lower confidence. Figure 4 shows two buildings that are likely to be under-modeled (i.e. modeled by simpler shapes) due to their complexity. These require additional search strategies designed to look for additional and possibly fragmented evidence, such as a large number of vertical or horizontal edge elements. Our system is not currently capable of determining these conditions, and thus the confidence values may be underestimated. Some of these conditions may require annotations in the site model to help the system process these appropriately.

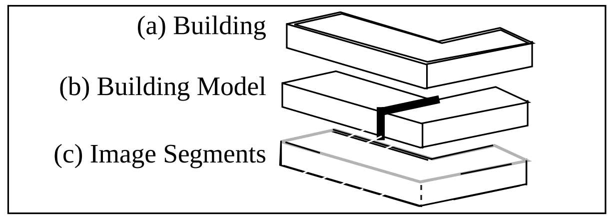


Figure 3. Missing match due to over-modeling.

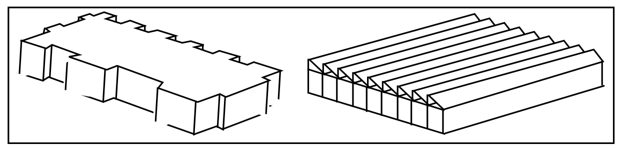


Figure 4. Some buildings may be under-modeled

3.2 Ambiguities in Matching

The system currently deals with two ambiguities inherent to the matching process: multiple or missing matches, and coincidental alignments due to viewpoint, illumination direction, or due to adjacent structures.

3.2.1 Multiple Matches

The model-to-image matcher corresponds each model element with one or more image elements (Figure 5) possibly involving more than one object. Allowing multiple matches is necessary to deal with expected fragmentation in the image elements. Fragmentation is due to inadequacies in the feature extraction process and due to actual image content, such as occluding trees, road boundaries and shadows. If a model segment matches multiple colinear image segments, all the image segments are considered to represent image support. If a model segments matches multiple parallel image segments, the overlap among these is considered to represent image support.

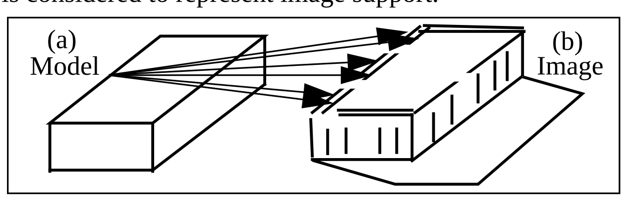


Figure 5. One-to-many correspondences.

3.2.2 Coincidental Alignments.

Some multiple matches are due to coincidental alignments of buildings with other structures (Figure 6). Some of these include roads, walkways, lawns, shadows and other adjacent objects. Nearby objects and shadows sometimes result in image features that have a larger extent than that predicted by the model features. These are explained by examining nearby shadows with knowledge of the direction of illumination, and by examining adjacent structures. Coincidental alignments due to nearby and adjacent structures are determined by looking for adjacent structures that help explain alignment, or a possible change in horizontal dimensions. An example of some ambiguities and alignments is shown in Figure 7.

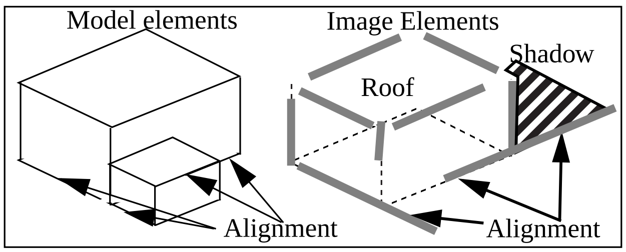


Figure 6. Coincidental alignments.

