A KNOWLEDGE BASED APPROACH TO RECONSTRUCT BUILDINGS IN DIGITAL AERIAL IMAGERY

Eberhard Gülch

Department of Photogrammetry, Royal Institute of Technology
100 44 Stockholm, Sweden
e-mail: ebs@fmi.kth.se

ISPRS Commission II

ABSTRACT

An approach for the reconstruction of buildings in multiple digital aerial images using high-level image analysis techniques is presented. Factual, judgemental and strategical knowledge is collected in a knowledge base. The input data or facts are the orientation parameters of the images and segmented features, like closed regions and lines with their properties. The reconstruction process is performed in steps according to the reasoning strategy, which selects the necessary judgemental knowledge. All segmented features in each image are given preliminary 3-D labels by applying region and line parsing. A consistency check is performed on the multiple parsing results, using a set of generic rules for buildings. The final description of a building is given in 3-D coordinates of the corners and labels for the building parts. The potential of the suggested parsing procedure is shown on experimental results assuming a perfect segmentation. It is shown that the classical feature extraction methods used are not capable to provide a useful input for the parsing process in real scenes. The segmentation process is regarded as a main bottle-neck for a successful high level image analysis. Instead of readjusting a potentially good and successful parsing procedure, several suggestions for extensions of non-optimal classical feature extraction methods are given. These are based on the requirements of the parser and the type of objects to be reconstructed.

KEY WORDS: Knowledge Base, Image Analysis, Feature Extraction, 3-D

1 INTRODUCTION

The automation of reconstructing objects from multiple images heavily rests on a rational procedure for the interpretation of extracted image features and a model of the object to be reconstructed. This interpretation process is called parsing. Several requirements must be fulfilled in order to be able to perform a parse on images. The image features derived by a segmentation process must be suitable for interpretation. Single edge elements allow no interpretation. Closed regions, T- or L-junctions have a high information content. A model of the type of object to be reconstructed must exist, it must be known what should be looked for. The model can be specific or generic depending on the variety of alterations of an object type. The parsing procedure itself must be able to handle errors and contradictions in the interpretation due to incomplete or erroneous segmentation results.

Several methods to parse man made objects have been developed. Walz (1975) used a parser for line drawings ascribing logical properties to line segments. This method strongly rests on the correct identification of object boundary lines. It is therefore very sensitive to errors due to missing lines and is probably not an appropriate method for parsing segmentations of grey level images. Dickinson (1990) introduces the idea of aspects, making the parsing procedure work at a higher description level. Also this method is sensitive to missing lines.

A method which combines parts of the earlier work done in this field is presented. The suggested method is designed for parsing multiple aerial images in order to reconstruct buildings. It assumes straight region boundary lines, but does not require an initial identification of object boundary lines. It works globally on sets of lines and not locally on single line junctions only. The approach integrates line with region parsing, by introducing aspect graphs as an alternative method for specific situations.

For the parsing process general rules, rules of thumb, heuristic experience, facts, strategic decisions etc. have to be collected. It is near at hand to use some type of knowledge base to collect all a priori information available. In this way new knowledge, improved rules etc. can be integrated in a predefined straightforward way.

The knowledge involved can be classified into three main types (cf. table 1), a division which is supported by many knowledge base system shells. The facts are the type of object to be reconstructed and one approximate global position, the given orientation parameters of the images and the results of various image segmentation procedures yielding closed regions with boundaries approximated by straight lines. The rules for the line and region parsers and a consistency check can be classified as judgemental knowledge. To this group also belong the generic rules for the object type 'building'. The collected strategic knowledge steers the whole reasoning process. Decisions on alternations between parsers and the consistency checks for the specific object type are provided.

The expected result of the reconstruction process are the parameters for shape, location and orientation of the building in object space as well as a labelling of building parts according to predefined categories given by the generic model for buildings.

Table 1 Types of knowledge used for the object reconstruction

Factual Knowledge
Judgemental Knowledge
Strategic Knowledge

2 THE RECONSTRUCTION PROCESS

The reconstruction of a building from multiple digital aerial images will be performed using a two step procedure. In a first step segmentation procedures are applied on the raw image data producing a set of closed polygons describing discontinuities in each image. In a second step, the segmentation output is interpreted in object space using a series of moderately successful parsers based on a generic model for buildings. 3-D labels are assigned to the segmented object descriptors. Corresponding object descriptors in different images are compared with each other. The procedure is terminated when a consistency check indicates an acceptable object description.

The knowledge based system shell NexpertObject provides the necessary tools to implement all kind of knowledge used in the parsing process and in the consistency check (Gülch, 1991). So far only parts of the whole process are implemented. The empirical tests (cf. section 3) were performed partly manually.

