# An Evidential Merit Function to Guide Search in a Semantic Network Based Image Analysis System

Franz Quint and Manfred Sties

Institute for Photogrammetry and Remote Sensing
University of Karlsruhe
D–76128 Karlsruhe, Germany
quint@ipf.bau–verm.uni–karlsruhe.de

**Abstract.** To analyse aerial images we are using a model based, structural, hierarchical process implemented in the system MOSES. Semantic networks are used as modeling tools. In a three stage scheme, the models are successively refined and for image analysis an automatically generated semantic network, specialized in the analysis of a concrete scene is used.

In this article we present the merit function used to guide search in the image analysis process. Relying on the Dempster–Shafer theory of evidence we extend approaches from the literature to combine subjective valuations at different hierarchical levels. We show how a merit function for informed search methods can be built using the induced basic belief assignment for the goal concept of the analysis. Making use of the capability of Dempster–Shafer theory to explicitly handle lack of knowledge, our formalism leads to an overestimate of the merit for the search path from the current node to the solution node. Experiments using the described merit function in an $\varepsilon$–$A^*$–algorithm are presented.

## 1 Introduction

Image analysis is implemented in our system MOSES[1] (*M*ap *O*riented *SE*mantic image under *S*tanding) [1], [11] as a model based, structural approach. We are using line segments gained from large scale color aerial images in a preceding feature extraction step as primitives for the structural analysis process. Context information is acquired from the German Topographic Base Map 1:5000.

Four different models are used in MOSES. For representing these models we use semantic networks as implemented by the shell for knowledge based analysis ERNEST [8]. Three of the models are scene independent and are specified by the system developer. The fourth is specific for the scene to be analysed and is generated automatically. It is the one actually involved in the task of image analysis.

The most general of our scene independent models is the *generative model*, which describes our environment in the scene domain. The knowledge contained

---

in this model is of declarative nature and is general, common sense knowledge we have about our environment. This model is further refined, resulting in two generic models: the *generic model in the map domain* and the *generic model in the image domain*. Besides the common properties inherited from the generative model, they reflect particularities of the representations in the map and image domain. Both generic models contain methods necessary for performing the analysis in their respective domain. These methods comprise functions for feature extraction, for the calculation of attribute values from the image or map data and functions for evaluating the preliminary and final analysis results.

The generic model in the map domain is used for map analysis. We gain this way a description of the scene as far as it is represented in the map. By combining this scene description with the generic model in the image domain we automatically generate a new semantic network, the *specific model*. This model, which is specific for the current scene, is used in the automatic image analysis process.

In the analysis processes we use the task independent control algorithm provided by the shell ERNEST. With a combined top–down and bottom–up strategy over the *part* and *concrete* hierarchy of the semantic network, restrictions are propagated and correspondences between primitives in the database and concepts in the knowledge base are searched [6].

Thus, analysis can be graphically represented as a search tree in which each node corresponds to a given state of the analysis. It is know that the problem of finding an optimal path in the search tree can be solved by the $A^*$–algorithm [9]. Its application is possible if one can evaluate the path from the root node to the current node and can give an estimate for the valuation of the path from the current node to the (not yet known) terminal node containing the solution.

## 2   Merit Functions for Informed Search Methods

Search methods which choose a path according to some information about the prospects of finding a solution at a node are called informed search methods. The $A^*$–algorithm belongs to this class of search methods and requires a merit function composed of two parts:

$$f^*(n) = g(n) + h^*(n) \ . \tag{1}$$

The term $g(n)$ is the merit function of the path from the root node of the search tree to the current node $n$, and $h^*(n)$ is an estimate for the merit of the path from the current node to the terminal node which will contain the solution. The $A^*$–algorithm always expands the node with the highest value of the estimated merit $f^*(n)$. The admissibility condition requires that the merit of the future successful path is overestimated.

