A Knowledge Based Approach to Automatic Image Registration

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

Abstract

The presented work addresses the problem of automatic control point matching for the registration of remotely sensed images. The inaccuracy of flight parameters and the sensor specific appearance of objects are the difficulties automatic registration suffers from. To overcome these problems the presented system uses prior knowledge to select appropriate structures for matching, i.e. control points, from a GIS and to extract their corresponding features from the sensor data. The knowledge is represented explicitly using semantic nets and rules. The best correspondence between the GIS data and the image is found by an A–Algorithm. The automatic control point matching is demonstrated for crossroads in aerial and SAR imagery.*

1. Introduction

The evaluation of remotely sensed images from multiple sensors requires the registration of all data in a common (geographic) coordinate system. This is especially true for multiple sensors that differ in geometry, spectrum, and time. Prerequisite for the image–to–map registration of remotely sensed images is the detection of corresponding points in the image and the map. The common approach uses manual control point matching. But the quantity of data and the short update periods ask for methods that automate the registration.

For precise registration the control points must be distributed equally and detected accurately. The different sensor specific appearances of the control points in the image data and the map require the segmentation of common features. The increasing availability of geoinformation systems (GIS) and digital landscape models (DLM) eases this task making the segmentation of maps abundant. The presented work exploits the German digital landscape model DLM 25 of ATKIS (Authoritative Topographic and Cartographic Information System) which corresponds to the contents of the 1:25000 map. ATKIS provides both semantics and geometry of the represented object classes.

The features used for control point matching have to be contained in the used digital landscape model and should be eminent in the sensor data to allow robust segmentation. The demand for equally distributed control points suggests the use of frequently present features. In the literature various features for control and tie point matching have been suggested like rivers, coastlines, roads, fields, or even manhole covers [1]. The extraction of feature points using a Gabor wavelet model for detecting local curvature discontinuities is proposed in [2].

The presented work uses roads to define control points by crossroads. However the system is not limited to linear shaped features. It could also handle areal features using for example the center of gravity as control point. The flight parameters taken from GPS and INS give an initial estimation of the sensor orientation. Nevertheless the orientation is inaccurate. The risk for matching wrong features is very high. To reduce the risk the matching must take structural relationships into account.

To exploit the relationships between the features for example relaxation labeling [3], and relational matching [4] have been suggested. In [5] a structural matching of polygonal features like buildings is proposed.

Here a knowledge based approach is used that provides explicitly a generic description of the structure used for matching. A further advantage is that this approach can also model the relationship between the objects and their expected appearance in the image data.

2. System Overview

2.1. Image Interpretation System AIDA

The task of feature extraction and matching is controlled by the image interpretation system AIDA [6] (Automatic Image Data Analyzer) that provides methods for explicit knowledge representation. The knowledge about object structure and their appearance in the image is represented by semantic nets (similar to ERNEST [7]). Rules exploit the knowledge to match the features in the GIS with their correspondences in the image. In AIDA the rules are problem independent and exploit only the syntax of the semantic net.

The input data for registration consists of the image to be registered, flight parameters, interior sensor parameters,

Published in the 1997 International Conference on Image Processing (ICIP'97), scheduled for October 26–29, 1997 in Santa Barbara, CA, USA

and a GIS. Matching proceeds in two steps. At first features are selected from the GIS to establish an initial scene description thus constraining position, radiometric and geometric appearance, and structure of the control points expected in the image. Consecutively the hint from the GIS is used to extract corresponding control point candidates from the image. The relationships between the features are exploited to increase or decrease the reliability of competing control point candidates. The best mapping between the features in the image and the features in the landscape model is found using graph search methods.

2.2 Knowledge Representation

Semantic Net: The knowledge about the object structure belonging to the control point and its relationship to the sensor specific appearance can be represented efficiently by semantic nets. Semantic nets consist of nodes and edges in–between (see fig. 1). The nodes are implemented as frames which contain a collection of attributes, like the width of a road or the intersection of crossroads. Further the object has methods, i.e. functions, at its disposal to compute the attribute values. There may also be a method available to segment the object in the image data, for example a contour finding algorithm to extract stripes from images. The edges or links of the semantic net form the relations between the objects.