Table 2 General strategy for object reconstruction

Step	**Results**
1) Segmentation	2-D polygons
2a) Parsing	2-D polygons with probable 3-D labels
2b) Consistency check	3-D description using multi image information

2.1 Segmentation

A set of segmentation algorithms is available to derive points, lines, polygons, and regions with their boundary lines and neighbourhood relations in the image (Gülch, 1990). So far the only implemented segmentation method to directly derive closed boundaries is the region growing technique. The boundaries, given as arcs in chain coding, are approximated by polygons. The resulting data set to be parsed consists of a set of interconnected polygons made up of straight line segments together with a window coarsely defining what part of the segmentation is to be considered. Both a line and a region representation of the segmentation is chosen to prepare for both parsers.

2.1.1 Line segments. The line segments given by a segmentation can in principal have both ends connected to other line segments or have one or both end points free. All connections are at end points, i.e. a T-junction is described using three line segments. All line segments are prepared for having the entries listed in table 3. The altitude level n is determined assigning the first point, assumed to be a ground point, the value n=1. Every time a new level is needed, n is raised one unit.

2.1.2 Region segments. An alternate representation of the segmentation result is given in the form of region segments. These are described by the entries given in table 4.

The complexity (cx) is the number of quantities necessary for specifying the region plus one unit for indexing the special kind of region. The interest index (ii) is low for highly interesting segments, e.g. rectangular structures indicating an artificial object. The group index (ki) is the sum of interest indices of the immediate neighbours plus the index of the region segment itself. The group index will be low when several neighbouring segments have low interest indices, a situation which will occur for man made objects such as buildings.

Table 3 Line segment feature list

n	line segment identifier
x_1,y_1 x_2,y_2	image coordinates of 1st resp. 2nd end point
p_i, q_i	line segments connecting to 1st resp. 2nd end point
S-label	line segment from segmentation
N-label	line segment pointing to nadir point
C-label	line segment of constraint
Hn-label	horizontal line seg. at altitude level n
V-label	vertical line segment
W-label	sloping line segment of roof
G-label	ground line segment
T-label	line seg. with unconnected end point

Table 4 Region segment feature list

n	region segment identifier
p_i	Identifiers of line segments generating region segment
q_i	Identifiers of neighbouring segments
cx	complexity of region segment
ii	interest index
ki	group index
P	label indicating that the segment when mapped to object space is planar

With the region representation it is possible to classify t-junctions. When a region segment contains two line segments having a common angle of 180 degrees they are assumed to be colinear, i.e. they make up one line segment, labelled (C). A third line segment of the junction lying outside the region segment is not connected to the concatenated line segment but introduced with a label (T) indicating an unconnected line.

Supersegments (label C) can be introduced as the envelopes of several segments, generating simple geometrical structures, like e.g. rectangles.

2.1.3 Window of attention. The input to the parsers consists of the segmentation and a window, identifying what part of the segmentation should be interpreted. The window is provided by the user or given by a simple recognition procedure. A scan of the region segments gives locations where the group index is minimum, indicating a probable location of a highly interesting object.

2.2 Projecting image features into object space

The parsers assume fully oriented images. As several different images are used to reconstruct the object in object space, a subobject which is present in several images will be localized approximately in the same place. Solving the matching problem amounts to introducing a suitable neighbourhood relation.

2.2.1 Geometric relations between image and object space. In a fully oriented image the camera constant (c) as well as the camera location (X_0,Y_0,Z_0) and orientation (φ,ω,κ) in object space are known. Points (x, y) in the images can be localized in object space as soon as their altitude (Z) is known:

$$X = X_0 + (Z-Z_0)\frac{x\,a_{11} + y\,a_{12} - c\,a_{13}}{x\,a_{31} + y\,a_{32} - c\,a_{33}}$$
$$Y = Y_0 + (Z-Z_0)\frac{x\,a_{21} + y\,a_{22} - c\,a_{23}}{x\,a_{31} + y\,r_{32} - c\,r_{33}} \qquad (1)$$

with (a_{ij}) as the coefficients of the rotation matrix defined by the rotation angles.

When line segments in the segmentation are said to be parallel, they are assumed to have a common point of convergency. This point is known, as all orientation parameters are known and a projective transformation is assumed.

2.2.2 Constructing a 3-D line skeleton. A first point of the building will be chosen in the image and assigned an arbitrary height. Using various criteria for line segments being horizontal or vertical, the building will successively be reconstructed as a 3-D line skeleton (cf. Stokes, 1992). Height differences between end points of vertical line segments are estimated using the relation

$$h = [\,(H/c)\,/\,(A/c)\,]\,a \qquad (2)$$

with the flying height (H), the distance (A) of the closest point on the line segment to the image nadir point and the length (a) of the line segment in the image and the camera constant (c).

A line segment with a start point connected to a point already present in object space is introduced in 3-D by determining the coordinates of the other end point, using relations (1) and (2). If this end point is located within a threshold distance to a point introduced earlier, this point is readjusted, else it is introduced as a new point. No further action due to the discontinuity is taken at this point. In the consistency check following the parsing it will be investigated if the discontinuity can be removed.

