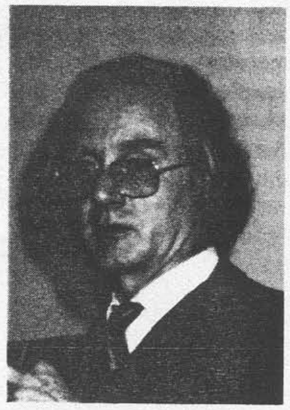
A future of photogrammetric research

by prof. dr.-ing. W. Förstner, Institut für Photogrammetrie, Universität Bonn.



SAMENVATTING

Een toekomst voor fotogrammetrisch onderzoek

De toepassing van computers voor de verwerking van beeldgegevens vereist een nieuwe definitie van de fotogrammetrische taken. De geometrische modellering bij aero-triangulatie en de fysische modellering in remote sensing moet worden ingebed in een semantische modellering van objecten die worden verkregen uit foto- en satellietbeelden. Dit artikel houdt een pleidooi voor een nieuwe theoretische basis en stipt enkele zaken aan bij de beeldinterpretatie die de kern vormt voor een wetenschappelijke evolutie in de fotogrammetrie.

SUMMARY

The computer's access to the image data requires a new definition of photogrammetric tasks. The geometric modelling in aerial triangulation and the physical modelling in remote sensing has to be embedded in a semantic modelling of the objects to be extracted from aerial and satellite images. This article wants to stress the urgent need for the development of a new theoretical basis and sketch some research issues in image interpretation being the key issue for a scientific evolution of photogrammetry.

Motivation

Photogrammetry is perceived as mensuration from photos. This holds both, for the internal perception, from surveying and geodesy, and for the external perception, from the various neighbouring disciplines. Photogrammetry and remote sensing cover two aspects of the same technology, photogrammetry being responsible for the geometric, remote sensing for the thematic part of information extraction. The inability to separate the used techniques in detail leads to a joint framework in research and in practice for using images of any kind for mensuration and mapping. Therefore the future perception of photogrammetry will replace the *photo* by a photon sensor, covering the classical analog photo, the scanned image, the video image but also other imaging sensors such as radar or laser range finders. The mensuration, *γραμειμ - gramein,* in photogrammetry is referring to, will be replaced by information extraction, covering the classical geometric information extraction in the form of object location and object reconstruction, but also object detection and the task of image interpretation. This generalization of the inherent tasks of photogrammetry and remote sensing is caused by the necessary technical tools, especially from computer science, pattern recognition or artificial intelligence, making photogrammetry and remote sensing a part of image understanding.
Now photogrammetry and remote sensing are mainly application driven disciplines. Mapping still is the central task. In spite of the digital techniques pushed by the availability of CCD-cameras, which make close range photogrammetry a booming area, this is valid due to the increasingly available image data, especially from satellites.
The information sources to be dealt with, when making maps, are manyfold: images, existing maps, data contained in Geographic Information Systems (GISs), object models etc. Taking the fascinating revolution in computer technology into account, the critical question therefore arises: 'If we had infinite computer resources at no costs, would we know how to make maps at no costs'? Certainly: 'No'. The reason is the *lack of a mapping theory* [79]. There are quite some indicators for this unfavourable situation:

— multi-spectral classification, being the only working automatic interpretation technique, still shows a too low accuracy;
— none of the commercially available Digital Photogrammetric Systems (DPS's) contains tools for semi-automatic or even automatic cartographic feature extraction;
— there is no commonly accepted definition of 'a map', 'the task of a DPS', or 'the functionality of a GIS'.

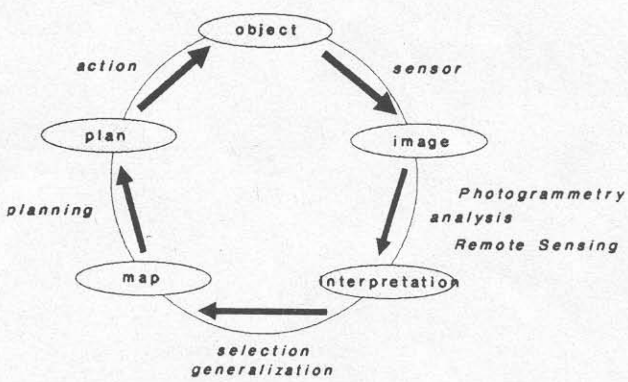


Fig. 1. The role of photogrammetry and remote sensing within surveying and mapping.

This article wants to give a framework for describing the questions to be answered when approaching the central tasks of photogrammetry and remote sensing. These may lead to a research program guiding the future photo-

[1]) Further processes are the *selection* of the parts of the interpretation in a GIS, the *generalization* in order to be able to adequately visualize the interpretation and the *planning* based on the interpretation leading to an action changing the object, the sensor (active vision, conform e.g. [3]) or other parts of the environment. These processes also need to be modeled and described, thus completing the loop in fig. 1. We will not discuss these aspects here.

grammetric research. This framework needs to be embedded into the activities of the surveying and mapping community. They can be described as an evolving cycle (fig. 1), where the object, namely our environment, including its topography, is observed, documented in maps used for planning, and changed to serve the ecological balance, or other needs. Photogrammetry and remote sensing here appear to be the technology for the automatic analysis [1]) of image data leading to an interpretation, thus uncovering the information about the objects for the use in (digital) maps.

Image interpretation

Image interpretation up to now still is mainly performed by human specialists. It is not well understood, though performed with great success. In order to be able (at least partially) to transfer this human capability to a machine a detailed modelling is required.

Machines only can handle symbols using some kind of specified formalism. *Semiotics* is the theory about handling symbols (fig. 2). It deals with the relations of symbols to other symbols *(syntax)*, to objects *(semantics)* and to the use of symbols *(pragmatics)*.

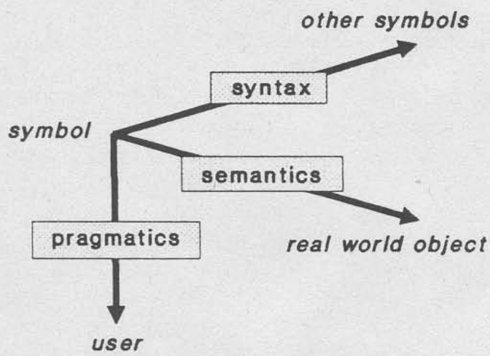


Fig. 2. Semiotics according to Bochenski [6]. Formalisms map semantics to syntax, requiring semantic integrity, i.e. proper meaning of all symbols and rules.

In our context the *pragmatics* reflects the different requirements of the users of photogrammetry and remote sensing and the different types of models to be used when analyzing image data. The same symbol, e.g. a bright spot, may have different context dependent interpretations: it may be treated as an outlier, as a building or as a factory depending on the model used for image analysis. The need to specify the purpose or goal of an interpretation task explains the urgent need for metadata, i.e. *explicit* description of the data. The necessity to describe context or relevance to a specific application is equivalent to specifying the pragmatics of an interpretation scheme.

Given a certain context specifying the *semantics* is identical with modelling the objects *and their appearance* [2]). This includes the description of all necessary relations between different objects, between objects and their parts and their change over time. This modelling has been realized to be the bottle neck for developing automatic interpretation systems. Existing interpretation systems contain more or less explicit object and image models, they, however, appear shallow compared to the models used by human analysts.

The mutual relations between the used symbols, i.e. the *syntax*, for describing objects, relations or images may be poured into grammatical or other rules. Such grammars may describe formalisms or (artificial) languages which we know from mathematics (theories) or programming (algorithms). The notion *semantics* is used in two closely related forms (fig. 3):

- the semantics of a symbol, a symbolic structure, a concept or a notion is their relation to a real object or structure. One also could say: the interpretation of a symbolic structure defines its semantics. This is semantics as part of semiotics [6];
- relations between concepts or notions, only existing in the model, are termed semantic relations as they restrict the semantic relation of the concepts or notions, which themselves are symbolic structures, to real objects or structures. This is semantics within knowledge representation [75].