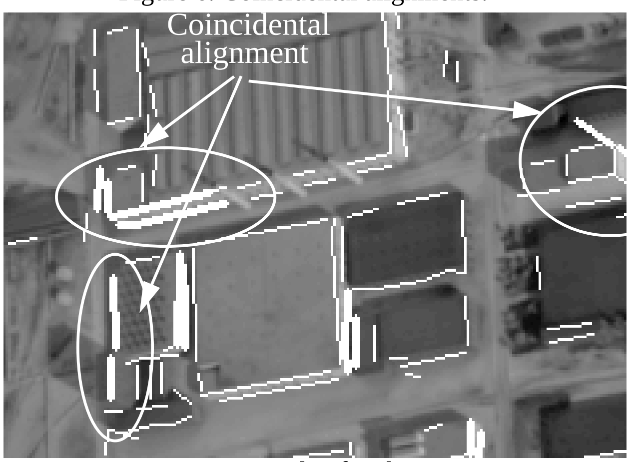


Figure 7. Example of ambiguities

3.3 Validation Confidence

We evaluate five kinds of evidence: edge visibility, edge presence, edge coverage, junction presence and shadow presence (these terms are explained below). Each evidence provides a score between 0 and 1 and a combined confidence score is computed by a linear weighted combination of them. We have chosen this method of combination for its simplicity. It works well in our tests but there may well be more optimal ways of combining such evidence. Here we have chosen a simple linear combination formulation with empirically assigned weights applied to the evidence terms. The weights however reflect the relative importance of the evidence but remain to be optimized over a larger sample of experiments. The confidence values derived take into account only visible elements from the particular viewpoint of the image after accounting for self and mutual occlusion. (Figures 8 and 9):

Let x be a model object defined by a set of vertices and a set of edges. For each object, x, we calculate a confidence value $C(x)$ as a contribution of the following terms:

Edge Visibility: $V(x)$ is given by the fraction of the model edges that are visible from the current viewpoint. Fewer visible edges result in lower confidence.

Edge Presence: $P(x)$ is defined as the fraction of the visible model edges that are matched to some image edges. In the schematic example shown in Figure 8 (a), all nine visible edges (dashed lines) have correspondences in the image (solid lines), giving a P value of 1.0. An object that is only 50% visible but that has the visible 50% corresponded to image edges has a P value of 1.0 also. P is calculated separately for **roof** elements, **vertical** wall elements and **base** wall elements which are given different weights (roof evidence is considered the most reliable, the wall base the least reliable).

Edge Coverage: $E(x)$ is defined as fraction of the *lengths* of the visible model edges that is actually covered by some image segments. Figure 8 (a) shows an object where all the model edges (dashed) have some, but small coverage; this object has good presence but poor coverage. Figure 8 (b) shows the opposite; a few model edges have good image edge support so the coverage is good but presence is not. $E(x)$ is also calculated separately for roof and wall elements. $E(x)$ is penalized by fragmented support.

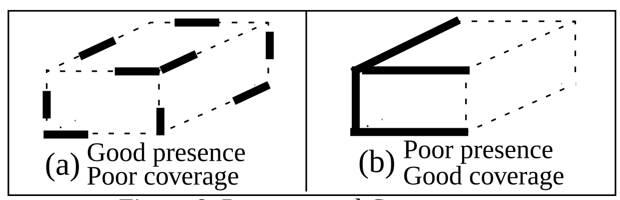


Figure 8. Presence and Coverage.

Junction Presence: $J(x)$ is defined as the ratio of the number of image L-junctions at locations predicted by the model (Figure 9) to the number of visible model vertices. Image junctions are extracted from the image from the line segments used for matching.

Shadow Presence: $S(x)$, is defined as the ratio of the number of shadow boundaries and junctions extracted from the image matched to predicted shadows, over the number of visible predicted shadow elements (boundaries and junctions) derived from the model (Figure 9). See [Lin et. al, 1995] for a description of our method to extract shadow boundaries and junctions from images.

