REALISTIC LANDSCAPE MODELLING WITH HIGH LEVEL OF DETAIL

Ralf Tönjes

Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung,
Division: Automatic Image Interpretation, Chair: Prof. C.–E. Liedtke
Universität Hannover, Appelstr. 9A, 30167 Hannover, Germany,
E–Mail: toenjes@tnt.uni–hannover.de

ABSTRACT

An approach for generating realistic 3D models of natural landscapes from overlapping aerial images is presented. Based on stereoscopic correspondence analysis a 3D terrain model is computed. The 3D reconstruction of a scene from 2D projections is underconstrained and details are lost due to occlusion and image resolution. The ambiguity and lack of detail is overcome by integrating a priori knowledge about the scene in the modelling process using image interpretation. Segmentation of linear features employs edge detection, and segmentation of areal features uses region growing based on a statistical texture model. Knowledge about object geometry is used to model continuous surfaces separately and to force height steps at discontinuities as forest edges. For visualization the landscape model is approximated by a polygon mesh with overlaid photo texture.

1. INTRODUCTION

Advances in computer hardware performance allow new applications in the domain of computer graphics and visualization. Computer generated landscapes gain importance because of their unrivalled vividness as proofed by flight and driving simulators. High realism is achieved by mapping aerial images on the geometric model. Thus landscape visualization is becoming an important tool for earth scientist, environmental researchers and civil engineers. A civil engineer could insert a CAD model of a building in the landscape and visualize the modified scene by simulation.

The generation of digital terrain models from aerial images using stereo matching techniques is described by various authors [1][2]. However this approach hardly works for complex scenes and fails when a high level of detail is required. Natural culture (wood, trees) and man made objects (buildings, roads) are not modelled or appear faulty. A human observer is in possession of a priori knowledge about the scene and expects height steps at the edges of forests and roads to run continuously. To obtain more realistic models this a priori knowledge shall be integrated in the modelling process using image segmentation and interpretation.

The following section gives an overview of conventional landscape modelling. Section three is devoted to segmentation of wood and road features. The fourth section describes the integration of the segmentation results in the modelling. The paper concludes with a presentation of the modelling results.

2. STEREOSCOPIC IMAGE PAIR ANALYSIS

The 3D modelling is based on photogrammetric stereo vision. Two overlapping aerial images are rectified in a way that the epipolar line coincides with the image scanline to ease search for homologous points. Camera orientation is calculated from 3D control points by bundle block adjustment.

The essential part of height measurement is the correspondence analysis that determines the height dependent parallax of homologous points. Search space is reduced by a parallax limit. Using normalized cross correlation as cost function for matching homologous points the optimum for each scanline is found by dynamic programming yielding a dense parallax map for each point in the image pair [3][4]. A Smoothness constraint is exploited by subsequently interpolating continuous regions [5].

Finally the parallax map is transformed to a height map using binocular camera geometry. For efficient visualization the height map is sampled and approximated by a polygon mesh in space. The aerial images are projected onto the polygon surfaces and stored in a texture map to synthesize natural looking views of the landscape model.

Stereoscopic image matching yields only a coarse terrain model without modelling details such as edges of forests, roads and buildings due to image resolution. These objects must be segmented and modelled separately.

3. SEGMENTATION

Three groups of scene features can be distinguished for segmentation and modelling: areal, linear and point features. In the context of aerial images woods and grassland represent areal, roads and rivers linear, and buildings and trees point features.

3.1. Segmentation of Linear Features

In the aerial images roads represent linear features of a few pixel width. Hence road segmentation employs edge detection based on orientation filters [6] with consecutive conditional local

