

KNOWLEDGE-BASED ROAD INTERPRETATION IN AERIAL IMAGES

John C. Trinder and Yandong Wang

School of Geomatic Engineering
The University of New South Wales
Sydney NSW 2052, Australia
j.trinder@unsw.edu.au
yandong@colo.gmat.unsw.edu.au

Commission IV, IC Working Group IV/ III.2

ABSTRACT

This paper presents a knowledge-based method for automatic road extraction from aerial images. The method uses a hybrid control strategy in which hypotheses of roads are generated in a bottom-up process, and a top-down procedure is applied to verify the generated hypotheses. In this paper, a road model is proposed, which includes the geometric and radiometric properties of roads and relationship between roads in low- and high-resolution images. These properties and relationship are formulated as rules in PROLOG and stored in the knowledge base. The structures and relationships of roads yielded from images are stored as facts in the knowledge base. The hypotheses of roads are generated by applying the corresponding rules to the derived facts. To remove the ambiguity of the generated hypotheses, structural information of road surface and topological information of road networks are used. The missing road segments are predicted in the process of verification using topological information. An image in Hunter Valley, New South Wales, has been tested in this study. The results show that the road network is successfully extracted using the proposed method.

1 INTRODUCTION

Automatic road extraction from remotely sensed imagery is an important subject of digital photogrammetry. It aims at automatic interpretation of roads from images and determining their positions automatically, which will provide geographical information systems with up-to-date and accurate spatial and semantic data. The first attempt in this area was made by Bajcsy and Tavakoli (1976), in which roads in Landsat-1 MSS images were detected by simply examining the grey values of pixels. The grey value of an object in images is not a stable entity as it depends on not only the spectral properties of the object, but also other factors such as the season, time and weather condition of photography, processing conditions of images, etc. The same object may have different grey values in different images, while different objects may have the same grey values in one image. Therefore, it can be predicted that reliable results may not be achieved by simply using spectral information of the objects. Zhu and Yeh (1986) developed a method for automatic road extraction from high-resolution aerial images based on the radiometric properties of the generated antiparallel pairs. They assumed that long antiparallel pairs with grey values similar to the average grey value of all antiparallel pairs correspond to road segments. Obviously this assumption is not always true. In some cases, some non-road objects may have similar shape and reflectance to a road segment in a local region. Thus, non-road objects may be interpreted as road segments by their method. Similar approaches are used in Wang and Howarth (1987) and Rustoné et al (1994).

To interpret objects correctly, artificial intelligence (AI) techniques should be used. In computer vision, a number of image understanding systems which incorporate AI techniques with image analysis have been developed during the past two decades (McKeown et al, 1985; Hwang et al, 1986; Hanson and Riseman, 1987; Strat and Fischler, 1991). These systems were devised for the interpretation of various scenes from images such as air-

ports, buildings, natural scenes, etc. Since they use different knowledge representation techniques and different control strategies, they exhibit different structures. Some of them use pure top-down or bottom-up control strategy, while others utilize a hybrid control strategy. In recent years, some knowledge-based systems for automatic road extraction have been developed (Cleynenbreugel et al, 1990; Gunst, 1996). Cleynenbreugel et al (1990) use information of terrain types in the area of interest derived from the existing geographical information systems, to guide the extraction of road networks. In their method, road models are formulated for different types of terrain. The system employs a bottom-up control strategy and uses structural and contextual constraints for recognition of roads. In Gunst (1996), a knowledge-based system for change detection of roads is developed. The system uses a multi-resolution semantic model of roads and works in a top-down mode. When a road change is detected, the system will invoke appropriate image processing algorithms to find the instances of road parts.

This paper presents a knowledge-based system for road network extraction. It will cover a semantic model for road extraction which includes the definition of roads and their relationship in low- and high-resolution images, knowledge representation in a knowledge base and the control strategy of the system. Some test results will also be included in this paper.

