Use of Explicit Knowledge and GIS Data for the 3D Evaluation of Remote Sensing Images

Stefan Growe, Ralf Tönjes
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
University of Hannover, Appelstrasse 9A, D–30167 Hannover, Germany
E–Mail: growe@tnt.uni–hannover.de, toenjes@tnt.uni–hannover.de

Abstract

The evaluation of 3D scenes observed from different sensors requires the co–registration of sensor images and the reconstruction of the 3D geometry. To solve both tasks the presented system exploits prior knowledge, represented explicitly by semantic nets, and uses a digital landscape model of a geoinformation system (GIS) as a hint for the object location. This is shown for the detection of control points for image registration and the extraction of objects (roads, buildings) for 3D reconstruction. For realtime visualization the 3D geometry is approximated by a polygon mesh with overlaid photo texture.

1. Introduction

Environmental and agricultural monitoring represents a major topic of remote sensing. Earth scientists, environmental researchers, and civil engineers ask for tools that help to evaluate the large amount of data. These days the images of visual, infrared, and radar (SAR) sensors from the same observation area are evaluated in parallel by displaying them on different monitors together with a map. Due to different image scales and orientations it is difficult for the human interpreter to explain the image content or to recognize corresponding structures.

To compare the different image data directly on one screen a perfect co–registration is required. Here, the images are registered, i.e. geocoded, in a common geographic coordinate system. To reveal the 3D geometry of the observed scene a 3D terrain model has to be reconstructed from the data which is finally visualized on a graphics computer for a three dimensional evaluation.

Automation of the processing pipeline shown in Figure 1 suffers from two obstacles: Firstly, co–registration relies on the accurate extraction of control points. Secondly, the reconstruction of the 3D geometry from 2D images is underconstrained. Stereo matching techniques provide only incomplete and erroneous height data. To tackle these two problems prior knowledge about the structure of the objects in the scene is used to select appropriate control point structures, e.g. crossroads, and to constrain the geometry of objects like roads and buildings in order to derive a more realistic 3D reconstruction. To exploit this knowledge an image interpretation is required that assigns a semantic like road or building to the data. The presented system exploits a digital landscape model (DLM) of a geoinformation system (GIS, here: German ATKIS DLM 25) as partial interpretation which provides both, semantics and geometry of the included object classes.

Several knowledge based image interpretation systems have been presented in the past. For the representation of the scene knowledge rules are often employed like in SPAM [1], SIGMA [2] or MESSIE [3]. ERNEST [4] uses semantic nets to exploit the object structure for interpretation. The presented system AIDA [5] adopts the idea to formulate prior knowledge about the scene objects with semantic nets. In addition the control knowledge is represented explicitly by rules.

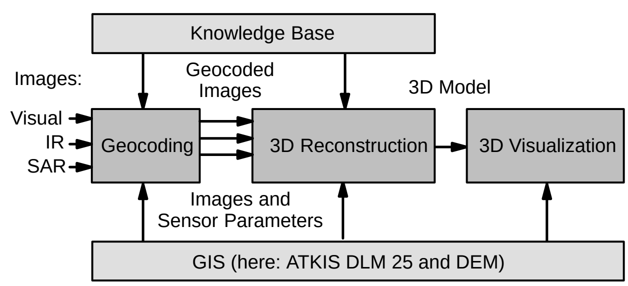

Figure 1: Processing Pipeline for the 3D evaluation of remote sensing data from multiple sensors

2. The Knowledge Based Approach

The prior knowledge used for the image interpretation can be divided into three types. The **3D scene domain knowledge** describes the semantics of the objects (e.g. road), the geometry (e.g. 3D stripe), and the material with its spectral reflectance (e.g. asphalt). The relationships between them decompose objects into their components or constrain the topology. (Crossroads are composed of roads. Buildings are sited close to roads.) The **2D image domain knowledge** contains the geometric and photometric properties of the 2D regions, lines, and points in the image and their relationships. Finally the **sensor knowledge** explains the mapping between the 3D scene domain and the 2D image domain. The transformation of visual or infrared cameras is described by a perspective projection. The SAR sensor is modelled by a

Published in the 1998 International Conference on Pattern Recognition (ICPR'98), 17–20 August, 1998, Brisbane, Australia

polynomial approximation of the range and doppler equations [6]. The geocoded images conform to a parallel projection. The sensor sensitivity describes the mapping of the material dependent reflection properties into image intensities.

