geoAIDA - A Knowledge Based Automatic Image Data Analyser for Remote Sensing Data

J. Bückner, M. Pahl, O. Stahlhut, C.-E. Liedtke
Institut für Theoretische Nachrichtentechnik
und Informationsverarbeitung
Hannover, Germany
geoaida@tnt.uni-hannover.de

This contribution describes a new approach towards an implementation of a semantic net with holistic sequential control in the field of remote sensing image data. A priori *knowledge about the scene, sensors and operators is represented explicitly by the semantic net. The system offers a flexible framework for the integration of a wide difference of image analysis operators and allows the interpretation of multisensor imagery. Efficiency and certainty are increased through incorporation of a geographic information system. Results of a region classification using multisensor input data will be shown.*

1 Introduction

Environmental monitoring tasks based on remote sensing images, especially the update of geographic information data, become more and more important. Due to the increasing number of such images and the labour-intensive manual evaluation there is much need for productive and robust techniques for automatic analysis and object extraction. Looking at low-level image segmenting algorithms they deliver incomplete, fragmented and erroneous information which has to be verified and validated by a high-level framework.

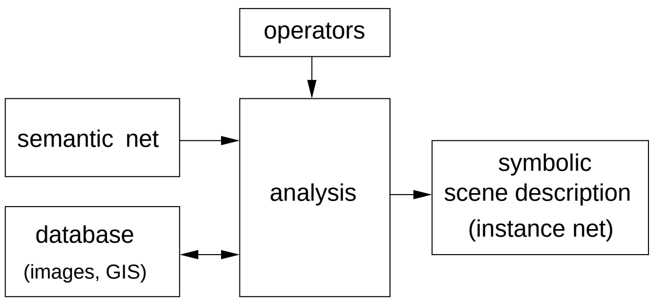

Figure 1: System components GEOAIDA

One possible way to obtain a high-level scene description is the use of a semantic net to model and evaluate *a priori* knowledge about the objects expected to be found in the scene. The semantic net describes structural relationships between the objects and the components the objects are made of. The interpretation process can make use of varying object properties seen in different data sources such as laserscan, radar (SAR), optical and thermal sensors (VIS/IR). Thereby the relationships between the image contents are exploited to gather additional information which can't be extracted from a single sensor alone. Available partial interpretations of the image such as data of a geographic information system (GIS) can be used to conduct and accelerate the interpretation.

Literature provides several references for the application of semantic nets for image interpretation such as ERNEST [Niemann *et al.*, 1990] [Kummert *et al.*, 1993] and AIDA [Tönjes *et al.*, 1999]. From the viewpoint of history the latter is the predecessor of the system described here, though a complete redesign was made.