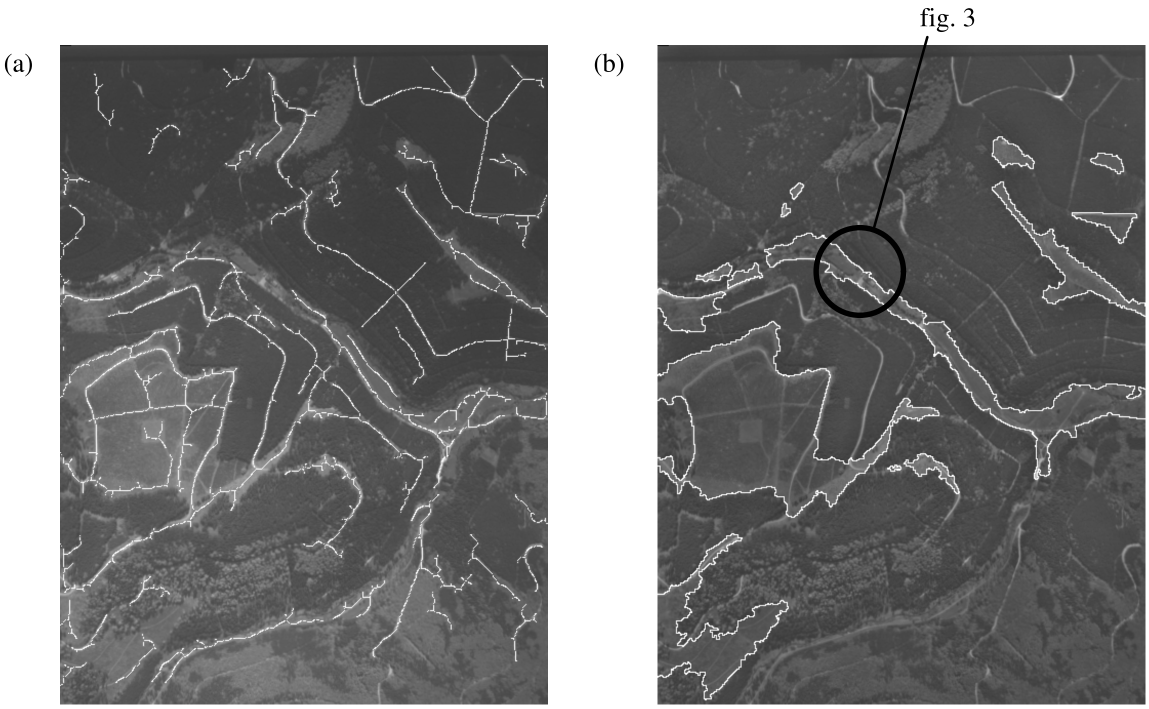

Figure 1: Segmentation results for (a) roads and (b) wood

ranking to enhance weak contours. Thresholding yields a binary image in which the long lines represent road segments. Figure 1a shows the results of road segmentation.

3.2. Segmentation of Areal Features

Aerial images of natural terrain are efficiently characterized by statistical models for texture and region form. Such textures represent different types of wood, grassland, fields, etc.. From literature various approaches to texture segmentation are known [7]. In the following a supervised segmentation with statistical models for texture and region form controlled by a morphological operator is proposed.

3.2.1. Texture Model for Statistical Independence

The texture of a region be generated by a stationary ergodic process. The Probability density p(y) of a texture process is calculated from the luminance histogram of all intensities y in the learn region.

For statistical independence of the texture processes within and between regions the joint probability for the realization of an image partitioning is given by the product of the probability for all pixels. Segmentation intends to maximize joint probability by maximizing the independent probabilities. For each pixel i the hypotheses H_k of class k membership is tested with the maximum–likelihood–ratio λ.

$$\lambda = \frac{(p(y_i)|H_m)}{(p(y_i)|H_n)} \tag{1}$$

The pixel is assigned to class m for $\lambda>1$ and to class n otherwise.

3.2.2. Gibbs Random Fields

Generally the probability for occurrence of a luminance value is not independent from its neighbors. Local mutual dependencies can be modelled by Gibbs random fields. Couples of neighbored pixels, named cliques, are inspected. Three measures of co–occurrence are computed for each clique [8]:

- the probability of common class membership,
- the probability of a luminance difference within a region,
- the probability of a luminance difference between different regions.

