LOCALIZATION AND GENERATION OF BUILDING MODELS

Ildiko Suveg, George Vosselman
Photogrammetry and Remote Sensing
Delft University of Technology, The Netherlands
Thijsseweg 11, 2629 JA, Delft
{I.Suveg, G.Vosselman}@geo.tudelft.nl

KEY WORDS: aerial images, 2D GIS map, building models, fitting, mutual information

ABSTRACT

This paper presents a knowledge based approach for 3D reconstruction of buildings from aerial images. The aerial images are combined with information from 2D GIS database and specific knowledge of the buildings.

This paper describes the generation of building hypotheses in case of large variations in the terrain height. First the possible locations of a building in the images are determined by using the ground plans of the building defined in the map and lines extracted from images. These possible locations are verified by the 3D building model generation process. The generated building hypotheses are improved by fitting them to the image data. An evaluation function based on information theory principles is used to select the best model.

Experiments are presented that demonstrate the approach by reconstructing 3D building models.

1 INTRODUCTION

A lot of approaches have been proposed for the 3D reconstruction of buildings from aerial images. However, despite of intensive research, the current state of automation in the 3D reconstruction of buildings from aerial images is quite low.

Most approaches have focused on the reconstruction of specific building models: rectilinear shapes [Noronha and Nevatia, 1997], [Roux and McKeown, 1994], flat roofs [Jaynes et al., 1997], [Lin et al., 1994] or parametric models [Fischer et al., 1998]. But buildings show a much wider variety in their shape. Other approaches employ a generic roof model that assumes planar roof surfaces [Bignone et al., 1996], [Moons et al., 1998], [Schmid and Zisserman, 1997]. These 3D roof planes are generated by grouping the coplanar 3D lines or corners computed by matching of image features extracted from stereo images. Hence, these approaches rely on the extraction of image features, which raises a lot of problems especially for aerial images. The feature extractors can fragment or miss boundary lines, due to low contrast, occlusions, and bad perspective.

A more robust approach should combine different data sources. The image data can be combined with scanned [Maitre et al., 1995] or digital maps [Haala and Anders, 1996]. These approaches represent the newest trend in 3D building reconstruction.

Our strategy for 3D reconstruction of buildings combines pairs of stereo images with large-scale Geographic Information System (GIS) maps and domain knowledge as additional information sources. The 2D GIS database contains the outline of footprints of the buildings. The knowledge about the problem domain is represented by a building library containing primitive building models. Although, buildings reveal a high variability in shape, even complex buildings can be generated by combining simple building models with flat, gable or hip roof.

This paper describes the localization and generation of building hypotheses using the ground plans of the buildings defined in the map.

The paper is organized as follows: Section 2 presents a brief overview of the steps involved in our approach for 3D reconstruction of buildings. The next section describes the approximate localization of buildings into images while section 4 presents the generation of building hypotheses. Section 5 presents experiments which demonstrate the effectiveness of the method. The conclusions and future work are discussed in the final section.

2 METHOD OVERVIEW

The complexity of the reconstruction process can be reduced by a large amount by focusing on one building structure. Therefore, it is desirable to localize the buildings in the images first and afterwards to do the actually reconstruction.

To cope with the complexity of aerial images, specific knowledge about buildings is integrated in the reconstruction process. Since most buildings can be described as an aggregation of simple building types, the knowledge about the problem domain can be represented in a building library containing simple building models (flat roof, gable roof, and hip roof building).

The building reconstruction process is formulated as a multi-level hypothesis generation and verification scheme and it is implemented as a search tree. The tree is generated incrementally by the search method.

The first step of the actual reconstruction process is the partitioning of a building in simple building parts, which might correspond to the building models defined in the building library. First, the partitioning is done using only the ground plan of the building defined in the GIS map. If the ground plan of the building is not a rectangle, then it can be divided in rectangles, called partitions. Then, a partitioning scheme can be defined as a subdivision of a building into disjoint partitions. Usually, a building can have multiple partitioning schemes.