When the parsing is initiated using the region segment parser, the region segments are matched to an aspect prototype and the respective line segments are introduced into object space using the known labels of the line segments in the prototype and a priori given arbitrary heights.

2.3 Parsing the segmentation output

The segmentation output has to be interpreted in terms of objects in 3-D. For the parsing step a generic model for possible objects is provided. This model

Table 5 Line parser - rules

Rule	Performance
1	Generation of bounded segmentation.
2	Grouping. Parallel line segments of a direction φ_i are put into groups 1...I-1 and an Ith group containing all line segments not parallel to any others.
3	Vertical lines. The line segments pointing at the nadir point are collected in group 1.
4	Search strategy. Lines in groups 2 to I-1 are sorted according to a predefined criterion which defines the overall search strategy.
5	Initial vertical line for parsing. Groups 2 to I-1 are searched for two line segments belonging to the same group, both of which are connected to a line segment in group 1. The set of three line segments found are introduced into object space. If no such a triple is found, the region segment parser is tried (section 2.3.2).
6	Connected horizontal and vertical line segments. All groups 1 to I are checked for line segments connected to those already introduced in object space.
7	Repeating rule 6), until no more lines are introduced.
8	Sloping roof line segments. Different combinations of three line segments connected to a pair of parallel, classified line segments labelled Hn_i are checked for being sloping roof line segments and eventually introduced into object space.
9	Repeating rules 6) - 8), until no more lines are introduced.
10	Colinear lines. The altitude of the end points of a line segment which is colinear to an already classified horizontal line segment is known. The line segment is introduced into object space under the assumption that both are colinear also in object space.
11	Repeating rules 6) - 10), until no more lines are introduced.
12	Miscellaneous. a) checking for windows and doors, b) introducing double connected line segments
13	Repeating rule 12), until no more lines are introduced.
14	Repeating rules 6) - 13), until no more lines are introduced.
15	Consistency check. The reconstructed building in object space is checked for consistency. If it is consistent, the parsing is terminated, if not, the process is restarted from rule 4).

describes the building in terms of roofs, walls, doors, windows and garden regions, represented as simple geometrical regions and relations between these. The parsing can so far be performed either using a line segment or a region segment parser. Sets of straight line segments are suitable for describing closed regions in a building, some of the lines being vertical and horizontal in object space. There are two advantages using line segments. First, the probability of a region segment being erroneous is high as one missing or misinterpreted line segment makes the whole region segment erroneous. Second, the fact that all vertical line segments point towards the nadir point in the image gives a simple criterion for the interpretation of these line segments. There are, however, also disadvantages. Vertical line segments of objects close to the nadir point are short or missing in the segmentation output. In these cases the parsing will start with the region parser.

2.3.1 The parser for line segments. The line parsing procedure starts by searching for a vertical line connected to several horizontal line segments and introducing these into object space at an arbitrary datum. Vertical and horizontal line segments connected to already introduced line segments are then successively introduced. Logically weaker rules are introduced when no more line segments can be introduced. As soon as a line segment has been introduced into object space, it can be used as a connection for new line segments. The various steps are described in detail in (Stokes, 1992). Here a summarizing sketch is presented (cf. table 5).

2.3.2 The parser for region segments. In the case that a building lies on or close to the nadir point the vertical (wall-) lines are generally invisible or very short in the segmentation output. In this case the parsing procedure is based on a small set of aspects, describing composite parts of the roof (cf. figure 1) rather than describing the complete roof in a single aspect. Also wall parts are included in these aspects.

The segments contained in the window of attention are tested for matching against this set of aspects, starting with the most complex aspects and continuing with matching on simpler ones. Only the structure of the aspects is important, i.e. only parallelity or orthogonality in the segments needs to be tested; lengths of line segments are not important.

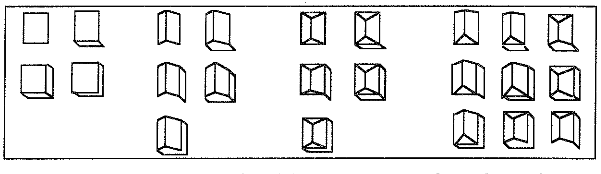

Figure 1 Aspects for building parts to be identified in the segmentation result. The aspects are divided in four groups with increasing complexity from left to right.

When a matching aspect has been encountered, the corresponding line segments in the segmentation are introduced into object space. Ground lines are assumed to be horizontal, with an altitude zero and introduced using relation (1). Vertical line segments are tested for passing through the nadir point and introduced into object space using relation (2). Horizontal line segments connected to points already introduced into object space are introduced at the appropriate altitude. Finally, obvious missing lines are introduced into object space and labelled C. If the prototype contains no vertical line at all, the horizontal lines are interpreted as roof lines and introduced into object space at a preassigned altitude.

The rest of the parsing is continued at rule 6 of the line segment parser (cf. table 5).