Because of the admissibility condition, the $A^*$–algorithm will usually expand many nodes not leading to the solution. An alternative is the $\varepsilon$–$A^*$–algorithm [10], which relaxes the admissibility condition. Instead of the merit function (1),

the $\varepsilon-A^*$–algorithm uses:

$$f^*(n) = g(n) + (1 - \varepsilon)h^*(n) \ , \tag{2}$$

where $\varepsilon$ is a small factor and $h^*(n)$ is an admissible function. Clearly, $f^*(n)$ may now underestimate the actual merit and one may miss the optimal solution when using this function. However, loss of optimality is in the most unfortunate case limited to $\frac{\varepsilon}{1-\varepsilon}$ percent of the merit of the best solution.

An approach to construct a merit function using the mutual information between parts of the model and the image primitives is presented in [15]. In [2], a probability based and a fuzzy based method for combining knowledge in expert systems and deriving a merit function are presented. In systems based on the shell ERNEST multicomponent judgment vectors have been used (see e.g. [7],[13]).

Generally, one faces the problem to estimate the merit of the future successful path appropriately. This is difficult since one has very little knowledge about it: at a given state of analysis, it is not known how many nodes the path from the current node to the solution node will contain. Due to noisy image data it is also not known, how many instances will have to be created until the solution node is reached, since for example a line segment in the model may be broken into several line segments in the image. The derived merit function has to be a common ground for both the comparison of paths developed to the same level, and for the comparison of paths advanced in the building of the solution with paths abandoned earlier. And of course, it has to fulfill the admissibility condition of the $A^*$–algorithm.

In our image analysis problem we relate the merit function to the analysis goal. The valuation of the analysis goal is calculated considering the valuations of already created instances and the estimates for the valuations of future instances. We embed these valuations in the Dempster–Shafer theory of evidence and we shall show how these valuations can be combined to achieve the valuation of the analysis goal. An estimate for the merit of the future search path is obtained in a natural way by modeling ignorance.


## 3   Propagation of Valuations in the Semantic Network

Like Bayesian approaches, the Dempster–Shafer theory of evidence [3],[14] aims to model and quantify uncertainty by degrees of belief. Dempster's rule of combination gives us a formalism for combining distinct pieces of evidence in the same frame of discernment (so called parallel combination). However, in expert systems one often also needs methods for propagating evidence along chains of expert rules, i.e. from one frame of discernment to another.

Ishizuka et al. [5] provide a method for sequential combination, assuming that the premises of the different expert rules are mutually exclusive. A major drawback of this approach is, that although the hypotheses established by the expert rules are not necessarily mutually exclusive, the method does not take

into account the fact that one hypothesis may be confirmed by different prerequisites of a rule. Ginsberg [4] developed a model for propagating belief through a semantic network, but he limits his approach to dichotomous frames, i.e. frames of discernment with only two events.

A quasi–probabilistic approach was introduced by Yen [16]. Extending Dempster's original multi–valued mapping to a probabilistic mapping, he measures certainty degrees of expert rules by means of conditional probabilities. Yen's model requires complete knowledge about all prior probability distributions. Propagation of belief along chains of expert rules is not obvious, since all the probabilities necessary in the prerequisites of the follow–on rule are generally not known on the basis of the belief masses calculated with the initial rule of the chain.

For our purposes we only postulate the existence of a basic belief assignment for our frame of discernment. The belief in an event induced by this assignment is a *subjective* measure assigned to the corresponding subset of the frame of discernment; the existence of an underlying probability distribution is irrelevant to our approach.

Let the frame of discernment be $\Omega_X = \{x_1, \ldots, x_K\}$ with attached basic belief masses $m_X(\cdot|E_X)$, where $E_X$ denotes some background evidential source. Events $X_j \subseteq \Omega_X$ condition hypotheses $H_i \subseteq \Omega_H$ from the frame of discernment $\Omega_H = \{h_1, \ldots, h_M\}$. The beliefs for these hypotheses with respect to the background evidential source have to be calculated.