2 A SEMANTIC MODEL FOR ROAD EXTRACTION

To extract roads from images, it is necessary to define a road model which can be implemented automatically. In model-based methods, the general properties of roads, eg, a road has a constant width, are used to describe them (Bajcsy and Tavakoli, 1976; McKeown and Delinger, 1988; Grün and Li, 1994; Trinder and Li, 1995; Vosselman and Knecht, 1995). They are formulated into a mathematical model or a process which can be executed by computer. As no specific properties of roads are used

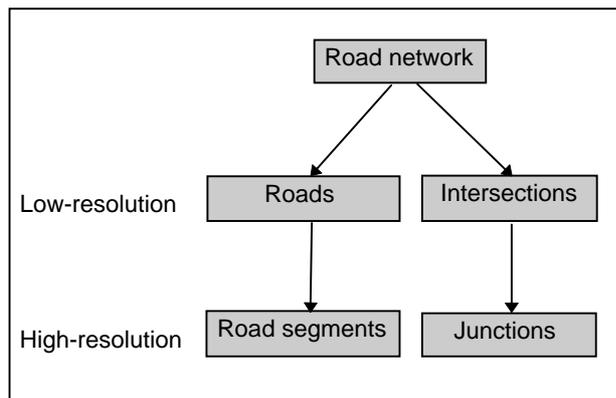


Figure 1 Semantic Road Model

in these models, it is impossible to recognize a road automatically. Instead, an operator is usually required to perform the task of interpretation.

In knowledge-based systems, a road model should include not only the general properties of roads, but also the specific properties of roads and their relationships with other objects and between road parts (Matsuyama, 1987; Cleynenbreugel et al, 1990; Gunst, 1996; Bordes et al, 1997; Baumgartner et al, 1997). In Matsuyama (1987), the road model contains the specific properties of a road and spatial relationships of roads with houses. The system mainly uses spatial relations for recognition of roads. Cleynenbreugel et al (1990) present different road models for different terrain types and use spatial and contextual constraints in recognition of roads. Since only generic knowledge about the structures of road networks is used, it is possible to generate ambiguous interpretation. Gunst (1996) developed a multi-scale road model for the change detection of roads. In this model, roads and road parts are defined by the specific properties which can be derived from the construction manuals of roads in three different levels, ie, small-, medium- and large-scale images. Road parts in different scales of images are related by specialization-generalization and spatial relationships. Baumgartner et al (1997) proposed a road model in two different resolutions, ie, low- and high-resolution, but how the road parts in low- and high-resolution images are related in their system is not clear. Bordes et al (1997) argue that most existing road models are limited to certain type of roads. They may not be suitable to other types of roads and other types of landscape. They build an image-independent semantic road model in which geometric properties, semantic attributes and contextual information of roads are derived from the existing geographical databases. But, this model is only used to determine the correct positions of existing roads in the databases. It cannot recognize roads in images automatically. In this study, the semantic model for road extraction is defined in two different resolutions, ie, low- and high-resolution images (Figure 1). In low-resolution images, a road appears as a line feature, while a road intersection has the characteristics of a point feature. They connect to each other to form a complete road network. In high-resolution images, a road segment is an elongated homogeneous area in which road boundaries (edge lines) are visible. Road markings, like lane lines on road surfaces, may also be visible, and the geometric and radiometric properties of road surfaces are measurable. Therefore, a road segment can be defined as an elongated area

bounded by two edges, with specific properties which can be obtained from manuals of road construction and spectral properties of road surface. Road junctions are also homogeneous areas, but have different shapes (T junctions and Y junctions). A road segment in high-resolution images can be considered as a part of a road in low-resolution images, thus, they can be related by part-whole relationship. A road junction in high-resolution images can be treated as the specialization of a road intersection in low-resolution images, and they are related by the relationship of specialization-generalization.

3 KNOWLEDGE REPRESENTATION AND CONTROL STRATEGY

In knowledge-based image understanding systems, the central part is the knowledge base. A knowledge base is the collection of knowledge of objects to be interpreted, including their properties and relationships, and knowledge related to the objects such as a priori knowledge on the imaging system. The knowledge is represented in the knowledge base in a certain form such as rules, frames, semantic networks etc. How the knowledge is represented in the knowledge base depends on the semantic models adopted. The form of knowledge representation determines not only the way the knowledge is expressed, but also the process of object recognition.

3.1 Knowledge Representation

In this study, the knowledge used to represent the semantic road model includes the specific properties of road segments, part-whole relationships between road segments in high-resolution images and roads in low-resolution images, spatial relationships between road segments and spatial relationships between road edges and centre line. The properties of road segments and these relationships can be easily formulated into rules since they have simple IF-THEN structure. As a rule is an explicit form of knowledge representation, it is easy to understand and to update. In addition, it can represent both declarative and procedural knowledge.

For the automatic interpretation of roads, the following types of rules are developed:

- rules for hypothesis generation. The geometric and radiometric properties of road segments are included in these rules. They are used to generate the hypotheses of road segments.
- rules for hypothesis verification. These rules include the part-whole relationships between road segments in high-resolution images and roads in low-resolution images and spatial relationships between road edges and centre line. They are used for verification of the generated hypotheses.
- rules for prediction of missing road segments. This type of rules contain the spatial relationships between the verified road segments. They are used to predict the missing road segments and search them.