We will use both aspects of semantics throughout this paper.

Within the semiotic framework using formalisms, realized as computer programs, needs a mapping of the semantics of the symbols to their syntax. This mapping requires a careful design such that all rules within the syntactic framework map to rules which hold in reality, in order to avoid inconsistencies and to keep semantic integrity.

The formal description of the used rules itself uses a language, i.e. a formalism with different symbols and its own rules. For example, in data bases this is the data description language. Therefore we at least have to distinguish three levels: the level of the real objects, which are no symbols, the level of the symbols (words) standing for the objects and their relations (object language) and the level of the formalism in which we describe the structure of the object language (meta-language).

Symbols, notions or words of the meta-language are not to be confused with symbols, notions or words of the object-language in order to avoid contradictions, conflicts or paradoxes. As the meta-language again has to be described we arrive at a multi-level description of the models we are using. The possibility to formalize knowledge obviously decreases with increasing level of abstraction finally leaving us with the natural language,

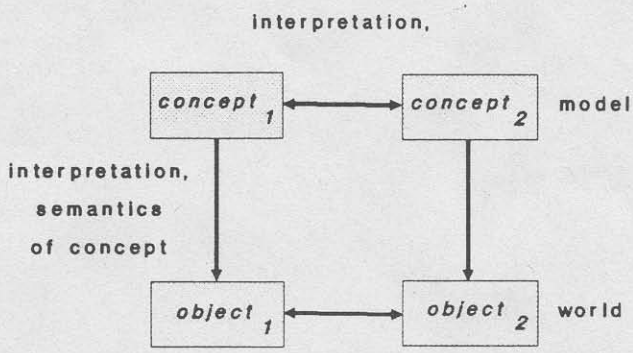


Fig. 3. Relations between concepts (notions) in a model and real world objects: relations between concepts are semantic relations, interpretations of concepts, i.e. their relations to real objects define the semantics of a concept.
***Warning:** In object oriented programming, but not only there, often concepts are treated as objects. Then objects in the model and objects in the real world have to be distinguished.*

[2]) This goes beyond the modelling in Geo-information systems where the appearance of objects is not required.

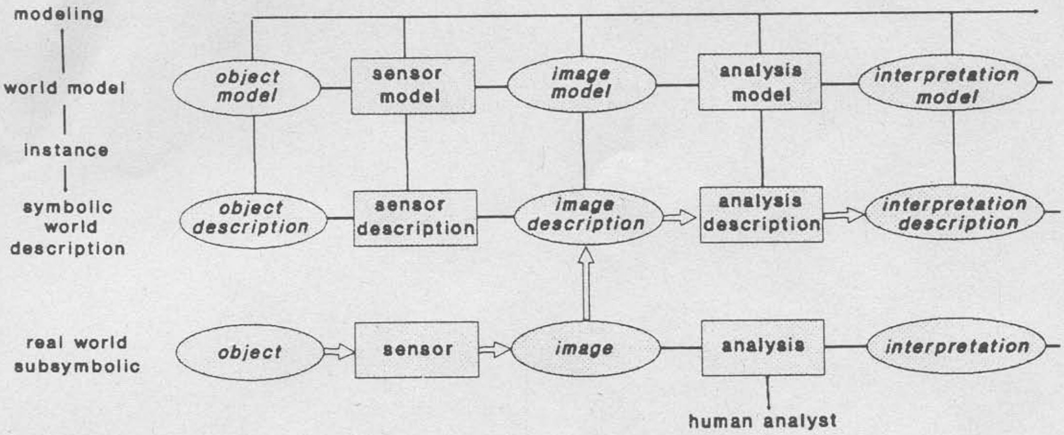


Fig. 4.
The different components of a system for image interpretation. The models (upper row) give rise to instances, i.e. descriptions (middle row) of the objects and processes in the real world (lower row).

which allows to construct new powerful models at every level, at the expense of weak possibilities of formalization.

The following generic model for the interpretation process wants to clarify the role of the different components of an image interpretation scheme. The image analysis part will be discussed in details below.

Generic model for image interpretation

The components of image interpretation are shown in fig. 4. There are three levels of description:
- real world, which is to be described. This level is sub-symbolic, in the sense that it is independent on whether we describe it by some means or not. Of course what is shown in fig. 2 is already a symbolic description on a meta-level as it describes the structure of the used models;
- symbolic description of the real world. Such descriptions are fixed numbers, symbols, relations, etc. which within a certain context defined by the models have an exact interpretation i.e. relation to certain objects in the real world.
- models of the real world. The symbolic descriptions are instances of models of the real world. The models thus can be seen to give the generic structure of the symbolic descriptions.

The observation process which is part of the loop of surveying and mapping activities discussed above now consists of the following three steps:
1. data taken with suitable sensors are collected in images. The design of sensors usually is based on some kind of model, especially with some application in mind. On the other hand — especially in remote sensing — often an analyst may be confronted with sensors or sensor data not specifically designed for his/her application, moreover where the description of the sensor may not sufficiently well be described. In this case the models of the sensor designer and the models of the analyst may not be coherent causing large efforts to find out the value of the sensor data for the specific task in concern.
   The images in all cases are analogue representations (in contrast to symbolic representations, see below), as the geometric relations between the objects are implicitly contained in the image and no whatsoever interpretation, i.e. labelling takes place.
2. the image data are processed to obtain a symbolic image description. *Symbolic* here stands for ,,symbolic representation'' [76] in contrast to the analogue (raster) representation of the sensor data. Any procedure for image processing or pattern recognition, often termed *segmentation* in fig. 4, may be used for feature extraction. In principle the symbolic description completely replaces the image. The structure and the complexity of the symbolic description depends on the image model which has to be derived from the object model and the sensor model. Therefore the image model is governed by the up to now not specified object model and the usually comparably simple sensor model.
3. the central part of the *analysis* consists of labelling the symbolic image description yielding an interpretation of the image [3]). The analysis process usually is hidden, both, when performed by a human or a machine, unless the process itself is analyzed. This analysis yields a description of the image analysis process whose structure is defined by the analysis model. This model, in addition to the object model, and the sensor model depends on the application defined by the user and the regularities which the developer or analyst assumes the system to follow.

Modelling the observation process thus consists of setting up the five models of fig. 4 and their mutual relations, namely the:
- object model;
- sensor model;
- image model;
- analysis model;
- interpretation model.

Object models

Object models contain geometrical, physical, biological, structural, semantical and possibly other elements. Relations between objects may be modeled by *specialisation* **(is-a)** or *containment* **(is-part-of)** hierarchies. Other relationships refer to time, or to causality between objects. Object models must be object centred in order to be invariant to the observation process used to infer the presence, form, class etc. of objects. This can only be

[3]) The notion ,,interpretation'' is ambiguous: meaning both, the process of labelling as well as the result of this process. We use the ,,analysis'' to name the process, though it suggests a too narrow and one sided process seemingly omitting the larger context. We use ,,interpretation'' for denoting the result of the analysis process, however, sometimes refer to both, the procedures as well as the whole analysis as ,,interpretation process''.

achieved in context with the used sensors which define the possible transformations when being observed. Thus invariants have to be identified, e.g. with respect to form, to topology, or other relations. Objects usually are assumed to be bounded by surfaces which are smooth nearly everywhere and which are linked by boundary lines which also are assumed to be smooth almost everywhere [23]. Many objects, as roads, fields, rivers, houses fall into that category. At special scales trees, river deltas or towns however are not bounded by smooth curves or surfaces, but show a fractal behaviour which then has to be transferred to the image model and taken into account when analyzing the images.

The above mentioned hierarchies play an essential role, as they are designed such that they are generic, i.e. invariant to viewpoint, sensor type etc. Different sensor types, especially having different spectral bands, observe different aspects of the same object. This on one hand does not necessarily influence the specialization hierarchy, but more likely the containment hierarchy. On the other hand in nearly all practical cases a physical, i.e. theoretical link between the responses of the different sensors can not be established, forcing the conceptual modelling to be phenomenological. Most problems occurring when fusing information of different sources therefore result from a lack of common invariants of the same object, mainly caused by the different classification schemes, i.e. specialization hierarchies within the different fields of application (see also the discussion in section about „meta-information and fusing objects models'', p. 380).