All the possible partitioning schemes of a building are represented on the first level of the search tree. To avoid a blind search method of the tree, partitioning schemes are ranked based on their support in the images. This ranking provides a means of giving higher priority to the partitioning schemes with a more support in the images. The second level of the tree contains the partitions corresponding to each partitioning scheme.

Next, the tree will be expanded with a level corresponding to the different building hypotheses generated for each building partition. Corresponding to each building primitive defined in the building library, a building hypothesis is generated. The estimation of the parameters of the building model is performed using a fitting algorithm, which fits the edges of the projected wire frame of the model to gradients of the pixels from both images simultaneously.

The building hypotheses can be verified by back projecting them into the images and then matching with the information extracted from the images. The matching defines a score function that will be used to guide the search in the tree. So, this function allows comparison and evaluation of different building hypotheses. The score function is based on the formulation of the mutual information between the building model and the images.

The CSG tree representing a building will be given by the best fit of the building models corresponding to the building partitions. In the final verification step the complete CSG tree will be fit to the image data. To improve the results, constraints, which describe geometric relationships between building primitives, are incorporated in the fitting algorithm.

If all the partitioning schemes are rejected by the tree search then the partitioning has to be refined using image information as well. This process will start up a new branch in the search tree and the whole process is repeated.

3 LOCALIZATION OF BUILDINGS INTO THE IMAGES

Most of the building reconstruction strategies have two main parts: the localization step and the actually reconstruction step, more or less connected. The localization of the building in the images means the detection of regions of interest where the buildings lie. By having the building localized in the images first, the reconstruction process can be focused on one building reduces the complexity of the reconstruction by a large amount.

In [Suveg and Vosselman, 2000] information about the ground plan of the building contained in the GIS database was used to delineate a building in an image. The uncertainties introduced by different knowledge sources were identified and a two step method was developed to quantify these uncertainties and determine the region where the building lies in the image. In the building localization process the uncertainties are due to: the unknown height of the buildings, the accuracy of the map, the roof extensions, and the feature extraction.

In the first step of the proposed method the uncertainty due to the unknown height of the buildings is handled, by assuming that height of a building is between two extreme values. By projecting the ground plan of the building into the image for each of these extreme values two contours are obtained. These contours are concatenated in order to get the area where the building is located. In the next step the contour obtained after the concatenation process is dilated taking in order to handle the other uncertainties mentioned above.

This method worked well in case of small height variations of the terrain, i.e. close extreme values. The method could handle variations of 15 m in height. Applying the method for larger height variations would increase the obtained region where the building lies. In case of a very oblique view this region becomes so large that it may include multiple building structures (Figure 1).

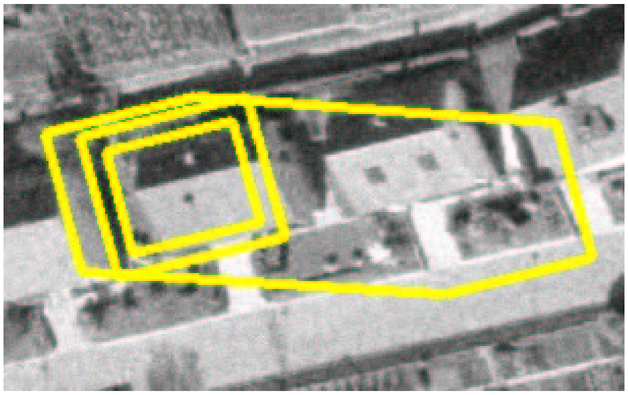


Figure 1: Localization of the building and the pattern used in the localization process

We have extended this method to be able to handle large variations in height. A large interval of height values can be reduced to a few individual height values which have to be checked afterwards by some higher level image understanding methods. Therefore the building localization can no longer be separated from the building reconstruction process anymore.

The main idea is to divide the interval of height variations into small intervals. On these small intervals the former method can be applied to determine the maximal region and the minimal region inside of which the building could lie. For each of these small intervals a function that looks for evidence in the image that the building lies in that region is defined. In order to compute this function lines are extracted from the image and the image lines which lie between the maximal and the minimal region are selected. Afterwards the length of the selected lines are summed up to compute the score for the given interval of height values. The score is integrated over both images. A high value indicates a high likelihood for the correct height of the building. This score is computed for each height interval.