3.2 Control Strategy

There are different control strategies for automatic inter-

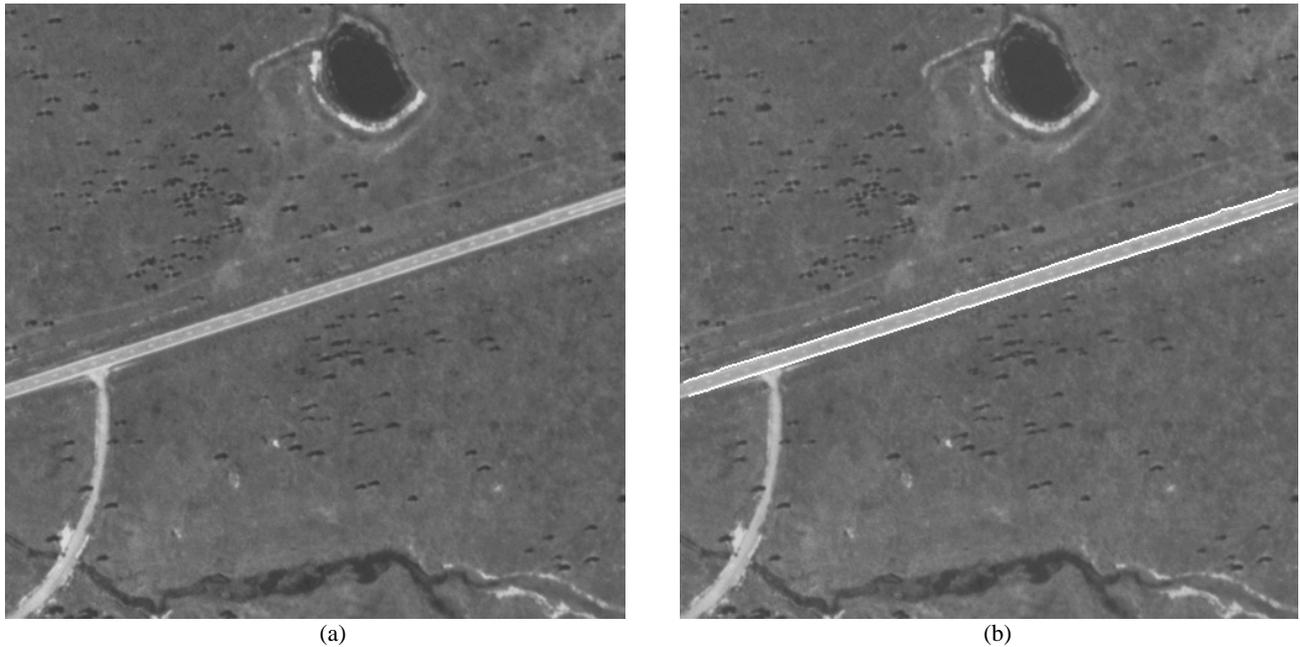


Figure 2 Generation of Hypotheses (a) Test Image (b) Hypothesized Road Segment (White Lines)

pretation of objects from images. The recognition of objects may proceed from image processing to interpretation, through formation of object structure and their relationships, ie, bottom-up control strategy, or commence from hypothesis generation to searching instances of objects, ie, top-down control strategy. What control strategy will be used depends on the task of interpretation. It is generally accepted that in most cases reliable results may not be achieved by pure bottom-up or top-down processes due to the complexity of images. A pure bottom-up or top-down process is used only when the objects to be interpreted have simple structures. In general cases, a hybrid control strategy, ie, integration of bottom-up and top-down processes, is applied to achieve more reliable results (Matsuyama, 1987; Hanson and Riseman, 1987). In this study, a hybrid control strategy is used in which hypotheses of road segments are generated in a bottom-up process and a top-down procedure is applied to verify hypotheses of road segments and to search the missing road segments. As a road segment is defined by a simple structure, it is believed that reliable results could be produced by a bottom-up process. Hypothesized road segments are verified using structural information of road surface and topological information of road networks. Centre lines on road surface are found based on the spatial relationship between road edges and centre lines. This is done in a top-down procedure. The searching for missing road segments is accomplished using spatial information between verified road segments. This is also done in a top-down process.