It is an open problem which representation is best suitable. This is because on one hand the semantic relations must be describable, the image descriptions must be derivable using the sensor model, but also the reconstruction of the form and the derivation of the relations must be realizable within the chosen representation. Multiple representation schemes seem to be unavoidable (e.g. raster and vector representations) but make a theoretical analysis at least cumbersome. Of course different object models may require different representations. The problem of linking different and possibly scarcely overlapping object models is the central problem also when linking heterogeneous data bases.
The object model also has to include the background, i.e. those parts of the scene which are not of primary interest (e.g. trees, when extracting highways). Object models for image analysis therefore have a much higher granularity than object models in GISs usually have [5] and also the footnote [2]).

Sensor models

Sensor models in our context solely contain geometric physical components [4]). Modelling the sensing process only requires a sufficiently dense geometrical/physical object description including all details namely form and reflectance properties, illumination sources, sensor models etc. Though the descriptions will be symbolic — because of the symbolic nature of the physical laws — the process itself is sub-symbolic.

[4]) In environmental monitoring of course chemical and biological sensors play a distinctive role.

Many problems in image analysis are caused by the discretization process. This in a first place refers to the raster structure of smooth boundaries which however in case of a suitable anti-aliasing may not influence the geometric recovery too much.
The main problem is the object-independent aggregation of the spectral information within one pixel (mixed-pixels). The pixel size usually is given. For satellite images its range is between a few meters (KWA 1000) and about 2 km (NOAA). The often spatially varying integration (blurring) effect of the point spread function needs to be known and explicitly modeled when performing feature extraction, as the measured intensities are the original observations for the whole interpretation process. Because of the aggregation being object-independent feature extraction without explicit reference to the geometric and physical components of the object model seems to be sub-optimal from the beginning. An approach trying to link sensor and object model for feature extraction is discussed in [23 and 54].

Image models

The image model forms the link between object model and interpretation model. It can be seen to be an appearance model [23, 50, 70 and 72]. The image model has to be derived from the object and the sensor model.
The above mentioned invariants (geometric, radiometric, texture parameters, topological) play a central role within the image model. As many objects are bounded by smooth surfaces or lines, the most commonly used image model consists of a segmentation of the image into non-overlapping partitions (see the collection in [77]). They show regular boundaries which themselves have to be modeled. The weakness of this image model reveals when the smoothness of the object surface and of the illumination interfere at abrupt albedo changes, causing the appearance of object edges, i.e. image edges to be broken, making a higher-level inference necessary for an appropriate interpretation in these cases.
The image model on one hand guides the feature extraction process thus should be of enough granularity, especially concerning the radiometric structures (shadows, shading, texture, specularities etc.). The hierarchical object structures usually are inherited by the image structures allowing to envoke grouping processes. On the other hand the image model requires a mapping onto a perceptual database [50], allowing an efficient access to image features of the usually numerous images.

Analysis models

The analysis model relies on the image model and strategic knowledge in order to arrive at an interpretation. Central point is the labelling of the image features which heavily depends on the semantic part of the object model, especially on the spatial relationships, the different materials and their appearance and of course the specialization and containment hierarchies set up in the object model.
The analysis may also require the use of sub-symbolic methods for describing shape from stereo [18, 37, 93], from shading [39] and from texture [8, 12, 45, 52, 55, 57, 64, and 88]. These methods often only work under special conditions (existence of texture, homogeneous reflectance) which even may contradict, requiring a link with symbolic methods for detecting structures of irregularity, thus interacting with the process of feature extraction.

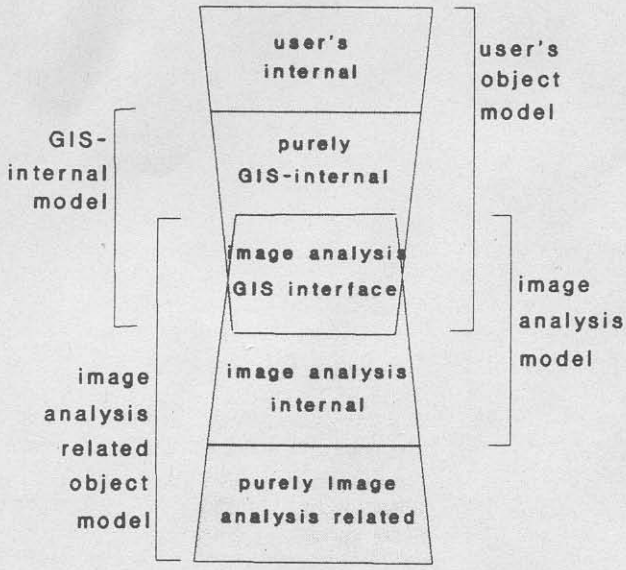


Fig. 5. The different parts of the object related models in image analysis and GIS-analysis: the image analysis related object model and the user's object model are linked by the Image Analysis-GIS-interface, it belongs both to the interpretation model and to the GIS-internal object model. The interpretation model as well as the GIS-internal object model both contain parts which are not relevant to the other model.

The analysis itself depends on user requirements which either are already (implicitly) contained in the object model, have to be provided during the analysis or result from the interpretation model, e.g. in form of a GIS-query. In these cases only a small part of the object model, e.g. the upper levels of a specialization hierarchy, may be available, i.e. formalized. This may prevent or at least hinder a smooth link with other information sources being available.
We will come back to this model.

Interpretation models

Interpretation models [5]) describe the result of the analysis step. It in principle contains all aspects of the object derivable from the images, and possibly other information sources, as far as the object model allows. Thus not only the user relevant parts of the scene are contained in the interpretation model but also all other parts which are necessary to complete the analysis, as e.g. parts of objects or background which the user is not interested in. This would imply all aspects of the object to be observable in principle, thus the interpretation model and the object model to coincide.
But much more likely the interpretation model will only cover parts of the object model namely those which a careful analysis of the design has found to be observable and excluding those parts of the object model which are only needed within the analysis procedures, e.g. all kinds of preprocessing steps. This model then can be seen to form the interface between image analysis and GIS-application (fig. 5).
This interpretation model moreover has to be distinguished from the (object) model used in a GIS, the whole analysis process is designed for. This GIS-internal object model on one hand will not contain irrelevant detail visible in the scene, which might be necessary for the analysis itself, and on the other hand will contain aspects of the scene which may not be relevant to the image analysis but only to the GIS-user. The different parts of the object related model are shown in fig. 5.
The description of the interpretation process up to now did not explicitly include prior knowledge not being of the form of an object model. Specifically existing maps or other knowledge sources are not mentioned. However maps or even non geometry related information may also be seen to be objects which have to be observed, from which features or signatures have to be derived which require a symbolic description and have to be analyzed [6]). All these information sources may be contained in a GIS. Therefore not only play photogrammetry and remote sensing the role of a special data source for GISs but also vice versa: GISs play the role of a general modelling tool for analyzing image data providing the interpretation tools for image analysis. This requires a smooth interaction between the different models discussed above and a much more detailed analysis of the task of a GIS.

Fusing information during image analysis

Zooming into the analysis model reveals several information and knowledge sources to take crucial parts in the analysis process (fig.6). Those are at least images, image models, non-image data, strategic knowledge and the human analyst.
Formally the fusion process to be discussed may be interpreted as inverting the observation equations: $o_i = f$ (unknown parameters p; image models; non-image data), $i = \ldots$ to yield $p = f^{-1}$ (o_i; image models; non-image data; strategic knowledge; human analyst), where the parameters include geometric, physical parameters, classes, relations etc. Whereas the relations between the different models have been discussed before this paragraph deals with the inherent complexity and therefore difficulty of the analysis process. It is caused by:

- the incomplete knowledge about the observation function f, e.g. when modelling the reflectance function;
- the incompatibility of the object models resulting in the incompatibility of the image models, e.g. when linking optical and radar images;
- the heterogeneity of the non-image data, specifically of maps, e.g. resulting from different purposes, different age or different scale;
- the hitherto unstructured strategic knowledge, e.g. when trying to reduce the inherent high computational complexity of matching algorithms and the nearly unformalizable capabilities of the human analyst.

