Recognition of Structured Objects in Monocular Aerial Images Using Context Information

Franz Quint

Institute for Photogrammetry and Remote Sensing
University of Karlsruhe
76128 Karlsruhe, Germany
quint@ipf.bau-verm.uni-karlsruhe.de

Abstract:
The purpose of our system MOSES is the recognition of objects in aerial images. It is a knowledge based system in which semantic networks are used as a repository for models. The models are automatically refined by using knowledge gained from topographical maps or GIS data. The analysis process is formulated as an optimization problem. After presenting the object models and the process model of our system we address the recognition of structured and compound objects in aerial images by using context information. Results are presented and performance aspects are discussed.

1 Introduction

Due to its complexity the interpretation of scenes in aerial images is one of the challenging problems in computer vision. Various applications like change detection, environmental monitoring, site modeling or map updating would greatly benefit from reliable automatic or semi-automatic procedures for aerial image interpretation. In this spirit, since several years the international community supports big efforts for developing efficient and reliable aerial image understanding procedures (see e.g. [10], [11], [12]). Recent research has mainly been focused on the analysis of urban scenes (man-made objects) from aerial images ([2], [3], [4], [5], [8], [14], [19], [21]).

Our research work aims at the determination of soil sealing in an urban environment. The degree of soil sealing has implications on the regeneration of ground water and on the micro-climate. This work is part of a research project [22] in cooperation with the Forschungsinstitut für Informationsverarbeitung und Mustererkennung (FGAN-FIM) Ettlingen and is funded by the Deutsche Forschungsgemeinschaft (DFG).

To achieve our goal we are analyzing color aerial images digitized to a pixel size of 30cm $\times$ 30cm on the ground. Aerial image analysis is based on knowledge gained from the large scale topographical map DGK 1:5000.

Our paper is organized as follows: after a short introduction to semantic networks we describe our system MOSES. Firstly the different object models involved in the representation of the knowledge are presented. In the following chapter we describe the process model of our system. The recognition of structured and compound objects in aerial images by using context information is demonstrated for the example of buildings and parking places. We finish our paper with a discussion of results.

2 Semantic Networks

The analysis of aerial images is performed in MOSES (*M*ap *O*riented *SE*mantic image under*S*tanding) [1],[16] as a model based, structural approach. For representing the models we use semantic networks as implemented by the shell ERNEST [6].

In semantic networks knowledge is represented using *nodes* and *links*. ERNEST provides three types of nodes: *concepts, instances* and *modified concepts*. Classes of objects, events or abstract conceptions are represented by concepts (for example: `parkPlace` represents parking places). Extensions of concepts found in the sensor data are represented by instances. Modified concepts are intermediate results representing constraints to concepts yet uninstantiated. Data structures called *attributes* are used to further describe the nodes. They enable us to specify and store properties of concepts, modified concepts or instances.

Links are used to represent relations between nodes. *Specialization* links and *part* and *concrete* links are of particular importance. Specialization links connect a concept with a more general concept (for example: $\texttt{sceneObject} \xrightarrow{spec} \texttt{parkPlace}$). Along this link type properties of the more general concept are inherited to the special concept, unless these properties are explicitly modified.

Part links represent the relations between a concept and its components (for example: $\texttt{parkPlace} \xrightarrow{part} \texttt{carRow}$). Concrete links connect concepts in different conceptional systems. For example, the roof of a building may belong to the conceptional system of objects. Its concretization in the geometrical conceptional system may be a parallelogram. These links induce a hierarchy over the concepts in a model, each hierarchical level representing a different degree of abstraction from the available visual information [9]. Part and concrete links may be multiple. One can specify for the parts and concretizations of a concept if they are obligatory, optional or inherent.

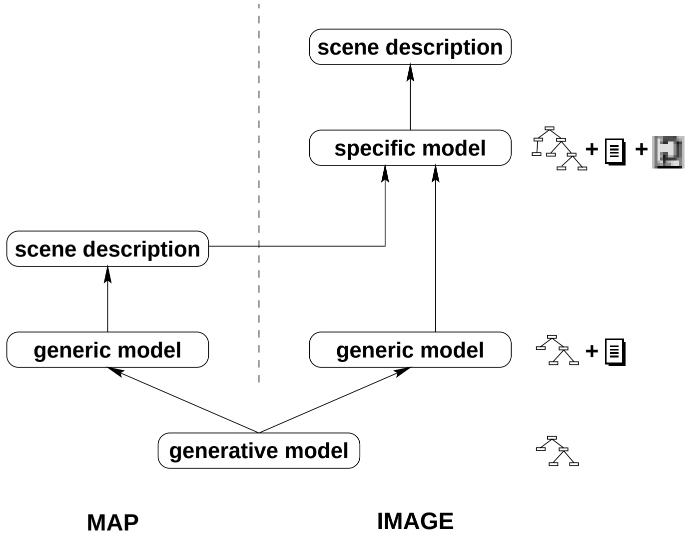


Figure 1: Architecture of MOSES

3 Object Models

To profit in our image understanding process from the general, common sense knowledge the system developer has about his environment as well as from the scene specific knowledge gained from the analysis of maps we are using a model hierarchy as presented in Fig. 1.

3.1 Generative Model

The most general object model used in MOSES is the *generative model*. This model contains general, common sense knowledge we have about our environment. Objects which may occur in a scene are described by specifying their parts and/or concretizations, their properties and their relations with other objects from the scene. For example the concepts `roof` and `hood` are specified as parts of a car. Attributes (properties) of the car are for example its position, its length and width and its color. For a car on a parking place we specify the relation of parallelism with other cars in the same parking row. By describing an object with help of its parts we construct a hierarchical, structural model. This model also has a parametric part, which is provided by the attributes of the concepts. If possible, we specify domains of validity for the attributes: e.g. the length of a car may vary between 3 and 7 meters.