4 INTERPRETATION OF ROADS FROM AERIAL IMAGES

Interpretation of roads includes hypothesis generation of road segments and verification of hypotheses. Hypotheses of road segments are generated in high-resolution images using the knowledge about the properties of road segments. They are verified using the structural and to-

pological information of roads. The missing road segments are predicted and found during this stage.

4.1 Hypothesis Generation

Hypotheses of road segments are generated in a bottom-up process which consists of low-level processing for extraction of road boundaries, intermediate-level processing for formation of road structure and high-level processing for recognition of road segments. In low-level processing, edge information is extracted from images using an edge detector. To generate elongated edge segments, some preprocessing such as edge tracking, edge linking are performed. Finally, short segments and segments with large curvature are treated as noise and removed. A road segment in high-resolution images is an elongated homogeneous area bounded by two parallel edges. It is described by a generalized antiparallel pair consisting of two edges which are parallel pointwise and have opposite gradient values. In intermediate-level processing, generalized antiparallel pairs are generated based on the edge segments extracted in low-level processing and their geometric and radiometric attributes are determined. A generalized antiparallel pair is represented symbolically in terms of a number of geometric and radiometric attributes. After grouping of antiparallel pairs, road-like features are formed, which are represented in the following form (Trinder and Wang, 1997):

```
feature(feature_no,  
        first_side(point(X,Y,Z), point(X,Y,Z)),  
        second_side(point(X,Y,Z), point(X,Y,Z)),  
        attribute(gradient, length, width,  
                 average_intensity)).
```

In high-level processing, the knowledge on the properties of road segments is applied to the derived road-like features to generate hypotheses of road segments. The