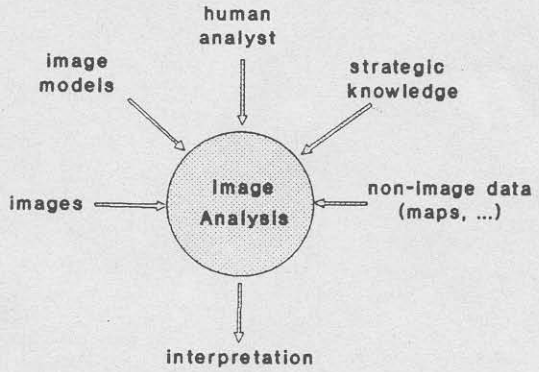


Fig. 6. Some of the information sources to be fused when analyzing image data.

[5]) Recall that we use the notion interpretation to describe the result of the analysis, thus meaning a state rather than a process.

[6]) Analyzing and interpreting existing maps actually follows the same steps and requires the same type of models.

The partial success of image interpretation systems results from the severe simplifications of the used models on one hand and from the excellent engineering capabilities of system designers on the other hand.
The different approaches for fusing these information sources may be classified depending on:
- whether the data is represented in raster or vector form, or equivalently, whether the data structure already reflects the structure of the object or not;
- whether the semantics of the fused information is used explicitly or only implicitly.

We thus may distinguish signal based, property based, feature based and object based information fusion (fig. 7).

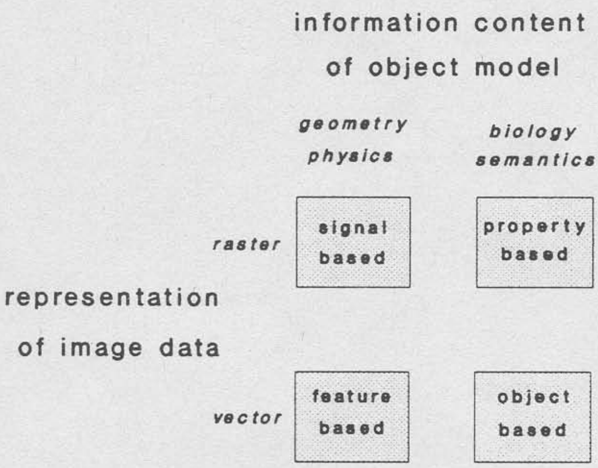


Fig. 7. The different types of fusing information.

Signal based fusion is characterized by the raster structure of the used data. Any type of image processing technique may be used here to obtain raster oriented attributes. No explicit reference to the semantics is assumed to be necessary.
Property based fusion uses the original or derived raster data together with their meaning derived by some statistical classification procedures. Thus local properties of the objects and possibly the relations between these properties are used.
Feature based fusion is characterized by the structural description of the image, e.g. lists, graphs or relational descriptions, including the attributes linked to the features or relations. No direct reference to the meaning of the features is assumed at this level, though the feature extraction in general will be guided by the scope of the interpretation.
Object based fusion relies on the semantics of real world objects or their models, either in the object model or in the interpretation model. Thus it is related to the relationships between the objects or between non-local properties of the objects to be extracted.

We do not want to discuss the techniques or the systems as such, which operate on these types of information. We will, however, concentrate on the ability of the differently structured approaches to fuse information dependent on their various types.

Signal based information fusion

Fusing information on the signal level is motivated by the possibility to formally invert the above mentioned observation process using some kind of least squares techniques in case geometric/physical properties are to be derived from the sensor data. This is the reason for the break through in digital photogrammetry with the object centred surface reconstruction schemes by [37 and 18]; also see [36 and 93]. The shape-from-techniques, being a special case of these approaches, therefore belong to this class as long as they lead to an iconic description of the object, e.g. a raster DTM.
As the information about the object used in these approaches is purely geometrical or physical, the resultant surface form and reflectance parameters may be the basis for a more deeper, property based information extraction in case the geometric and physical information can be related to the objects to be recovered.

Property based information fusion

Fusing information for image interpretation based on properties of the objects visible in the images is the most common and intuitive technique [19]. It is motivated by:
- ease of maximum likelihood classification applied to the channels of multi-spectral images which due to the design of the sensors (approximately) refer to the same object position;
- dominance of the spectral features for identifying object classes in case of low resolution images ($\geq$ 30 m pixel size) where geometric features play a secondary role.

The reason for the still great success of pixel based classification schemes are the increasingly advanced techniques to reduce the signal i.e. data values to parameters describing the object related, i.e. invariant reflectance properties. Reductions include sensor calibration, atmospheric corrections, influence of terrain aspect and illumination direction etc. Rastered maps may be used as additional channel [49]. An increase of the classification accuracy (probability of correct classification) is expected from taking the local context into account. This on one hand may be achieved by using parameters derived from a certain neighbourhood, reflecting texture parameters such as variance, average orientation, gradient magnitude or local power spectra. In contrast to the use of the (reduced) intensity values themselves no strong physical models are available which motivate the selection of proper texture parameters.
The result of the pixelwise classification usually shows unfavourable irregularities especially at the borders of otherwise homogeneous areas [40, 58 and 82]. In order to achieve cleaner results often a post processing is applied e.g. by replacing the class of a pixel by the majority in a 3 by 3 neighbourhood, which of course is an ad hoc procedure [42 and 59].

A more rigorous model based technique are hidden Markov random fields (HMRF) [10, 27 and 69]. The relation of neighbouring class labels is modeled using their conditional probabilities. Also line processes may be included to obtain smooth region boundaries. The techniques have been extended to allow the integration of multiple sensor data, of data from different times or even of map data.
An example is the approach given in [60]. There an integration multi-temporal images on the pixel level, by actually using surface elements, are derived from rectification. The model explicitly applies the temporal relations using probabilities of crop rotation, i.e. about the temporal relation of the local object properties between the fields the surface elements belong to. Therefore explicit refer-

ence to the semantics of the classified elements is made via the appearance model of the fields. The increase in classification accuracy was significant.
In all cases a registration with respect to a common geometric frame is necessary. This requires some type of resampling both for image and for map data. Common practice is to use nearest neighbour resampling in order to leave the measured data unchanged.

A closer analysis reveals this approach to be inappropriate, moreover it uncovers the deficiencies of the pixel based classification techniques:

- the underlying model is image, not object based thus not natural. Modelling line processes in the HMRF approach relates to the crack edges which are only crude approximations of the edges of the objects;
- it is by no way clear how to model the resampled intensities derived from sensors having different resolutions, as the original intensities approximately are integrations of the object centred reflectance function with the point spread function of the sensor. Finally there is no way to solve the „mixed pixels'' problem without explicit reference to some true or estimated object boundaries;
- pixel based classification schemes do not allow to introduce a larger context. Though the HMRF approach theoretically is able to model dependencies of classes over a larger range the modelling is done implicitly. E.g. the straightness of boundaries can not be expressed. Even larger contexts such as hierarchical (containment) or topological (road) structures cannot be handled at all. With increasing resolution (≥ 10 m pixel size) the amount of information contained in geometric structures increases which cannot be captured by a pure pixel based approach;
- the fusion of images taken at different times has to track the spectral responses of each pixel over time. This introduces an additional instability in the classification as the causing (e.g. growth) processes usually are correlated between neighbouring pixels. Though this also may be modeled using Markov random fields the computational effort is high in case a certain rigour in the modelling is aimed at.

These critics of course only hold for automatic interpretation schemes. The methods may very well be used for getting approximate classification results, or for supporting manual interpretation, e.g. [2, 35 and 90] or in case the pixel resolution compared to the size of the objects is sufficient and the objects are highly homogeneous.
The discussion, however, wanted to show the pixel based methods to have severe disadvantages in case information of different sensors or maps has to be merged. One way to avoid these deficiencies is to base the interpretation on larger, aggregated units.