The knowledge contained in the generative model is of declarative nature exclusively. The model is independent of a particular scene to be analyzed and the description is performed in the scene domain. This means that the scene is described as perceived by a human, independent from the representation of the scene in an image or a map. However, in order to save effort, the model is goal oriented. Since its goal finally is the recognition of objects in aerial images, we will not describe for example the engine of a car and all of the engine's parts.

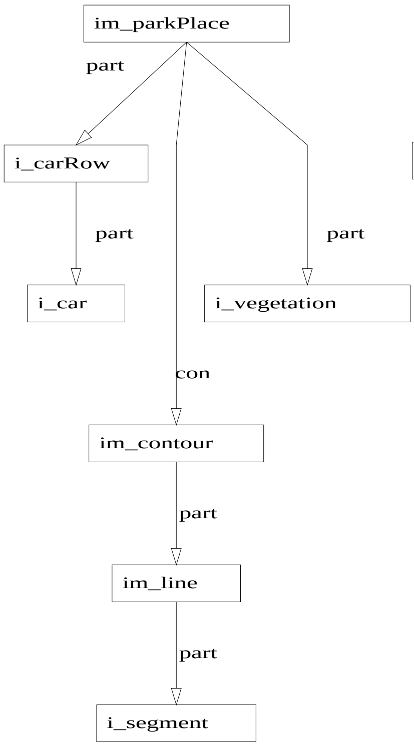
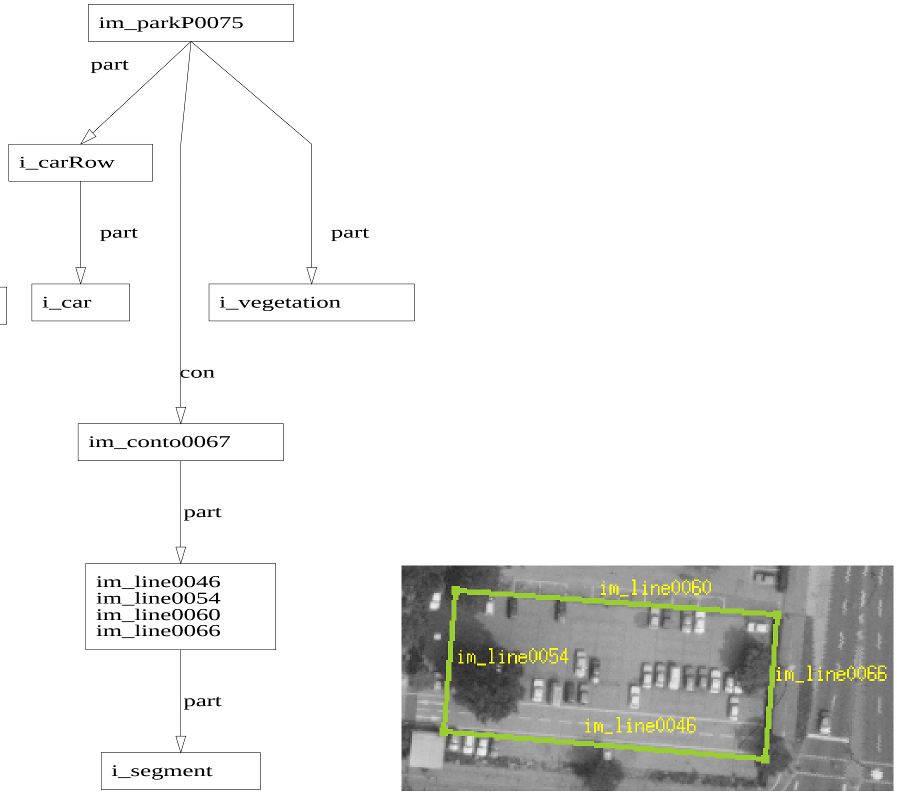


Figure 2: Detail of the generic model in the image domain: part- and concrete-hierarchy for a parking place



(a) Part- and concrete-hierarchy

(b) Concepts as overlay on the aerial image

Figure 3: Detail of the specific model for a parking place

3.2 Generic Models

The knowledge of the generative model is further refined by constructing the *generic model in the map domain* and the *generic model in the image domain*. These models describe the projection of the scene in a topographic map and in an aerial image, respectively.

The concepts in the generic models are specializations of the corresponding concepts in the generative model. Due to the inheritance mechanism of the semantic network it is not necessary to specify again all the parts, relations and properties of a concept. However, particularities of the projection of the scene in the map and in the image are represented by the generic models. For example, since cars are not represented in maps, all the part links in the generic model in the map domain pointing towards concepts `car` are marked as deleted.

In addition, procedural knowledge is added to this step of model refining. This comprises methods for feature extraction, for the calculation of the attribute values from the image data or from the map data and functions for evaluating the preliminary and the final analysis results. These methods are called during the analysis process by the control algorithm.

A detail of the generic model in the image domain representing the part- and concrete-hierarchy for a parking place is given in Fig. 2. Concepts are represented in boxes and links are represented as arcs between two concepts. The links in the generic model are multiple links since it is not yet known for example, how many rows of parking cars there are in a parking place. The actual multiplicity count of such links will be determinated only in the analysis phase. Some of the links are optional links: for example the part link connecting the concepts `im_parkPlace` and `i_carRow`. The multiplicity of a link and whether it is optional or not are not represented in Fig. 2.

The generative model and both generic models are build by the system developer. They are independent of a particular scene to be analysed.

3.3 Specific Model

The *specific model* is automatically generated by the system. This model is specific for the current scene to be analysed; it is generated by combining the scene description obtained after the map analysis with the generic model in the image domain.