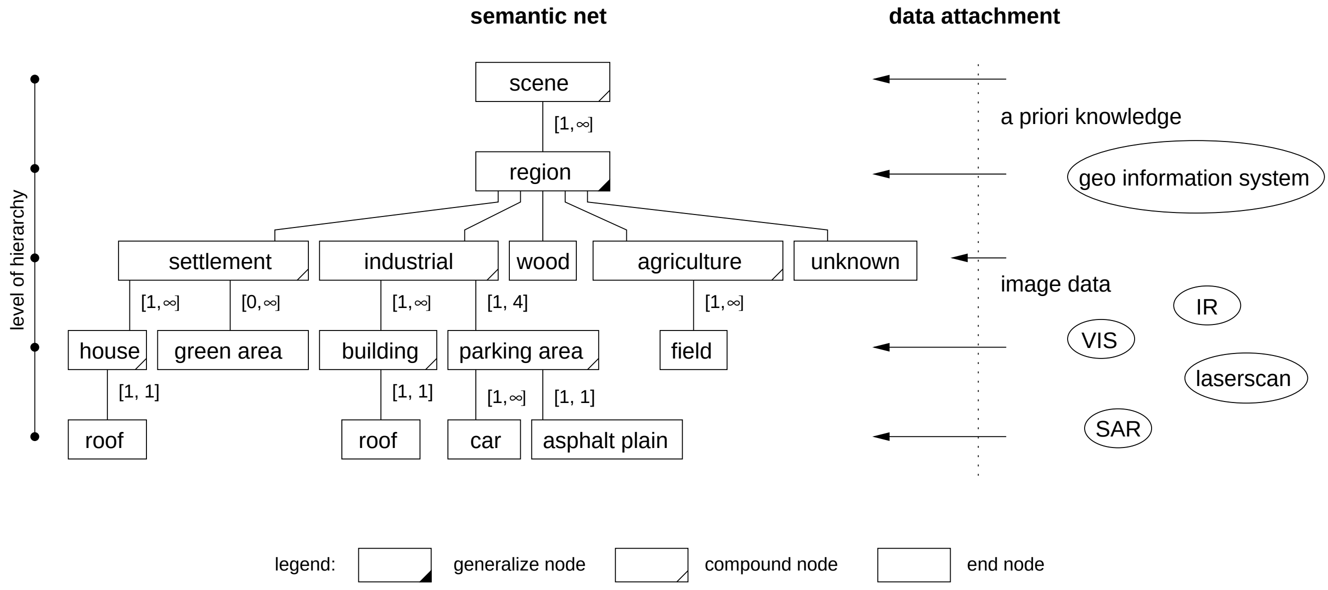

Figure 2: Example for the modelling of a scene

2 Knowledge-Based Interpretation System

For automatic interpretation of images the knowledge-based system GEOAIDA was developed. AIDA is short for **A**utomatic **I**mage **D**ata **A**nalyser. GEOAIDA separates the knowledge represented explicitly as semantic net from the control of the analysis process. Multi-sensor imagery, supplement knowledge stored in a GIS and intermediate analysis results are administered in a database. The analysis provides interpretation results as a symbolic scene description, as shown in figure 1.

In general a semantic net consists of nodes which are linked by edges. Semantic nets are directed acyclic graphs. The nodes represent objects which are expected to be found in the scene, the edges describe the relations between the nodes. An initial description of the expected scene content including nodes and edges is called a concept net. In practice it showed up that a hierarchical arrangement of the nodes and edges is suitable for the description of the objects in remote sensing data. Such a semantic net, which will serve as an example later, is shown in figure 2.

During the interpretation of the image, hypotheses are constructed on the existence of particular objects. These hypotheses are represented by nodes. In case of validation acknowledgement nodes, so called instances, are generated. Thus in the course of the interpretation process a symbolic description of the scene contents is formed.

GEOAIDA utilizes two different types of nodes. The generalization node (comparable to logic *xor*) is used to split up and branch into alternative interpretations of a scene - in figure 2 "region" is a node of this type. The compound node (comparable to logic *and*) represents objects which are recognized by the existence of their components. For example it is necessary to find at least one or more buildings and one to four parking areas to validate the node "industrial", as shown in figure 2. In opposite to other image analysis systems using semantic nets GEOAIDA doesn't put any special functionality into the edges. Edges are merely bidirectional links between two nodes.

Every node "knows" his parents and children, carries information about the object it represents and possesses dynamic administered attributes and methods. An important attribute which is also a part of the knowledge base is an interval for the number of possible instance nodes. For example it can be assumed, that a road junction consists of three to six intersecting roads. The child node "street" of a node "road junction" would obtain an interval attribute for the expected node instances with the limits $[3, 6]$. Optional objects such as "green area" in figure 2 are modelled by an attribute with the interval $[0, \infty]$.

3 Structural vs. Holistic Interpretation

Former image interpretation systems based on semantic nets [Niemann *et al.*, 1990] [Tönjes *et al.*, 1999] propagated only a purely structural analysis. For example they tried to recognize a settlement by detecting houses, which again were identified by walls and roofs, which were made up by lines and polygons. The connection to the image data was accomplished in the lowest level of the net, i.e. by operators delivering geometric primitives with additional attributes extracted from the image data.