To each clique a potential is assigned. The potential V_c of a clique c should be smaller for higher probabilities of realization. To ease computation the logarithm of the probability is chosen.

$$V_c(y) = -\ln(p(y)) \tag{2}$$

Gibbs random fields describe the joint probability P(y) of a possible segmentation for selected types of cliques.

$$P(y) = \frac{1}{Z}\, e^{-\sum_c V_c(y)} \; ; \; Z = \text{normalizing constant.} \tag{3}$$

Due to dependencies from the class membership of the neighboring pixels the segmentation problem must be solved iteratively using maximum–a–posteriori estimation.

3.2.3. Usage of Morphological Operators

So far segmentation does not take advantage from prior knowledge about form, location and number of regions. To use this

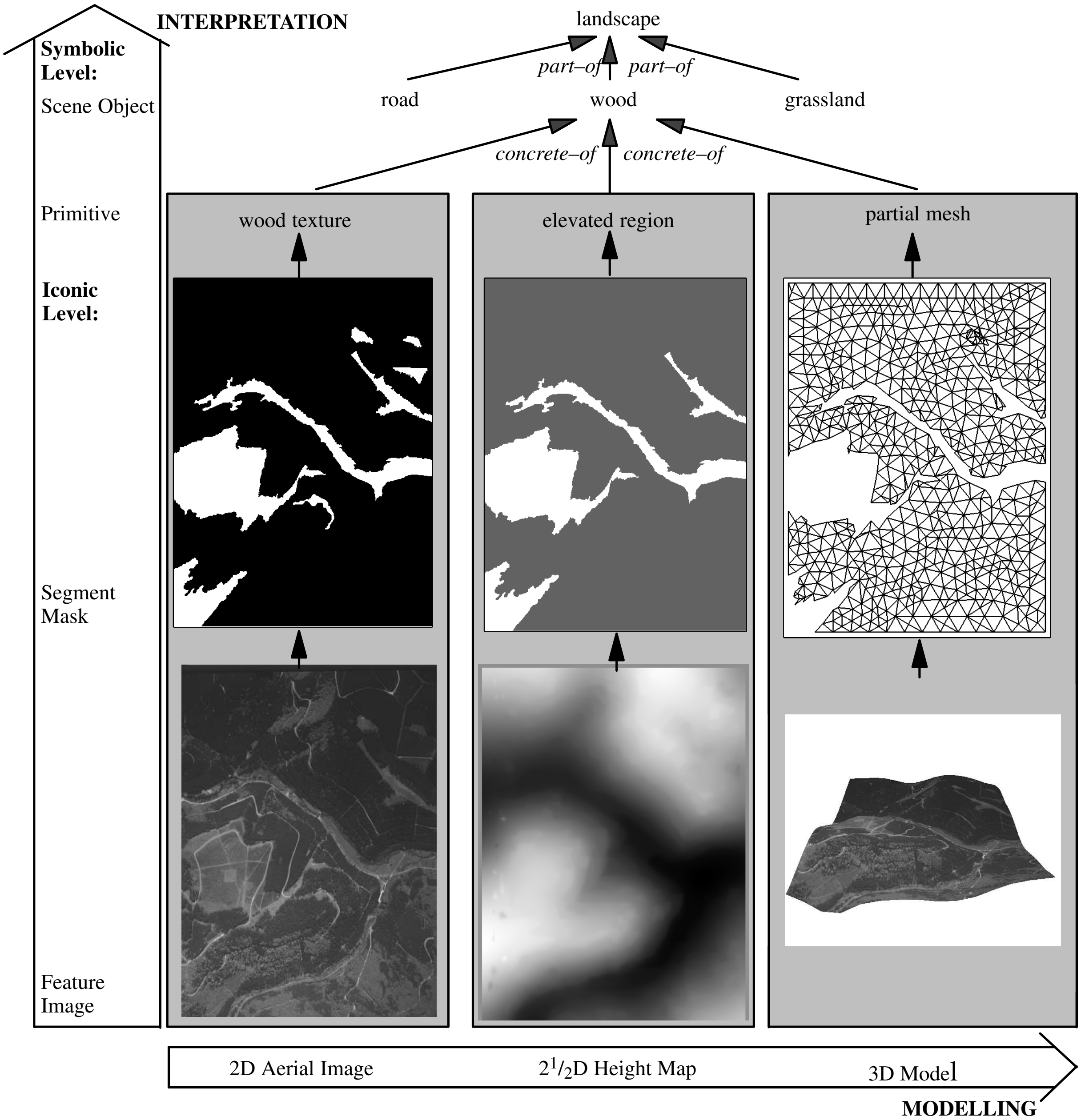

Figure 2: Integrating a priori knowledge in the modelling using image interpretation

information the segmentation with statistical models is controlled by morphological operators described in the following.

