Knowledge-driven Information Mining in Remote-Sensing Image Archives

M. Datcu
Remote Sensing Technology Institute, German Aerospace Centre (DLR), Oberpfaffenhofen, Germany

K. Seidel
Computer Vision Laboratory, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

S. D'Elia & P.G. Marchetti
Earth Observation Applications Department, ESA Directorate of Earth Observation Programmes, ESRIN, Frascati, Italy

Information mining/knowledge discovery and the associated data management are changing the paradigms of user/data interaction by providing simpler and wider access to Earth Observation (EO) data archives. Today, EO data in general and images in particular are retrieved from archives based on such attributes as geographical location, time of acquisition and type of sensor, which provide no insight into the image's actual information content. Experts then interpret the images to extract information using their own personal knowledge, and the service providers and users combine that extracted information with information from other disciplines in order to make or support decisions.

In this scenario, the current offering, which is 'data sets' or 'imagery', does not match the customer's real need, which is for 'information'. The information extraction process is too complex, too expensive and too dependent on user conjecture to be applied systematically over an adequate number of scenes. This hinders access to already available or new data (data accumulation rate is increasing), penalises large environmental-monitoring type projects, and might even leave critical phenomena totally undetected. Emerging technologies could now provide a breakthrough, permitting automatic or semi-automatic information mining supported by 'intelligent' learning systems.

The EO information burden and the technological challenges

Users in all domains require information or information-related services that are focused, concise, reliable, low cost and timely and which are provided in forms and formats compatible with the user's own activities. In the current Earth Observation (EO) scenario, the archiving centres generally only offer data, images and other 'low level' products. The user's needs are being only partially satisfied by a number of, usually small, value-adding companies applying time-consuming (mostly manual) and expensive processes relying on the knowledge of experts to extract information from those data or images.

In the future, these processes will become even more difficult to perform and to manage because of the growing diversity of the user communities, the greater sophistication of user needs requiring, for example, the fusion of multi-sensor or EO and non-EO data, and the exponential increase in the volume and complexity of the data archives, due to the rapid increases in:
- number of missions (even constellations)
- number of sensors
- kinds of sensed data
- sensor resolution
- number of spectral bands
- number of data formats
- number, type and size of distributed archives.

Today's Synthetic Aperture Radar (SAR) and optical sensors generate 10 – 100 Gbytes of data per day, so that in a multi-sensor spacecraft scenario the volume of data to be archived annually easily reaches 10 Tbytes. However, this figure can sometimes be at least one order of magnitude larger: the Shuttle Radar Topography Mission (SRTM) provided about 18Tbytes of SAR data in just 11 days, and ESA's Envisat spacecraft launched on 1 March 2002 is going to collect about 80 Tbytes of multi-sensor data per year! Future European programmes like GMES (Global Monitoring for Environment and Security) will be even more challenging, unless major progress is achieved soon. Emerging technologies for the automatic extraction, classification and easy provision of information, from EO data alone or after fusion with data and information from other fields, could provide this breakthrough.

After 30 years of remote sensing, for almost any site on Earth there are data takes piling up. They contain valuable information that is not being fully exploited because of the lack of automated tools. New technologies are required to automatically analyse such data and data series to detect changes and trends, for example, which could otherwise remain hidden forever or be detected only by chance.

From data to information

In recent years, our ability to store large quantities of data has greatly surpassed our ability to access and meaningfully extract information from it. The state-of-the-art of operational systems for remote-sensing data access, particularly for images, allow queries by geographical location, time of acquisition or type of sensor. This information is often less relevant than the content of the scene, i.e. structures, objects or scattering properties. Meanwhile, many new applications of remote-sensing data require knowledge of the complicated spatial and structural relationships between objects within an image. This knowledge is 'hidden' in the image's structure and must be 'mined' to retrieve meaningful spectral or polarimetric signatures or objects of higher-level abstraction, such as cities, roads, rivers, forests, etc. The hidden information can relate to very localised phenomena, such as subsidence or even to the structural stability of individual buildings, but can also include phenomena related to global change.

Knowledge-driven information mining from EO archives requires the exploitation of a family of methods for knowledge discovery, learning and automatic information extraction from large amounts of data. It may be performed with the identification of a specific feature and application in mind, such as the high density of strong scatterers and structures in SAR images to detect settlements, hot spots in ATSR products to detect fires, etc. Alternatively, it may be used to identify key features without having a specific application in mind at that very moment.