Unlike in previous approaches where conditional probabilities or conditional belief masses only for events occurring in expert rules are necessary, we require the definition of a complete conditional basic belief assignment $m_{H|X}(H_i|X_j)$ on the frame of discernment $\Omega_H$ with respect to every subset $X_j \subseteq \Omega_X$ It has to be mentioned that it eventually can be difficult for the system designer to establish the conditional belief assignments in the case of large frames of discernment since a total number of $(2^K - 1) \times (2^M - 1)$ conditional belief masses has to be specified.

For propagating the belief masses in a hierarchical environment, i.e. for calculating the basic belief assignment for a frame of discernment $\Omega_H$ with respect to the evidential source $E_X$, we use the equations:

$$m_{H_X}(\emptyset|E_X) = 0 \tag{3}$$
$$m_{H_X}(H_i|E_X) = \sum_{X_j \subseteq \Omega_X} m_{H|X}(H_i|X_j) \, m_X(X_j|E_X) \ .$$

It is easy to verify that the assignment induced by (3) satisfies the necessary conditions to be a basic belief assignment. With known belief masses for the frame of discernment $\Omega_H$ with respect to $E_X$, we can calculate the belief for the hypotheses $H_j$:

$$Bel(H_j|E_X) = \sum_{H_k \subseteq H_j} m_{H_X}(H_k|E_X) \ .$$

Equation (3) is formally similar to Yen's formula for sequential combination. However, our conditional belief masses $m_{H|X}(H_i|X_j)$ are not subject to other restrictions than the one imposed by the conditions to be a belief mass. Especially, they also may be zero.

The approaches also differ in the fact that in our combination formula (3), belief masses $m_X(X_j|E_X)$ for *subsets* of the frame of discernment are used whereas in Yen's approach probabilities $P(x_j|E_X)$ for *elements* of the frame of discernment are used. The use of belief masses has as a consequence that in (3) a term containing the belief mass of the whole frame of discernment is added:

$$m_{H_X}(H_i|E_X) = \sum_{X_j \subset \Omega_X} m_{H|X}(H_i|X_j)\, m_X(X_j|E_X) + m_{H|X}(H_i|\Omega_X)\, m_X(\Omega_X|E_X) \ .$$

With $m_X(\Omega_X|E_X)$ our lack of knowledge with respect to events from $\Omega_X$ is modeled. Thus, to the belief mass of a hypothesis $H_i$ also contributes a term which stands for the belief we have exactly in $H_i$ (not the total belief in $H_i$) when we have no knowledge about events from $\Omega_X$. The term $m_{H|X}(H_i|\Omega_X)$ can be considered to be a measure for our exact belief in $H_i$ when it is known, that no evidence at all is available from $\Omega_X$. Propagation of ignorance is thus possible. Also the propagation of evidence in a hierarchical environment over several stages (chains of expert rules) is possible since output basic belief masses of one stage can serve as input basic belief masses for the following stage.

Given several independent sources of evidence in $\Omega_X$ which condition hypotheses from $\Omega_H$, we first perform sequential combination according to (3) for each of the sources separately. Thereafter we use Dempster's rule of combination to calculate the belief mass in $\Omega_H$ regarding all available sources of evidence. Because of associativity, Dempster's rule is applied repeatedly on pairs:

$$m_H(H_i|E_X, E_Y) = \frac{\sum_{H_j \cap H_k = H_i} m_{H_X}(H_j|E_X)\, m_{H_Y}(H_k|E_Y)}{1 - \sum_{H_j \cap H_k = \emptyset} m_{H_X}(H_j|E_X)\, m_{H_Y}(H_k|E_Y)}.$$

## 3.1 An Example

We demonstrate the application of our approach for propagating the valuations in the hierarchy of our semantic network by means of a simple example. The contour of a building is represented in our generic model in the image domain by the concept i_polygon. This concept has a multiple *part*–link to a concept i_line. After the map analysis and the generation of the specific model, the inner contour of the building shown in Fig. 1 is represented by the concept i_polyg0319. Its parts are the concepts i_line0297, i_line0304, i_line0311 and i_line0318. Due to the previous map analysis, their locations in the image are approximatively known and stored in the semantic network.