Compactness of regions is forced by a region growing process. Starting from interactively selected region kernels, border pixels are assigned to the region, when their membership probability exceeds a threshold. Finally to smooth region contours border pixels are relabeled using Gibbs random fields with clique potentials that favour smooth contours [9].

Figure 1b shows the results of wood segmentation employing a region growing process that considers only first order statistics of the texture.

4. MODELLING USING IMAGE INTERPRETATION

To exploit prior knowledge about the scene objects for 3D modelling an image interpretation is required. The goal of image interpretation is to assign meanings to image regions and thus to distinguish different objects. While object recognition only requires a correspondence between an object and a certain image region, object reconstruction implies the exact geometric modelling of the 3D object. The presented approach determines both object geometry and object semantic.

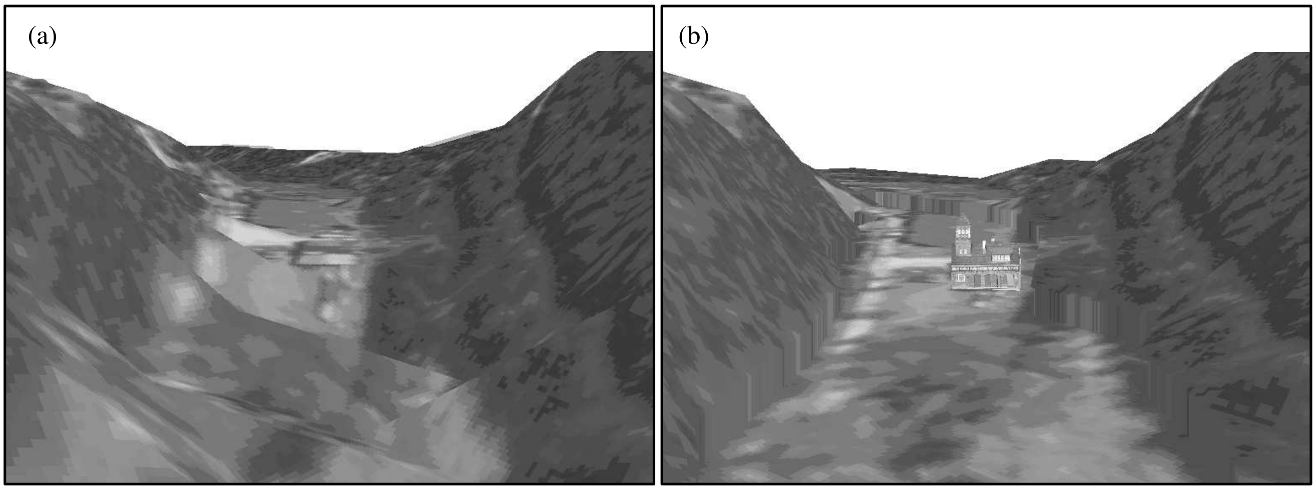

Figure 3: Synthesized close up view of landscape section marked in fig. 1 (a) without and (b) with usage of image interpretation.

Conventional landscape modelling ignores the object meaning and applies only stereo matching techniques to construct a terrain model. This data controlled approach fails to model details like edges of forests and the continuous course of roads. Their modelling has to be taken from a priori knowledge about the scene. Such common knowledge states that

– wood areas exhibit a height discontinuity, i.e. step, at their edge,
– roads run continuously without breaklines.

Knowledge representation employs semantic nets [10]. The knots of the semantic net are implemented as frames, which represent a collection of attributes. The edges or links of the semantic net form the relations between the nodes. Image interpretation uses especially two relations: decomposition and concretion.