A substantial deficiency of structural analysis is the fact, that more high grade information which could speed up finding a (partial) interpretation cannot be flexibly attached to the system.

Therefore GEOAIDA introduces a so-called "Holistic Operator" that analyses a scene the holistic way. In literature, different approaches to utilization of holistic principles can be found, e.g. in [Kummert *et al.*, 1998] or [Büker *et al.*, 2000] - the latter is driven by a different motivation. The topic is imbedding into autonomously mobile systems with concern to the aspects of Active Vision and learning of structures.

Inside GEOAIDA any node can bind a holistic operator. An operator for the recognition of settlements could be the evaluation of aerial images with affiliated laserscan data. The operator searches for typical height signatures of houses and performs a clustering of the results. A structural analysis which searches for the polygons of roofs and lines of accommodation ways would have to be used only in situations where the holistic analysis cannot make a unique predication.

The binding to the image data is no longer made strictly in the lowest level of the semantic net, but by the holistic operator at any node. Therefore the data sources are conceptually put aside the net, see figure 2.

4 Sequence of the Scene Analysis

In the process of image interpretation the semantic concept net is transferred into an instance net, that represents the objects of the scene and their connotation. During the analysis a concept node K produces an instance node $I(K)$ with the following possible states: hypothesis $I_H(K)$, missing instance $I_M(K)$, partial instance $I_P(K)$, complete instance $I_V(K)$, complete holistic instance $I_V^H(K)$. Starting point is the hierarchically highest node of the concept net, which is initialized with the complete to be analysed scene. This concept node creates the root instance node in the state $I_H(K)$. The following state transitions are possible:

- $I_H(K) \longrightarrow I_P(K)$, once a child node is in the state $I_V(K)$, i.e. it is completely instantiated.
- $I_P(K) \longrightarrow I_V(K)$, once all obligate child nodes are in the state $I_V(K)$.
- $I_H(K)$ or $I_P(K) \longrightarrow I_M(K)$, if at least one obligate child node is in the state $I_M(K)$, respectively the hypothesis cannot be verified by means of the image data.
- $I_H(K) \longrightarrow I_V^H(K)$, after the holistic operator of the node verified the hypothesis by means of the image data.

The analysis is finished as soon as a complete instantiation of the goal concept, i.e. the root node, is achieved.

The interpretation strategy rests upon fixed control rules of the analysis interpreter. These rules include the automatic call of holistic operators during the initial top-down progression through the semantic concept net, the creation of new instance nodes according to the concept, the binding of image data to the necessary (low-level) image processing algorithms, the acknowledgement of instance nodes during the bottom-up progression of the net, and many more.

Every possible interpretation of an investigated region is stored as search-tree node including all values generated during interpretation. Competing interpretations lead to a split of the search tree - the leaves are judged and an A*-algorithm [Winston, 1989] selects the most promising path. Input values for the judgement and ranking of the interpretation paths are the probability attributes (confidence values) [Dempster, 1967] calculated in the course of the analysis.

Results of classification are always afflicted with uncertainty and inaccuracy. Thus the selection of competing partial interpretations requires valuation. Possible valuation methods for the interpretation results can be the application of Bayes nets as shown in AIDA [Growe, 2000] or "Possibility Theory" [Dubois, Prade, 1988]. Both methods allow to model uncertainty and inaccuracy numerically. GEOAIDA integrates the basic ideas of "Possibility Theory" as a ruleset for the valuation and selection process.