Companies and research centres around the World are devoting a large effort to the second approach through the design and production of Content-Based Image Retrieval (CBIR) systems. Several attempts have already been made to apply the CBIR approach to EO archives, but difficulties have been encountered in applying searching by global image similarity (the basic concept of CBIR tools) because of the predominance in the EO domain of grey-scale and false-colour imagery. The problem of how the system might remember particular features so that it gets 'smarter' with increasing use is also being addressed.

In Europe, IMF-DLR has been working since 1993 in collaboration with ETH Zurich on developing and refining a novel image information mining concept. Unlike traditional feature-extraction methods that rely on analysing pixels and looking for a predefined pattern, it is based on extracting and storing basic characteristics of image pixels and areas, which are then selected (one or more and weighted) by users as representative of the feature being searched for. This approach has a number of advantages:

- there is no need to re-scan the entire image archive to detect new features
- the selected feature can be closer to the user's expectations and perceptions (the same feature can have different meanings for different users: e.g. a forest for an environmentalist, a forest ranger, a geologist, or an urban planner)
- the system can learn from experts' knowledge.

Drawing upon the research experience of IMF-DLR and ETH Zurich and the systems engineering competence of Advanced Computer System SpA, an ESA Technology Research Programme (TRP) project has been started in ESRIN with the title: 'Knowledge-driven Information Mining in Remote-Sensing Image Archives'. This KIM project takes all of the above background into account, as well as the facts that:

- The huge and exponentially growing volumes of existing and new EO data archives need to be more fully exploited in terms of their true information potential.
- Human-centred computing will play an increasing role in the design of EO data exploitation, i.e. intelligent man/machine interfaces, systems that infer and adapt to user needs, etc.
- Fusion of sensor data with non-EO data and information will be used to better understand the identities of the observed scenes and the Earth cover structures.
- Information mining, knowledge discovery and other exploratory information-retrieval methods should be used to try to fully understand highly complex data, phenomena or global observations.
- There is a need to enlarge and reinforce the reconnaissance/surveillance applications spectrum.
- It is necessary to migrate (adopt and promote) from data to information management and dissemination.

KIM tries to satisfy the requirements of the various communities including:

– End users (access to basic information in a simple way).
– The EO value-adding industry and service providers (access to data and information for enhancing existing and providing new services).
– The scientific community (access to large information sets, e.g. for the analysis of global change).
– Civil protection agencies (access to specific information in support of their operational activities, directly or through service providers).
– Institutions involved in education (access to various data and information types to be used as examples or training cases).