The concepts of the specific model are specializations of the concepts of the generic model in the image domain. Thus, they inherit the procedural knowledge specified in the generic model. The attributes of these concepts are modified in order to reflect the knowledge gained after map analysis. In particular, after map analysis, expectations for the location and the ground-plan of the objects contained in the map are known and are incorporated automatically into the specific model.

In Fig. 3(a) a detail of the specific model representing a parking place is shown. This parking place now represents not just any parking place in a scene, as it is the case in the generic model, but a particular parking place which is depicted in the map. Having the ground-plan and the location of the object gained after map analysis, we can transform these data into the image domain and represent that specific object and its parts as overlays to the aerial image. This is shown in Fig. 3(b).

As long as it can be done automatically (and this is the case in our procedure to incorporate map information), one must specialize the model as much as possible when starting from a generic model. This will dramatically reduce the necessary effort to interpret the image and will lead to more reliable results.

In conclusion, an approximative characterization of the models used in MOSES is given by: the generative model contains the structural information, in the generic models the domain specific analysis procedures are added and finally the specific model is refined by adding constraints (see Fig. 1).

4 Process Model

At the beginning of an analysis task the system consists of the generative model and the two generic models. On the way to a scene description gained by the analysis of an aerial image, several steps are performed automatically:

- structural analysis of the map,
- generation of the specific model,
- extraction of image primitives,
- structural analysis of the image.

4.1 Structural Analysis of the Map

Digital maps are used as input data for this step. Therefore a feature extraction process is not necessary for the map; the digitally available line segments are used directly as primitives for the structural analysis. This analysis process is performed similarly for the structural analysis of the image and will be described there.

The result of this step is a structural description of the scene, as complete as far as the scene is depicted in the map. Instances representing more and less complex structures in the map are arranged in a graph. The arcs of the graph are *part* and *concrete* links which connect instances being in the corresponding relation. The *attribute* data structures of the instances are filled with values calculated from the map data.

4.2 Generation of the Specific Model

The scene description gained after the analysis of the map is automatically combined with the generic model in the image domain to yield the specific model which will be used for the image analysis.

Each instance in the scene description is connected with an instance link to the belonging concept. Following this link back we find the concept in the generic model in the map domain for which the instance has been created. The found concept is a specialization of a concept from the generative model. Following this specialization back, we reach the concept in the generative model. From this concept we can follow another specialization link, which leads to a concept in the generic model in the image domain. We have now found a correspondence between an instance in the scene description after map analysis and a concept standing for a similar structure, but represented in the image. We call this concept *partner-concept*.

For each instance of the scene description we create a new concept in the specific model as a specialization of the partner-concept. Starting with the top-most concept this can be done recursively over the part- and concrete-hierarchy of the model. Constraints are derived for the attributes of the newly created concepts by using the attribute values of the corresponding instances.

The newly created concepts of the specific model form a structure which mirrors the structure of the scene description after map analysis. Therefore, the instantiation of this part of the specific model in the image analysis process is equivalent to a verification of the map contents in the image. To be able to recognize in the image also objects which are not depicted in the map, the concepts of the generic model in the image domain are copied into the specific model. However, the structural analysis of the image will start with the newly created concepts of the specific model. Thus, when attacking the problem of recognition of objects which are not depicted in the map one can benefit from the context assembled by the objects already instantiated.

4.3 Extraction of Image Primitives

Before the structural analysis of the image using the specific model is started, primitives are calculated from the image data. We are currently using line segments and regions as primitives.

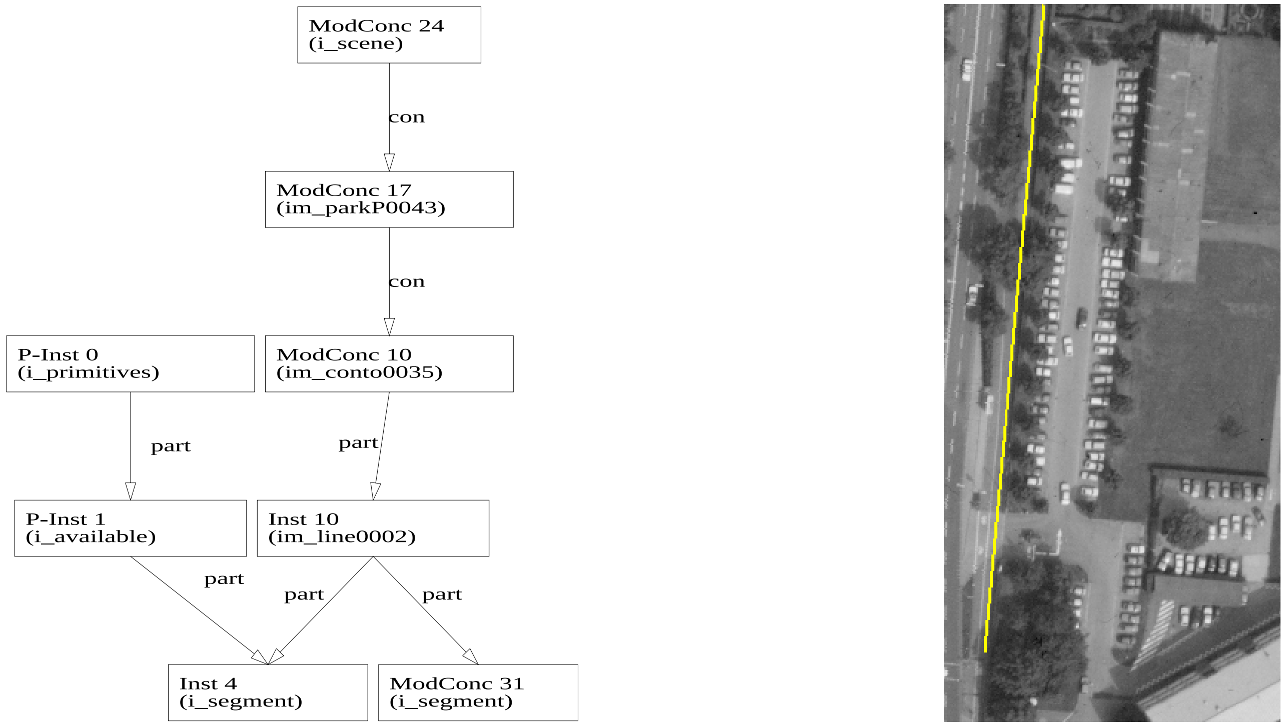
The line segments are gained by detecting contour points with a Canny-like operator and approximating the collected contour point chains with straight line segments. The line segments are additionally attributed with features like the mean magnitude and the mean angle of the gray value gradient along the line and with a measure for the goodness of approximation of the contour point chain by a line segment.

