Image Information Mining: Exploration of Image Content in Large Archives

Mihai Datcu[1] and Klaus Seidel[2]
[1]German Aerospace Center (DLR), German Remote Sensing Data Center (DFD)
D-82234 Oberfaffenhofen
+49 8153-28 1388
Mihai.Datcu@dlr.de
[2]Computer Vision Group, Communication Technology Laboratory, ETHZ
CH-8092 Zürich
+41 1 632 5284
Klaus.Seidel@vision.ee.ethz.ch

Abstract--We present an intelligent satellite information mining system, a next generation architecture to help a user to rapidly gather information about courses of action, a tool to add value and to manage the huge amounts of historical and newly acquired satellite data-sets by giving to experts access to relevant information in an understandable and directly usable form and to provide friendly interfaces for information query and browsing. In a remote sensing image archive the data access by geographical position, time of acquisition or type of sensor is often less important than what is the content of the scene, i.e. structures, objects, scattering properties. Interesting applications involve complicated spatial and structural relationships among image objects.[1]

TABLE OF CONTENTS

1. PRELIMINARIES

The Earth coverage has high complexity, thus making difficult their recognition and characterization using observations provided by a single sensor. More accurate interpretation of remotely sensed scenes is obtained by synergistically combining signals, information or knowledge from different sources.

The information extraction from remote sensing images requires the signal modelling as a realization of a stochastic process, e.g. the model of speckle for Synthetic Aperture Radar, texture models for forest in high resolution images, or the statistics of gradients for edge detection.

The land cover structures, due to their complexity are not uniqeuely characterized by a single model. For example the "forest" and "grass" can have the same spectral signatures, but they differ in texture parameters.

Thus, the recognition of image structures requires the aggregation of information obtained in the assumptions of variuos models.

The data fusion at signal level shows limited success for the automatic interpretation of remote sensing data. The reason is the incommensurability of the information in its raw format, the image samples. However it shows promising results as a visualization technique, e.g. the presentation of a scene, in computer graphics environment, using synthetic aperture radar, interferometric terrain models and surface appearance derived from optical images. Dealing with heterogeneous sources of information is of key importance for the design of appropriate information representation and its levels of abstraction. Thus data fusion is replaced by information fusion. Methods showing promising results are based on the knowledge of the physical model of the candidate scene. In situations of poor knowledge of the scene and image formation models, fuzzy methods and evidential reasoning or other vague information representations have been successfully applied.

In situations when good models are known, the Bayesian inference allows a precise description of the problem, and thus, accurate scene information retrieval. The power of the Bayesian approach consists in the possibility to represent uniformly the

1. 0-7803-5846-5/00/$10.00©2000 IEEE

incertitude over a certain scene parameter in data acquired from heterogeneous and incommensurable sources. The incertitude has the form of the posterior distribution of the desired scene parameter given the observations from several sources. The posterior distribution is inferred from the likelihood of each data source and the a priori model of the parameter. The likelihood distribution takes into consideration the forward model of the specific sensor and its noise. The a priori information models describe the behaviour of the desired parameter under the observation conditions which are specific for each sensor type.

As an example, the accuracy of the classification of the land cover from synthetic aperture radar can be improved using information from optical images. The posterior distribution of the land cover parameters is expressed as being conditioned by both the synthetic aperture radar and optical data.

One of the difficulties in many data fusion applications for remote sensing image classification is the identification of the data sets relevant for the user. Here the information retrieval task becomes an information mining task. The Bayesian inference can be applied again, but the goal is to find the set of images, generally acquired from different sensors, which might contain the information desired by the user. The inference is used to find the posterior distribution over the set of available images according to the user hypothesis in a given application.

Thus, we developed a system based on an intensive preprocessing of received satellite images to extract relevant features, structures and objects and automatically record and analyse their interrelationships to learn their behaviour so as to be able to detect relevant information. The system is operated using an intelligent interface able to correlate the information content of the satellite images with the relevant goals of the application. The user has at his disposition tools for definition of specific goals using his own semantics. The problem of large dimensionality, which for computationally efficient data analysis is a primary concern, is solved using pre-extracted representative features instead of raw images. Methods presently being developed make use of SAR, InSAR, and optical data of the missions operated now, and can be refined and extended for the high resolution data (0.5m -1m) that will soon be on the market.

2. Present Status in the Field of Image Content Representation

Starting in 1993 the leading groups in the field of query by image content from image archives, were Media Lab at MIT, IBM Almadem and U.C. Santa Barbara [1,2,3].

The first journals dedicated to image content representation and related topics in data mining and information visualization, have been published in the period from 1995-1997:

- a special issue on data bases (IEEE Computer, 1995)
- a special issue on data mining (IEEE Knowledge and Data Engineering, 1996)
- a special issue on digital libraries (IEEE PAMI, 1996)
- a special issue on visualization (IEEE Computer Graphics, 1997)
- a special issue on image data bases (Pattern Recognition, 1997)

In the meantime the growing interest in managing very-large distributed information systems is reflected also in the progress of image archiving technology. It is now difficult to enumerate the groups involved in this field. A measure of the progress is the number of demonstrators available on different web sites. A synopsis can be found from [42]. Almost all of the previously enumerated systems rely on the integration of elements from image recognition in the query system. Not so many groups are involved in investigations of knowledge representation and perception for adaptive queries specification [4]. Commercial or operational systems are not so often and are based on very simple principles [5,32].

The operational application in Remote Sensing data management is still based on "classical" technologies based on inventories and catalogues indexed by geographical location, sensor type or time of acquisition. The table from the web site [43] lists several of the operational systems.

Recently, IBM in collaboration with NASA designed the "Satellite Image Explorer". The system is based on a multiresolution approach for finding similar areas of coregistered images and for finding similar patterns within the same image. Presently the emphasis is on developing real-life applications of the previously developed techniques. A description of the project can be found on the web site [44]. Similar systems, both for query by image content and data mining are under development at NEC, NASA-JPL, and NASA-Ames Research Center.

3. Data and Information Retrieval

Next sections present an operational system in exploitation at DFD, and the results of two years of research of DFD in collaboration with ETH Zurich.