Actually this process can be seen as moving the pattern defined by the maximal and minimal region along a line in the image and counting for evidence for each position. In this way we get the graphic of the likelihood of having the building correctly localized for different height values (Figure 2 corresponding to the building from Figure 1).

Generally, a threshold of 70% of the largest peak can be used the select the peaks which might correspond to the correct building height. With this thresholding a large interval of height values is reduced to some individual values.

Unfortunately, not all peaks in the graph correspond to correct building locations. For instance, an edge corresponding to a road might introduce a false peak in the graph. Therefore, each peak has to be verified. This verification can be done by actually trying to find the building model which best describes the image data.

Since the building model generation process was designed to

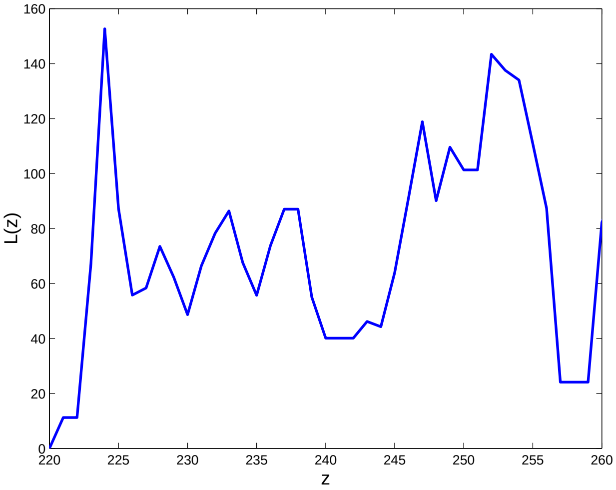


Figure 2: Likelihood of building location

be able to deal with small height variations, there is no need to check every individual height value. If these height values can be grouped in small intervals then it suffices to check a value from each group. Therefore a k-nearest neighbor algorithm is applied for creating clusters of height values. The value of k is determined by according to the variations that can be handled by the reconstruction procedure.

For the building from Figure 1 by thresholding we get 7 height values. These 7 values can be further reduced to 4 clusters of height values by applying k-nearest neighbor classification.

The height corresponding to the highest value of each cluster is used to generate some building hypotheses which are going to be verified by fitting them to image data. Hence the localization of the building in the image will only be completed when the correct building model is found.

4 GENERATION OF BUILDING HYPOTHESES

4.1 Library of Building Primitives

Most buildings can be described as aggregation of simple building types. Starting from this observation, the knowledge about the problem domain can be represented in a building library containing the simple building models. As basic building models, we can consider a flat roof, a gable roof and a hip roof building. The approach of modelling buildings using a set of basic building models (primitives) suggests the usage of CSG representation for building description. In this way, a complex building can be seen as a CSG tree, where the leaf nodes contain primitive building models and the internal nodes contain boolean operations such as union, intersection, difference. The basic building primitives of the CSG representation can be described by parametric models having shape and pose parameters.

The first of the primitives is a flat roof building (Figure 3a). A rectangular volume encodes the geometrical properties of this type of building. To describe a flat roof building primitive 6 parameters are necessary: 3 shape parameters and 3 pose parameters. The shape parameters are: width (w), length (l), height (h). The pose parameters are: x, y coordinates of the buildings' reference point and the orientation in the xy-plane. The height h is actually the sum of the height of the terrain and roof.