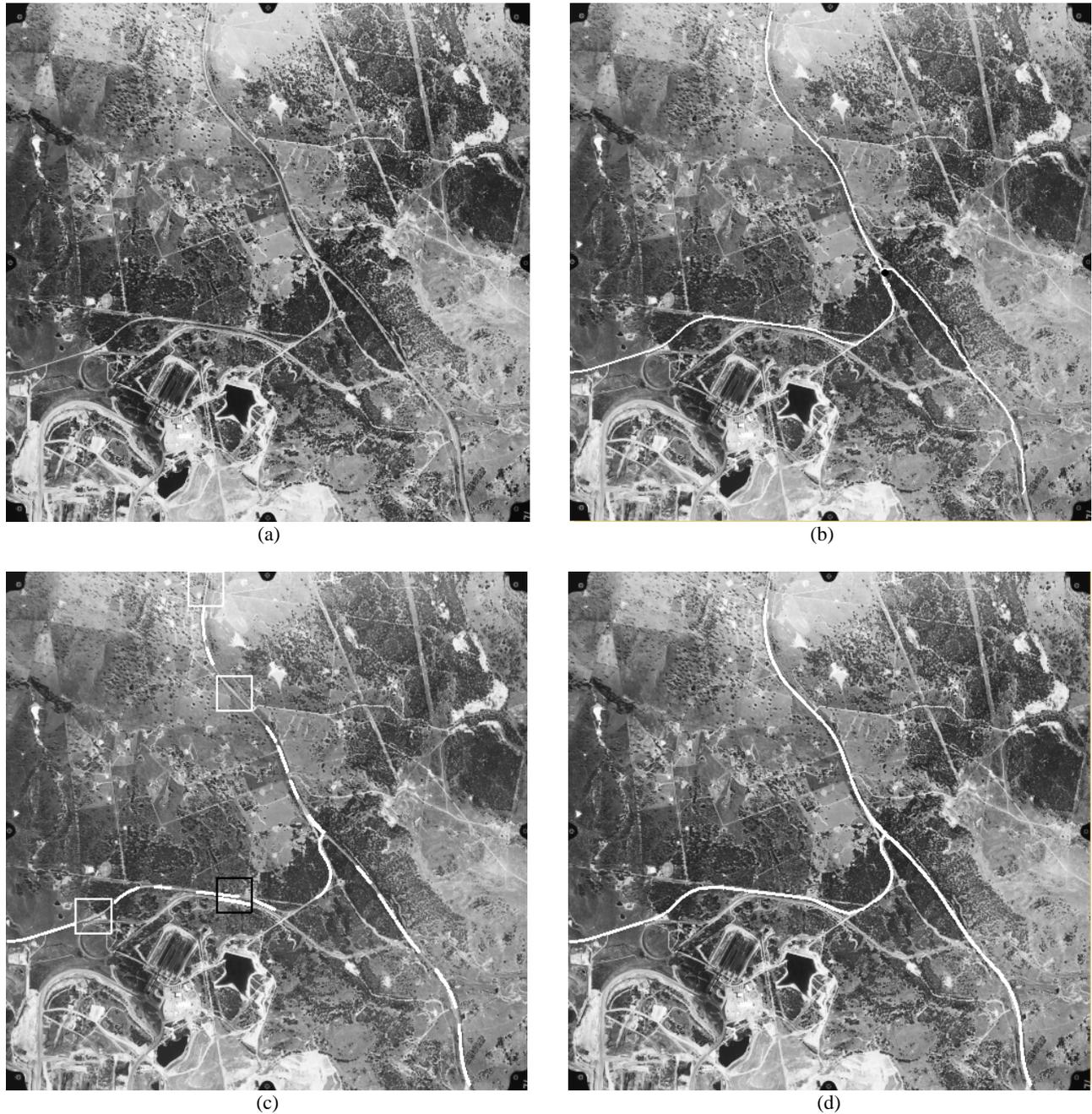


Figure 3 Verification of Hypotheses (a) Test Aerial Image (b) Generated Network (c) Hypothesized Road Segments (d) Road Network After Verification

properties of road segments include their geometric and radiometric attributes, which are formulated into one rule. The rule for the recognition of road segments has the following form (Trinder and Wang, 1997):

```
road(X) :- feature(X, _, _, attribute(gradient, length,
                                     width, average_intensity),
                gradient is -1,
                length > Tl,
                width > Wd - Td,
                width < Wd + Td,
                average_intensity > G0 - Ti,
                average_intensity < G0 + Ti.
```

Where T_l , T_d , and T_i are predefined thresholds for the length, width and intensity of road segments respectively. Their values can be obtained based on the knowledge of road width, size of test image, accuracy of computation of road width in the image and spectral properties of road surface. W_d is the width of road which can be determined from the manual of road construction. G_0 is the average intensity of road segments which is determined through a large number of experiments. As the grey value of a road segment is dependent on not only the spectral properties of surface materials, but also other factors, a large threshold for this variable is usually given. A road usually has higher reflectance than its background, therefore, -1 is chosen for the gradient value.

An example of hypothesis generation in high-resolution images is given in Figure 2. Figure 2(a) is a test window of an aerial image in Hunter Valley, New South Wales (Figure 3(a)). The image has a scale of 1:25,000 and a pixel size of 30 μm which yields a ground resolution of 0.75m. The generated hypothesis of the road segment is shown in Figure 2(b) (white lines). In this example, the threshold for segment length is set as 200 pixels in image.

4.2 Verification of Hypotheses

After hypotheses of road segments are generated, they are verified using the structural information of road surface and the part-whole relationship between road segments in high-resolution images and roads in low-resolution images. Finding instances of centre lines on hypothesized road segments is a top-down process which is formulated as follows:

```
find_line(X) :- road(X)
```

find_line is a procedure which is interfaced with the knowledge base. When a road segment is hypothesized, the system will invoke the procedure to find the instances of the centre line on it. The hypothesized road segment is accepted only when a centre line is found.

The part-whole relationship is determined by projecting hypothesized road segments onto low-resolution images and is expressed as (Wang and Trinder, 1998):

```
part(h, edge)
```

where 'h' is a hypothesized road segment and 'edge' is a road in low-resolution images which is expressed as an edge in a graph. When the percentage of the length of all hypothesized road segments belonging to an edge, to the length of the edge, is larger than a given value, they are accepted.

When road segments are verified, they are related by a spatial relationship which has the following form (Wang and Trinder, 1998):

```
neighbour(h1, h2, d)
```

'h1' and 'h2' are two verified road segments belonging to the same road and close to each other. d is the distance between them. If the distance d is larger than the defined threshold, a missing segment is predicted and a top-down procedure is applied to find the instances of road segments.

Figure 3 shows the result of hypothesis verification using the part-whole relations and searching of missing road segments. Figure 3(b) presents the network extracted in low-resolution images. The hypothesized road segments are shown in Figure 3(c). Using the rules defined above, 22 out of 23 hypothesized road segments are accepted and 16 shorter road segments are found. A complete road network is formed after all verified road segments are connected according to their spatial positions (Figure 3(d)).