To gain the regions we are using a segmentation procedure based on a Bayesian homogeneity predicate [15]. To describe the shape of the regions we are using moment invariants [7]. Furthermore the chromatic information and the list of neighbors is calculated for each region.

To support perceptual grouping we arrange the primitives in a parametrized graph. The nodes of the graph are attributed with regions. Nodes containing neighbouring regions are connected with arcs. The arcs are attributed with those line segments which form the border between neighbouring regions.

4.4 Structural Analysis of the Image

The image primitives serve as input data for the structural analysis of the aerial images with the specific model. The analysis process is conducted by a task independent control algorithm and is essentially model driven. Performing the analysis with our model stored in a semantic network



(a) Structural description

(b) Highlighted instances

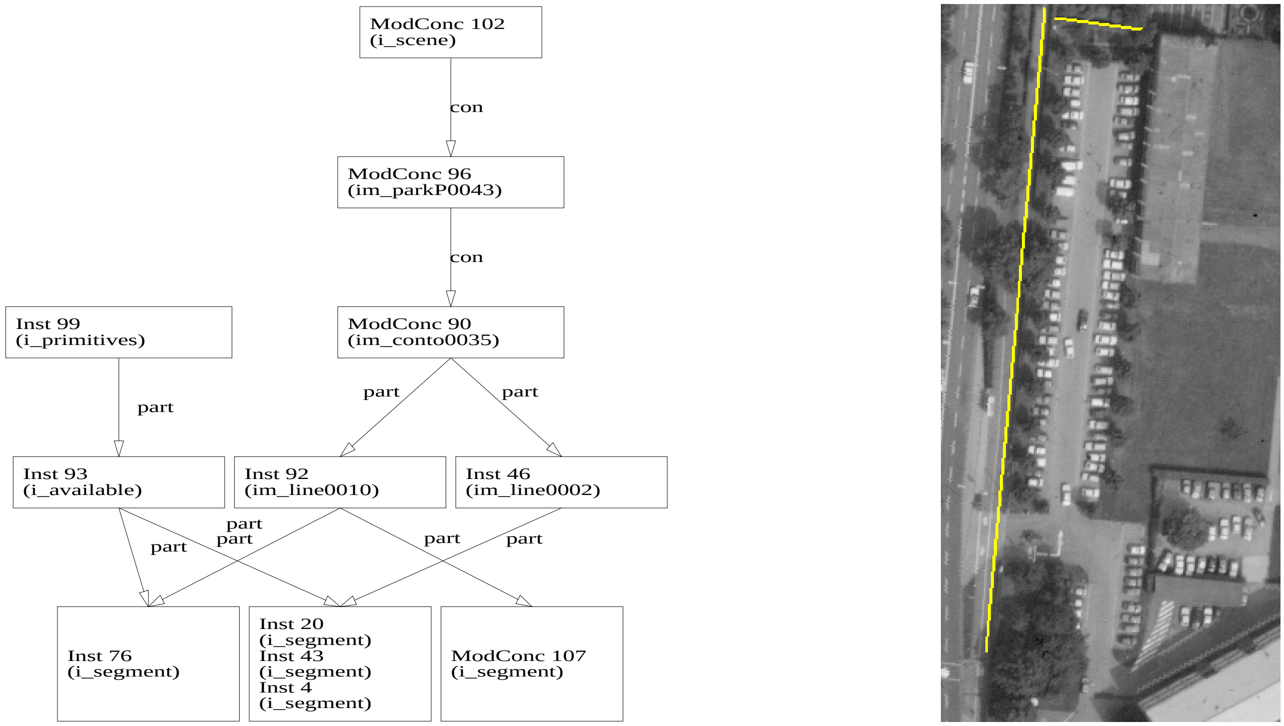
Figure 4: Intermediate analysis result at node #16

means to find correspondences between the concepts in the model and single primitives or groups of primitives from the database. Establishing these correspondences is called instantiation.

Analysis starts by creating a modified concept for the top-most concept in the hierarchy of the semantic network. This modified concept is used to reflect the constraints which will be acquired during the instantiation of concepts on lower hierarchical levels. It can also be regarded as a hypothesis for the existence in the data of the object or structure represented by this concept.

Thereafter, analysis moves down in the part- and concrete-hierarchy of the semantic network and creates modified concepts, until it reaches a minimal concept. A minimal concept is a concept without parts and concretizations. This concept is "close" to the signal and can be instantiated. After this instantiation the acquired knowledge is propagated bottom-up into the modified concepts on higher hierarchical levels. As a next step, this knowledge is also propagated top-down. To store this knowledge, a new modified concept on the low hierarchical level will be created, preparing a new instantiation.