In AIDA [5] this prior knowledge is represented by semantic nets. These nets consist of nodes describing objects and links in between (see fig. 2). The decomposition of objects in their parts is indicated by the *part–of* link. Thus the detection of a complex structure is simplified to the search for its parts. The transformation of an abstract object to its more concrete realization is represented by the concrete–of link, abbreviated *con–of*. The specialization of an object is described by the *is–a* link introducing the concept of inheritance. The *data–of* link establishes a relation to the features segmented in the image data or contained in the GIS. Relevant object properties are modelled by attributes, like for example the width of a road.

Problem independent rules exploit the knowledge to generate successively hypotheses for object parts, concrete realizations, etc. and verify them in the image data. The final scene description contains a node for each road found in the image. Competing interpretations are judged and an A*–algorithm selects the most promising interpretation for further investigation.

3. Geocoding

The *geocoding* process splits into three steps. At first corresponding points in the image and in the map have to be detected. Based on these control points the parameters of a sensor specific mapping model are estimated describing the exterior sensor orientation. Finally the calculated mapping model is used to resample a geocoded image. While the last two steps are automated already the control points are mostly defined manually. We developed a knowledge based approach based on AIDA for the automatic control point detection using GIS data [7].

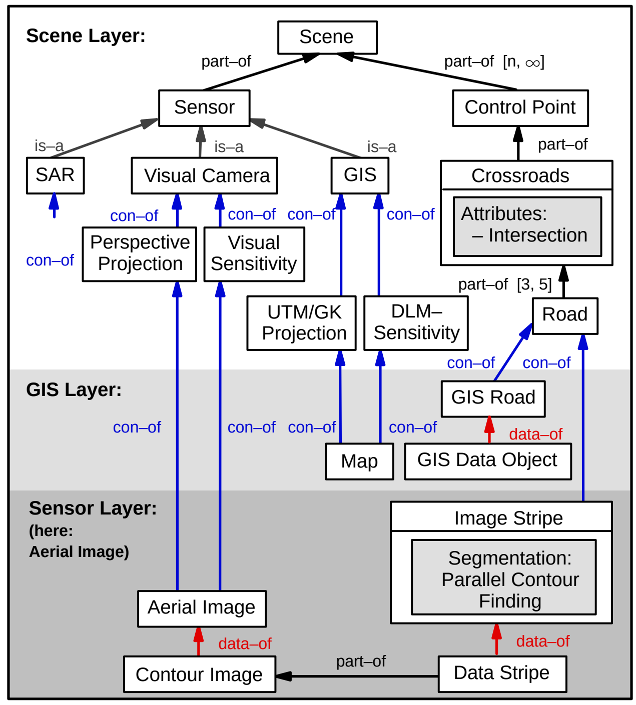

Figure 2: Simplified semantic net for the automatic detection of control points (here: crossroads)

Figure 2 shows a simplified semantic net representing the required prior knowledge. Beside the *scene layer* and the *sensor layer* (here: aerial image) the *GIS layer* is used to represent the GIS data. Crossroads consisting of three to five intersecting roads are used as control points. Each of these roads has a realization in the GIS database and in the image as an *Image Stripe* segmented by road detection algorithms. During interpretation the system extracts roads from the GIS, groups them to crossroads and matches them with the road net in the image. Results for a SAR image are shown in fig. 3.