3.1 Conventional Satellite Data Ground Segments

Conventional satellite data ground segments comprise sensor mode planning and data acquisition systems, data ingestion interfaces, processing capabilities, a catalogue of available data, a data archive, and interfaces for queries and data dissemination. The transfer of data mostly relies on common electronic networks or high capacity tapes. DFD's Intelligent Satellite Data Information System (ISIS), provides data catalogue search and retrieval. In order to support a variety of hardware platforms ISIS operates different interfaces and is available via common communication lines. The graphical interface GISIS provides the most comfortable access to the data of the DFD

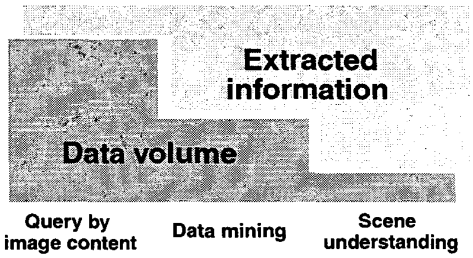

Figure 1 Diagram of the scalable system for information retrieval from very large image archives. Due to the problem complexity the system scales-down the amount of data to be analysed such that the detail of extracted information can be increased.

catalogue offering the following services: distributed catalogue retrieval from more than 200,000 data entries with about 150,000 digital images, supported by a map browser and geographical names; display of data-set footprints on a map; automated network transfer of digital quick-looks (compressed images) for visual inspection; transfer of full resolution data (full or interactively defined subsets) for some data products; guide information about sensors, data products and distribution modalities; on-line ordering; on-line placement of user comments. The system can be accessed from http://isis.dfd.dlr.de/.

3.2 Future Ground Segments

Future ground segments have to be capable of handling multi-mission data acquired by SAR and optical sensors and have to provide easy user access to support the selection of specific data sets, fuse data, visualize products and compress data for transmission via Internet. In particular, the search for data sets has to support individual queries by data content and detailed application areas as well as capabilities for automated extraction of relevant features and the application oriented representation of results. In the case of high resolution SAR and optical images, we face an enormous volume of data combined with very specific analysis requests posed by the users. In order to reconcile these conflicting aspects we suggest the development of tools and a scalable system for information retrieval [6]:

a) query by image content from large image archives,
b) data mining in the set of images obtained as a query response, and
c) scene understanding from the images found to be relevant for the user's application.

The diagram in Fig. 1 presents the system concept. The three levels of the scalar system are designed to accommodate the huge volume of data containing high complexity structures. They perform a progressive data reduction with simultaneous increase of the extracted image details.

3.2.1 Query by image content

The goal of query by image content is to enable an operator to *see* into a large volume of high complexity images. The incorporation of content-based query and retrieval techniques into the RS image archives will meet the requirements expressed by today users and at the same time lay foundations for the efficient handling of the huge data-flow that is expected from the next generation of satellites. Drawing the connection from the image signal to the image content is the key step in providing an intelligent query from large RS data bases, thus making the user able to *see* into the data base before actually retrieving data.

The *design concept* for query by image content in large image archives aggregates three components [7,8,9]:

i) image content representation and definition of similarity functions,
ii) a multiscale approach with two goals, retrieval at different image scales and fast progressive discarding of non-interesting images,
iii) automatic re-clustering of image archive according to the user query.

The *elaboration of the system* is based on a Bayesian approach for the hierarchic information representation in large image archives [10,11,12,13]:

i) image signal modelling and feature extraction
ii) hierarchic clustering in the joint model&feature space of the image archive,
iii) semantic analysis of the query and user conjecture modelling.

The method aims at obtaining a better selectivity of the query process for different subjective user requests, and it is a basis for the design of optimal multi-dimensional signals representation in large over-the-net distributed archives.

At ingestion in the physical archive the image content is extracted and presented in form of signal features and stored in the inventory. The indexing mechanism of the data base management system allows to cluster together all images containing *similar* content. The *similarity* is defined adaptively depending on the user conjecture, i.e. the hypothesis he uses in a certain application.

Many of the RS image interpretation techniques are oriented to the extraction of radiometric or multispectral features. These approaches are some times hindered by the limited information available in the observations, e.g. by the small number of spectral channels. Additional information can be extracted using spatial (structural) analysis. We put emphasis on structural information extraction as texture features [11,12].

3.2.2 Data mining

The expected answer of a query is a set of images presumably containing similar structures. Further refinement of the information content in the now reduced data set, can be obtained by data mining. The goal is to explore the information content of the images and decide which ones are relevant for the user's application. The tool we propose is mainly based on interactive decision using data visualization, e.g. image browsing integrated with image segmentation, or information visualization, e.g. exploration of the clusters structure in the feature space and the interactive design of the decision surfaces.

The first experiments led to the development of the image epitome: integrated image restoration, feature selection and compression in the wavelet transform domain. The image epitome is one of the tools for mining in image data sets (37).

The data and information mining tool will be integrated in the visualization module of the remote sensing ground segment.

3.3 Scene understanding

The last step of highest detail information extraction is scene understanding. Scene understanding is the attempt to extract information about the physical characterization, the structure and geometry of the three-dimensional scene from a two-dimensional image which is explained elsewhere [38, 11]. The task of the scene understanding process is to find the scene which best explains the observed data. Scene understanding is an ill-posed inverse problem. The solutions we adopt are model based formulations in the general frame of Bayesian inference. The implemented algorithms allow accurate texture segmentation in adverse noise conditions, very high spatial and radiometric detail extraction from remote sensing images, super-resolution of targets, and surface reconstruction, automatic segmentation and target detection in mountainous areas [14, 39, 15].

4. THE BAYESIAN INFERENCE AND HIERARCHIC DATA MODELLING

The Bayesian approach consists of interpreting probabilities based on a system of axioms describing the incomplete information rather than randomness. Hence, Bayesian approaches are very well suited for the extraction of information out of observed data. Using Bayesian techniques one can extract the most probable scene parameters explaining the observed data under the assumption of some prior knowledge. The prior knowledge is expressed in form of stochastic models. The Bayesian methods allow also to choose the most plausible model from a given class. Two levels of inference are introduced [16]:

Model fitting: the first level of inference assumes that the used models are true. The task is to fit the assumed model to the data estimating the most plausible model parameter values Θ and a measure of error. The inference requires the knowledge of the prior model $p(\Theta|M_i)$ and the data prediction term or forward model $p(D|\Theta, M_i)$ pursues Bayes' rule and is applied separately for each of the models $\{M_i\}$:

$$p(\Theta|D, M_i) = \frac{p(D|\Theta, M_i)p(\Theta|M_i)}{p(D|M_i)} \tag{1}$$

The information extraction is a Maximum A Posteriori estimation (MAP):

$$\hat{\Theta} = \arg max p(\Theta|D, M_i) \tag{2}$$

The evidence term $p(D|M_i)$ in the denominator of Eq. 2 is generally neglected in model fitting, but is important in the second level of inference, and here lies the novelty of the Bayesian approach in data inversion.