A snapshot of such a situation is shown in Fig. 4(a). The concept on the lowest hierarchical level is `i_segment`, representing a primitive from the database. One instance (instance #4) for `i_segment` has already been created; it is highlighted in Fig. 4(b). The modified concept #31 is the next one to be instantiated. The location of instance #4 known so far has been propagated



(a) Structural description

(b) Highlighted instances

Figure 5: Intermediate analysis result at node #54

to the modified concept #31 and provides an expected location for the next instance to be created.

When all the obligatory parts of a concept are created, the concept itself can be instantiated: the analysis moves one step upwards in the hierarchy. An example for this is shown in Fig. 5(a). After creation of all the parts of the concept `im_line002`, the instantiation of this concept has been completed. Analysis moved one step upwards to the modified concept for `im_conto0035` and than again downwards to instantiate the next part of this concept, namely `im_line0010`. This process continues in a similar way for all concepts of the model and it is usually finished as soon as an instance for the top-most concept can be created. In this case the goal of the analysis has been reached.

In conclusion one can say that in our case analysis is an active construction of complex structures by composing less complex structures. Step by step interim results of increasing degree of abstraction are generated. The further the analysis proceeds, the more the structure of the interim results resembles the specific model. Generation of the analysis results follows the hypothesize and test paradigm.

Each time a new instantiation or concept modification takes place, analysis reaches a new state. All of these states are valid intermediate results and are logically consistent. Since the preceding

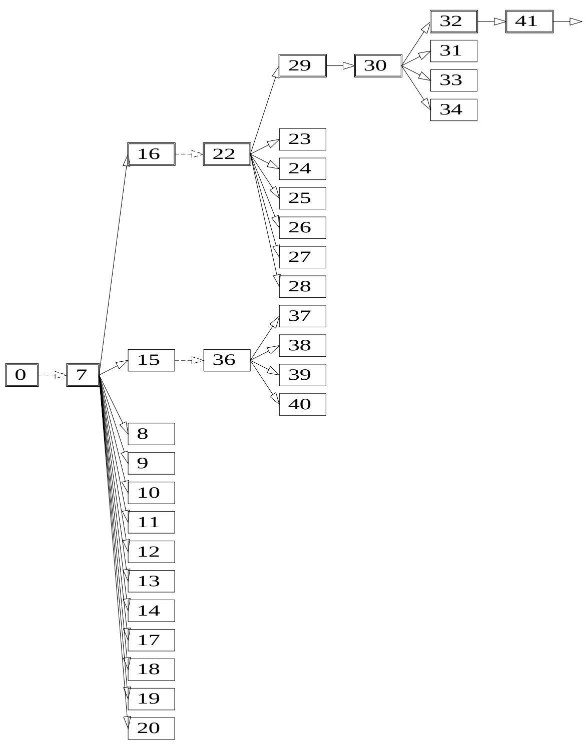


Figure 6: Detail of a search tree

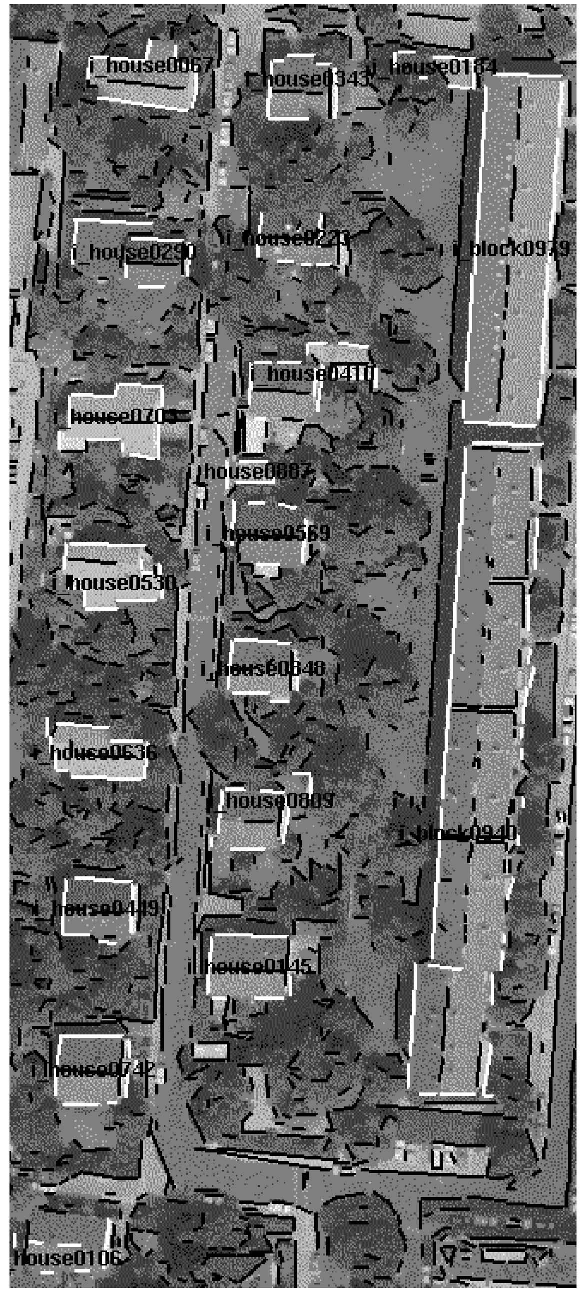


Figure 7: Recognized buildings

states are known, in each state of the analysis it is possible to explain which decisions have been taken to reach the current state and one can determine the subset of image primitives which form the (intermediate) analysis result in this state.

We can graphically represent an analysis state by a node and can connect a state with its preceding state by an arc. Thus, the progress of the analysis process is graphically represented by a tree, the *search tree*. A detail of a search tree is presented in Fig. 6.