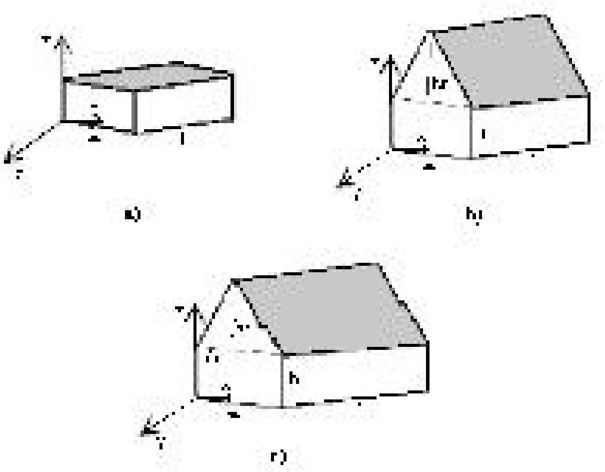


Figure 3: Parametric building models a) Flat roof building b) Symmetrical gable roof building c) Non-symmetrical gable roof building

Another primitive is a symmetrical gable roof building composed from a rectangular volume and a triangular volume (Figure 3b). Therefore, for a gable roof primitive an extra parameter, the height of the ridge has to be added to the parameters of a flat roof primitive.

A more general gable roof primitive, the non-symmetrical gable roof has an additional parameter the distance from the roof reference point to the ridge base point (Figure 3c).

4.2 Fitting of Building Models

To find the best fit of an instance of a building model to an image, we must find the shape parameters and the position parameters, which best match the model to the image. This can be done using a fitting algorithm.

The approach for fitting 3D building models to an image is based on projecting the model into the image and finding the parameters of the model that maximizes some measure of the goodness-of-fit between model projection and image. In most cases it is possible to solve for all unknown parameters of a building model from matches to a single image. However, the accuracy of the parameter estimation can be substantially improved by simultaneously fitting the model to images taken from different viewpoint. The method presented here can be used in either situation. This method is a modification of Lowe's fitting algorithm [Lowe, 1991]. It was developed by Vosselman and used for semi-automatic reconstruction of 3D buildings [Vosselman and Veldhuis, 1999] and industrial piping installation [Ermes, 2000].

The application of this fitting algorithm requires approximate values for the model parameters. These approximate values can be obtained from the map and from images.

In the initial approximation the x, y coordinates and the orientation of a building primitive are given by the ground plan of the building. The width and length parameters are the width and the length of the rectangle corresponding to the ground plan of the building partition. The height is given by the localization procedure. In case there are more possible heights, then a fitting is done for each of them.

For a gable roof primitive, the height of the ridge is considered as the height of the reconstructed 3D top line if the top lines were detected in both images and the 3D line could be reconstructed. Otherwise, in case of of a symmetric gable roof, the approximate positions of the projected ridge in the images can be deducted by taking into account the symmetry of the gable roof. Then the 3D ridge can be reconstructed by matching these two approximate line segments. In case of a non-symmetric gable roof if no image lines are found, default values are adopted for the roof parameters.

To improve the parameters of a building model, the hypothesized model is fit to image data. The edges of the projected wire frame of the model are fit to gradients of the pixels from both images simultaneously. The fitting algorithm is described as an iterative least-squares algorithm. It estimates the changes of the parameter values that have to be applied in order to minimize the square sum of the perpendicular distances of the image pixels to the nearest wire frame edge ([Suveg and Vosselman, 2002]).

The resulting 3D building models are evaluated by computing a score function based on the formulation of the mutual information between the building model and the images. This score combines two measures. One of them counts for evidence along the projected building model contour in the images. If the hypothesis is correct, we expect to find changes in image gradients along the projected model contours. The second one measures image intensity similarity over two images. Given an object hypothesis and a pose, a point to point mapping can be defined between images. If the hypothesis is correct then the intensities at corresponding pixels will be highly correlated. The building model with the highest score is taken as the correct hypothesis ([Suveg and Vosselman, 2001]).