Model comparison. The task of the second level of the Bayesian inference is to find the most plausible model given the data. The inference applies in the space of models:

$$p(M_i|D) \propto p(D|M_i)p(M_i) \tag{3}$$

The inference relies on the evidence of M_i carried by $p(D|M_i)$ and the subjective prior over the assumed hypothesis space $p(M_i)$. $p(M_i)$ shows how plausible we thought the alternative models were before the data arrived.

Inference of probability distributions from observations of sensory data aims at finding the best stochastic models able to consistently characterize classes of images [11,17]. The Bayesian approach for data modelling is used. The information contained in a data set (provided by a unique sensor) is extracted in different assumptions. The assumptions are represented by different prior models (Fig. 2). In the case of a multispectral sensor the assumed prior models can characterize either the spectral components or the texture structures. The

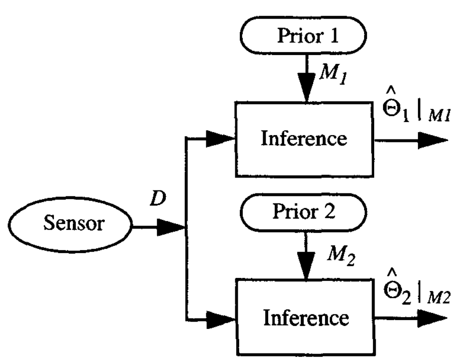

Figure 2 Information extraction using qualitatively different models.

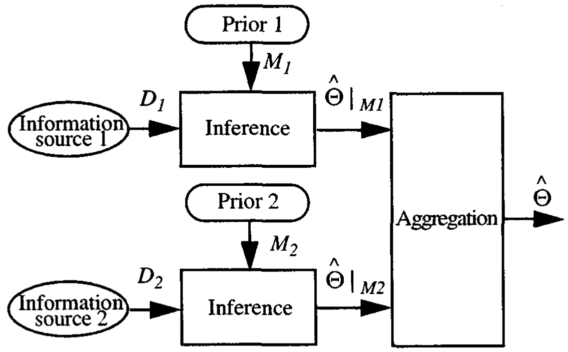

Figure 3 First paradigm for sensory information extraction and data fusion

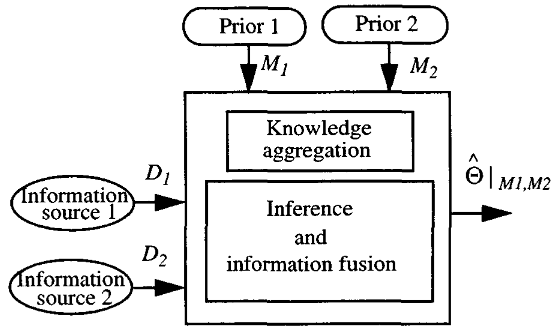

Figure 4 Paradigm of Bayesian information and knowledge fusion.

extracted information according to these two models is not commensurable, it represents different qualities.

We call the process to extract information from sensory data using different prior models, *data fission*.

The objective of this work is to use the Bayesian inference in order to obtain an estimate of a given physical parameter using observations acquired with different sensors. We assume a model for the desired physical parameter and try to estimate it by fitting the model to the data.

If we take into consideration two data sets D_1 and D_2 separately for observations, the information extraction can be splitted into two separated problems with solutions given by the maximum of the following a posteriori probabilities (MAP):

$$p(\Theta|D_1, M_1) = \frac{p(D_1|\Theta, M_1) \cdot p(\Theta|M_1)}{p(D_1, M_1)} \quad (4)$$

$$p(\Theta|D_2, M_2) = \frac{p(D_2|\Theta, M_2) \cdot p(\Theta|M_2)}{p(D_2|M_2)} \quad (5)$$

where Θ is the desired physical parameter, and $p(\Theta|M_i)$ encapsulates our a priori knowledge. The measure of fidelity to the observed data is given in terms of the conditional probabilities $p(D_i|\Theta, M_i), i=1,2$.

In Fig. 3 we introduce a first paradigm for data fusion. It refers mainly the extraction of image content information.

Here, we can identify three cases. The first case, the case of a unique source of information, was treated as data fission. A second situation is the extraction of information from different sources using the same prior model. The estimated parameters will have identical representation, however their meaning can be different. A simple example is texture parameter estimation from data with different resolution. The scale plays the role of meta information, thus the direct interpretation of the estimated parameters is not consistent. The third case assumes information extraction from different sources using different prior models. The resulted information has incommensurable representations.

A full Bayesian approach for information fusion can be formulated as maximizing the following a posteriori probability:

$$p(\Theta|D_1, D_2, M_1, M_2) = \frac{p(D_1|\Theta, M_1)p(D_2|\Theta, M_2) \cdot \boldsymbol{P}\{p(\Theta|M_1)p(\Theta|M_2)\}}{p(D_1, D_2|M_1, M_2)} \quad (6)$$

where $\boldsymbol{P}$ is an operator representing the prior information in the assumption of two different models. We observe that the problem of fusion of information from two data sets is extended with fusion of knowledge, in form of the specification of the a priori models M_1 and M_2. The paradigm of Bayesian information and knowledge fusion is presented in Fig. 4.

In the second paradigm we address mainly the fusion of information about the physical characterization of scenes, e.g. estimation of terrain heights derived jointly from image intensities and SAR intereferometric phases [18].

5. Stochastic Models

A "universal" model for stochastic processes is not mathematically tractable. In applications, we are faced either with simple case studies, e.g. the data is precisely described by a low complexity stochastic model, as in laser speckle images, or the data is of high complexity, we use an approximate model. A challenging research task is to find a "quasi-complete" family of models for a certain class of signals, for example all images provided by one sensor.

In this spirit, we concentrate on the following stochastic models: Gibbs random fields, multidimensional stochastic processes, cluster based probability modelling.

5.1 Gibbs random fields

In many situations signals satisfy predetermined constraints. We can restrict the modelling of these signals by considering only probability measures that fulfil these constraints. Here we have to choose the appropriate probability measure: those satisfying the set of constraints. Applying a maximum uncertainty principle, the probability measure that satisfies all relevant constraints should be the one that maximizes our incertitude about what is unknown to us. The probability measure resulting from such a principle is a Gibbs distribution:

$$p(x) = \frac{1}{Z}e^{-\frac{1}{T}H(x)}$$
$$Z = \sum_x e^{-\frac{1}{T}H(x)} \tag{7}$$
$$H(x) = \sum_{allcliques} V_{clique}(x, \theta)$$
$$\theta = \{\alpha_0, \alpha_1, \ldots\alpha_4, \beta_1, \ldots\beta_4, \gamma_0\}$$

where $\alpha_0, \ldots\alpha_4$, $\beta_1,\ldots\beta_4$, γ_0 are the model parameters associated to the corresponding cliques, V is the potential function characterizing the interaction between the samples of the random field inside the clique, H represents the energy function for the corresponding neighbourhood [11,19,12].