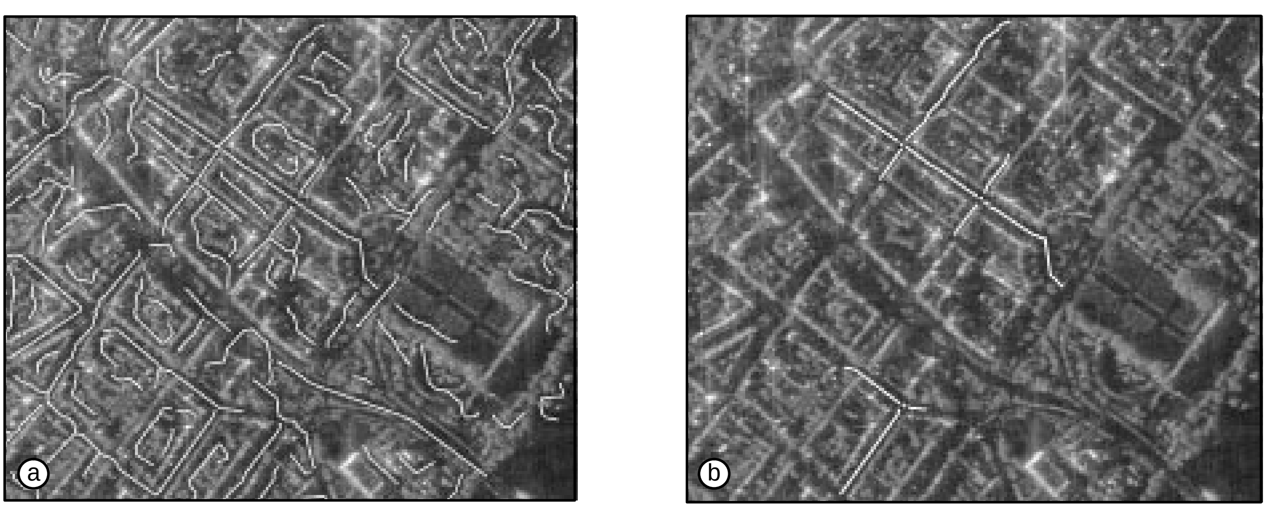

Figure 3: (a) Segmented roads and (b) selected crossroads with interpolated control points in a SAR image

Published in the 1998 International Conference on Pattern Recognition (ICPR'98), 17–20 August, 1998, Brisbane, Australia

4. 3D Reconstruction

For a 3D evaluation of the remote sensing data a realistic reconstruction of the landscape is required. But existing DEMs contain only the height at ground level and ignore vegetation and buildings. Thus a height map is computed by matching corresponding points of a stereo image pair [8]. However, the height map is erroneous and incomplete. To overcome this problem model expectations are employed to constrain the object geometry [9]. The prior knowledge suggests smoothness criteria for the course of roads and geometrical constraints for buildings. Figure 4 shows a part of the used knowledge base.

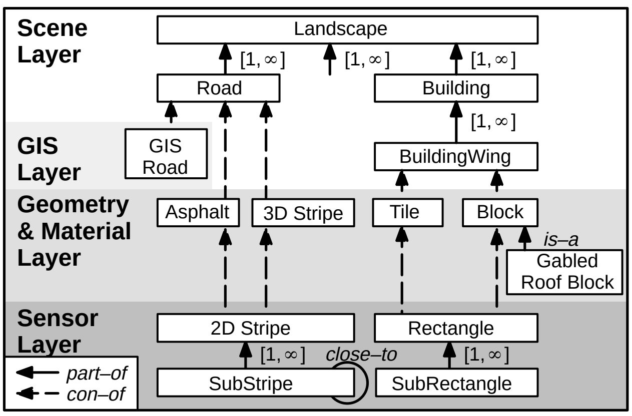

Figure 4: Simplified semantic net representing the generic landscape model

The roads extracted from the GIS are matched with the segmented image stripes as described in chapter 3. The prior knowledge is used to model a continuous course of the road in 3D.

The buildings are composed of several building wings which are searched for one after the other generating hypotheses for blocks, rectangles, and sub rectangles. The initial length, width, and height of a block is computed from the height map and adjusted consecutively to the grey level edges in the image. The analysis tries to specialize the block to a gabled roof block and selects the better approximation. Figure 5b shows the resulting building composed of a gabled roof and a flat roof building wing.