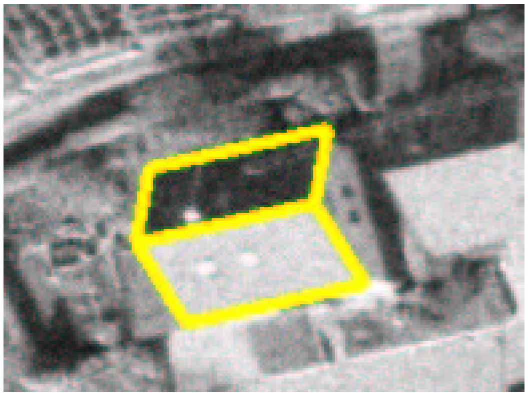
5 RESULTS

The test data consists of high-resolution aerial images. The scale of the images is 1:1300 and they are scanned at 15 mikrons. Two images with 60% overlap are used. The interior orientation parameters of the camera and also the exterior orientation parameters of the images are known. A 2D GIS database containing the ground planes of the buildings is given.

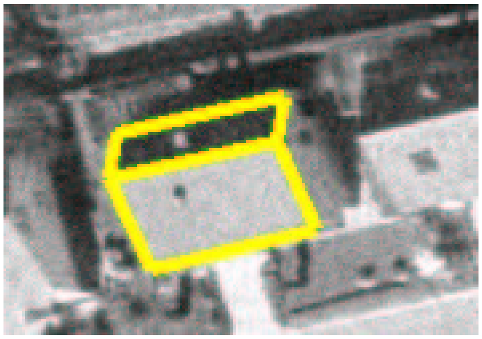
In our current implementation five hypotheses are generated corresponding to a flat roof building primitive and two symmetric and two non-symmetric gable roof primitives with different orientations. Therefore we can reconstruct only flat roof buildings, gable roof buildings or buildings formed by combining these two building types. However the building library can be easily extended with other primitive building models.

The first experiment was to generate and evaluate building hypotheses for simple buildings composed by only one building primitive. First the possible locations of the building in the images are determined. Afterwards, each of these possible locations are verified by generating building hypotheses. The building hypotheses derived from outlines of building footprints from the map are generated corresponding to the building models from the building library. Next, the building hypotheses are fit to the image data. The scores computed for matching the hypotheses against the images are used to choose the best model.

The resultant building models projected back into one of the images are presented in figure 4. The tables show the scores computed for different height values and for 4 building primitives (flat roof, two symmetrical gable roofs with different orientations and a non-symmetrical gable roof with the orientation along the larger edge of the building base). If a gable roof model oriented along the larger edge of the building is found then the search stops and the other gable roof is not generated at all. This is the case of the first building from 4, when for the last two height values only a gable roof model is generated.



height	flat	gable1	gable2	nonsymg2
223	-249.1	-376.6	-300.8	-193.4
231	-236.7	-279.5	-308.8	-154.0
235	-288.1	-299.7	-99.6	-173.8
250	-186.8	-	**138.7**	-
254	-214.8	-	37.4	-



height	flat	gable1	gable2	nonsymg2
220	-242.6	-324.3	-369.0	-254.1
224	-230.2	-226.9	-228.1	-154.3
250	-168.2	-230.6	-265.2	**58.7**
255	-162.9	-267.9	-295.4	-65.3

Figure 4: Reconstructed 3D models of simple buildings

Next, we tested our approach on complex buildings. First, the partitioning of the building into building primitives based on the ground plan is performed. Then, for each resultant building part, hypotheses are generated and the best fit hypothesis is selected as building model. The final CSG tree describing the building is further refined by simultaneous fitting of the building models contained in the CSG tree. Constraints, which describe geometric relationships between building models, are incorporated in the global fitting algorithm. The us-