Images and other multidimensional signals satisfy the local requirement that neighbouring sites have related intensities. On the other hand, a model should also be able to represent long-range interactions. This is a global requirement. Gibbs random fields are able to satisfy both requirements.

The equivalence of Gibbs and Markov random fields gives the appropriate mathematical techniques to treat these behaviours. A pragmatic problem is to fit optimally a Gibbs random field model to real data. It was expected that a maximum likelihood estimate gives the desired result. That is not generally possible due to the requirement to evaluate the partition function.

Several alternative solutions have been proposed; the coding, the maximum pseudo-likelihood, however none of these is efficient. Recently, a consistent solution for the maximum likelihood was introduced: Monte Carlo maximum likelihood. This algorithm requires no direct evaluation of the partition function, it is consistent and converges to the maximum likelihood with probability one [20,21].

5.2 Stochastic pyramids

The wavelet transformation of images in its operator formalism suggests the decomposition of the signal into two components: the approximation and the detail, and induces a pyramidal data representation similar to a quadtree. The accurate modelling of wavelet coefficients requires the representation of both intra-scale and inter-scale dependencies. The intra-scale model captures the local structural behaviour of the image. The information within a neighbourhood in the original image is distributed in each "quad" (brothers) and at the corresponding coordinates of the different orientations in the detail space. The inter-scale relationships are modelled as Markov chains. A more general formalism is the multiscale stochastic modelling. In the graph below a random field is defined having as support a tree structure.

The inter-scale causalities are described by transition probability density functions **P(*|*)**, see Fig. 5 .

These models allow synthesis of finer scales $\mathbf{X^k}$ of a random field beginning with its coarser scales $\mathbf{X^{k-1}}$. The stochastic process represents the new information, in a way similar to the detail signal in the wavelet transform domain. The models are inspired from the multigrid techniques in numerical analysis and define consistent Multiscale Markov Random Fields starting from a single resolution. They can model a variety of image configurations and are implicitly the basis for picture parameter estimation. The estimation principle consists of solving a sequence of global optimization problems defined on a sequence of embedded configuration subspaces accepting constraints in form of prior distributions [22,23,24].

6. The Basic Concepts

Due to the incommensurability of the images obtained from different sensors and due to the high complexity of the imaged scenes, data fusion systems demands high level representation of information. Thus, scene interpretation is done by augmentation of the data with meaning [25]. Based on the Bayesian principles of inference and on data fission and fusion paradigms, we present in Fig. 6 the simplified architecture of a scene interpretation system using multiple data sources.

The information source is assumed to be a *collection* of multidimensional signals, e.g. airborne or satellite images, acquired from optical or SAR sensors. Due to the incommensurability of the data provided by heterogeneous sources, the information fusion process is split into two steps: i) information fission and ii) information aggregation.

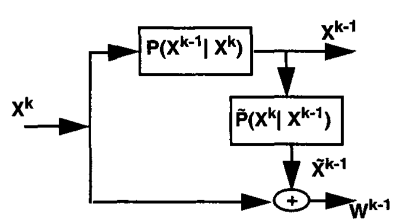

Figure 5 Characterization of a multiscale stochastic process

Information fission is an analysing step enabling to deal with heterogenous sources of data, and also to cope with different and incommensurable types of information extracted from individual sources. The information source is partitioned in elementary sources.

The information fission requires the signal modelling as a realization of a stochastic process. A library of stochastic and deterministic models is used to infer the signal model. The information extraction is a model fitting task. The information content of data sets acquired by individual sensors is extracted.

The resulting objective features are aggregated according to the user conjecture. Thus, the information fusion process relies on restructuring (using a certain syntax) the signal feature space according to the user semantic models. At this step the information from incommensurable sources is ordered and augmented with meaning, thus providing a scene interpretation [25,26].

7. Levels of Abstraction of Information Representation

In the case of modelling high complexity signals, e.g. collections of multi-sensor images, a large number of sources coexists within the same system. The solution of information and knowledge fusion as stated in Eq. 6 becomes a very difficult task. However, in many practical applications the candidate models are likely to be analysed hierarchically. Thus, it is desirable to integrate probabilistic models, i.e. they should store common parts for efficiency of the model representation, and they must be represented hierarchically in order to capture the class structure and to provide computational advantages. Starting from remotely sensed images of a scene a hierarchy of information is defined.

- **Image data:** the information is contained in the pixel intensities of the raw data. It is the lowest level of information representation. However there are useful applications, e.g. image classification based on image intensity. Data fusion at this level of representation is limited by the incommensurably of the measurements. Typical applications are fusion of multiresolution data of the same type of sensor.
- **Image features:** the information is extracted in form of parameters characterising the interactions among spatially distributed pixels, or different spectral channels. Additionally to the parameter values, characterization of parameters incertitudes can be obtained. Popular examples of image features are: texture, multi-spectral features or geometrical descriptors. Data fusion at this level is possible in the case of parameters representing the same quantity, e.g. geometric precision enhancement of an edge using information from multiple sensors.
- **Physical parameters:** the image features reflect the physical parameters of the imaged scene, thus, assuming the availability of certain models, the scene parameters can be extracted. For example, image texture carries information about the size of tree crowns, or the SAR backscatter of ocean surface contains the wind speed information. Data acquisition using different type of sensors can be fused to increase the estimation accuracy of physical parameters, or to complement missing observations.
- **Meta features:** estimation of both image features, and physical parameters requires the assumption of some data models. The type of model used, its evidence and complexity, plays the role of meta information, i.e. describing the quality of the extracted parameters. From a data aggregation perspective, a meta feature is an indicator of information commensurability, e.g. estimated texture features using coocurence matrices are not comparable with parameters of Markov random fields.
- **Cluster model:** image features or estimated physical parameters have n-dimensional representations. Due to observation noise or model approximations the feature space is not occupied homogeneously. Thus, another level of information abstraction, and fusion at the same time, is the type of feature grouping, i.e. the cluster models, and the associated parameters. Clustering can represent information only for one category of the features.
- **Syntactic modelling:** stochastic grammars are defined to represent the order of joint spaces, image features and models. The induced syntax can be interpreted as a clustering of incommensurable information representations. Thus, information in incommensurable representations can be aggregated. In the case of fusion of physical parameters extracted from different data types, the result of information fusion is equivalent to a supervised classification, otherwise, for the case of abstract features, the classification is unsupervised.
- **Semantic representation:** due to the impossibility to estimate the physical parameters, an unsupervised classification of the data is produced. Augmentation with meaning requires a higher level of abstraction. Prior information in form of training data sets or expert knowledge is used to create semantic networks. Thus, the observations are labelled and the contextual meaning is defined.