As explained, instantiation means establishing a correspondence between a primitive from the database and a concept of the model. In a given analysis state, situations may occur where

Object	Qual.	Object	Qual.	Object	Qual.	Object	Qual.
i_house0067	0.98	i_house0106	0.36	i_house0145	0.97	i_house0184	0.93
i_house0223	0.84	i_house0290	0.88	i_house0343	0.98	i_house0410	0.86
i_house0449	0.95	i_house0530	0.87	i_house0569	0.90	i_house0636	0.64
i_house0703	0.95	i_house0742	0.94	i_house0809	0.88	i_house0848	0.85
i_house0887	0.98	i_block0940	0.79	i_block0979	0.90		

Table 1: Model fidelity for the buildings of the scene in Fig. 7

several correspondences seem to be possible (e.g. in state #22 of the tree in Fig. 6), but only one of them being the correct one. In such a case one has to account for all these correspondences. Thus, a given analysis state may have several successors and the search tree is splitted.

From this point of view, analysis in our context is nothing else than search for an optimal path in the tree of the analysis states. There exist several well known graph-search methods. In our system we are using the A^*-algorithm [13]. To construct the merit function needed by this algorithm we evaluate the instances and modified concepts in an analysis state [17]. The valuations are computed on the basis of the Dempster-Shafer theory of evidence [20]. Using an extension to this theory [18], the valuations are propagated through the hierarchy of the semantic network and are combined to result in a merit function for the A^*-algorithm.

The valuations for the instances are not only used to calculate the merit function of an analysis state, but they are also used as a quality measure for the verification and recognition of objects. An example for this is given in Figure 7 and Table 1. After calculation of the image primitives near 5000 line segments have been detected. These are displayed as dark lines in the image. The line segments which have been selected by the analysis to compose the buildings of the scene are drawn in white color. For each building a label is printed in black color at its location in the image. The computed quality measures are listed in Table 1. In a verification task one would apply a threshold on these quality measures to decide if the verification was successful. Choosing for example this threshold to be 0.5, in the presented scene all buildings except the one labeled `i_house0106` (situated in the lower left corner of the image) have been verified successfully.

5 Recognition of Parking Places

There are several reasons why a parking place is interesting to be addressed in an object recognition task. Parking places are composites containing objects of different classes (e.g. cars, vegetation). As opposed to other man-made objects like for example buildings, the borders of the parking area are not always visible in the image. Parking places can often be recognized as such only in the context of parking cars.

Our approach for the recognition of parking places differs from the approach used to recognize other objects as for example buildings. The contours of a parking place as predicted by the map are less reliable than the contours of a building. This is considered in the evaluation functions, where the valuations for the contours of a parking place contribute in a smaller amount to the total valuation of the instance than this would be the case for buildings. Nevertheless, as a first step an attempt is made to detect the contour of the parking place in the image. In this step we also identify the image regions which may be part of the parking place. The predicted location serves now as a region of interest to search for typical configurations of cars.

Since the image primitives (regions and line segments) are stored in a neighbourhood graph, we can easily find the regions which are enclosed by or are near to the regions predicted to be part of the parking place. The regions selected in this step of the process are candidates for the detection of cars. Cars may be formed by one, two or more regions. The evaluation of the car-hypothesis is performed considering the dimensions of cars as specified in the generative model, the shape of the regions, their relative position and the similarity in color of the regions.

After the detection of one car, hypotheses for rows of cars can be established. Further cars fitting into these predicted rows are now searched. The evaluation functions for car-hypothesis inside a row are dynamic: with growing size of the car rows, the stringency of the evaluation is gradually released. However, it is again drastically enforced as soon as the car rows reach a maximum size, which is calculated from the dimensions of the parking place in context.

6 Results and Discussion

An example for the recognition of buildings and parking places in an aerial image is given in Fig. 8. After the calculation of the primitives, more than 6300 line segments and 1640 regions are stored in the neighbourhood graph and are presented to the structural analysis process. These primitives are not displayed in Fig. 8. The white lines displayed in the Figure are the representations of the instances to concepts `im_line` which form the contours of the objects. The solid white lines have been built by aggregating image primitives. The broken white lines are *wildcards*. They are inserted by the system in the case that the evidence given by the image primitives was not sufficient to detect a contour in the image at that location.

Contour lines have not been reliably detected where they are occluded by vegetation or by the cars (the contour on the right side of the big, L-shaped parking place) or in the case of missing contrast in the image. This can be observed for an edge of the H-shaped building and also for the left-side edge of the parking place in the lower left area of the image. The contours of this parking place as given by the map are only administrative border lines and are not visible in the image: in reality, the entire area in front of the H-shaped building is asphalted and only a part of it is used as a parking place.The map does not provide information about the type of