2.3.3 Alternation between parsers. The general procedure begins with trying a line segment parser (cf. table 5). If the parse produced is unsuccessful, other versions of the same parser are used trying the various strategies. If no parse passed the consistency test, the region segment parser is used. If also this attempt is a failure, the user must be asked for help. Additional parsers can be easily introduced to this general strategy.

A typical result of an initial parsing of the segmentation of a single image is a partial building in object space and a set of line segments in the segmentation, where the parsing has failed, or which were not accessed by the parser. These line segments might belong to the building or to the ground. Such line segments which are connected to a ground point of the reconstructed building can be introduced in 3D, assuming to be ground line segments in a horizontal plane.

In order to complete the parse, the procedure is started again, disregarding the line segments already included in object space and selecting a new vertical line segment connected to horizontal line segments. The sequence of line segment parsers and region segment parser is carried through until no more line segments are parsed. In this way weakly connected parts of the same building, and also different buildings contained in the window to be parsed, are introduced into object space at approximately correct locations.

If there still are line segments left, they are unconnected to anything parsed and must be labelled by the user.

2.4 Consistency check in object space

Complex objects like buildings are difficult to model using specific models. Here, a generic model is specified, defining buildings as objects present in object space and consisting of subobjects (e.g. roofs, walls, etc.) with certain properties and relations. This model is the basis for the consistency check.

Both parsers are tried until a consistency check indicates an acceptable object description. The consistency check is done in two main parts, first each parse is checked for itself. The consistency check includes procedures for removals of inconsistencies which might occur due to erroneous parses of individual line segments and also missing lines which sometimes must be introduced in order that the 2-D regions in object space be plane. Second the inconsistencies between parsing results from different images are resolved. Substructures, in the form of 2-D regions in object space, which have passed the consistency test, are finally ascribed 3-D labels (wall, roof, etc).

2.4.1 Removal of discontinuities. When a line segment as been given an erroneous parse, it has been introduced into object space in the wrong direction. Consequently, the end point of the line segment will not connect to its neighbour in object space, although it does in the segmentation. Such discontinuities are removed by relabelling those line segments, which were found to be erroneous and reperforming the complete parsing with the important modification that all available labels are used without renewed checking.

2.4.2 Introduction of missing lines. All parsed region segments are checked for planarity in object space. Segments satisfying this requirement are labelled P. The rest are assumed to be composed of several segments divided by line segments missing in the segmentation. Subsegments are produced and tested for planarity by introducing line segments. When a planar subsegment has been found, the line segment is introduced into both object space and the region segment list with a label P. The remaining subsegment is again tested for planarity, continuing the procedure until no segment is left to be tested.

2.4.3 Interpretation. All polygons being closed and planar, are labelled as walls, roofs, floors, windows, doors, chimneys or ground according to the properties of these subobjects as specified in the generic model. This amounts to checking surface orientation and neighbour relations, using multi image information. Vertical region segments are checked for generating horizontally connected structures. All region segments participating in such structures are labelled walls. Sloping regions above walls are labelled roofs. Horizontal regions above walls are labelled floors. Unclassified vertical regions located in the interior of walls or roofs are interpreted as doors or windows, depending on them sharing the bottom line segment with the parent region or not.

2.4.4 The final decision. After running the consistency check on a given parse, the final decision of accepting the parse has to be made. Local failures can be accepted as well as local errors. Both these defects only make the parse incomplete. The kind of parsing errors which should lead to rejection are global. A line segment pointing at the nadir point can be interpreted as if it were vertical and chosen by the line segment parser as the place to start. If it actually is horizontal, the complete parse of the building will be erroneous and should be rejected. Suitable criteria for this and other fatal mistakes can not be completely foreseen in advance, but are expected to be derived from extensive empirical testing.

3 EMPIRICAL TESTS

The results of two empirical tests are presented. In a first test the input for the parsers consists of an optimal manual segmentation without any background lines in four different views of a complex building. In this way the performance of both parsers can be demonstrated undisturbed by segmentation problems. In a second test the basis is a real segmentation using the region growing technique. The images chosen are taken from four scanned aerial photographs showing a quite simple building as a part of a more complex building structure.

3.1 Empirical test on an optimal segmentation

In a first empirical test a set of four views (ART1-ART4) of a synthetic building with highly complex structure is used. An optimal manual segmentation is performed resulting in perfectly closed boundaries. To these segmentation results the reconstruction procedure described above is applied. The reasoning process is performed manually. In figures 2-1 through 2-4 the manually segmented lines and the result of the parsing steps in all four images are given. In one view (cf. figure 2.3) a combination of region and line parser was used. Lines are labelled as horizontal, vertical or as sloping roof line and estimates of the 3-D height values of horizontal lines are derived. Occlusion areas are marked specially. In one view (figure 2-2) several lines remain unlabelled. The height of the building is introduced formally in each image parse. As the approximate absolute height is known in each image, relational matching procedures are suitable to adjust the different parses to each other, resp. to locate the building in object space. In this test, this matching was performed manually. Figure 2-5 shows the inconsistencies in object line labelling after merging the results of all four views. 12 line segments show contradictions, i.e. different labels or different heights.