The performance of information extraction depends critically on the descriptive or predictive accuracy of the probabilistic model employed. Accurate modelling typically requires high-dimensional and multi-scale modelling. For heterogeneous sources, accuracy also depends on adaptation to local characteristics.

8. Assessment of New Image Information Extraction Technologies

The classical task related to the interpretation of remote sensing data generally assumes that the source of information is just *one* image. The methods applied for information extraction are image enhancement, image segmentation, feature extraction, fitting physical models to the data, etc. A more accurate inter-

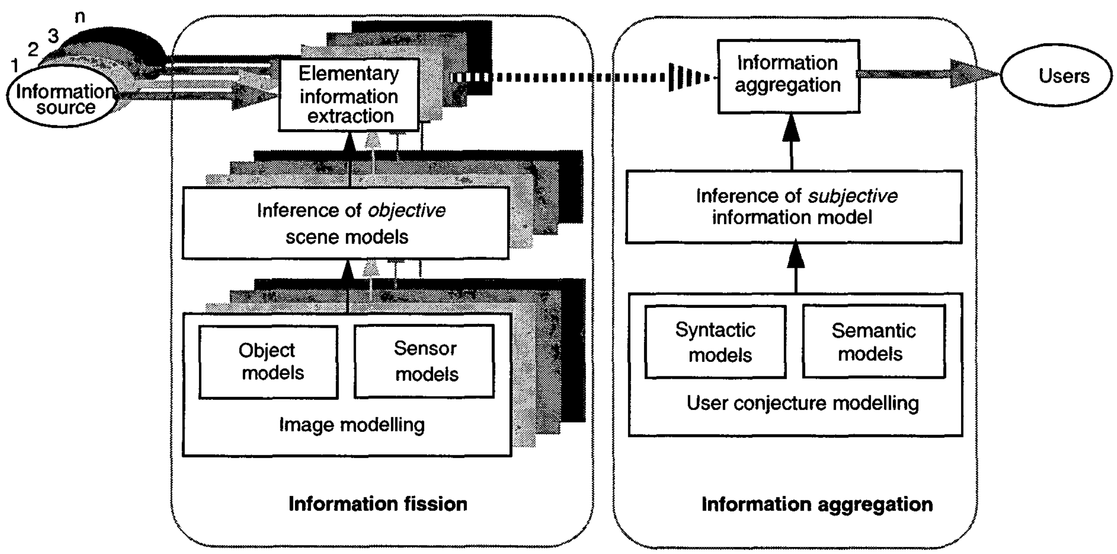

Figure 6 Simplified diagram of a information fusion system. Heterogenous sources of information are splitted into elementary sources: the information fission. The information aggregation completes the fusion process.

pretation of remotely sensed scenes is obtained by synergistically combining signals from different sources, e.g. fusion of data acquired from different sensors or multitemporal analyses. One of the difficulties in many data fusion applications for remote sensing image classification is the identification of the data sets relevant for the user. Here the information retrieval becomes a data and information mining task.

The explosion in sensor technology, in both high resolutions and frequent repeat passes, results in an increasing number of large multimission remote sensing archives, e.g. DFD stores more than 80 data product types adding up to more than 200,000 images and increasing in 3 years up to 1,500,000 images! Thus, the problem of image content extraction should be reformulated such to take into consideration the *new source of information*: the image archive. The state of the art remote sensing archives are managed by relational data base management systems able to query for the data with no reference to the information content, except the visual browsing using quick-look images. This technology is limited to applications involving a reduced number of images and also the accuracy is restricted by the mostly poor quality of the quick-look images.

The query by image content methods we develop are intended to overcome the informational bottle-neck of classical approaches and also to stimulate the user in finding new scenarios for data interpretation, e.g. *find all images containing cities surrounded by forest*. It is known that the distinction between the perception of information as signals and symbols is generally not dependent on the form in which the information is presented but rather on the conjecture in which it is perceived, i.e. upon the hypothesis and expectations of the user. Thus, the new technology requires a different attitude of the user and also a careful design of an intelligent, adaptive query engine.

In the field of remote sensing, the research interest was recently stimulated by the new generations of high and very high resolution sensors, Synthetic Aperture Radar (SAR), interferometric SAR, optical and hyperspectral [41]. There is also a clear trend to enlarge the existing system for data distribution with new tools for information distribution [41,40]. However, only few projects are carried at high scientific level to solve the difficult problems encountered in the field of remote sensing image information retrieval, mining, or query by image content, many other approaches are still driven by classical methods.

The concept we elaborated for content-based query and information retrieval from Remote Sensing Image Archives (RSIA) was founded on a hierarchical Bayesian learning model and is demonstrating a system with two levels: i) interactive training of the desired image content in terms of image features, followed by, ii) query by image content using as content the image features defined in step i [41]. The logical diagram of the system is shown in Fig. 7.

Both levels make use of the pre-extracted image parameters. For computational complexity reasons, the image parameters are extracted off-line at data ingestion in the archive. The parameters are extracted for different image scales [11,27,28,12]. In the next processing step the image parameters are clustered, and further a signal content index is created using the cluster description, the scale information, and type of stochastic model assumed for the image parameters [29,30]. A Bayesian hierarchical decision algorithm (naive Bayes) allows a user to visualize and to encapsulate interactively his prior knowledge of certain image structures and to generate a supervised classification in the joint space of clusters, scales, model types. The user is enabled to attach his meaning to similar structures occurring in different images, thus adding a label in the archive inventory. This label is further used to specify queries. This concept was implemented and successfully demonstrated with an experimental system [28,31,32].

During the RSIA project we were facing difficult problems. Some of them have been solved with original solutions, e.g. texture features estimation in presence of noise (Walessa and Datcu, 1999), hierarchy of information representation for image content characterization [30], supervised classification and interactive training using Bayes networks [31]. Other solutions implemented in the system followed the most advanced results obtained until now, e.g. image features extraction from optical multispectral data [33]data clustering, and part of the user adaptation techniques [17].