Figure 3: Aerial Image

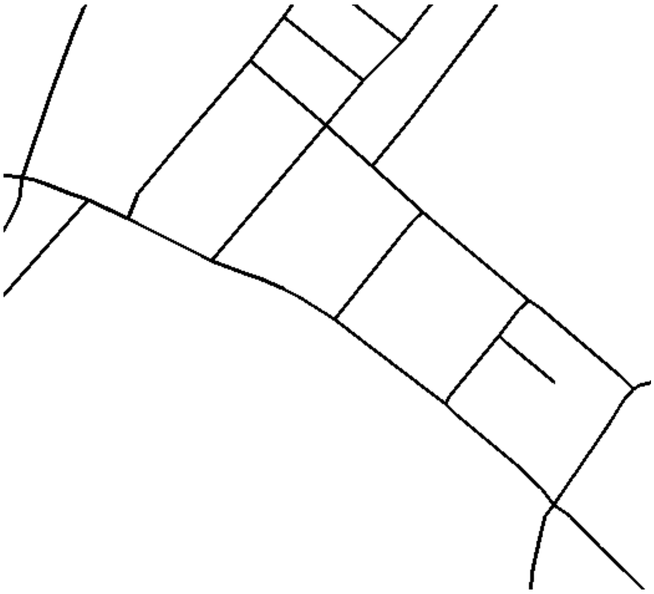

Figure 4: Streets from GIS Data

5 Analysis of Remote Sensing Data

Referring to the semantic net as shown in figure 2, a detailed overview of the analysis process will be now given. Figure 3 shows a part of an aerial image in the visual wavelength spectrum which is to be interpreted. The scene and the according geographic information and laserscan data reside in the GEOAIDA database.

The analysis interpreter is initialized with the geographic coordinates of the scene and a description of available sensors. This leads to creation of a root instance node "scene" in the state hypothesis. The aim of the interpretation is to completely instantiate this node. "Scene" consists of different regions, therefore the semantic net lists a node "region" which generates an instance "region" in the state hypothesis as a result of the initial top-down progression through the concept net.

At this point a holistic operator is used for the first time. "Region" uses the previous knowledge of the streets positions and directions stored in the geographic information system (see figure 4) to split up the scene into areas which are bounded by broader streets. The label image of this initial region separation is shown in figure 5.

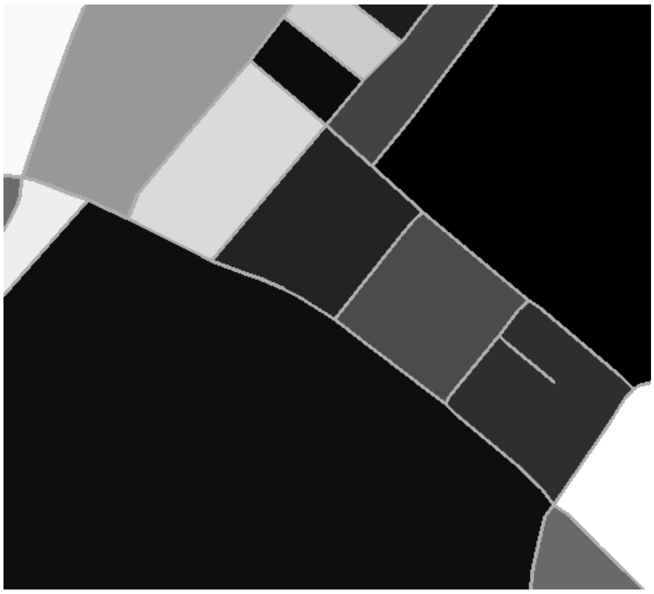

Figure 5: Initial Region Separation

Each of these sixteen new regions is equivalently treated and creates hypothetic instances of the child nodes of "region" for determination of the regions' type, see figure 2. The type is tested for settlement, industrial area, wood and agricultural land. All regions that cannot be classified during the interpretation process are assigned to the rejection class "Unknown".

The holistic operators of the nodes "settlement" and "industrial" use laserscan data of the scene (see figure 6) to perform a special segmentation of building structures.