Objects are composed of parts, indicated by the part–of link. Object search can be reduced to a more simple task, the detection of its components.

Objects can often be detected based on their geometric or photometric appearance, that can directly be segmented in the image data. This transformation of an abstract symbol to a concrete realization is represented by the concrete–of link. Generally three different aspects or concretes of the model can be distinguished: culture, terrain and model representation. They are obtained from the aerial image, the height map and the 3D model respectively.

Figure 2 shows the integration of a priori knowledge into the modelling process using image interpretation. The bottom level represents the steps of a solely data driven modelling, transforming one iconic representation, i.e. image, directly into another, i.e. textured polygon mesh, disregarding semantics. The prior knowledge about the scene features is represented by the semantic net at the symbolic level. The landscape is composed of roads, wood and grassland. To use the knowledge the symbolic description must be assigned to image data using interpretation. More concrete wood is a region with tree texture, continuous besides a height discontinuity at the edge and thus represented by a partial mesh. By interpretation of the aerial image a mask for the wood region is generated. Neglecting small regions this mask is used for elevating the wood region and considering object boundaries for triangulation. Controlled by the symbolic level a priori knowledge is integrated top–down in the modelling process by

– applying height interpolation only to continuous regions, inserting a height displacement between neighboring regions at height steps, and
– representing continuous regions by a partial mesh .

The semantic modelling yields a symbolic description of the parts of the iconic model.

5. RESULTS

Figure 3 compares the close up views of the landscape model neglecting and using image interpretation. Figure 3b shows the improved realism by modelling continuous regions with a separate polygon mesh and inserting a height step at the edge of the forest. The symbolic model description eases the postprocessing with computer graphic techniques. Thus the inserted edge of the forest is textured and the flat building is replaced by a CAD model. For visualization a camera trajectory can be positioned in the model to synthesize a realistic looking image sequence.