In figure 2-6 the result after the consistency checks is shown. Inconsistencies in single parses are removed and the interpretation from all four views is given. In this case all visible line segments and building parts are labelled correctly. The rest of the building was not visible in either of the views.

3.2 Empirical test on a non-optimal segmentation

The results obtained with the synthetic material were promising enough to proceed with the implementation of the prototype reconstruction process in order to test it on a larger set of images with relaxed a priori constraints. This means at the end the application to a real segmentation with missing and wrong features.

Four image patches (cf. figure 3-1) were chosen for the second test. They have a patch size of 1024x1024 pixels and a pixelsize of 42 µm and they are taken from digitized aerial images (camera constant ≈ 153 mm, flying height ≈ 650 m). All four views K1 - K4 contain a building complex were the lower right building is the desired one (cf. figure 3-2).

So far the only method implemented to derive closed boundaries and thus connected boundary lines is the region growing method. The region growing yields nodes, where several regions meet, and between them arcs which describe the common boundary segment between two regions (cf. figure 3-3).

Two different segmentation parameter sets for the region growing were used which were chosen ad hoc. The used region growing version assumes constant grey value in a homogeneous region with a maximal difference of 10 grey levels for the first parameter set (1) leading to the segmentations K1_1 - K4_1 and a difference of 20 grey levels for the second parameter set (100) resulting in the segmentations K1_100 - K4_100. The result for the view K1 with both parameter sets is shown in figures 3-4 and 3-5.

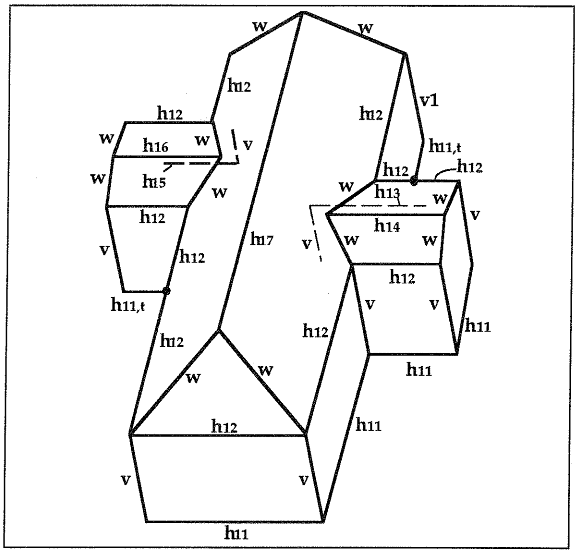

Figure 2-1 ART1*) - Result of a line parsing.

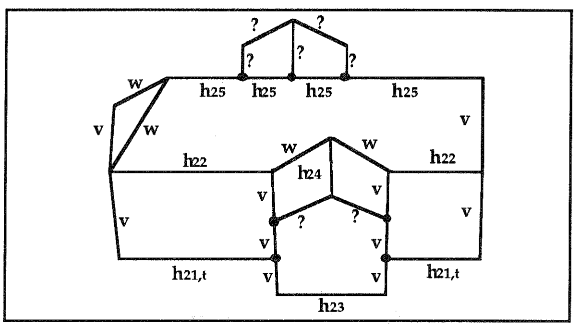

Figure 2-2 ART2*) - Result of a line parsing.

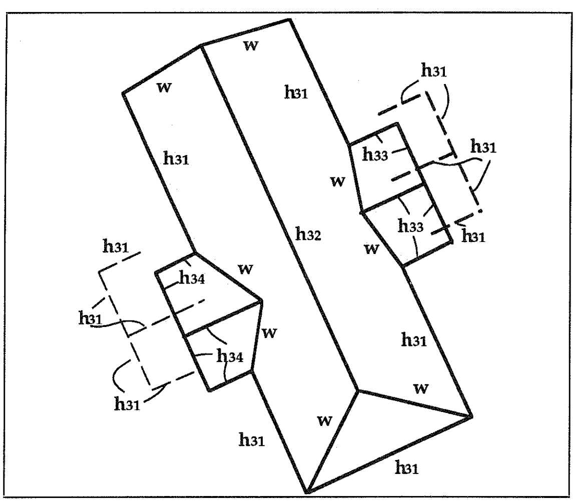

Figure 2-3 ART3*)- Result of a region parsing followed by a line parsing.