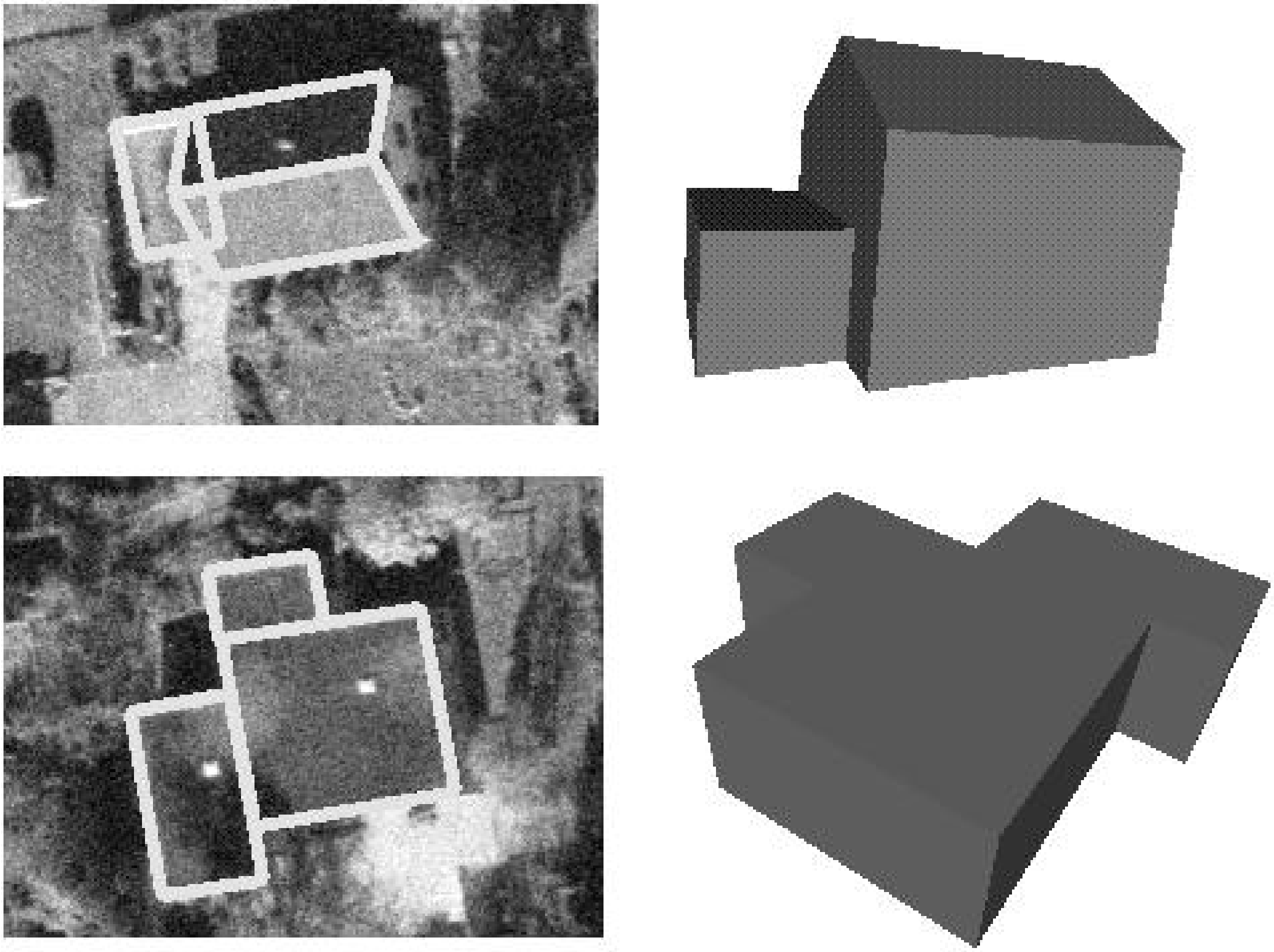


Figure 5: Reconstructed 3D models of complex buildings

age of constraints reduces the degree of freedom of some parameters, therefore precision of the parameter estimation is increased.

In our building reconstruction system the following types of constraints are used:

- Parameter constraints: establishes a relation between two parameters of two building models. For example, two building models have the same orientation.
- Connection constraints. One edge of a building model lies on one of edges of the other building model
- Corner constraints. Two building models share a common corner
- Extension constraints. Two building models share a common edge

The constraints are implemented in the least-square adjustment as weighted observations with standard deviations. The weight specifies the strength of the constraint in the adjustment.

Evaluating the partition schemes we found that the partitions presented in Figure 5 are the best ones. The 3D building models were obtained by adding artificial vertical walls to the reconstructed roofs.