As a first step in the analysis process correspondences between a concept and one or more image primitives are established. Thereafter these correspondences are evaluated. A correspondence is interpreted as a hypothesis of match between a concept and an image primitive. We evaluate our belief in this hypothesis and also the belief against this hypothesis.

At the level of line segments this belief is evaluated using a compatibility measure between a line segment $a$ and its model $L_i$ with help of a metric defined on a parametric space for line segments [12]. The quantity $m_{L_i}(L|E_a)$ measures
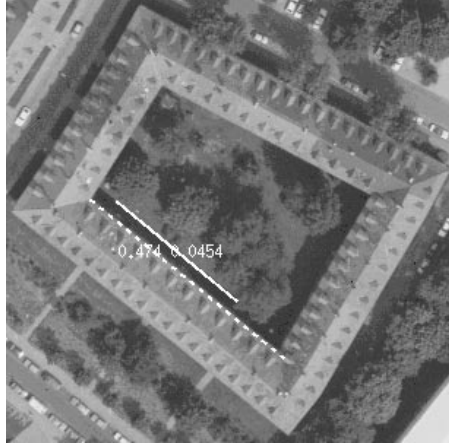
**Fig. 1.** Compatibility measures for a line segment (white line) and its model (dashed white line)



**Fig. 2.** Propagation of the line segment compatibility measure in the valuation of a polygon

our subjective belief in the hypothesis of match $L$. The quantity $m_{L_i}(\overline{L}|E_a)$ measures our subjective belief in the hypothesis of erroneous match $\overline{L}$. The sum of these two quantities may be less than 1. The difference up to 1 is assigned to $m(\Omega_{L_i}|E_a)$ and models our ignorance. We have thus established a basic belief assignment on the frame $\Omega_{L_i}$.

An example is given in Fig. 1, where these measures are displayed for an image primitive (white line) with respect to its model (dashed white line). In another search path, where for the same model another line segment $k$ was chosen (Fig. 2), the belief masses are: $m_{L_{0297}}(L|E_k) = 0.961$, $m_{L_{0297}}(\overline{L}|E_k) = 0.000743$ and $m_{L_{0297}}(\Omega_{L_{0297}}|E_k) = 0.038257$ .

The belief assignments for the hypotheses established at the other lines are calculated in a similar way. However, for the concepts which in a given analysis state are not yet addressed, we explicitly model our lack of knowledge by assigning a vacuous belief function to the corresponding frame of discernment. Thus, in the analysis state of the example of Fig. 2 the belief assignments for the frames $\Omega_{L_j}$ would be: $m_{L_j}(L|E_x) = 0$, $m_{L_j}(\overline{L}|E_x) = 0$ and $m_{L_j}(\Omega_{L_j}|E_x) = 1$ for $\mathsf{j} \in \{0304, 0311, 0318\}$.

The hypotheses established at the level of line segments jointly are a source of evidence for hypotheses at the decision level of polygons. Thus, the frame of discernment $\Omega_{\{L_i\}}$ which conditions the event of match $P$ or erroneous match $\overline{P}$ in the frame of discernment $\Omega_{P_{0319}}$ at the polygon level is given by the Cartesian product of the frames of discernment $\Omega_{L_i}$, with $\mathsf{i} \in \{0297, 0304, 0311, 0318\}$. Since the correspondence between a line model and a primitive is established independently of the correspondence between another line model and a primitive, we can calculate the belief mass for a joint event $\{l_i\} \subseteq \Omega_{\{L_i\}}$ as a product of

the belief masses of its constituents $l_\mathtt{i} \subseteq \Omega_{L_\mathtt{i}}$:

$$m_{\{L_\mathtt{i}\}}(\{l_\mathtt{i}\}|E_k) = \prod_\mathtt{i} m(l_\mathtt{i}|E_k) \ . \tag{4}$$