Figure 6: Laserscan data

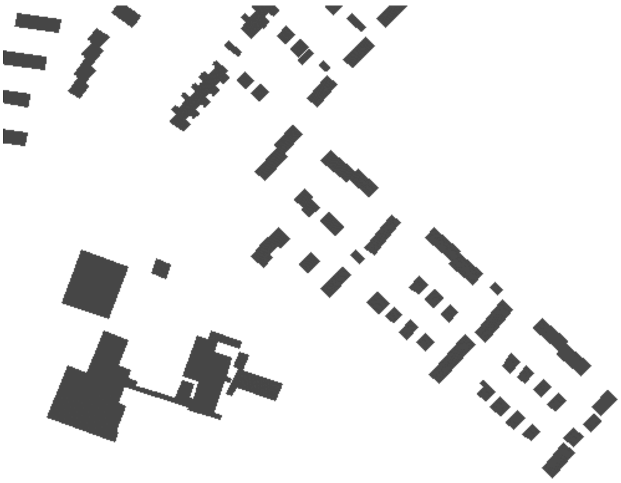

Figure 7: Segmented houses

The operators use results of a low-level segmentation process (see figure 7) together with general assumptions about buildings' shapes to obtain a classification and confidence values for this disposition.

If the holistic analysis doesn't produce any results, for example because of missing laserscan data, the structural interpretation is continued. The instance nodes of the type "settlement" then create hypothetic instances for their child nodes "house" and "green area", which are tried to be detected by holistic or even further structural analysis. At some point of the top-down progression through the semantic net end nodes are reached, which are completely instantiated or set to missing instance by texture analysis of the sensor images [Rosenfeld, 1976] or by extraction of simple geometric objects produced by a low-level segmentation. In particular the nodes "wood", "agriculture", "green area", "roof" and "asphalt plain" utilize texture analysis operators.

As described in section 4 any completely instantiated node leads to partial instantiation of its parent node as a result of the bottom-up progression. Once all child nodes are completely instantiated the parent node also changes its state to complete instance. If the parent is a generalization node it has to evaluate the interpretation results of its child nodes and splits up if there are competing classifications. An example for this behaviour is the division of the large region seen in figure 5 (left bottom) into three subregions industrial area, wood and agricultural land, because all partial interpretations propagated from bottom to top are valid for certain areas of the region and concurrently they share no common area. If different interpretations for identical areas exist the generalization node has to decide according to the confidence values which partial interpretation is going to be validated and passed on to the bottom-up progression.

The result of the scene interpretation is shown in figure 8. The system classified the scene according to the know ledge base modelled by the semantic concept net. The figure shows a **s**ettlement consisting of several regions, an **i**ndustrial area, **a**gricultural land, **w**ood and two regions classified as "**u**nknown". The delimitation of the wood areas was extracted exclusively from the texture features. The boundary region of the agricultural land was extended up to the neighbouring regions (grey zones along the black field borders). The delimitation of industrial areas and settlements was realized by the initially selected region splitting with streets, which leads to classification errors. For example the little piece of wood and the field seen bottom right in the figure are surrounded by a seam with allocation to the industrial area. Possible improvements for the operational application can be achieved by more sophisticated region-grow algorithms in conjunction with the building segmentation and better description of the neighbourhood relations for the region splitting done by the generalization node.