High confidence $C(x)$ values indicate good image support while low values denote low support. Low values may signify change as lack of image support may be due to missing buildings, or buildings that have undergone significant change with respect to their current model. However, model buildings that have strong image support, may have

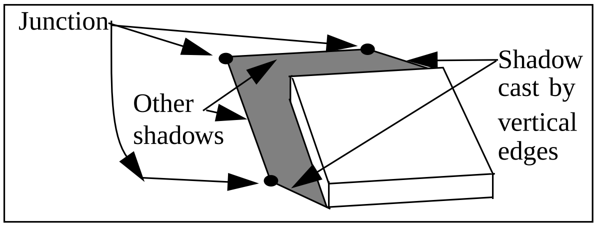


Figure 9. Shadows cast by a rectangular building.

changed also as additions to structures, such as a new wings, may not affect significantly the appearance of the previously modeled portions.

Figure 10 shows the results of the validation step applied to the image shown in Figure 2 earlier. The labels indicate the validation confidence level as a function of *C* values: high (**H**) for *C >= 0.5*; medium (**M**) for *0.4 <= C <0.5* and low (**L**) for *C < 0.4*. Note that a building indicated as having low (**L**) confidence has actually drastically changed (or was grossly mis-modeled) whereas the ones marked high (**H**) are in fact, unchanged. Medium level (**M**) typically denote buildings with moderate or "acceptable" image support. These assignments are arbitrary however, and would have to be set as a function of the task at hand. Some applications may require detailed explanations of possible change that require higher discrimination. Here we show only three for simplicity.

4 Structural Change Detection

The validation step makes available information that is used to start analyses to determine structural changes. Two cues are used to investigate structural changes:

Validation Confidence values: These values reflect image support for a model object. Although low support may be due to poor image quality, lack of contrast, occlusion, or viewpoint, they can signify missing structures, substantially altered structures or incorrect modeling. Medium level support denote "acceptable" indication of presence with reduced support due to poor image quality, lack of contrast and other image dependent characteristics. High values clearly denote strong presence and image support, at least for the modeled portions.

Extra Image Elements: Model elements that correspond to image elements having greater extent than that of the model element provide preliminary indication of possible changes in dimensions. This situations occur regardless of confidence levels assigned.

The above cues are used to further analyze whether one of the following three classes of structural change has occurred. The three classes are: missing (or mis-placed) buildings, dimensional changes, and new buildings. The methods to infer these changes are described next.

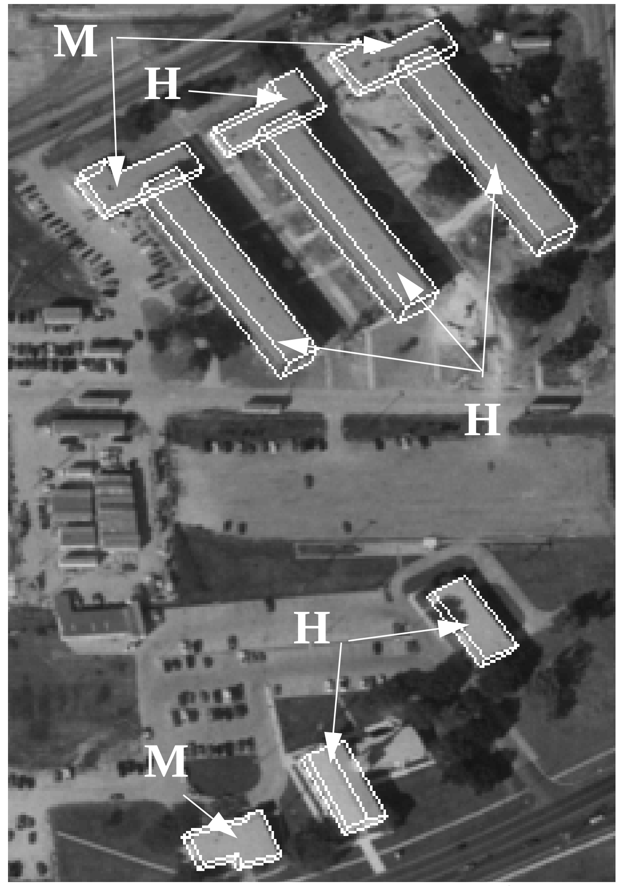


Figure 10. Validation result (partial) and confidence levels.