We also have to assign the conditional belief masses $m_{P|\{L_\mathtt{i}\}}(p|\{l_\mathtt{i}\})$, with $p \subseteq \Omega_P$. This is done by assuming that each hypothesis at line level contributes with a fraction equal to the ratio between the length of the edge and the polygons perimeter to a corresponding hypothesis at polygon level:

$$m_{P|\{L_\mathtt{i}\}}(p|\{l_\mathtt{i}\}) = \frac{\sum_\mathtt{i} \delta_{PL}(p, l_\mathtt{i})\, s_\mathtt{i}}{\sum_\mathtt{i} s_\mathtt{i}} \ , \tag{5}$$

with $s_\mathtt{i}$ the length of the model line $\mathtt{i\_line_i}$. The function $\delta_{PL}(p, l_\mathtt{i})$ is a Kronecker–like function which takes the value 1 if the hypotheses $p$ and $l_\mathtt{i}$ are in concordance (for example if $p$ is the hypothesis of match at polygon level and $l_\mathtt{i}$ is the hypothesis of match at line level etc.) and takes the value 0 otherwise.

Substituting the assignments (4) and (5) in (3), we obtain for the belief masses at the decision level of polygons induced by an instantiation which has taken place at the level of lines for the example of Fig. 2:

$$m_{P_{0319}}(P|E_k) = 0.295 \qquad m_{P_{0319}}(\overline{P}|E_k) = 0.000228$$

In the analysis state of Fig. 2 only for the concept $\mathtt{i\_line0297}$ an instantiation has been performed. Having at this time no knowledge about the other edges of the polygon, we choose for their belief assignment vacuous belief functions, explicitly modeling our ignorance. As the analysis progresses, the vacuous belief functions are replaced with the actual evidence calculated at line level.

After the propagation of the belief masses from the line level to the polygon level these are combined using Dempster's rule of combination with belief masses obtained from other, independent evidential sources for polygons.

## 4   An Evidential Merit Function

Various valuations are calculated for each instance and modified concept created during the analysis process. These valuations refer to two aspects: the compatibility of a match and the reliability of the instances. The compatibility of a match describes the quality of correspondence between the primitives extracted from the image data and the model. An example for this is the line segment compatibility described in the previous section. The reliability of the instances is calculated regarding only the primitives extracted from the image data. Examples for the calculation of the reliability are a function which evaluates the strength of a line through its mean gradient magnitude or a function which evaluates the gradient angle consistency for the edges of a polygon. All these valuations are propagated in the hierarchy of the semantic network according to the presented method, finally resulting in a basic belief assignment at the level of the analysis goal.

Let $S$ be the hypothesis at the level of the analysis goal that the instances attached to concepts in the current node of the search tree match the model and $\overline{S}$ be the hypothesis, that they do not match. For these hypotheses we obtain after the propagation of the valuations the subjective belief masses $m(S|\{E_x\}_n)$ and $m(\overline{S}|\{E_x\}_n)$. They are with respect to the background evidential sources $\{E_x\}_n$ of the instances in the current node $n$.

Our propagation scheme also provides us the measure $m(\Omega_S|\{E_x\}_n)$ which models the lack of knowledge we have at the current search node. There are two sources for this lack of knowledge. Firstly, besides our subjective belief in favor and against a hypothesis of match there may remain an unassigned quantity. This is often the case when an object has optional parts, where their presence confirms the hypothesis of match, but their absence does not necessarily deny this hypothesis. And secondly, our formalism for propagating the belief masses assumes a vacuous belief assignment for the (yet unknown) instances in the future path. Part of this vacuous belief can support or deny at future instantiation time the established hypotheses.