Figure 1 shows a simplified semantic net for registration. The knowledge base distinguishes three conceptual layers. The scene layer describes the scene specific semantic. The GIS layer contains the objects as represented in the GIS. The bottom layer is sensor related and describes the sensor specific radiometric and geometric appearance of the objects. According to the chosen sensor (visual, infrared, or SAR) a specific sensor layer is used. Each layer possesses a common appropriate vocabulary. For instance the scene layer uses a cartographic coordinate system (e.g. German Gauss–Krueger or UTM) while the sensor coordinate system is referred to the image raster. The layer specific level of abstraction is modelled by the concrete–of link, abbreviated *con–of*. The decomposition of an object into its parts is described by the *part–of* link. Here the control point consists of a crossroads which is composed of intersecting roads. The *data–of* link establishes a relation to the features segmented in the image data or contained in the GIS.

The syntax of the semantic net distinguishes between two types of nodes: concepts and instances. Concepts describe the generic model of the objects. The instances $I(n)$ are realizations of the concept n in the observed scene. Thus they are related via *instance–of* to their concepts. During interpretation the state of the instances changes from hypothesis $I_H(n)$ to complete instance $I_C(n)$ or missing instance $I_M(n)$. Counters in the links of the concept net state the

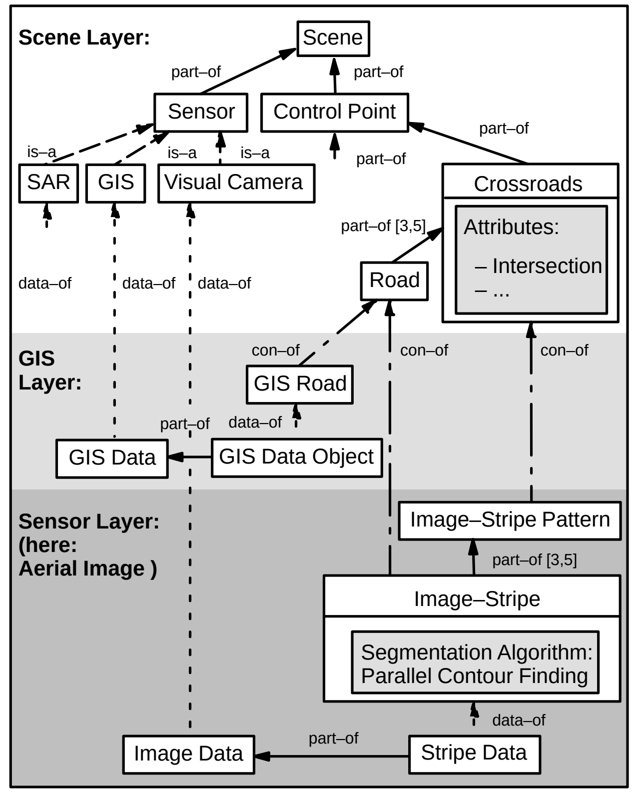

Fig. 1: Simplified semantic net for automatic search of control points in aerial images

minimum and maximum number of links required in the instance net.

Rules: To make use of the knowledge represented in the semantic net procedural knowledge for control is required that states how and in which order scene analysis has to proceed. The control knowledge is represented explicitly by a set of rules. A rule is composed of a condition and an action part. The condition checks for a new interpretation state of neighboured nodes in the semantic net. The action part adapts the interpretation state of the focused node. Furthermore the attributes are restricted top–down by new expectations and bottom–up by new measurements.

For example a hypothesis $I_H(n)$ can be denoted as complete instance $I_C(n)$, if all obligatory subnodes have been instantiated completely. This can be expressed by the following rule:

Rule–complete–instantiation:

CONDITION: $I_H(n) \wedge I_C(m)\ \forall m \in \mathcal{M}$
$\mathcal{M} = \{m \mid m = \text{part–of}(n) \vee m = \text{con–of}(n)\}$

ACTION: Change status to $I_C(n)$,
compute bottom–up all attributes,
return $I_C(n)$.