* Labelling: (contradictory, multiple labelling is possible)

hij	=	horizontal, view i, height j	v	=	vertical
w	=	sloping roof line	?	=	undetermined
t	=	open ended line	•	=	Occlusion area

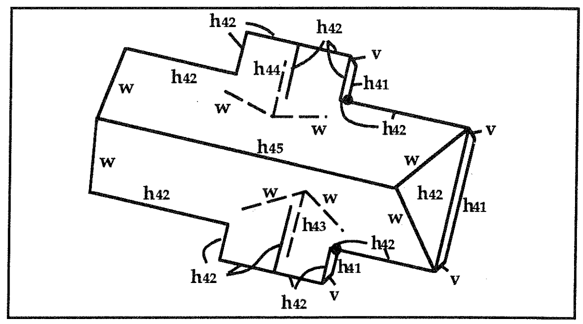

Figure 2-4 ART4*) - Result of a line parsing.

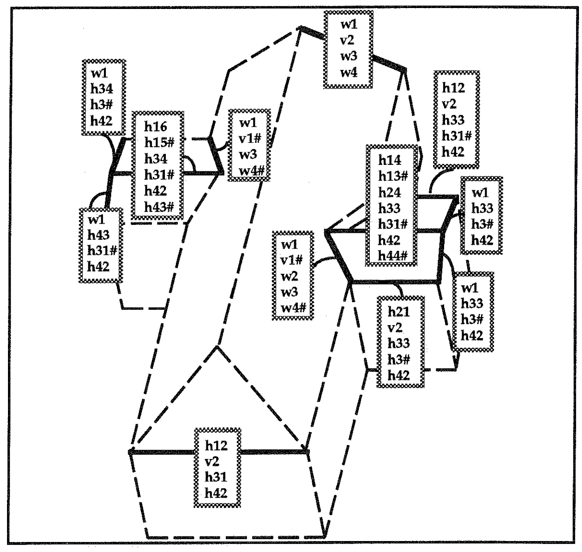

Figure 2-5 Merging the parsing result of the views ART1 - ART4*) in object space. Inconsistencies in line labelling are indicated. Double labelling in a single view is marked by (#).

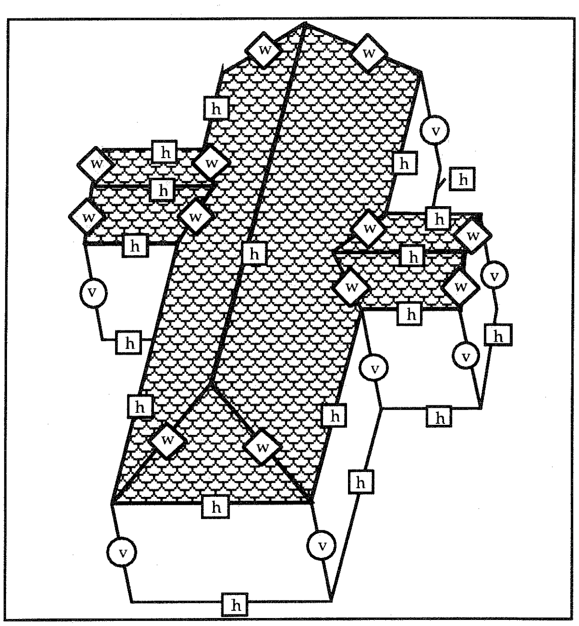

Figure 2-6 Final result after consistency check*) for building ART. Lines and building parts in object space are correctly found.

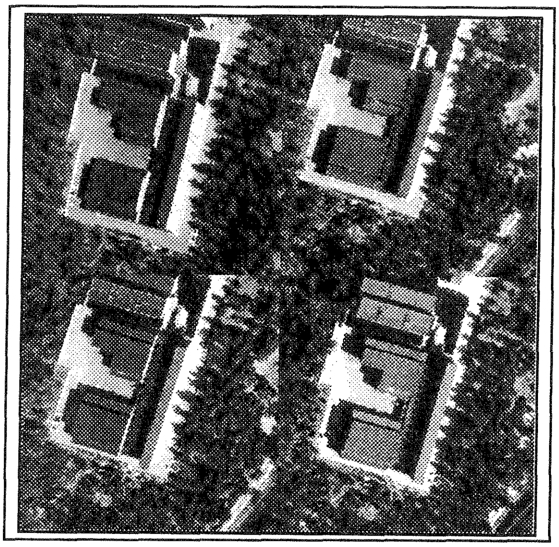

Figure 3-1 Four aerial views of a building complex. Upper: views K1, K2 lower: views K3, K4

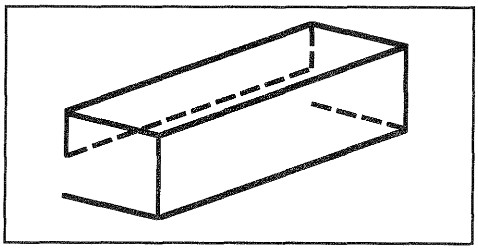

Figure 3-2 Desired object boundary lines of the lower right building in the views K1 - K4.

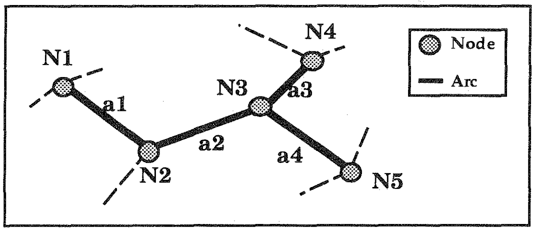

Figure 3-3 Segmentation node and arc structure.