As a merit of the path from the root node to the actual node we use the basic belief mass calculated at the level of the analysis goal for the hypothesis $S$: $g(n) = m(S|\{E_x\}_n) = Bel(S|\{E_x\}_n)$.

As an estimate for the merit of the path from the current node to the solution node we choose: $h^*(n) = m(\Omega_S|\{E_x\}_n)$. This assignment fulfills the admissibility condition overestimating the actual merit since only a part of $m(\Omega_S|\{E_x\}_n)$ will be assigned in the future to the belief in hypothesis $S$.

The merit function $f^*(n)$ for the $A^*$–algorithm finally results in:

$$f^*(n) = m(S|\{E_x\}_n) + m(\Omega_S|\{E_x\}_n) = 1 - m(\overline{S}|\{E_x\}_n) = Pl(S|\{E_x\}_n).$$

Thus the plausibility computed at the level of the goal concept for the hypothesis that instances of the current search node match the model is used as a merit function for the $A^*$–algorithm. When using the $\varepsilon$–$A^*$–algorithm, the merit function results in:

$$f^*(n) = \varepsilon\, Bel(S|\{E_x\}_n) + (1 - \varepsilon)\, Pl(S|\{E_x\}_n). \tag{6}$$

The merit function (6) is used to direct search in our image analysis system. Although the results depend of course from the scene to be analysed and from the specific model of the scene, all the experiments show a tendency which will be presented by the example of the scene in Fig. 3. After preprocessing, a total number of 2591 image primitives (dark lines in Fig. 3) are presented to the model based image analysis process. For 57 concepts of the specific model instances have to be created.

Tests made for different values of $\varepsilon$ show that for values $\varepsilon \in [0.05, 0.3]$ the instances belonging to the solution are identical. The instances belonging to the solution for these values of $\varepsilon$ form the white polygons in Fig. 3. For higher values of $\varepsilon$, a slightly different solution was found.

The total number of nodes and the number of leaf nodes of the search tree are measures for the search effort. The smaller the value of $\varepsilon$, the more the search
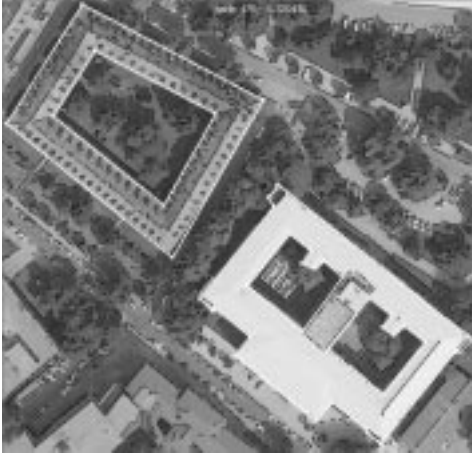
**Fig. 3.** Solution for $\varepsilon = 0.2$. White lines: solution; dark lines: image primitives.
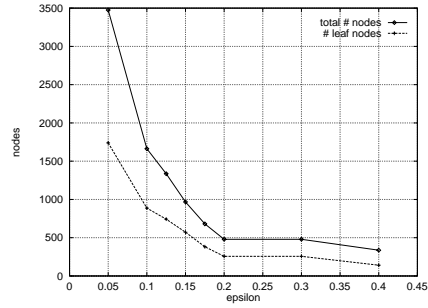
**Fig. 4.** Search tree characteristics as a function of $\varepsilon$

has a tendency to breadth: the number of nodes increases. The dependency of the total number of nodes and of the number of leaf nodes from $\varepsilon$ is graphically shown in Fig. 4. It is interesting to observe, that for values $\varepsilon \in [0.2, 0.3]$ the characteristics of the search path did not change. This behavior was also observed for other scenes.

## 5  Summary and Conclusion

We presented a method to derive a merit function for guiding search in a model based image analysis system. The Dempster–Shafer theory of evidence serves as a theoretical background. We have extended proposals found in the literature to suit our needs for propagating valuations calculated at different hierarchical levels of our model.