4.1 Missing Buildings

Model buildings having very low confidence values denote poor image support. The possible causes for this condition are that either the model is incorrect, the structure is heavily occluded or that the building has been removed or destroyed (assuming that images are of sufficient quality), or that its position is grossly incorrect. A low confidence is sufficient to report a missing building, if additional images were available, they could be examined for confirmation. Figure 11 shows the result in a small window containing two modeled buildings. The model building on the right was added to the model by hand to test this condition. It is reported as having low confidence correctly as evidence for its presence can not be found in the image.

4.2 Validated Buildings

Model buildings having moderate (**M**) to high (**H**) levels of support are considered validated. That is, their presence in the image is verified. The buildings labeled (M) typically require verification in another image to increase, if possible, their confidence level. An example is shown in Figure 12. The L-shaped building in the small window corresponds to the L-shapes building, labeled M on the top left

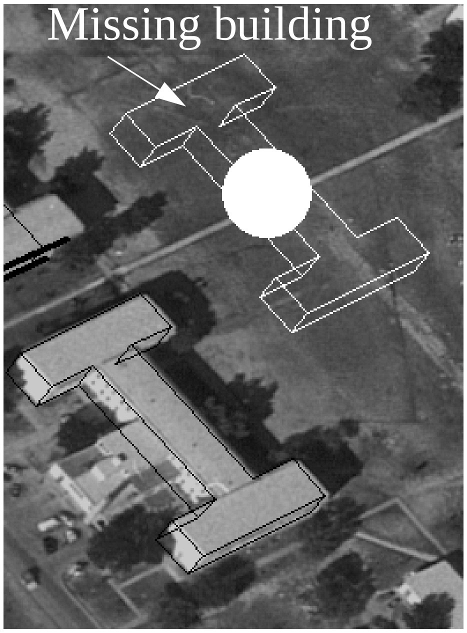


Figure 11. Missing building (white outlines)

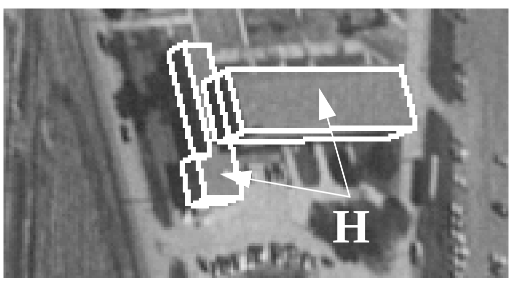


Figure 12. Validation from another viewpoint.

of the image shown in Figure 10. From this viewpoint the confidence value of this building increased to 0.5661, thus becoming validated with high confidence.

4.3 Dimensional Changes to Modeled Structures

In some cases, regardless of confidence, the system is able to cue possible structural changes based on the fact that model edges are matched to *longer* image edges. Those cases where these conditions arise due to accidental alignments have been handled earlier, as explained in section 3.2. The remaining cases therefore represent cues for possible change. We apply a building finding tool at this stage to confirm a change (and to derive a description of the change) as shown later in section 5, however, this relies on the building finder's ability to find new buildings. A less stringent criterion may be to only search for some additional evidence such as the extra edges casting a shadow and/or having other evidence of being above ground (from multiple images). We have not implemented such partial analysis though components of it are available in the building finding tool.

Figures 13 (a) and (b) show two examples of evidence that support cueing dimensional changes. The matching and fine registration step correctly registers the modified models to the structures in the image. The thick white gray lines are the extended image segments that matched the corresponding model edges thus triggering the cue.