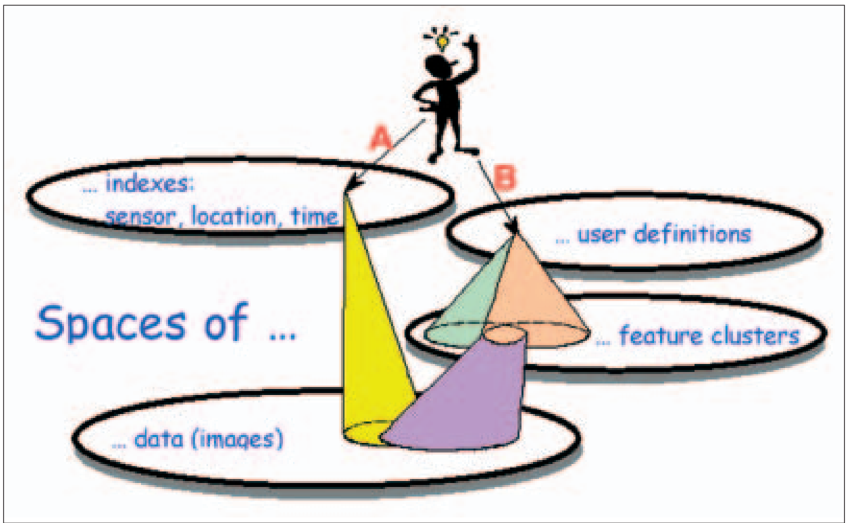


Figure 1. Schematic of current (A) and future (B or A&B) search methods

Table 1. Classification of query attributes

Attribute Type	Description
Basic attributes	Sensor, time, latitude, longitude (or location name), directly related for example to raw data or geocoded products
Primitive features	Intensity, texture or shape, for example: attributes that are both objective, and directly derivable from the images themselves, without the need to refer to any external knowledge base
Derived features	Sometimes known as logical features, these are objects of a given type (e.g. 'mountains') or specific objects (e.g. 'Jura-like mountains'). This level involves some degree of logical inference about the identity of the objects and permits searches in user semantic terms (which can be assigned during system training to weighted combinations of primitive features)
Abstract attributes	Involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted (e.g. illegal plantation). They are outside the scope of the current research activity.

A better grasp of user interaction

In today's EO ground segments, the retrieval of images is mainly based on sensor, location and time criteria (Case A in the schematic of Fig. 1). A more user-friendly interaction model would also permit searches using other attributes of the image or its parts (Cases B or A&B in Fig. 1). Access to a desired image from the archive might thus involve a search through:
– mission attributes such as sensor, time and location
– the presence of a particular combination of intensity, texture and shape
– the presence of specific 'object' types, e.g. forest, rice field, etc.
– the presence of a particular type of event, e.g. burnt forest, flood.

This list of possible queries represents an increasing level of abstraction, complexity and answering difficulty (requiring more and more reference to some body of external knowledge), corresponding to the increasing complexity of the related attributes, which can be classified as shown in Table 1 (from low to high complexity).

A model based on primitive and derived features would also require semi-automated extraction of primitive features, image annotation in terms of primitive and derived features, and the capability to select, weight and combine primitive or derived features during the query process.

A logical representation of the above classification of attributes is reflected in the model of Figure 2, where the arrows represent logical flows. The upward arrows represent unsupervised algorithmic flow, while the downward one describes the creation of the basic attributes or the user's subjective classification. In fact, the identification of derived features is supposed to be performed only via user subjective classification of objects of a given type. Figure 2 appears to include an additional level with respect to the above classification: the 'primitive feature clusters'. This level is necessary only to reduce the data volume, by grouping into clusters primitive features that show similar behaviours (values). In reality, 'primitive features' and 'primitive feature clusters' pertain to the same level.

First step: extraction of primitive features

The IMF-DLR/ ETH Zurich concept applied in KIM is aimed at building a system free from application specificity, so as to enable its open use in almost any scenario, and also to accommodate new scenarios required by the development of new sensor technology or growing user expertise. Its first step is the extraction of primitive features and the reduction of the resulting data into primitive feature clusters.

The primitive features to be extracted need to be carefully selected, since they mainly determine the quality and capabilities of the resulting system. SAR and optical images, for example, will have to be handled differently and texture primitive features will have to be extracted at various resolution levels, since different textures can dominate at different scales.

The steps necessary to properly extract primitive features in the EO context are shown in Table 2. The steps need to be repeated iteratively for each band of the image.

Primitive feature extraction generates a huge amount of data, which cannot be handled in practice and therefore has to be compressed somehow. This process is represented in the left part of Figure 3, which depicts the result of the scanning of two images (or of two bands of the same image). Each pixel of the image will be located in n-dimensional space in the position determined by the values of the contributing primitive features (their units are non-commensurable, e.g. texture and spectral features). The pixels will tend to group themselves into specific regions of this space. Through clustering (right part of Fig. 3), the 'clouds' of image primitive features are replaced by parametric models of their groups, which can be represented in more compact forms. This reduces the precision of the system, similar to a quantisation process, but permits its practical use thanks to the huge data reduction obtained. The primitive features are compressed into clusters using the K-means approach.

The clusters (condensed representation of primitive features) have no direct meaning, since they group points in an n-dimensional space of non-commensurable variables. Still they represent characteristics of the image seen as a multi-dimensional signal. It is possible to associate meaning with these clusters through training. A user can tell the system that a specific, weighted combination of some clusters represents a derived feature of the image. By making this association, it is possible to select all images in the database that have that specific combination and may therefore contain the feature that the user is searching for. This step is discussed below in more detail.

Second step: information mining
The second step in KIM is aimed at assigning physical meaning to the primitive features, i.e.