Feature based information fusion

Symbolic descriptions of the image may have any level of abstraction and then may always be made equivalent to the aggregation level of the object model, provided a high enough resolution of the images is available. We may distinguish at least three levels of abstraction:

- lowest representation level is characterized by lists of basic elements, namely attributed points, edges and/or regions;
- medium representation level in addition contains attributed relations between the basic elements;
- highest representation level consists of further aggregated basic elements which may result from a grouping process.

In all cases we do not assume the semantic aspect to play the central role, i.e. no interpretation has taken place. However the selection of the criteria for extracting features, their relations and possibly their grouping may very well be guided or at least motivated by the image model, which itself contains information about both the structure and the meaning of the different components of an image. Thus feature extraction may be performed bottom-up on the complete data set or top-down depending on a request of the analysis system or the image analyst. The main task of fusion on the feature level again is the registration and rectification of the image data.

List of points, lines and regions

The features easiest to extract in remote sensing images are points, lines or regions and their attributes such as type of point (T-, Y-, or L-junction, blob) or type of line (edge, dark/light line, strength, contrast) or type of region (round, rectangular, polygon shaped) etc. The advantage already of these low-level features is their invariance to a great variety of transformations, i.e. observational situations. Especially their geometric properties and a great number of their attributes remain invariant to lighting and sensor conditions, e.g. edges (between fields), lines (representing roads) or homogeneous regions (caused by lakes) appear very similar over time and may even be linked with map data [31]. This enables at least a geometric link between different images and possibly map data.
In case the data are already geocoded a link with map data appears to be promising, as the aggregation especially of lines and regions may be guided by map information. In case no or poor map data are available, geocoded features of different image sources may be used for grouping, thus supporting each other to arrive at a higher level structured description, which is richer than the ones derived from the individual images. The drawback of these features is obvious: Stable features only cover a small percentage of the scenes to be analyzed. Most features to be extracted are unstable in existence, geometry, and attribute. Except for very clear lines (representing line type objects or boundaries of regions) a fusion of different information sources is extremely difficult in case no relations between these features are used to increase mutual consistency with the image model.
An example for the fusion of image data using these low level features is the registration based on straight line segments for lining image and map data [80] for orientation.

Relational descriptions

Relations between the low level features increase the strength of the representation. The same relations as in GIS-modelling could be used [65] in case a complete representation of the images can be achieved. As this may not be feasible without knowledge about the object classes shown in the image, also weaker relations (near to, possibly connected etc.) may be used.
The fusion of different images may rely on the similarity of the relational descriptions in order to arrive at a more complete relational description. There seem to be only few examples of using relational descriptions without

reference to the semantics of the descriptions (see the paragraph on object based information fusion). An example is given in [89] using relational description for locating control points, derived from a topographic map. In case no map data are available any type of segmentation and the corresponding relational description will have to cope with the problem of observability of complete segments: many object parts do not show in image segments due to reason of lighting, common use, common or similar reflection properties or weak radiometric resolution of the sensor. Other problems are caused by textures where boundaries are even more unstable/uncertain. In case a generic model about the region boundaries is available (e.g. land-use units have polygon-like boundaries) this might be used to improve segmentation and thus fusion of different information sources.

Higher level aggregates

The containment hierarchies of most objects to be extracted from images give rise to several levels of grouping, lines and regions representing the lowest level. Often such groupings can be defined without explicit reference to the semantics of the objects. Examples are field structures, ribbons of streets, symmetries of buildings, sequences of trees at avenues or in plantations. The feasibility of such groupings has been shown in [64] for matching and for building extraction. An example for extracting the structure of land-use areas is given in [70]. It is essential that such structures can be derived without explicit object model but of course could be guided by such models. Obviously the boundary to techniques which perform the fusion on the object level is not clear cut, as it is the geometric and topological structure of the objects which is mapped to the structure of the image features.

Object based information fusion

The strength of fusing information on the signal and the feature level results from the rigour of the used geometrical, physical and/or structural models. The fusion is performed bottom up possibly leading to hypothesis of high level aggregates. The semantic relations between the parts of the objects in concern are only used implicitly, e.g. when searching for antiparallel edges for road extraction in case this search is part of the standard feature extraction procedure.

In case the *semantics* of the objects, not only of their local properties, is used *explicitly for fusing information* we refer to object based fusion. We may distinguish different knowledge sources, namely the human analyst, interpreted map data and generic models. They obviously are different with respect to flexibility, adaptability, automation potential, quality or suitability. All of them may be used within the analysis process either individually or integrated.

In addition we therefore have to distinguish two cases:

- there is only a single knowledge source containing the object model. Then this model normally is used to guide the image analysis top down;
- there are several knowledge sources containing the object model or parts of it. Then besides guiding the image analysis it is necessary to fuse the parts of the object model into one coherent model.

Both cases are discussed separately.

Image analysis with a single object model

In case we only use a single object model as the basis for image analysis we in a first instance can assume the model defining the context of the analysis in a unique way. This of course does not imply the resultant interpretation to be unique as e.g. sensor information may not be sufficient.

Having a *human analyst* defining the context of the analysis still provides highest flexibility and adaptability when integrating different information sources. This is due to the background and common sense knowledge which is used, especially when coping with long term monitoring where explicit object models are not available, e.g. in case political, social or historical knowledge is required, or this knowledge is just too weak. Humans are also very flexible in selecting analysis strategies, especially when coping with incomplete or seemingly contradicting information. The coding and use of common sense knowledge is a classical topic in AI but still waits for results applicable in practice [34 and 53].

The least flexible but useful and available object models are coded in *maps.* We assume a map interpretation is performed, either off-line in advance like in a GIS or available during image analysis, such that the necessary attributes and relations are accessible. Whereas the legend of a map contains information about object classes, the map itself describes an *instance of a complex object.* It does not contain generic information about topographic features, say.

This special model, however, may already be very useful, as sensor and map data are in some way complementary: maps describe mainly geometric aspects of the objects, including topological relations. They only give access to the highest level of the specialization hierarchy and contain names, i.e. user readable keys which may provide links to other information sources. On the other hand sensor data provide up to date information about object parts and — in case of multi-temporal data — time series of object attributes which may directly be linked to the objects in the map [43 and 44]. Thus using maps as the only knowledge source containing an object model may very well be useful for monitoring in case the attributes of well defined object (instances) have to be observed over time.

The inflexibility of maps reveals when changes of the objects occur, be it geometrical, physical, or semantic changes, or when the context the map was made for is no longer valid. Then a new map is required, at least partially. Thus the problem of map update cannot be solved without reference to an object model containing semantic relations between classes and their change over time, which either requires a human analyst or generic models.

The term *generic model* here is used for object models which are formalized statements about object classes and possibly instances. Generic models thus describe the objects, their possible attributes and their mutual relations (see the paragraph on object models).

Typical representations of generic models are the entity-relation-model [83] in the area of data bases, semantic networks [68] and frames [61] in the area of AI, grammars in formal languages [25] or object oriented techniques [15]. Rules may be used to represent actions, changes or other operations on the above mentioned representations.

An example where a generic model in the form aggregated features are used as a basis for an interpretation is given in [26], describing buildings with flat roofs by homogeneous areas with boundaries consisting of polygons with rectangles only. Then local groups of edges, their mutual relations (rectangular) and their relation to the interior of a region is used to generate building hypothesis. The model is local, in the sense that the relations to neighbouring segments are not incorporated. The final decision on the classification is left to the human analyst. We do not want to evaluate these techniques with respect to object modelling in image analysis. Actually all representations are used frequently [9, 11, 16, 32, 38, 50, 67, 70, 85, 86 and 87]. Moreover they may serve as a sufficiently good description of the object model in not too complex circumstances.