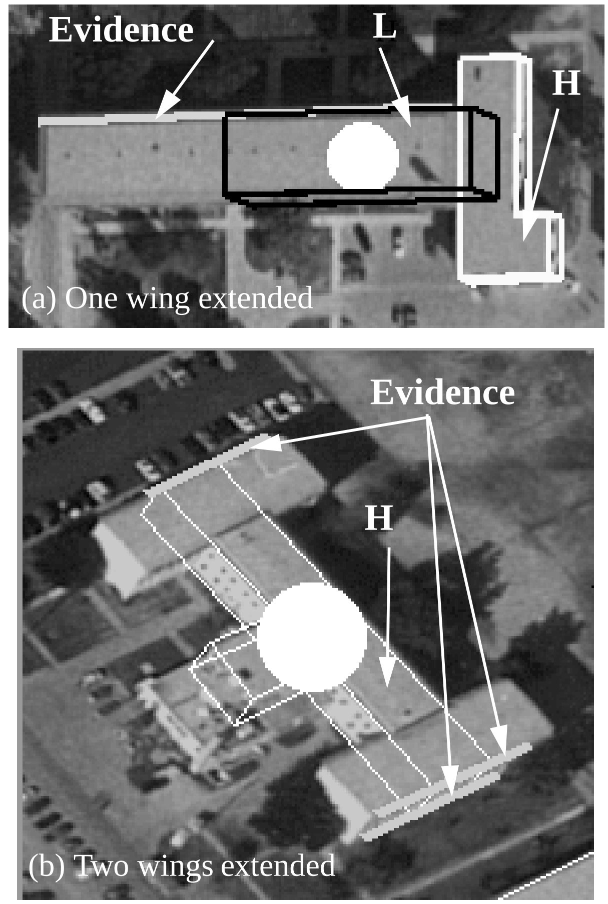


Figure 13. Actual change in dimensions.

4.4 New Buildings

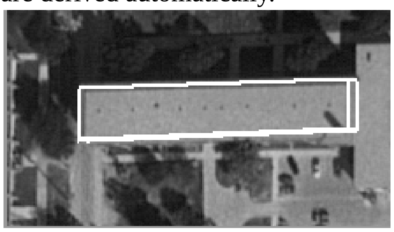
One important type of site change is the introduction of new structures. For such changes, the previous model is less useful but can still be relied on to provide some context. Such context can consist of areas of interest and characteristics of existing buildings (to check if buildings similar to existing ones have been constructed). Our system only uses the site model to mask out the modeled areas and a building finding tool is applied to all other areas, or all other areas containing large number of unmatched image features such as corners. The camera models and terrain models associated with the site are used to derive viewpoint and illumination parameters automatically. We have experimented with focus of attention mechanisms to select the areas where automated detection should be applied.

5 Model Updating

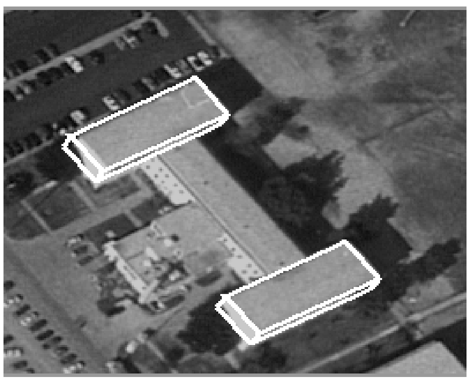
The next task is to make a model for the detected structural changes and to incorporate them in the site model. We describe two situations: where changes are made to existing structures and where new structures are detected.

5.1 Modeled Buildings

Changes in the dimensions of modeled structures that have been cued by the previous step need to be analyzed further, possibly using more than one view of the scene. We use a monocular building detection system [Lin et al, 1995] to return the highest rated building hypothesis that can be formed in the location of the cued change. These are shown in Figure 14 for the two examples shown earlier in Figure 13. The 3-D models are derived automatically.



(a) 3-D model generated by building detector for the extended wing.



(b) 3-D models suggested for changed wings.

Figure 14. Suggested updating of cued changes

5.2 New Buildings

Detection of new buildings is a more difficult task as the site model is less helpful. Our method consists of applying one of our building finding tools (see for example [Lin et al, 1995]) to areas of interest and reporting the results that are of sufficiently high confidence. Typically the system would be instructed to locate new buildings in designated areas that are of interest, such as functional areas. The buildings detected automatically become candidates to be added to the site model.