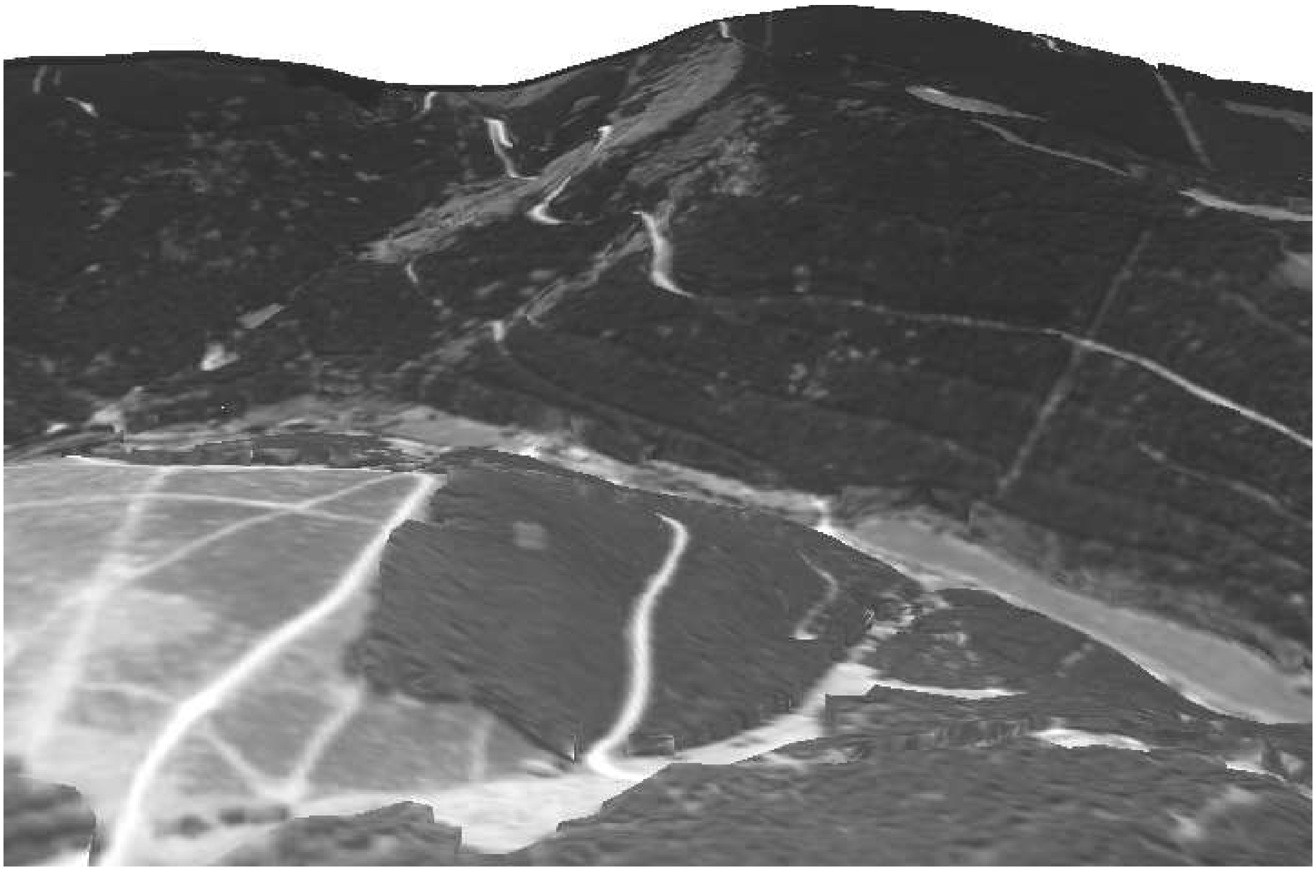

Figure 4: Synthesized view of landscape model with continuous roads and elevated woods

Figure 4 shows a reconstruction of the scene Sieber valley. To enhance realism the wood is elevated non–uniformly. Roads passing woods are not elevated and appear partially occluded.

6. CONCLUSION

The 3D reconstruction of landscapes from 2D projections is ambiguous and model detail is lost. To improve realism 3D modelling exploits a priori knowledge about the scene objects. This is achieved by image interpretation. The symbolic description of the 3D model establishes an interface to control postprocessing with computer graphic techniques.

7. REFERENCES

[1] **Ackermann, F., Krzystek, P.** , "MATCH–T: Automatic Mensuration of Digital Elevation Models", *Proceedings of Technical Seminar of the Sociedad Espanola de Cartografia Fotogrammetria y Teledeteccion*, pp.67–73, Barcelona, 12th April 1991.

[2] **Lemmens, M.**, "A survey on Stereo Matching Techniques", *ISPRS Conference*, Kyoto, Japan, 1988.

[3] **Cox, I.J., Hingorani, S., Maggs, B.M., Rao, S.B.**, "Stereo without Regularisation", *British Machine Vision Conference*, Leeds, Sep. 21 – 24, Eds. David Hogg & Roger Boyle, Springer Verlag, 1992.

[4] **Koch, R.**, "Automatic Reconstruction of Buildings from Stereoscopic Image Sequences", *Eurographics '93*, Barcelona,Spain, 1993.

[5] **Terzopoulos, D.**, "The computation of visible–surface representations", *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol 10, pp.417–438, 1988.

[6] **Nevatia, R., Babu, K.R.**, "Linear Feature Extraction and Description", *Computer Graphics and Image Processing*, Vol. 13, pp. 257–269, 1980.

[7] **Reed, T. D., Buf, J. M. H.**, "A Review of Recent Texture Segmentation and Feature Extraction Techniques", *CVGIP: Image Understanding*, Vol. 57, No. 3, May, pp. 359–372, 1993.

[8] **Gimmelfarb, G. L., Zalesny, A. V.**, "Low–Level Bayesian Segmentation of Piecewise–Homogenous Noisy and Textured Images", *International Journal of Imaging Systems and technology*, Vol. 3, pp. 227–243, 1991.

[9] **Mester, R., Franke, U.**, "Statistical model based image segmentation using region growing, contour relaxation and classification", *Proc. of SPIE Conf. on Visual Communication and Image Processing*, Cambridge Mass., Vol. 1001, November 1988.

[10] **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, pp. 883–905, Sept. 1990