We only want to comment on two problems which one are faced when applying generic models to image analysis:

1. there is no commonly accepted *methodology to formalize knowledge,* here to establish object, image, analysis or interpretation models. The problem of knowledge acquisition is as old as AI. There is no clear way to avoid the impression of ad-hoc-ery (which may be not so different also in hard sciences). The appropriateness and the predictive power of the model, being a kind of theory, still is the only way to evaluate the quality of a chosen model. The advantage of rule-based descriptions, also useful for representing semantic networks of frames, is their flexibility and their expressive power, especially in case non-specialists have to use image analysis systems [7].
   The problem how to represent uncertainty of data, relations or classes in a coherent way is not solved yet, especially if hard decisions during the analysis should be avoided and the quality of image analysis procedures has to be evaluated by comparing the result of empirical tests with theoretical predictions. This is due to the lack of an observation theory for fuzzy measures and belief functions though their mathematics can be related to the classical probabilistic techniques [1, 47 and 84].
2. the *computational complexity* of the analysis procedures in principle is prohibitively high, as the task of labelling image features is NP-complete [7]). All techniques to circumvent this problem therefore are suboptimal. Classical examples are decision trees [4], which require hard decisions, clique formations, which try to balance the influence of the local and the global context [85] or techniques based on perceptual grouping, which exploit the containment hierarchies within the object models [14 and 64]. Parallel techniques are relaxation [41 and 62] which under certain conditions guarantee to find the global optimum of an interpretation.

Meta-information and fusing object models

The use of a single object model may be sufficient in some specialized cases. If two or more models of the object are used within the same image analysis procedure the generic problem arises how to fuse different object models. This problem is even more complex than using object models in a knowledge based interpretation system, as the *semantic consistency* between the different object models has to be provided. In our context there are several cases of pairing object models (including examples):

— analyst[1], analyst[2] (say a geologist and a soil scientist);
— analyst, map [42 and 48];
— analyst, generic model [26 and 32];
— map[1], map[2] (e.g. having different scale and/or age);
— map, generic model [11 and 63];
— generic model[1], generic model[2].

Having two generic models is the one which is equivalent to having two data-models in federated data bases and is the case where the *problems* of model fusion are best visible [46, 81 and 88]:

1. semantic equivalence. Two concepts refer to the same real world object. This is the most simple problem, E.g. is the point type object in one map the same real world object as the area type object in the larger scale map of the same area? Scale differences usually lead to different granularity of concepts (see below). Sometimes it may be possible to establish semantic equivalence automatically. Another example is the differently looking but semantically equivalent result of two analysis processes;
2. semantic heterogeneity. This will be the most frequent situation. Examples for semantic heterogeneity are manyfold:
   — different context of concepts, e.g. the location of an object may be described by 2 or by 3 co-ordinates;
   — different representation, e.g. an area may be represented in raster or vector format [66 and 92];
   — different cardinality of relations, e.g. the relation *visible-from* may be defined as a 1 : n or as a m : n relation;
   — different granularity, e.g. a containment hierarchy is only designed for GIS-analysis not for image analysis purposes, neglecting details being not essential for GIS-applications but essential for automatic image interpretation;
   — unit differences, e.g. feet versus meters.
   Semantic heterogeneity thus is related to different representations which, however, refer to similar aspects of the same object;
3. semantic discrepancy. This is the hardest case as the same concept/class in different knowledge/information sources refers to different real world objects. A typical example is the notion 'road' used in different analysis procedures, e.g. in case the minimum and maximum widths differ. This example also reveals that such discrepancies may occur within the same environment, in case it changes over time, which is the normal case in program development: The notion 'road' today is not the notion 'road' tomorrow in case the algorithm changed.

Semantic consistency only can be achieved if *meta-data* are available. Meta-data describe the interpretation of the data thus are part of the representation (see above). But meta-data usually are not accessible by the program which works with the data but (hopefully) contained in the documentation. A first step to improve semantic consistency is the paradigm of object oriented programming, where changes of objects only can be performed within specified ranges.
The necessity to explicitly store the semantics of knowledge, however, was already stressed in the 70th [13 and

[7]) A problem is called NP-complete (non-polynomial-complete) if its algorithmic complexity is larger than polynomial, i.e. the computation time grows faster than any polynomial, e.g. exponentially.

33]. The recent developments in worldwide environmental information systems make meta-data an urgently necessary tool to exploit the value of these data [74].

The advantages of an explicit representation of the data- and program structures are obvious [91]:
— simplification of program control;
— ability for abstraction, due to the availability of specialization hierarchies;
— ability to explain reasoning processes;
— simplification of knowledge acquisition;
— better software maintenance.

In our context the problem of achieving semantic integrity may be solved by several means [88]:
— making meta-information explicit already increases the transparency of image analysis procedures. Meta-information in a first step consist of data about: time of measurement, accuracy, source, derivation formula, history, etc. It may be stored as text or explicitly coded which then allows direct manipulation by an analysis program;
— data - meta-data synchronization aims at updating meta-data together with changes in the data- or object model;
— meta-data comparison may be used to find semantic equivalences or heterogeneities and possibly resolve conflicts. This of course requires a formalized representation of the meta-data e.g. [28 and 29];
— meta-data generation may help keeping the meta-data consistent and up to date. This immediately is available for meta-information such as time of creation or accuracy of results;
— relevance evaluation may be used to assess the sensitivity of reasoning results with respect to certain model assumptions and may be performed on the meta-data. This is equivalent to algebraic (i.e. theoretical) derivations of the sensitivity of mensuration designs, here performed on more complex image analysis procedures.

This discussion wanted to stress the importance of explicitly storing the object, image, analysis and interpretation models in a way such that they are accessible by computer programs [76]. This may be a long term goal, but it needs to be approached soon in order to increase the transparency of the image interpretation results and to make them more useful for other contexts.

A future

No conclusions can be drawn, no future can be predicted: both would suggest solutions to be available for solving the difficult problems sketched in this paper. We, however, can take possibilities and opportunities future essentially consists of. A possible future of photogrammetric research may be guided by the following goals:
— image interpretation can be taken to be the central issue in photogrammetric research. This involves the extensive use of tools from computer graphics for forward modelling the imaging process, from image understanding for modelling the analysis process and from the research in man-machine-interfaces, for modelling the cognitive aspects of computer programs and human analysts and their interaction;
— semantic models seem to replace the geometric/physical models dealt with during the last 100 years of photogrammetric research. They need to explicitly contain knowledge about user needs and of meta-information of all kind. Languages for smoothly interacting with object, analysis and perception models seem to be a possibility to adequately represent this knowledge;
— the integration with the neighbouring disciplines appears to give the right boundary conditions for a fruitful development. All aspects need to be covered: applications (geography, geology, ecology, . . .), basic sciences (physics, computer science, cognitive science, control, . . .) and techniques (pattern recognition, computer vision, artificial intelligence, . . .). It seems to be high time to realize this integration not only in scientific conferences or meetings, but in joint projects, interdisciplinary research teams and — last but not least — in the corresponding curricula.

Our field of research becomes more and more exciting. We should grasp the fantastic opportunities in order to provide tools for solving the huge and urgent problems.