6 Validation and Change Cueing Results

We have tested our system on several images of a modelboard and Fort Hood sites. In this section we show part of a representative example only, due to lack of space. Fort Hood images are typically 7775x7720 pixels in size. The 3-D site model contains 79 objects representing building structures. Processing time is about 15 seconds per structure on a Sun sparc-10 workstation, running under the RCDE. The results are summarized in table 1. It shows the number of building objects visible in the image and the distribution of validation confidence values (**H**, **M** and **L)**. The label codes are also shown in Figure 15. The confidence values are dependent on the image content and may not necessarily reflect structural changes but generally there is a high correlation between confidence level and the number of buildings changed, not changed, or missing. All matching ambiguities, with one exception (not shown), are correctly handled. This case involves an alignment with a ground feature not present in the model, a situation, not currently handled by the system. The building involved is the only one, of the 54 non-changed buildings, cued incorrectly to have changed. This situation is however likely to be corrected by confirmation of the change using another view of the building. Fourteen buildings that are actually present in the image had changes. Thirteen of these are found to be changed correctly. Representative changes are shown in the Figure 15 with a circle on top and thick white lines cueing evidence of dimensional changes. The remaining changed building is an L-shaped building (not shown). The building has two wings, both of which have changed. Only the left wing is detected to have changed. This situation is likely to be corrected by reconciling the output from more than one view.

Table 1: Summary of Results

Image (fhov927)	Visible Buildings	Validation Confidence			Non-changed Buildings			Changed Buildings			Missing Buildings		
		High (**H**)	Medium (**M**)	Low (**L**)	Number of buildings	Reported non-changed	Reported changed	Number of buildings	Reported changed	Reported non-changed	Number of buildings	Reported missing	Validated
No.	79	54	13	12	54	53	1	14	13	1	11	11	0
%	100	68.3	16.4	15.1	100	98.1	1.9	100	92.8	7.3	100	100	0.0

Buildings that changed considerably or are missing have a low validation confidence (labeled **L** in Figure 15). There are 12 of these, 11 of which were added by hand to test the "missing building" detection capability. The remaining one, represents a significantly changed building (The cross-shaped building in Figure 15). All these are labeled correctly as changed or missing.