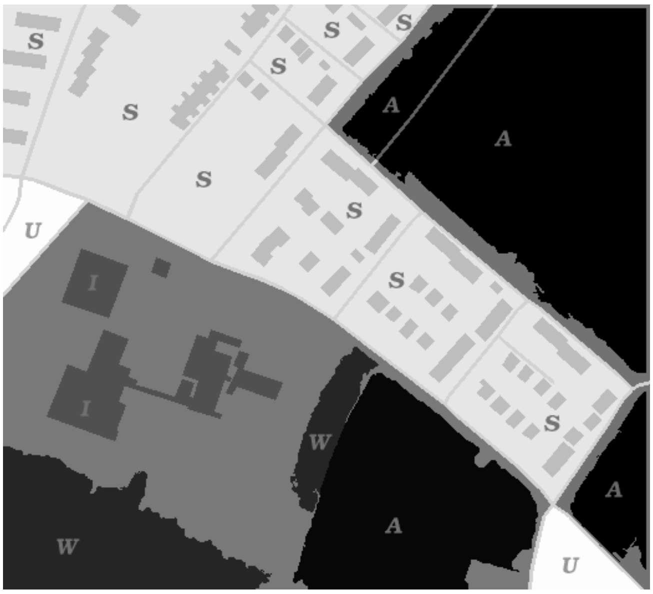

Figure 8: Results of the Interpretation

6 Conclusion

The knowledge-based automatic image interpretation system GEOAIDA was presented and the functionality and modes of operation were described using an example from the field of region classification. GEOAIDA utilizes a semantic net for the representation of scene knowledge, a rule based analysis interpreter, with evaluation based on possibility theory and segmentation algorithms for the interpretation of images. The interpretation transfers the observed scene into a symbolic description in the shape of an instance net.
Due to its capability of combining holistic and structural analysis within a semantic net, the system can be used as a flexible tool for the classification of objects and regions using multisensor remote sensing imagery. By adding external, high grade image segmentation methods via holistic operators, and because of the possibility of modelling scene features pre-analytically by linking geographic information systems, the areas of application are vastly extended, comparing GEOAIDA to a system that works purely structurally. Also, the procedure of designing the knowledge base is simplified. GEOAIDA is a promising approach towards production in the field of image analysis and verification of geographic information systems.

References

[Büker *et al.*, 2000] U. Büker, S. Drüe, N. Götze, G. Hartmann, B. Kalkreuter, R. Stemmer, R. Trapp, "Vison Based Control of an Autonomous Disassembly Station", Robotics and Autonomous Systems (Elsevier), 2000.

[Dempster, 1967] A. Dempster, "Upper and lower probabilities, inferences based on a sample from a finite univariate population", *Biometrika*, Vol. 55, 1967, pp. 515-528.

[Dubois, Prade, 1988] D. Dubois and H. Prade, "Possibility Theory: An Approach to Computerized Processing of Uncertainty", Plenum Press, New York and London 1988, p. 263

[Growe, 2000] S. Growe, T. Schröder, C.-E. Liedtke, "Use of Bayesian Networks as Judgement Calculus in a Knowledge-Based Image Interpretation System", 19th Congress of the Intl. Society of Photogrammetry and Remote Sensing, Amsterdam, July 2000

[Kummert *et al.*, 1993] F. Kummert, H. Niemann, R. Prechtel and G. Sagerer, 1993, "Control and explanation in a signal understanding environment", SIGNAL PROCESSING, Vol. 32, No. 1-2, May 1993

[Kummert *et al.*, 1998] F. Kummert, G. A. Fink, G. Sagerer and E. Braun, "Hybrid object recognition in image sequences", 14th International Conference on Pattern Recognition, volume II, pp. 1165-1170, Brisbane, 1998.

[Niemann *et al.*, 1990] H. Niemann, G. Sagerer, S. Schröder, F. Kummert, "ERNEST: A Semantic Network System for Pattern Understanding", IEEE Trans. on Pattern Analysis and Machine Intelligence, 12(9):883-905, 1990

[Rosenfeld, 1976] Rosenfeld; *Digital Picture Procesing*, Academic Press, New York, 1976

[Tönjes *et al.*, 1999] R. Tönjes, S. Growe, J. Bückner and C.-E. Liedtke, "Knowledge-Based Interpretation of Remote Sensing Images Using Semantic Nets", Photogrammetric Engineering and Remote Sensing, Vol. 65, No. 7, July 1999, pp. 811-821

[Winston, 1989] Winston, Patrick Henry, "Artificial Intelligence", Massachussetts, Addison-Wesley, 1989