5 CONCLUSIONS

This paper presents a knowledge-based method for automatic road extraction from aerial images. The method

uses a hybrid control strategy in which hypotheses of roads are generated in a bottom-up process and a top-down procedure is applied to verify the generated hypotheses. In this paper, a road model is proposed, which contains the geometric and radiometric properties of roads and part-whole relationships between roads in low- and high-resolution images and spatial relations between road segments. These properties and relationships are formulated as rules in PROLOG. The structures and relationships of roads yielded from images are stored as facts in the knowledge base. Road segments are hypothesized by applying the recognition rules to the produced road-like features. To remove the ambiguity of the generated hypotheses, structural information of road surface and topological information of road networks are used. The missing road segments are predicted in the process of verification using spatial relationships between verified road segments. The results show that the road network is successfully extracted using the proposed method.

REFERENCES

- Bajcsy, R. and Tavakoli, M., 1976. Computer Recognition of Roads from Satellite Pictures. *IEEE Transactions on Systems, Man, and Cybernetics*, **6**(9), 623-637.
- Baumgartner, A., C., Steger, H., Mayer and W. Eckstein, 1997. Semantic Objects and Context for Finding Roads. *SPIE Proceeding: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III*, Orlando, Florida, 98-109.
- Bordes, G., G. Giraudon and O. Jamet, 1997. Road Modeling Based on a Cartographic Database for Aerial Image Interpretation. In Förstner and Plümer (eds): *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag, Basel, 123-139.
- Cleynenbreugel, J. van, Fierens, F., Suetens, P. And Oosterlinck, A., 1990. Delineating Road Structures on Satellite Images by a GIS-Guided Technique. *Photogrammetric Engineering and Remote Sensing*, **56**(6), 893-898.
- Grün, A. and Li, H., 1994. Semi-Automatic Road Extraction by Dynamic Programming. *International Archives of Photogrammetry and Remote Sensing*, **30**(3), 324-332.
- Gunst, M. de, 1996. Knowledge-Based Interpretation of Aerial Images for Updating of Road Maps. *Publications on Geodesy*, No. 44, Netherlands Geodetic Commission.
- Hanson, A.R. and E.M. Riseman, 1987. *VISIONS Image Understanding Systems*. In *Advances in Computer Vision*, C. Brown (ed.), Hillsdale, NJ, 1988.
- Hwang, V.S.S., L.L.S. Davis and T. Matsuyama, 1986. Hypothesis Integration in Image Understanding Systems. *Computer Vision, Graphics, and Image Processing*, **36**: 321-371.
- Matsuyama, T., 1987. Knowledge-Based Aerial Image Understanding Systems and Expert Systems for Image Processing. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. **GE-25**, No. **3**, 305-316.
- McKeown, D.M., W.A. Harvey and J. McDermott, 1985. Rule-Based Interpretation of Aerial Imagery. *IEEE Transactions on Pattern Analysis and Machine Perception*, **7**(5): 570 - 585.
- McKeown, D.M. and Delinger, J.L., 1988. Cooperative Methods for Road Tracking in Aerial Imagery. *IEEE Pro-*

ceeding on Computer and Computer Recognition, 662-672.

Ruskoné, R., Airault, S. And Jamet, O., 1994. Road Network Interpretation: A Topological Hypothesis Driven System. International Archives of Photogrammetry and Remote Sensing, **30**(3), 711-717.

Strat, T.M. and Fischler, M.A., 1991. Context-Based Vision: Recognizing Objects Using Information from Both 2-D and 3-D Imagery. IEEE Trans **PAMI-13**, 1050 - 1065.

Trinder, J.C. and H. Li, 1995. Semi-Automatic Feature Extraction by Snakes. In A. Grün, O. Kübler, P. Agouris (Eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (I), Birkhäuser Verlag, Basel, 95-104.

Trinder, J.C. and Wang, Y., 1997. Automatic Road Extraction from Aerial Images. Proceeding of First International Work-shop on Image Analysis and Information Fusion, Adelaide, Australia, 61-71.

Vosselman, G. and Knecht, J. de, 1995. Road Tracing by Profile Matching and Kalman Filtering. In A. Grün, O. Kübler, P. Agouris (Eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (I), Birkhäuser Verlag, Basel, 265-274.

Wang, J.F. and Howarth, P.J., 1987. Automatic Road Network Extraction from Landsat TM Imagery. ASPRS-ACSM Annual Convention, Baltimore, USA, Vol. **1**, 429-438.

Wang, Y. And J. Trinder, 1998. Use of Topology in Automatic Road Extraction. International Archives of Photogrammetry and Remote Sensing, Com. III, ISPRS, Columbus.

Zhu, M.L. and Yeh, P.S., 1986. Automatic Road Network Detection on Aerial Photographs. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 34-40.