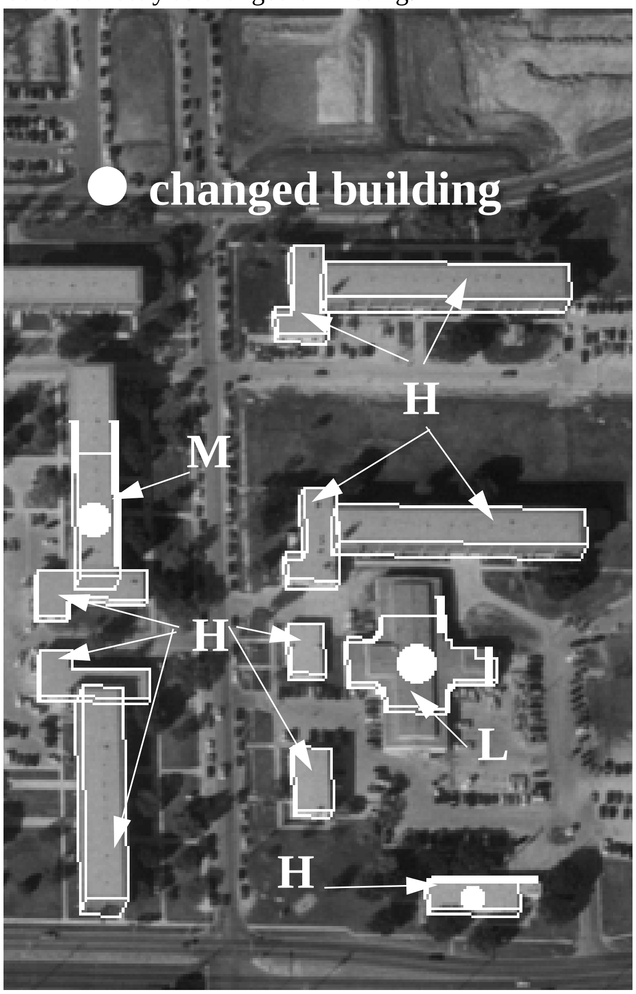


Figure 15. Validation and change detection result for a portion Fort Hood, Texas

7 Conclusion

We have shown some results and capabilities of our system for detecting and describing structural changes. It has been tested on real images (though with simulated changes to the model) and seems to be quite effective at finding significant changes in rather complex images. It is able to find missing (or misplaced) buildings, buildings with changed (or incorrectly modeled) dimensions and new buildings (or previously unmodeled buildings). This system relies on use of a single image to find changes. We anticipate that its performance would be significantly enhanced by use of multiple images as they would provide independent evidence of changes and also allow the reasoning to proceed more directly in 3-D space. Several multiple image building finders are becoming available [Noronha & Nevatia, 1996; Jaynes et al, 1994; Collins et al 1995] and could be easily incorporated in our method. Our system has been ported to an industrial laboratory for possible use in current applications.

References

[Collins et al, 1995] R. Collins, A. Hanson, E. Riseman and H. Schultz, *Automatic Extraction of Buildings from Aerial Images*, in A. Grun, O. Kubler and P. Agouris, editors, Automatic Extraction of Man-Made Objects from Aerial and Space Images, Virkhauser Verlag, Basel, pp. 169-178.

[Fua, 1996] P. Fua, *Cartographic Applications of Model-Based Optimization.* Proceedings of the DARPA Image Understanding Workshop, Palm Springs, CA, Morgan Kaufman, Publisher, Feb., pp 409-419.

[Gerson and Wood, 1994] D. Gerson and S. Wood, *The RADIUS Testbed System*, Proceedings of the DARPA Image Understanding Workshop, Monterey, CA, Morgan Kaufman, Publisher, Nov., pp 231-237.

[Huang et al., 1994] C. Huang, J. Mundy and C. Rothwell, *Model Supported Exploitation: Quick look, Detection and Counting, and Change Detection*, Proceedings of the Second IEEE Workshop on Applications of Computer Vision, Sarasota, FL, pp 144-151.

[Huertas et al., 1995] A.Huertas, M. Bejanin and R. Nevatia, *Model Registration and Validation, In Automatic Extraction of Man-Made Objects from Aerial and Space Images,* A. Gruen, O. Kuebler, and P. Agouris, Editors. Birkhauser Verlag, Switzerland, pp 33-42.

[Huttenlocher & Ullmann, 1990] D. Huttenlocher and S. Ullmann, *Recognizing Solid Objects by Alignment with an Image,* International Journal of Computer Vision, Vol. 5, No.2., Nov., pp. 195-212.

[Jaynes, et al., 1994] C. Jaynes, F. Stolle, and R. Collins, *Task Driven Perceptual Organization for Extraction of Rooftop Polygons*, Proceedings of the 1994 ARPA Image Understanding Workshop, 359-365.

[Lillestrand, 1972] R. Lillestrand, *Techniques for Change Detection*, IEEE Transactions on Computers, No.7, Jul., pp 654-659

[Lin et al., 1995] C. Lin, A. Huertas and R. Nevatia, *Detecting Buildings from Monocular Images*. In Automatic Extraction of Man-Made Objects from Aerial and Space Images, A. Gruen, O. Kuebler and P. Agouris, Editors, Birkhauser Verlag, Switzerland, pp 125-134.

[Medioni et al., 1991] G. Médioni, A. Huertas and M. Wilson, *Automatic Registration of Color Separation Films*, Machine Vision and Applications, Springer-Verlag, New York, Vol. 4, pp 33-51.

[Noronha & Nevatia, 1996] S. Noronha and R. Nevatia, *Detection and Description of Buildings from Multiple Aerial Images*, Proceedings of the 1996 DARPA Image Understanding Workshop, Palm Springs, CA, pp 469-478.

[Strat et al, 1992] T. Strat, et al., *The RADIUS Common Development Environment*, Proceedings of the DARPA Image Understanding Workshop, San Diego, CA, Morgan Kaufman, Publisher, Jan., pp 215-226