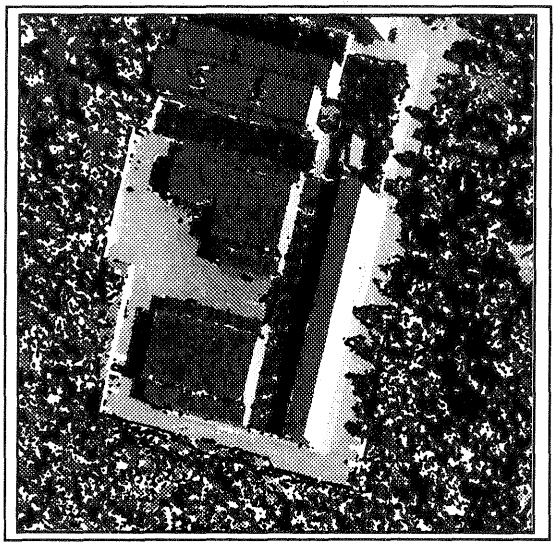

Figure 3-4 Segmented regions (K1_1) by region growing in view K1 with parameter set 1.

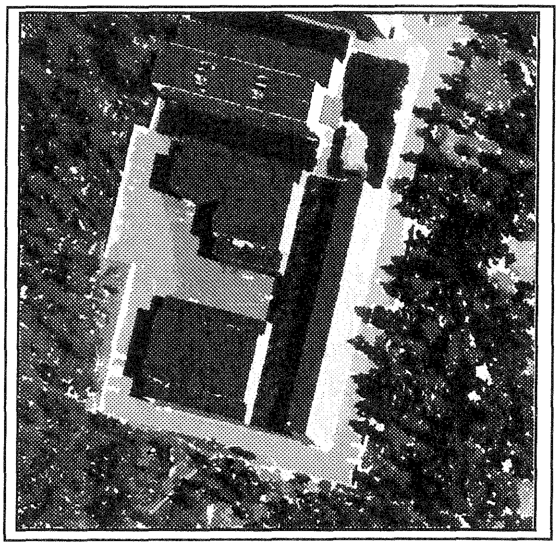

Figure 3-5 Segmented regions (K1_100) by region growing in view K1 with parameter set 100.

A window of attention around the building was chosen by the user with a border approximately 50 pixels wide. All arcs cutting the window borders were eliminated, accepting some eventually open regions. In addition single arcs which were closed, thus indicating e.g. windows were eliminated for this test. All other arcs were accepted for further processing. Figure 3-6 shows as an example the accepted arcs as black overlay lines in the given quadrilateral window of attention for the view K1.

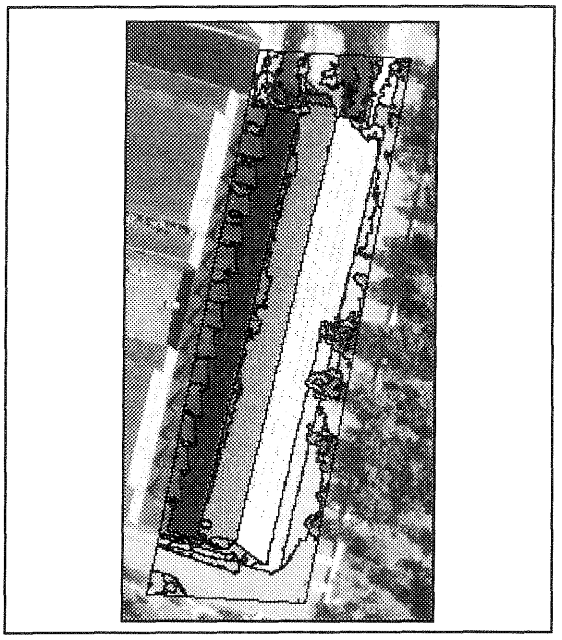

Figure 3-6 Accepted arcs (black overlay) in the window of attention in segmentation K1_1.

A closer look on these segmentation results shows that the used segmentation is not suited as input for the described parsing procedure. It was checked how many lines in each window of attention actually lie on an object boundary (one of those in figure 3-2). The result of this investigation for all segmentations is shown in figure 3-7. Only between 7.0% and 31.8% of all accepted arcs in the window of attention actually

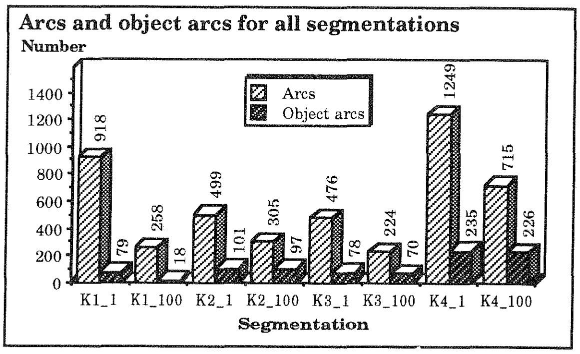

Figure 3-7 Number of arcs in the segmentation and arcs on object boundaries for all views.

are desired ones, i.e. object arcs, which is an extremely large, unacceptable handicap. The percentage decreases even further if the approximation for the window of attention is not as good.