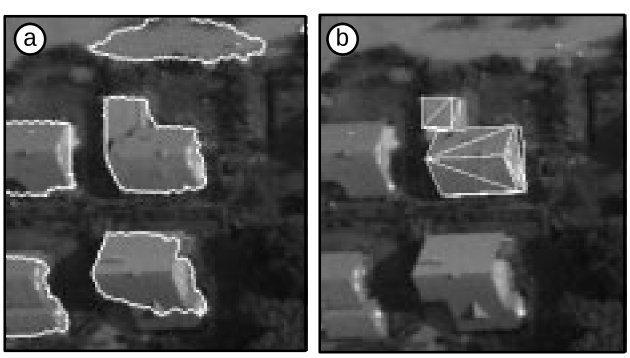

Figure 5: (a) Segmentation, (b) reconstructed building

5. 3D Visualization

The 3D landscape model consisting of the digital terrain model textured with the geocoded images is visualized stereoscopically and in realtime to the human interpreter. Therefore a high–end graphic engine renders two views for the left and right eye respectively. For the 3D evaluation of the data we developed a graphical user interface [10]. Beside the realtime navigation in the virtual landscape it allows separate activation and blending of the different textures. The user is able to mark interactively regions of interest within the 3D model to create arbitrary image mosaics. An on–line access to the GIS allows the visualization of GIS data for a selected region.

6. Conclusions

A system for the 3D evaluation of remote sensing images from multiple sensors was presented. The images are geocoded, projected onto a 3D terrain model of the observed scene and visualized stereoscopically. The approach benefits from two things: general knowledge about the objects and scene specific GIS data. The image interpretation system AIDA exploits both to detect control points, here crossroads, for the geocoding and to constrain the parameters for the realistic 3D reconstruction of the terrain surface. During the interactive evaluation the user can manipulate the 3D model and visualize additional GIS information.

7. Literature

[1] McKeown, et al., *Rule–Based Interpretation of Aerial Imagery*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI–7, No. 5, pp. 570–585, Sept. 1985

[2] Matsuyama, T., Hwang, V.S.–S., *SIGMA : A Knowledge–Based Aerial Image Understanding System*, Plenum Press, New York, 1990

[3] Clement, V., Giraudon, G., Houzelle, S. , Sadakly, F., *Interpretation of Remotely Sensed Images in a Context of Multisensor Fusion Using a Multispecialist Architecture*, IEEE Trans. on Geoscience and Remote Sensing, Vol. 31, No 4, July 1993

[4] Niemann, H., Sagerer, G., Schröder, S., Kummert, F., *ERNEST: A Semantic Network System for Pattern Understanding*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 12, No. 9, September 1990.

[5] Liedtke, C.–E., Bückner, J., Grau O., Growe S., Tönjes, R., *AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data*, 3. Int. Airborne Remote Sensing Conference & Exhibition, Copenhagen, pp. 313–320, July 1997

[6] Raggam, J., Strobl, D., Hummelbrunner, W., *Product Quality Enhancement and Quality Evaluation*, in: Schreier, G.: *SAR Geocoding: Data and Systems*, Wichmann, Karlsruhe, 1993, pp. 187–206

[7] Growe, S., Tönjes, R., *A Knowledge Based Approach to Automatic Image Registration*, Proc. of International Conference on Image Processing (ICIP'97), October 1997, Santa Barbara

[8] Falkenhagen, L., *3D Object–Based Depth Estimation from Stereoscopic Image Sequences*, International Workshop on stereoscopic and three–dimensional imaging, September 6–8 1995, Fera Congress Center, Santorini, Greece, 1995

[9] Tönjes, R., *3D Reconstruction of Objects from Aerial Images using a GIS*, ISPRS Joint Workshop on "Theoretical and Practical Aspects of Surface Reconstruction and 3–D Object Extraction", Haifa, Israel, September 9–11, 1997

[10] Growe, S., Schulze, P., Tönjes, R., *3D Visualization and Evaluation of Remote Sensing Data*, Computer Graphics International CGI'98, Hannover, June 22–26 1998

Published in the 1998 International Conference on Pattern Recognition (ICPR'98), 17–20 August, 1998, Brisbane, Australia