3. Automatic Search of Control Points

As mentioned before, the coarse orientation of the remote sensing data is known from flight parameters. The exact image–to–map registration is performed in a block-wise manner. The registration searches for one control point, i.e. crossroads, in each block of the image to be registered by scene interpretation.

In a first step the semantic net is initialized by loading the knowledge base represented by a semantic net (see fig. 1) and a set of instantiation rules. According to the image data (aerial image or SAR data) the suitable sensor is instantiated. Since the GIS provides the most reliable data the aerial image/SAR layer is neglected at first and the scene analysis exploits only the GIS data. As crossroads themselves are not registered in the DLM roads have to be grouped to crossroads by the interpretation system.

Because the goal is to instantiate a control point the initial hypothesis I_H*(Scene)* generates a hypothesis I_H*(Control Point)*. The condition for the complete instantiation is not fulfilled because the obligatory part *Crossroads* is not present. Hence the rule for top–down propagation of hypotheses is executed generating I_H*(Crossroads)*. The rule is fired repetitively until the initial concept *GIS–Road* is reached resulting in the hypothesis I_H*(GIS Road)*. The GIS database is asked for roads within the current image region. The objects returned are added to the semantic net as *Ic(GIS Data Object)*. An appropriate data node is chosen to verify a complete instance I_C*(GIS Road)* and to compute its attributes like width in meters, material, and 3D coordinates. The new information is propagated consecutively bottom–up instantiating I_C*(Road)*. Because crossroads consist of at least three intersecting roads a second hypothesis I_H*(Road)* is generated. From the 3D coordinates of the first road an estimation for the position of the second one can be computed resulting finally in the instantiation of an adjacent road. The process is repeated until no further road belonging to the crossroads is found in the GIS data. The instances I_C*(Road)* are grouped, complete instances I_C*(Crossroads)* and I_C*(Control Point)* are generated and the 3D coordinate of the control point is computed from the intersection of all roads.

The found scene consisting of a crossroads and several roads has to be validated in the image to find the corresponding 2D coordinates of the objects. Therefore the knowledge base is extended by the sensor layer, here the aerial image layer. A hypothesis I_H*(Image–Stripe)* is instantiated top–down to search for the pictorial representation of the first road. From the coarse orientation of the image we can compute an estimation for the position and the appearance of the searched road. Transforming the 3D coordinates into image coordinates the search area for the road segmentation is restricted significantly. From the known road width (in meter) the expected width of the image stripe (in pixel) is computed using again the world–to–image coordinate transformation. The luminance of the image stripe can be predicted exploiting the known material of the road, knowledge about the sensor type and reflection properties. The attribute ranges of I_H*(Image–Stripe)* are restricted according to these estimations. All expectations of the semantic net are finally used for the configuration of the following segmentation algorithm.

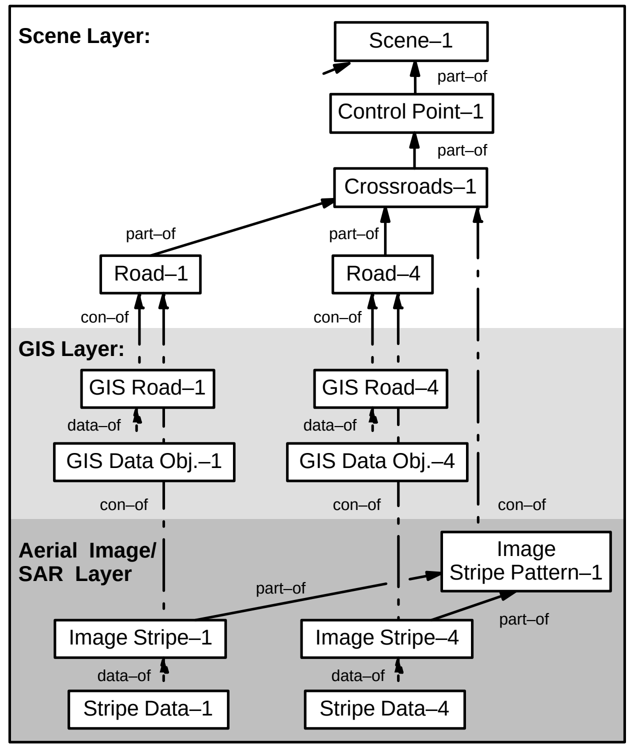

Fig. 2: Simplified instantiated semantic net matching GIS data with image stripes. The 3D/2D control point is stored in node *Control Point–1.*