The system was integrated in a server client configuration as presented in Fig. 8.

The reader can interactively generate illustrations of information extraction accessing the demo Information mining system on [45]. Presently it manages 160 Landsat TM subscenes representing the full covarage of Switzerlamnd, 110 X-SAR subscenes covering ~50% of the Switzerland, and a selection of 66 high resolution images acquired with an airborne optical camera.

9. Conclusions

The management of high and very high resolution SAR and optical images requires the integration of the existing information and knowledge systems with methods of signal analysis emerging in new methods for information access and dissemination, e.g. query by image content, information mining, scene understanding, synergetic classification and data compression, data and information visualization, user adaptation and semantic modelling of queries. Data and information fusion and mining have as common tasks the information extraction and representation. The differentiation of the two fields is in the way how information is treated. Information fusion has as goal the aggregation of incommensurable pieces of information trying to enhance the quality of data interpretation. Data and information mining has as a goal the exploration of the unexpected relationships among the elementary items of information extracted from the abbreviations.

We presented a Bayesian perspective of the two fields, and propose several new approaches. Part of the methods or similar techniques are integrated in a demonstrator system for query by image content from large image archives [34].

The classical tasks in the interpretation of remote sensing data generally assume that the source of information is just *one* image. The methods applied for information extraction are image enhancement, image segmentation, feature extraction, fitting physical models to the data, etc. A more accurate inter-

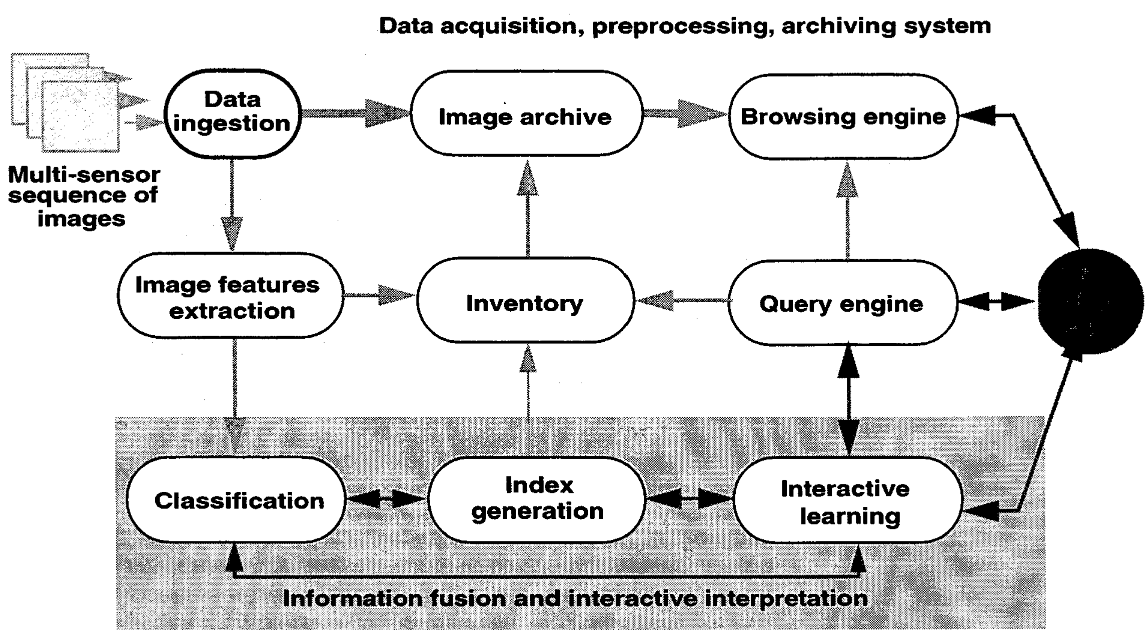

Figure 7 The system consists of two main modules: a *data aquisition, preprocessing, archiving system,* responsible for extraction of the image information content, storage of the raw images, and their management using query by image features, and an *information mining system,* which enables the user to rapidly investigate the information pre-extracted from multi-mission satellite images, explore their meaning, detect relevant events and relate them to the goal of the application.

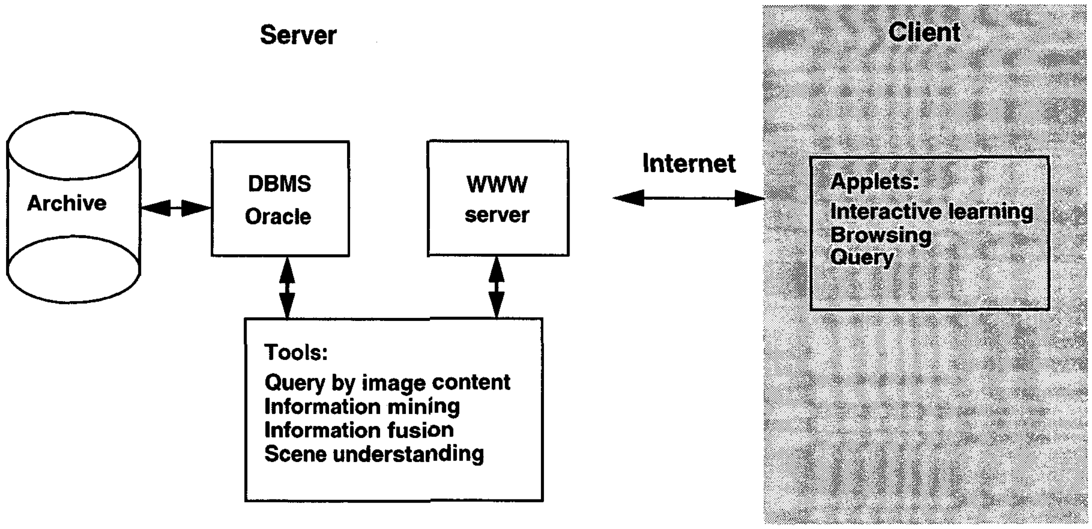

Figure 8 Figure www: configuration diagram of the information miming system. A server-client architecture is a two folded the solution: the off line image preprocessing on the server site gives acces to the image features in real time relative to the user needs, and the system is distributed with interactive operation over-the-net.

pretation of remotely sensed scenes is obtained by synergistically combining signals from different sources, e.g. fusion of data acquired from different sensors or multitemporal analyses. One of the difficulties in many data fusion applications for remote sensing image classification is the identification of data sets relevant for the user. Here the information retrieval task becomes a data information mining topic [34,30,13,17].