The results from the proposed approach are encouraging. The method worked well even in difficult conditions where feature based approaches would have failed.

6 CONCLUSIONS AND FUTURE WORK

A knowledge-based approach for automatic 3D reconstruction of buildings from aerial images was presented. The generation of building hypotheses in case of large variations in the terrain height was described. In this case the building localization process can no longer be separated from the building reconstruction process. Some possible locations of a building in the images were determined by combining information from the map with image data. These locations were verified by the building models generation process. The robustness of this method was shown by the presented experiments. The correct building model was found for different height values. Future work will be directed towards refining the partitioning by image information in case all the partitioning schemes are rejected by the tree search.

REFERENCES

[Bignone et al., 1996] Bignone, F., O.Henricsson, P.Fua, and M.Stricker (1996). Automatic extraction of generic house roofs from high resolution aerial imagery. In *Com-*

puter Vision - ECCV'96, volume 1, pages 85–96.

[Ermes, 2000] Ermes, P. (2000). Constrains in cad models for reverse engineering using photogrammetry. In *Proceedings of the XIXth congress of ISPRS, Amsterdam*, pages 215–221.

[Fischer et al., 1998] Fischer, A., Kolbe, T. H., Lang, F., Cremers, A. B., Foerstner, W., Plumer, L., and Steinhage, V. (1998). Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. *Computer Vision and Image Understanding*, 72(2).

[Haala and Anders, 1996] Haala, N. and Anders, K. (1996). Fusion of 2D GIS and image data for 3D building reconstruction. In *International Archives of Photogrammetry and Remote Sensing, ISPRS, Vienna, Austria*, volume 31, pages 289–290.

[Jaynes et al., 1997] Jaynes, C., Marengoni, M., Hanson, A., Riseman, E., and Schultz, H. (1997). Knowledge-directed reconstruction from multiple aerial images. In *Proceedings of ARPA Image Understanding Workshop, New Orleans*, volume 2, pages 971–976.

[Lin et al., 1994] Lin, C., Huertas, A., and Nevatia, R. (1994). Detection of buildings using perceptual grouping and shadows. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle*, pages 62–69.

[Lowe, 1991] Lowe, D. (1991). Fitting parameterized 3D models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):441–450.

[Maitre et al., 1995] Maitre, H., Bloch, I., Moissinac, H., and Gouinaud, C. (1995). Cooperative use of aerial images and maps for the interpretation of urban scenes. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona*.

[Moons et al., 1998] Moons, T., Frere, D., Vandekerckhove, J., and Gool, L. (1998). Automatic modeling and 3d reconstruction of urban house roofs from high resolution aerial imagery. In *Computer Vision - ECCV'98*.

[Noronha and Nevatia, 1997] Noronha, S. and Nevatia, R. (1997). Detection and description of buildings from multiple aerial images. In *Proceedings of DARPA IUW 97*.

[Roux and McKeown, 1994] Roux, M. and McKeown, D. M. (1994). Feature matching for building extraction from multiple views. In *Proceedings of ARPA IUW, Menterey*, pages 331–339.

[Schmid and Zisserman, 1997] Schmid, C. and Zisserman, A. (1997). Automatic line matching across views. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*.

[Suveg and Vosselman, 2000] Suveg, I. and Vosselman, G. (2000). 3d reconstruction of building models. In *Proceedings of the XIXth congress of ISPRS, Amsterdam*, pages 538–545.

[Suveg and Vosselman, 2001] Suveg, I. and Vosselman, G. (2001). 3d building reconstruction by map based generation and evaluation of hypotheses. In *12th Proceedings of the British Machine Vision Conference, Manchester*, pages 643–652.

[Suveg and Vosselman, 2002] Suveg, I. and Vosselman, G. (2002). Automatic 3d building reconstruction. In *Electronic Imaging, SPIE Photonics West, San Jose*.

[Vosselman and Veldhuis, 1999] Vosselman, G. and Veldhuis, H. (1999). Mapping by dragging and fitting of wireframe models. *Photogrammetric Engineering and Remote Sensing*, 65(7):769–776.