For road segmentation we developed an algorithm which searches for parallel contours by tracking edges in the gradient image perpendicular to the gradient direction [6]. It returns one or several candidates I_C*(Stripe Data)* for the given hypothesis I_H*(Image–Stripe)* as a polyline. For each candidate one I_C*(Image–Stripe)* is instantiated, each describing one possible interpretation. These competing scene interpretations are represented by separate leaf nodes of a search tree. Each leaf node is judged by comparing the expected with the measured attributes. An A*–algorithm selects the most promising, i.e. best judged, interpretation alternative for further investigation.

Finally the crossroads is instantiated completely by generating hypotheses top–down for all roads and subsequently verifying them in the image. From the image coordinates of the roads their intersection representing the 2D

Fig. 3: Expected (black), segmented (grey) and selected (white) roads with computed control point in an aerial image block

Fig. 4: Expected (black), segmented (grey) and selected roads (white) in the corresponding SAR–image block

coordinate of the control point is computed by extrapolation. The found 3D/2D correspondence is one of the control points used later to estimate the sensor orientation and to resample the image accordingly exploiting a DEM.

4. Results

The knowledge based approach to automatic control point matching was tested successfully for aerial images and SAR data. Figure 3 shows the expected position of crossroads derived by projecting the GIS data into the aerial image (black lines). The system selects the four white road candidates and computes the marked intersection used as control point in the later rectification process. The results for the corresponding SAR–image are depicted in figure 4.

The control algorithms of the system AIDA are implemented in C++ with a Tcl/Tk interpreter and graphical user interface. The procedural knowledge is represented by Tcl scripts while the road segmentation is implemented in C++. The ATKIS interface is also realized in Tcl/Tk to allow easy access to the used GIS (SICAD/open) from AIDA.

5. Conclusions

The presented approach exploits prior knowledge about scene objects and a GIS to compute control points for the automatic registration of remote sensing data. Therefore suitable GIS objects are selected, grouped, and interpreted by assigning a symbolic meaning to them. Top–down propagation of object specific expectations is used to constrain the search of the image features and to configure the segmentation algorithms. Due to the flexibility of the knowledge representation the system can be extended easily to other object classes representing control points.

The knowledge based scene interpretation with semantic nets is a promising approach in the field of image understanding. Therefore the system AIDA is currently tested in other applications like land–use analysis, map update, and 3D reconstruction of landscapes and buildings.

6. References

[1] **Drewniok, C., Rohr, K.,**"Automatic Exterior Orientation of Aerial Images in Urban Environments", *XVIII Congress of Int.. Soc. of Photogrammetry and Remote Sensing*, Vienna, Austria, July 1996, pp. III: 146–152

[2] **Zheng, Q., Chelappa, R.**, "Automatic Registration of Oblique Aerial Images", *IEEE Int. Conference on Image Processing*, Austin, Texas, pp. 218–222, November 1994

[3] **Li, S.Z., Kittler, J., Petrou, M.**, "Automatic Registration of aerial photographs and digitized maps", *Optical Engineering*, Vol. 32, No. 6, pp. 1213–1221, June 1993

[4] **Haala, N., Vosselman, G.**, "Recognition of Road and River Patterns by Relational Matching", *Int. Archives of Photogrammetry and Remote Sensing*, 17. Congress of ISPRS, Vol. 29, Part B3, Com. III, pp. 969–975, 1992

[5] **Perlant, F.P., McKeown, D.M.**, "Scene Registration in Aerial Image Analysis", *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 4, pp. 481–493, April 1990

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

[7] **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, Sept. 1990.