The derived merit function gives a common ground for the comparison of paths developed further with paths abandoned earlier in the search tree. A property of our merit function is, that by explicitly modeling the lack of knowledge with the methods offered by the Dempster–Shafer theory for the yet unknown instances and modified concepts, our formalism provides in a natural way the required overestimate for the merit of the yet unknown path from the current node to the solution node of the search tree.

The experiments have shown that our merit function can be used successfully to guide search with an $\varepsilon$–$A^*$–algorithm. The merit function is robust with respect to the parameter $\varepsilon$ and leads to a good solution for values of $\varepsilon$ up to a problem dependent upper bound. Higher values of the parameter $\varepsilon$ lead to a considerable speed up and smaller memory requirement of the analysis process.

Several other factors also contribute to the success of the analysis process,

like for example the valuations computed for the instances and modified concepts at the different levels of the hierarchical model. For defining these valuations we take advantage of having a specific model for the objects to be recognized. This specific model is automatically build by our system through the analysis of the available map of the scene. We are extending our system to recognize objects in the image which are not represented in the map, i.e. for which a specific model is not available.

# References

1. H.-P. Bähr, F. Quint, and U. Stilla. Modellbasierte Verfahren der Luftbildanalyse zur Kartenfortführung. *Zeitschrift für Photogrammetrie und Fernerkundung*, 63(6):224–234, 1995.
2. M. de Mathelin, C. Perneel, and M. Acheroy. Probability vs fuziness for knowledge combination in expert systems. In N. Piera Carreté and M.G. Singh, editors, *Qualitative Reasoning and Decision Technologies*, pages 652–661, Barcelona, 1993.
3. A.P. Dempster. A generalization of Bayesian inference. *Journal of the Royal Statistical Society, Series B*, 30:205–247, 1968.
4. M.L. Ginsberg. Non-monotonic reasoning using Dempster's rule. In *Proceedings of the AAAI-84*, pages 126–129, Austin, 1984.
5. M. Ishizuka, K.S. Fu, and J.T.P. Yao. Inference procedures under uncertainty for the problem reducing method. *Information Science*, 28:179–206, 1982.
6. F. Kummert, H. Niemann, R. Prechtel, and G. Sagerer. Control and explanation in a signal understanding environment. *Signal Processing*, 32:111–145, 1993.
7. H. Niemann, H. Brünig, R. Salzbrunn, and S. Schröder. A knowledge-based vision system for industrial applications. *Machine Vision and Applications*, 3:201–229, 1990.
8. H. Niemann, G. Sagerer, S. Schröder, and F. Kummert. ERNEST: A semantic network system for pattern understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):883–905, September 1990.
9. N. Nilsson. *Principles of artificial intelligence*. Springer, Berlin, 1982.
10. J. Pearl. *Heuristics: Intelligent search strategies for computer problem solving*. Addison-Wesley, Reading, MA, 1984.
11. F. Quint and M. Sties. Map–based semantic modeling for the extraction of objects from aerial images. In A. Grün, O. Kübler, and P. Agouris, editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 307–316. Birkhäuser, Basel, 1995.
12. F. Quint and M. Sties. Evaluating model fidelity in an aerial image analysis system. In *Proceedings of the 18th ISPRS-Congress*, Vienna, 1996.
13. G. Sagerer, U. Ehrlich, F. Kummert, H. Niemann, and E.G. Schukat-Talamazzini. A flexible control strategy with multilevel judgements for a knowledge based speech understanding system. In *9th International Conference on Pattern Recognition*, pages 788–790, Rome, 1988.
14. G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.
15. G. Vosselman. *Relational Matching*. Springer, Berlin, 1992.
16. J. Yen. GERTIS: a Dempster-Shafer approach to diagnosing hierarchical hypotheses. *Communications of the ACM*, 32(5):573–585, 1989.