In addition there exist many small arcs which can not contribute to the parsing very much. From the histogram in figure 3-8 it can be derived that around 80% of all accepted arcs have a length below 20 pixels in chain code representation. Between 20% and 60% are shorter than 5 pixels. These huge number of small arcs will make the parsing process extremely difficult. It is not possible to just eliminate them, because they are needed to keep up with all geometric relations among the arcs, an essential requirement for both parsers. With the applied segmentation it is in addition not possible to get a complete coverage of all desired object boundaries, i.e. that at least 75% of an object boundary line is covered by arcs in the segmentation. This is evident from the figures in table 6 where for all four views the coverage of visible object lines is given. The segmentation with the less tight threshold (parameter set 100) gives a lower total number of arcs, but results in less coverage of object boundary lines and more very small arcs. More tight thresholds give higher coverage but also an enormous increase in the number of arcs (cf. figure 3-7).

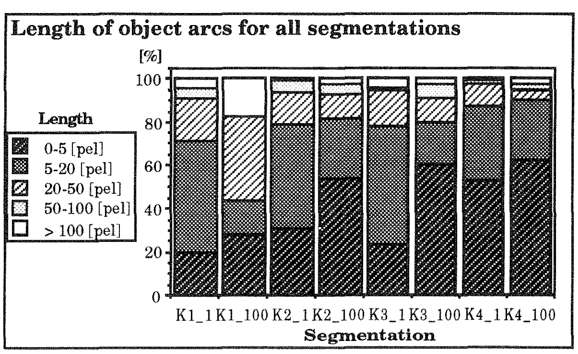

Figure 3-8 Histogram on length of arcs (in pixel) belonging to desired object boundaries.

Table 6 Coverage of visible object boundary lines

View	# of visible object lines	# covered Segment. 1	# covered Segm. 100
K1	9	*8*	*8*
K2	7	*7*	*6*
K3	9	*7*	*7*
K4	9	*9*	*8*

DISCUSSION

Based on these unsatisfying segmentation results an application of the described parsing procedure was not regarded as meaningful. Instead of readjusting the procedure the attention is now focused on improving the segmentation according to the requirements of the parsers.

Given a perfect segmentation the suggested parsing procedure is capable to handle quite difficult views of a complex building. In the first case discussed it was possible to reconstruct all visible object lines and object parts. Nevertheless the description of the whole building is still not complete. This means after the parsing procedure an additional step, an extrapolation of the hidden parts should be applied using e.g. symmetry relations. It is clear that these hypotheses can be not checked from the data and must be treated with care.

Nevertheless the whole reconstruction mainly rests on the quality of the segmentation procedure. If it would be a matching against known structures, i.e. if a specific model would exist then already fragmented boundaries can be sufficient to extract an object. This is not the case for the described problem, where only a generic model of the objects in question is available. The parsing depends on long, connected line segments forming mainly closed polygons, with a high coverage of object boundary lines. It is believed that the knowledge that buildings are to be reconstructed must be used already in the segmentation process and not in the high-level interpretation phase. In the last two, three years several approaches on segmentation have been published which try to include such knowledge in a procedure, warming up the old idea of combining region segmentation with boundary detection. Following these ideas first attempts have been made to improve the segmentation results according to the requirements of the described parsers (Stokes, 1992). Amongst others texture information and constraints on boundaries like pairwise parallelity, length etc. are possible extensions which have to be incorporated in a way or other to existing or new segmentation methods in addition to a data and knowledge driven choice of thresholds.

Having solved the segmentation problem in a more suitable way it is believed that the described parsing procedure can be successfully developed further.

REFERENCES

Dickinson, S. J., Pentland, A. P., Rosenfeld, A., 1990. Qualitative 3-D Shape Recovery using Distributed Aspect Matching. CAR-TR-505, CS-TR-2484, AFOSR-86-0092, University of Maryland, USA

Gülch, E., Axelsson, P., Stokes, J., 1990. Object Description and Precise Localization of Prescribed Objects in Digital Images. In: IAP Vol 27-3.

Gülch, E. 1991. Using a Knowledge Based System in the Development of a Generic Model for Buildings. In: Proceedings IGARSS'91, Espoo, Finland.

Stokes, J., 1992. Parsing segmented images. In: Proceedings XVII ISPRS Congress, Washington D.C., USA.

Walz, D. L., 1975. Understanding Line Drawings of Scenes with Shadows. In: Winston P.H. (Ed.): The Psychology of Computer Vision, MacGraw-Hill.