Literature

1. Abdulghafour, M., J. Goddard, M. A. Abidi, *Non-deterministic approaches in data fusion — A review.* SPIE vol. 1383, 1990, Sensor Fusion III: 3-D Perception and Recognition.
2. Albertz, J., K. Zelianeos, *Enhancement of satellite image data by data cumulation.* ISPRS Journal of Photogrammetric Engineering and Remote Sensing, 1990, p. 161 - 174.
3. Aloimonos, J. Y., A. Bandyopadhyay, *Active vision.* International Conference on Computer Vision, London, 1987.
4. Anuta, P. E., L. A. Bartolucci, M. A. Dean, *LANDSAT-4 MSS and Thematic Mapper data quality and information content analysis.* IEEE Transactions on Geo-science and Remote Sensing, vol. 22, 1984, no. 3, p. 222 - 236.
5. ATKIS, *Amtliches Topographisches-Kartographisches Informationssystem.* Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV), Hannover 1989.
6. Bochenski, I. M., *Die zeitgenössischen Denkmethoden.* Francke Verlag, Berlin, 1954.
7. Bolstad, P. V., T. M. Lillesand, *Rule-based classification models: flexible integration of satellite imagery and thematic spatial data.* Photogrammetric Engineering and Remote Sensing, vol. 58, no. 7, July 1992, p. 965 - 971.
8. Bovik, A., M. Clark, W. S. Geisler, *Multichannel texture analysis using localized spatial filters.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 1, 1990, p. 55 - 73.
9. Bruce, A., R. T. Draper, J. B. Collins, A. R. Hanson, E. M. Riseman, *The schema system.* International Journal on Computer Vision, vol. 2, no. 3, 1989, p. 209 - 250.
10. Busch, A., K.-R. Koch, *Reconstruction of digital images using bayesian estimates.* Zeitschrift für Photogrammetrie und Fernerkundung 1990, no. 6, p. 173 - 181.
11. Cleynenbreugel, J. van, F. Fierens, P. Suetens, A. Oosterlinck, *Delineating road structures on satellite imagery by a GIS-guided technique.* Photogrammetric Engineering and Remote Sensing, vol. 56, no. 6, p. 893 - 898.
12. Conners, R. W., C. A. Harlow, *A theoretical comparison of texture algorithms*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 2, no. 3, 1980, p. 204 - 222.
13. Davis, R. J., B. G. Buchanan, *Meta-level knowledge: overview and applications.* Proceedings of the International Conference on Artificial Intelligence, Cambridge MA, 1977, p. 920 - 927.
14. Dickinson, S. J., A. P. Pentland, A. Rosenfeld, *From volumes to views: an approach to 3D-object recognition.* Computer Vision, Graphics and Image Processing: Image Understanding, vol. 55, no. 2, 1992, p. 130 - 154.
15. Dittrich, K. R., U. Dayal, A. P. Buchmann, *On object-oriented database systems.* Springer, 1991.
16. Dori, D., *A syntactic/geometric approach to recognition of dimensions in engineering machine drawings.* Computer Vision Graphics and Image Processing, vol. 47, 1989, p. 271 - 291.
17. Driesch, M., M. Westhoff, *Zuordnung von Fernerkundungs- und Kartendaten.* Diplomarbeit, Institut für Photogrammetrie der Universität Bonn.
18. Ebner, H., D. Fritsch, E. Gillessen, Ch. Heipke, *Integration von Bildzuordnung und Objektrekonstruktion innerhalb der digitalen Photogrammetrie.* Bildmessung und Luftbildwesen, vol. 55, no. 5, p. 194 - 203.

19. Ehlers, M., *Multisensor image fusion techniques in remote sensing.* ISPRS Journal of Photogrammetric Engineering and Remote Sensing, vol. 46, 1991.
20. Förstner, W., *Quality assessment of object location and point-transfer using digital image correlation techniques.* International Archives of Photogrammetry and Remote Sensing, Rio de Janeiro, Comm. III, vol. 25, part A3a, p. 197 - 219.
21. Förstner, W., *Object extraction from digital images.* Schriftenreihe des Institut für Photogrammetrie der Universität Stuttgart, Heft 15, p. 269 - 279.
22. Förstner, W., *Least squares matching.* Haralick/Shapiro (Editors) Robot and Computer Vision, vol. 2, Addison Wesley, 1992.
23. Förstner, W., *Feature extraction for digital photogrammetry.* Photogrammetric Record, vol. 14, no. 82, october 1993.
24. Förstner, W., St. Ruwiedel, *Robust computer vision.* Wichmann, Karlsruhe, 1992.
25. Fu, K. S., *Syntactic pattern recognition and applications.* Prentice Hall, Englewood Cliffs, NJ, 1982.
26. Fua, P., A. J. Hanson, *Objective functions for feature discrimination: theory.* Proceedings DARPA Image Understanding Workshop, 1989.
27. Geman, S., D. Geman, *Stochastic relaxation, Gibbs distribution and the bayesian restoration of images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, p. 721 - 741.
28. Geoffrion, A. M., *The SML language for structured modelling: levels 1 and 2.* (to appear in Operations Research).
29. Geoffrion, A. M., *The SML language for structured modelling: levels 3 and 4.* (to appear in Operations Research).
30. Günther, O., H. J. Schek, (Editors), *Advances in spatial databases.* Lecture Notes in Computer Science 525, Berlin, 1991.
31. Guindon, B., *Multi-temporal scene analysis: a tool to aid in the identification of cartographically significant edge features on satellite imagery.* Canadian Journal of Remote Sensing, vol. 14, no. 1, May 1988, p. 38 - 45.
32. Hanson, A. J., L. H. Quam, *Overview of the SRI cartographic modelling environment.* Proceedings Image Understanding Workshop, Cambridge, Mass., 1988, p. 576 - 582.
33. Hayes, P. J., *Some problems and non-problems in representation theory.* Proceedings AISB, Summer Conference, University of Sussex, 1974, p. 63 - 79.
34. Hayes, P. J., *The second naive physics manifesto.* Formal Theories of the Common. (J. R. Hobbs and R. C. Moore, Editors).
35. He, G., J. Jansa, *Eine radiometrische Anpassungsmethode für die Mosaikherstellung aus digitalen Bildern.* Zeitschrift für Photogrammetrie und Fernerkundung, 1990, vol. 58, no. 2, p. 43 - 49.
36. Heipke, Ch. *A global approach for least squares image matching and surface reconstruction in object space.* Photogrammetric Engineering and Remote Sensing, vol. 58, no. 3, p. 317 - 323.
37. Helava, U. V. *Object-space least-squares correlation.* Photogrammetric Engineering and Remote Sensing, vol. 54, no. 6, p. 711 - 714.
38. Herman, M., T. Kanade, *The 3D-mosaic scene understanding system: incremental reconstruction of 3D scenes from complex images.* Readings in Computer Vision, (Fischler and Firschein, Editors) Kaufmann, 1987, p. 471 - 482.
39. Horn, B. K. P., M. Brooks, *Shape from shading.* The MIT-Press, Cambridge, Massachusetts, 1990.
40. Hoyano, A., Y. Komatsu Y., *Influence of mixels on land cover classification in residential areas using airborne MSS data.* International Archives of Photogrammetry and Remote Sensing, Kyoto 1988, p. VII - 399.
41. Hummel, R. A., S. W. Zucker, *On the foundations of relaxation labelling processes.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5., 1983, p. 267 - 286.
42. Hutchinson, C. F., *Techniques for combining LANDSAT and ancillary data for digital classification improvement.* Photogrammetric Engineering and Remote Sensing, vol. 48, no. 1, p. 123 - 130.
43. Janssen, L. L. F., M. N. Jaarsma, E. T. van der Linden, *Integrating topographic data with remote sensing for land cover classification.* Photogrammetric Engineering and Remote Sensing, vol. 56, no. 11, November 1990, p. 1503 - 1506.
44. Janssen, L. L. F., J. van Amsterdam, *An object-based approach to the classification of remotely sensed images.* Proceedings IGARSS-Symposium 1991, Espoo, Finland.
45. Jarczyk, G., *Untersuchungen zur Texturanalyse.* Diplomarbeit, Institut für Photogrammetrie der Universität Bonn.
46. Kent, W., *The breakdown of the information model in multi-database systems.* Sheth 1991, p. 10 - 15.
47. Klir, G. J., T. A. Folger, *Fuzzy sets, uncertainty and information.* Prentice Hall, 1988.
48. Konecny, G., *Der Einsatz von Fernerkundungsdaten in GIS.* Fernerkundung in der Forstwirtschaft (Oesten, Kuntz and Gross, Editors), Wichmann, Karlsruhe.
49. Laan, F. B. van der, *Improvement of classification results from a satellite image using context information from a topographic map.* International Archives of Photogrammetry and Remote Sensing, Kyoto 1988, VII - 98.
50. Lawton, D. T., T. S. Levitt, C. McConell, J. Glicksmann, *Terrain models for an autonomous land vehicle.* Readings in Computer Vision, (Fischler and Firschein, Editors) Kaufmann, 1987, p. 483 - 491.
51. Leclerc, Y. G., *Constructing simple stable description for image partitioning.* International Journal on Computer Vision, vol. 3, p. 73 - 102.
52. Leclerc, Y. G., *Region grouping using the minimum-description-length principle.* Proceedings Image Understanding Workshop, 1990, p. 473 - 481.
53. Lenat, D. B., R. V. Guka, *Building large knowledge-based systems: representation and inference in the CYC-project.* Addison Wesley Publishing Corporation, Reading, 1989.
54. Löcherbach, Th., *Reconstruction of land-use units for the integration of GIS and remote sensing data.* International Archives of Photogrammetry and Remote Sensing, TC7 Workshop on Multi-source Data Integration with Respect to Land Inventory Applications, September, 1992, Delft.
55. Lowe, D.G., *Perceptual organization and visual recognition.* Kluwer Academic Publisher.
56. Malik, J., P. Perona, *Detecting and localizing edges composed of steps, peaks and roofs.* Proceedings 3rd International Conference on Computer Vision, Osaka, Japan, 1990, p. 52 - 57.
57. Malik, J., P. Perona, *Pre-attentive texture discrimination with early vision mechanisms.* Journal of the Optical Society of America.
58. March, S. E., P. Switzer, R. J. P. Kowalik, *Resolving the percentage of component terrains within single resolution elements.* Photogrammetric Engineering and Remote Sensing, vol. 46, no. 8, p. 1079 - 1086.
59. Merchant, J., *Using spatial logic in classification of LANDSAT TM data.* Proceedings IEEE IX Pecora symposium on spatial information technologies for Remote Sensing today and tomorrow, p. 378 - 385.
60. Middelkoop, H., L. L. F. Janssen, *Implementation of temporal relationships in knowledge-based classification of satellite images.* Photogrammetric Engineering and Remote Sensing, vol. 57, no. 7, p. 937 - 945.
61. Minsky, M. L., *A framework for representing knowledge.* The Psychology of Computer Vision, (P. H. Winston, Editor) 1975 New York, McGraw-Hill, 1975.
62. Modestino, J. W., J. Zhang, *A Markov random field model-based approach to image interpretation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14., no. 6, 1992, p. 606 - 615.
63. Møller-Jensen, L., *Knowledge-based classification of an urban area using texture and context information in LANDSAT-TM imagery.* Photogrammetric Engineering and Remote Sensing, vol. 56, no. 6, p. 899 - 904.
64. Mohan, R., R. Nevatia, *Perceptual grouping for the detection and description of structures in aerial images.* IRIS, University of South California, Report no. 225, Los Angeles, 1987.
65. Molenaar, M., *Formal data structures, object dynamics and consistency rules.* Digital Photogrammetric Systems (Ebner/Heipke/Fritsch, Editors) Wichmann, Karlsruhe, 1991.
66. Molenaar, M., D. Fritsch, *Combined data structures for vector and raster representation in Geographic Information Systems.* GIS, vol. 4, no. 5, Karlsruhe, 1991, p. 26 - 32.
67. Niemann, H., G. F. Sagerer, St. Schröder, F. Kummert, *ERNEST: A semantic network system for pattern analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 9, 1990, p. 883 - 905.
68. Nilsson, N. J., *Principles of artificial intelligence.* Palo Alto, California, Tioga, 1980.
69. Palubinskas, G., *A comparative study of decision making in images modeled by Gaussian-Markov random fields.* International Journal of Pattern Recognition and Artificial Intelligence, vol. 2 no. 4, p. 621 - 639.
70. Pan, H.-P., *A spatial structure theory in machine vision and its application to structural and textural analysis of remotely sensed images.* Ph.D.-thesis, University Twente, ITC, Enschede.
71. Pan, H.-P., W. Förstner, *Stochastic polygon map grammers: a generic model for understanding landuse maps and images in remote sensing.* (Förstner/Ruwiedel, Editors), 1992.