Thus, the problem of image content extraction should be reformulated to take into consideration the *new source of information*: the image archive. The state of the art remote sensing archives are managed by relational data base management systems able to query for data with no reference to the information content, except the visual browsing using quick-look images. The technology is limited to applications involving a reduced number of images of medium resolution.

Inclusion of existing and newly acquired *ground truth* data and also models applied for RS data interpretation in systematic archives with appropriate access to the information will enable any RS user to access and mine the necessary models, using previous experience, thus reducing the requirements of new *ground truth* data. Information mining technology could offer a new interpretation methodology speeding up the interpretation process and increasing the amount of analysed data.

The image information mining and query by image content methods recently developed intend to overcome the informational bottle-neck of classical approaches and also to stimulate the user to find new scenarios for data interpretation, e.g. find all images containing cities surrounded by forest [35,36].

It is known that the distinction between the perception of information as signals and symbols is generally not depending on the form in which the information is presented but rather on the conjecture in which it is perceived, i.e. upon the hypothesis and expectations of the user. Thus, the new technology requires a different attitude of the user and also a careful design of an intelligent, adaptive query engine.

ACKNOWLEDGEMENT

The authors would like to acknowlege the contributions of Hubert Rehrauer, Michael Schröder, Gintautas Palubinskas, Marc Walessa, Sorel Stan and Andrea Pelizzari to the *ETH/ DLR Remote Sensing Image Archive (RSIA)* demonstrator.

REFERENCES

[1] D. Swanberg and R. Jain. *Knowledge guided parsing in video data bases.* In Proceedings of IS&T/SPIE Symposium on Electronic Imaging: Science and Technology, page 214. IS&T/SPIE, February 1994.

[2] A. Gupta, T. E. Weymouth, and R. Jain. *An extended object oriented data model for large image data bases.* In O. Günther and H.-J. Schek, editors, Advances in Spatial Databases, 2nd Symposium SSD 91, Zuerich, pages 45-61. Springer-Verlag, August 1991.

[3] A. Pentland, R. W. Picard, and S. Sclaroff. *Photobook: Tools for content-based manipula-tion of image databases.* In W. Niblack and R. C. Jain, editors, Storage and Retrieval for Image and Video Databases II, volume 2185, pages 34-47. SPIE, February 1994.

[4] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos. *Target testing and the PicHunter bayesian multimedia retrieval system.* In Forum on Research and Technology Advances in Digital Libraries (ADL), Washington, DC.

IEEE, 1996.

[5] R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, D. Petkovic, J. Ashley-J, T. Mc-Connell, J. Ho, J. Jang, D. Berkowitz, P. Yanker, M. Vo, D. Haas, D. Lassig, S. Tate, A. Chang, P. van Houten, J. Chang, T. Petersen, D. Lutrell, M. Snedden, P. Faust, C. Mat-teucci, M. Rayner, R. Peters, W. Beck, and J. Witsett. *Ultimedia manager: Query by image content and its applications.* In Digest of Papers. Spring COMPCON 94, San Jose, CA, USA, 1994. IBM Almaden Res. Center.

[6] K. Seidel, R. Mastropietro, and M. Datcu. *New architectures for remote sensing im-age archives.* In T. I. Stein, editor, A Scientific Vision for Sustainable Development, IG-ARSS 97, volume 1, pages 616-618, 1997.

[7] U. Benz, M. Datcu, G. Schwarz, and K. Seidel. *Image communication: New features of a remote sensing data ground segment.* In CEOS SAR Workshop, ESTEC, February 1998.

[8] K. Seidel, M. Schröder, H. Rehrauer, and M. Datcu. *Query by image content from re-mote sensing archives.* In IEEE Intern. Geoscience and Remote Sensing Symposium IG-ARSS 98, volume 1, pages 393-396, 1998.

[9] H. Rehrauer, K. Seidel, and M. Datcu. *Multiscale markov random fields for large image datasets representation.* In T. I. Stein, editor, A Scientific Vision for Sustainable Develop-ment, IGARSS 97, volume 1, pages 255-257, 1997.

[10] M. Schröder, K. Seidel, and M. Datcu. *Gibbs random field models for image content characterization.* In T. I. Stein, editor, A Scientific Vision for Sustainable Development, IGARSS 97, volume 1, pages 258-260, 1997.

[11] Mihai Datcu, Klaus Seidel and M. Walessa, 1998, *Spatial Information Retrieval From Remote Sensing Images: Part A. Information Theoretical Perspective*, IEEE Tr. on Geoscience and Remote Sensing, Vol. 36, pp. 1431-1445.

[12] M. Schröder, H. Rehrauer, K. Seidel, and M. Datcu. *Spatial information retrieval from remote sensing images Part II: Gibbs Markov random fields.* IEEE Trans. on Geoscience and Remote Sensing, 36(5):1446-1455, Sept. 1998.

[13] K. Seidel, M. Schröder, H. Rehrauer, M. Datcu, 1998, *Query by Image Content from Remote Sensing Archives*, Proc. of 1998 International Geoscience and Remote Sensing Symposium (IGARSS'98), volume 1, pages 393-396.

[14] M. Datcu and F. Holecz. *Generation of synthetic images for radiometric topographic correction of optical imagery.* In P. S. Chavez and R. A. Schowengerdt, editors, SPIE Proceedings on Recent Advances in Sensors, Radiometric Calibration and Processing of Remotely Sensed Data, Orlando, USA, volume 1938, pages 260-271. SPIE OE/ Aerospace Sensing, 1993.

[15] M. Walessa and M. Datcu. *Texture reconstruction in noisy images.* In Proceedings of the 18th International Workshop on Maximum Entropy and Bayesian Methods (MaxEnt 98), pages 65-72. Kluwer Academic Publishers, 1998.

[16] MacKay D. J. C. (1991) *Maximum Entropy and Bayesian Methods in Inverse Problems*, chapter in Bayesian Interpolation, pp. 39–66. Kluwer Academic Publisher.

[17] Minka T.P.,.Picard R.W (1997) *Interactive learning with a society of models.* Pattern Recognition, vol. 30, pp.565–581.

[18] G. Nicco, M. Walessa, M. Datcu, 1998, *Topographic Mapping: A New Approach*, Proc. of 1998 International Geoscience and Remote Sensing Symposium (IGARSS'98).

[19] Geman S., Geman D. (1984) *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.* IEEE Trans. on PAMI, vol. 6, pp. 721–741.

[20] Cowles, M.K., Carlin, B.P. (1996) *Markov Chain Monte-Carlo convergence diagnostics: a comparative study,* J. Amer. Statist. Soc. 91, pp. 883-904.