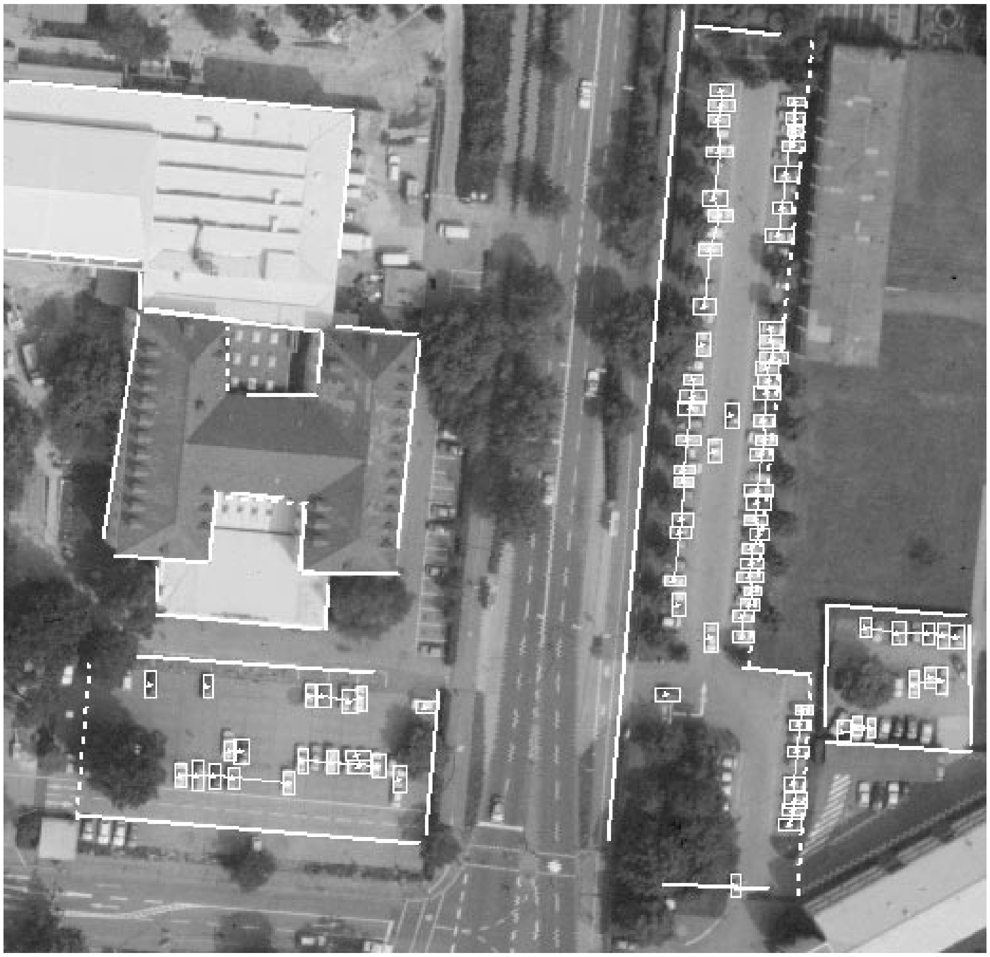
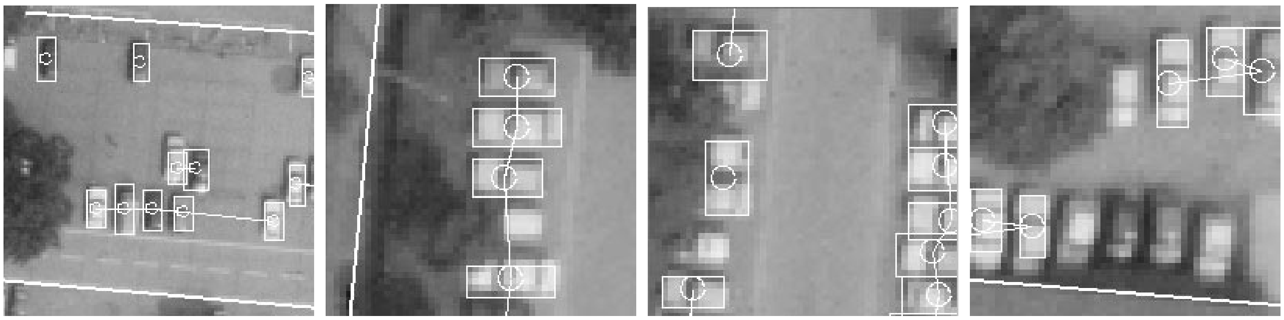


Figure 8: Result of the image analysis

the surface outside the area marked as parking place.

Lane markings on the asphalt are found as upper and lower borders of this parking place. The upper border was detected at an offset of about 4m from its location predicted by the map. The marking lane at the lower border fits well to the predicted location for the border. However, as it can be observed in Fig. 8, below this border there are also parking cars and in fact one could regard all the area down to the stripe of vegetation as belonging to the parking place. Due to the well fitting, the analysis state containing the lane marking as the low border has a high valuation and thus search paths containing other competing borders are not followed. As well, we do not yet search for cars outside the detected area of a parking place. This is the reason why the four cars below the detected parking place are still missed. It is part of our future work to extend our system such, that also in neighbouring areas of a parking place rows of cars are searched and in the case of success an expansion of the parking place is performed.

The instance representing the edge on the left side of the H-shaped building is erroneously aggregated from a mixture of line segments from the edge of the roof and from the ground-plan of the building. This could have been avoided if 3D data would have been available.



(a) (b) (c) (d)

Figure 9: Details of the recognition results of Figure 8

In Fig. 8 the detected cars are surrounded with their bounding box in white color. A marking in form of a circle is drawn in the middle of this bounding box. A white line connects the markings of all cars which have been found to be in the same row. As it can be observed in Fig. 8, the recognition has been successful for most of the cars on the parking places.

No structures other than cars have been detected as such. This is not only due to the parametric part of our model (size of a car etc.), but it is also a merit of the structural part of the model. Cars are optional and context dependent parts of a parking place, but not of buildings, playgrounds or other objects in the scene (streets are not yet incorporated in the system). Thus, the system does not even make an attempt to search for cars e.g. on the roof of a building – although there exist regions on top of buildings which would satisfy the parametric model of a car. This is an advantage of an essentially model driven approach over a data driven approach. Inconsistencies as mentioned above are likely to be generated in a data driven analysis process and have to be removed later on during a consistency checking; they are unlikely to be generated in a model driven approach like the one we are using.

There remain some cars on the parking places which have not been detected or where some parts have been missed. We will only focus on the problems and present them at the details shown in Fig. 9(a)-(d).