72. Pan, H.-P., *Production, inversion and learning of spatial structure: a general paradigm to generic model-based image understanding.* International Archives on Photogrammetry and Remote Sensing, Commission III, Washington, p. 930 - 937.
73. Pertl, A., *Digital image correlation with the analytical plotter Planicomp C100.* International Archives of Photogrammetry, Rio de Janeiro, Comm. III, p. 874 - 882.
74. Radermacher, F. J., *Die Rolle von Metawissen in Umweltinformationssystemen.* Advances in databases (Günther/Schek, editors), Lecture Notes in Computer Science, 525, Springer, 1991, p. 35 - 44.
75. Reimer, U., *FRM: Ein Frame-Repräsentationsmodell und seine formale Semantik.* Informatik Fachberichte 198, Springer Verlag, 1989.
76. Reimer, U., *Einführung in die Wissensrepräsentation.* Teubner Verlag, Stuttgart, 1991.
77. Rosenfeld, A., *Image modelling.* Academic Press Inc., 1980.
78. Ruwwe, Th., Th. Scherer, U. Tempelmann, *Abschlußbericht zum Forschungsvorhaben: Multitemporaler Vergleich der vom Satelliten multispektral registrierten Reflexion von Pflanzenbeständen mit zeitgleich durchgeführten In-Situ-Messungen.* Institut für Photogrammetrie, Bonn.
79. Schenk, T., *Algorithms and software concepts for digital photogrammetric workstations.* Invited Paper, ISPRS Congress, Washington D.C., Intercommission II/III, 1992.
80. Schickler, W., *Matching for outer orientation of single images using 3D-wire frame control points.* International Archives on Photogrammetry and Remote Sensing, B3, Washington D.C., 1992, p. 591 - 598.
81. Sheth, A., *Semantic issues in multi-databases.* Special Issue of SIGMOD RECORD, vol. 20, no. 4, 1991.
82. Shibasaki, R., H. Shimamura, M. Shizukuishi, *Improvement of classification accuracy with consideration of the mixture of land-covers.* International Archives Photogrammetry, Kyoto, VIII-4, p. VII - 444.
83. Spaccapietra, St. (Editor) *Entity-relationship approach.* North Holland 1986.
84. Srinivasan, A., J. A. Richards, *Knowledge-based techniques for multi-source classification.* International Journal for Remote Sensing, vol. 11, no. 3, p. 505 - 525.
85. Strat, T. M., M. A. Fischler, *Context-based vision: recognizing objects using information from both 2-D and 3-D imagery.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 10, 1991, p.1050 - 1065.
86. Strat, T. M., G. B. Smith, *Information management in a sensor-based autonomous system.* Proceedings Image Understanding Workshop, Los Angelos, California, 1987, p. 170 - 177.
87. Ton, J., J. Sticklen, A. K. Jain, *Knowledge-based segmentation of LANDSAT images.* IEEE Transactions on Geoscience and Remote Sensing, vol. 29, no. 2, 1991, p. 222 - 232.
88. Ventrone, V., S. Heiler, *Semantic heterogeneity as a result of domain evolution.* Sheth, 1991, p. 1 - 20.
89. Vosselman, G., N. Haala, *Recognition of road and river pattern by relational matching.* International Archives on Photogrammetry and Remote Sensing, Commission III, Washington, p. 969 - 975.
90. Welch, R., M. Ehlers, *Merging multi-resolution SPOT HRV and LANDSAT TM data.* Photogrammetric Engineering and Remote Sensing, vol. 53 no. 3, p. 301 - 303.
91. Weyhrauch, R. W., *Prolegomena to a theory of mechanized formal reasoning.* AI 13, 1980, no.1 and 2, p. 133 - 170.
92. Worboys, M. F., S. M. Deen, *Semantic heterogeneity in distributed geographic databases.* Sheth 1991, p. 30 - 34.
93. Wrobel, B., *Digitale Bildzuordnung durch Facetten mit Hilfe von Objektraummodellen.* Bildmessung und Luftbildwesen, vol. 55, no. 3, p. 93 - 101.