[21] Li S. Z. (1995) *Markov Random Field Modeling in Computer Vision,* Computer Science Workbench, Springer.

[22] Baldwin D. G., Emery W. J., Cheesman P.B., *Higher resolution earth surface features from repeat moderate resolution satellite imagery. IEEE Tr. GSRS* **36** (1998), pp. 244-255.

[23] Bouman, C., B. Liu (1991) *Multiple resolution segmentation of textured images,* IEEE Tr. PAMI, Vol. 13, pp. 99-113.

[24] Luettgen, M.R., at al. (1993) *Multiscale Representation of Markov Random Fields,* IEEE Tr. on Signal Processing, Vol. 41, pp. 3377-3395.

[25] Jumarie G. (1991) *Relative information*, Springer Verlag.

[26] Lauritzen S. L., Spiegelhalter D.J. (1988) *Local computations with probabilities on graphical structures and their applications to expert systems (with discussion),* J. Rag. Statist. Soc. B, Vol. 50, pp. 157-224.

[27] H. Rehrauer, K. Seidel, and M. Datcu. *Characteristic scale detection in remote-sensing data.* In IEEE Intern. Geoscience and Remote Sensing Symposium IGARSS 99, volume I, pages 116--118, 1999.

[28] H. Rehrauer, K. Seidel, and M. Datcu. *Multi-scale indices for content-based image retrieval.* In IEEE Intern. Geoscience and Remote Sensing Symposium IGARSS 99, volume V, pages 2377--2379, 1999.

[29] M. Schröder, K. Seidel, and M. Datcu. *Bayesian modeling of remote-sensing image content.* In IEEE Intern. Geoscience and Remote Sensing Symposium IGARSS 99, volume III, pages 1810--1812, 1999.

[30] K. Seidel, M. Schröder, H. Rehrauer, M. Datcu, 1998, *Meta Features for Remote Sensing Image Content Indexing*, Proc. of 1998 International Geoscience and Remote Sensing Symposium (IGARSS'98), volume II, pages 1022--1024.

[31] Michael Schröder, Hubert Rehrauer, Klaus Seidel and Mihai Datcu, *Interactive Learning and Probabilistic Retrieval in Remote Sensing Image Archives*, IEEE Trans. on Geoscience and Remote Sensing, 2000 (submitted}

[32] R. Finn. *Query by image content.* IBM Research, (3):22-25, 1996.

[33] G. Healey and A. Jain. *Retrieving multispectral satellite images using physics-based in-variant representations.* IEEE Trans. on PAMI, 18(8):842-848, 1996.

[34] Mihai Datcu, Klaus Seidel and Gottfried Schwarz, *Elaboration of advanced tools for information retrieval and the design of a new generation of remote sensing ground segment systems*, in Machine Vision in Remote Sensing, Ed. I. Kanellopoulos, Springer (in print).

[35] U. M. Fyyad et. al. (1996) *Advances in knowledge discovery and data mining*, AAAI Press/ The MIT Press.

[36] R. S. Michalski, I. Bratko (1998), *Machine learning and data mining*, Wiley.

[37] M. Datcu, G. Schwarz, M. Walessa, *Wavelet Feature Coding for Quicklook Synthetic Aperture Radar Images: An Image Epitome,* Proc. of EOS/SPIE Vol. 3217, The European Symposium on Aerospace, Remote Sensing-Sept'97, London, pp. 58-66, 1997.

[38] M. Datcu, *Scene Understanding from SAR Images,* Proc. of 1996 International Geoscience and Remote Sensing Symposium (IGARSS'96), pp. 2054-2056, 1996.

[39] M. Datcu, D. Luca, and K. Seidel. *An inverse problem approach for the segmentation of the snow cover in satellite images.* In P. Gudmandsen, editor, Future Trends in Re-mote Sensing , 17th EARSeL Symposium 1997, Lyngby, Denmark, pages 321-328. A. A. Balkema Rotterdam/ Brookfield, 1997.

[40] M. Datcu, G. Schwarz, *The Application of Forward Modelling as a Data Analysis Tool, in Information Processing for Remote Sensing,* Ed. C.H. Chen, World Scientific Publishing Company Pte Ltd, 1999 (in print).

[41] M. Datcu, K. Seidel, and G. Schwarz. *Information mining in remote sensing image archives.* In I. Kanellopoulos, G. Wilkinson, and T. Moons, editors, Machine Vision and Advanced Image Processing in Remote Sensing (MAVIRIC), pages 199-212. Springer, 1999.

[42] http://www.cs.cmu.edu/afs/cs/project/cil/ftp/html/v-demos.htm.

[43] http://www.vision.ee.ethz.ch/npoc/Links/browser.html

[44] http://www.research.ibm.com/networked_data_systems/9auc/

[45] http://www.vision.ee.ethz.ch/~rsia

Mihai Datcu *received his M.S. and Ph.D. degrees in electronics and telecommunications from the University "Politehnica" of Bucharest (UPB), Romania. in 1978 and 1986, and in 1999 the title "Habilitation a diriger des recherches" from University "Louis Pasteur" of Strasbourg, France.*

He holds an image processing professorship with UPB since 1981. Currently he is scientist with the German Remote Sensing Data Center, DFD, of the German Aerospace Center, DLR. He held visiting professor appointments from 1991 to 1992 at the Department of Mathematics of the University of Oviedo, Spain, from 1992 to 1993, 1996 to 1997 and 1998 at the Swiss Federal Institute of Technology, ETH Zurich, Switzerland. His current interest is in multidimensional signal processing, model-based scene understanding, information theory, Bayesian inference, and multiresolution signal analysis, with applications in Synthetic Aperture Radar (SAR) remote sensing.

Klaus Seidel *received his B.S. degree in experimental physics in 1965 and the Ph.D degree in 1971, both from the Swiss Federal Institute of Technology (ETHZ). He is with the Image Science Group at ETHZ and is head of the Image Processing Group for remote sensing applications. He is working in different organisations and committees: since 1985 as National Point of Contact (NPOC) in Switzerland in cooperation with the Swiss Federal Office of Topography. 1987-1997 Swiss delegate at the European Association of Remote Sensing Laboratories (EARSeL). He was chairing EARSeL's Special Interest Group on Land Ice and Snow. Since 1987 Swiss Delegate and Expert in ESA Working Groups. A most recent activity is related to the Swiss Remote Sensing Image Archive.*

He also teaches courses in digital processing of satellite images and has published several papers concerning snow cover monitoring, geometric correction and multispectral analysis of satellite images.