Fig. 9(a) gives a detail of the parking place in the lower left part of the image shown in Fig. 8. All the cars have been recognized correctly except the leftmost car in the middle row. For this car the region representing the motor-hood has been missed. It is because the other two regions of the car were already satisfying the parametric model and because without the motor-hood the evaluation of the fitting inside a row together with the car nearby provided a higher evidence than in the case the motor-hood was also considered.

The Figures 9(b) and 9(c) are details of the big L-shaped parking place. A typical example of failed detection due to occlusion is presented in Fig. 9(b). A big part of the second car from the lower border of the detail is occluded by trees. The remaining visible part of the car is too small to satisfy the model requirements. The planned recognition of the vegetation should help us to overcome this kind of problems.

A similar situation is encountered in Fig. 9(c). On this section of the parking place the cars are arranged in rows having North-South direction in the image. Substantial occlusions occur for four cars in the left row. Two of these cars have been completely missed. The other two cars are close one to the other and have the same color. The roofs are the only visible part for each of these two cars and they are interpreted by the system as being the roof and the motor-hood of one and the same car (the vertically oriented box in the Figure).

The Figure 9(d) is a detail of the parking place in the right part of the image. Here the requirement for a similar color of the regions belonging to a car leaded to the failure of the detection of the four cars in the lower part of Fig. 9(d). This condition is violated because of the spot-like reflections of sunlight on the cars.

For a part of the presented problems it will be difficult to develop a recognition procedure which will solve them in any situation. However, the other part can be eliminated by enlarging the knowledge base of our system. Already now the automated recognition of structured and compound objects in aerial images with our system MOSES is successful in most of the cases.

Acknowledgment

Discussions with Dr. Manfred Sties have been most helpful for improving this article.

References

[1] H.-P. Bähr, F. Quint, and U. Stilla. Modellbasierte Verfahren der Luftbildanalyse zur Kartenfortführung. *Zeitschrift für Photogrammetrie und Fernerkundung*, 63(6):224–234, 1995.

[2] R.T. Collins, A.R. Hanson, E.M. Riseman, and H. Schultz. Automatic extraction of buildings and terrain from aerial images. In A. Grün, O. Kübler, and P. Agouris, editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 169–178. Birkhäuser, Basel, 1995.

[3] W. Förstner. Mid-level vision process for automatic building extraction. In A. Grün, O. Kübler, and P. Agouris, editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 179–188. Birkhäuser, Basel, 1995.

[4] N. Haala and K.-H. Anders. Fusion of 2D GIS and image data for 3D building reconstruction. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 285–290, 1996.

[5] O. Henricsson, F. Bignone, W. Willuhn, F. Ade, O. Kübler, E. Baltsavias, S. Mason, and A. Grün. Project AMOBE: Strategies, current status and future work. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 321–330, 1996.

[6] F. Kummert, H. Niemann, R. Prechtel, and G. Sagerer. Control and explanation in a signal understanding environment. *Signal Processing*, 32:111–145, 1993.

[7] Y. Li. Reforming the theory of invariant moments for pattern recognition. *Pattern Recognition*, 25(7):723–730, 1992.

[8] C. Lin, A. Huertas, and R. Nevatia. Detection of buildings from monocular images. In A. Grün, O. Kübler, and P. Agouris, editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 125–134. Birkhäuser, Basel, 1995.

[9] D. Marr. *Vision.* H. W. Freeman and Co., San Francisco, 1980.

[10] T. Matsuyama and V. Hwang. *SIGMA: A Knowledge-Based Aerial Image Understanding System.* Advances in Computer Vision and Machine Intelligence. Plenum Press, New York, 1990.

[11] D.M. McKeown, W.A. Harvey, and J. McDermott. Rule based interpretation of aerial imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(5):570–585, 1985.

[12] B. Nicolin and R. Gabler. A knowledge–based system for the analysis of aerial images. *IEEE Transactions on Geoscience and Remote Sensing*, 25(3):317–329, 1987.

[13] N. Nilsson. *Principles of artificial intelligence.* Springer, Berlin, 1982.

[14] M. Pasko and M. Gruber. Fusion of 2D GIS data and aerial images for 3D building reconstruction. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 257–260, 1996.

[15] F. Quint and S. Landes. Colour aerial image segmentation using a Bayesian homogeneity predicate and map knowledge. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 663–668, 1996.

[16] F. Quint and M. Sties. Map–based semantic modeling for the extraction of objects from aerial images. In A. Grün, O. Kübler, and P. Agouris, editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 307–316. Birkhäuser, Basel, 1995.

[17] F. Quint and M. Sties. Evaluating model fidelity in an aerial image analysis system. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 669–674, 1996.

[18] F. Quint and M. Sties. An evidential merit function to guide search in a semantic network based image analysis system. In P. Perner, P. Wang, and A. Rosenfeld, editors, *Advances in Structural and Syntactical Pattern Recognition*, pages 140–149. Springer, Berlin, 1996.

[19] F. Sandakly and G. Giraudon. Multispecialist system for 3D scene analysis. In A. Cohn, editor, *11th European Conference on Artificial Intelligence, ECAI 94*, pages 771–775. John Wiley & Sons, Ltd., 1994.

[20] G. Shafer. *A mathematical theory of evidence.* Princeton University Press, 1976.

[21] U. Stilla and K. Jurkiewicz. Structural 3D-analysis of urban scenes from aerial images. In *ISPRS, XVIIIth Congress, Intern. Arch. of Photogr. and Rem. Sens.*, volume 31, Part B3, pages 832–838, 1996.

[22] U. Stilla, F. Quint, and M. Sties. Analyse von Luft- und Satellitenbildern zur automatischen Ermittlung der Bodenversiegelung städtischer Siedlungsbereiche. Zwischenbericht II, FIM-FGAN/IPF-Universität, Ettlingen/Karlsruhe, 1995.