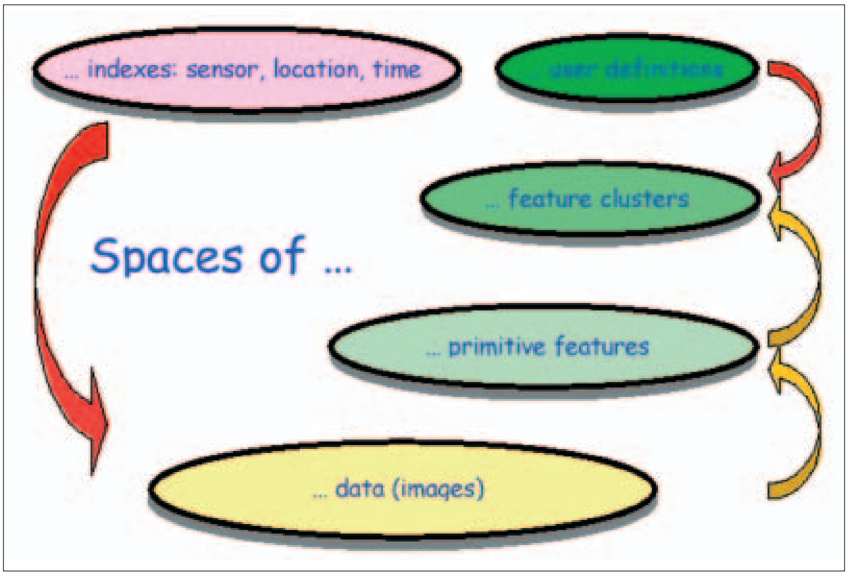


Figure 2. Logical representation of data and derived attributes

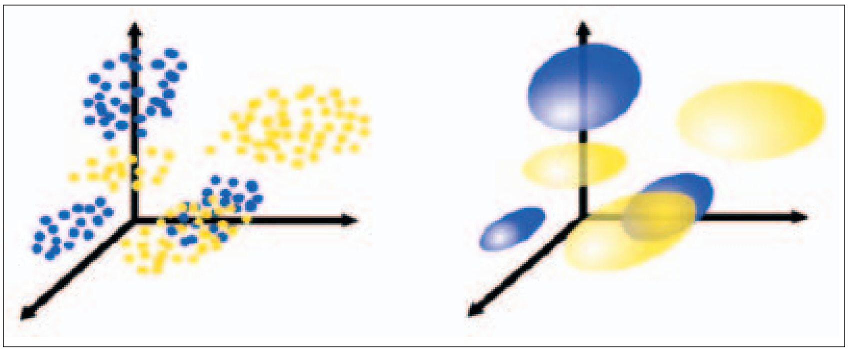


Figure 3. Unclustered and clustered primitive features

Table 2. Steps for primitive feature extraction in the EO context

Step	**Objective**
Image geo-coding	Permit co-registration of different images and absolute geographical reference of the features
Segmentation by geometry	The homogeneous areas that can be detected in the full geo-coded image are assimilated to reference shapes and related to an absolute coordinate system
Sub-scenes	Large images are split into sub-scenes to reduce the probability that an image item contains all the primitive features and therefore that it is always retrieved during a search
Sub-sampling	Perform progressive sub-sampling of each sub-scene to ensure that the various textures are identified at the related scale
Texture analysis:	
Optical image	Extract primitive features using the Gibbs Random Field approach
SAR image	De-speckle the image and extract primitive features using the Gauss-Markov Random Field approach

at identifying 'derived features'. This information-mining step involves a learning phase. The system presents sample images in which the user marks areas with positive and negative traits, refining the definition of the derived feature through an iterative process. Once this process/system training has been satisfactorily completed, the definition can be saved and used by other users, who then will have only to request images containing the derived features corresponding to that definition.

The information-retrieval process is divided into two steps:
– objective information extraction
– semantic representation.

The objective information extraction requires signal modelling as a realisation of a stochastic process. A library of stochastic and deterministic models is used to infer the signal model. The resulting objective features are interpreted according to user conjecture. The interpretation process relies on restructuring (using a certain syntax) of the signal feature space according to the semantic models of the user. Augmentation of the data with meaning can be seen as a data-encoding task including the modelling of the user's understanding.

The source manipulates the information input and provides a filter, which may however also add some process-induced noise. If the user is an expert, he/she plays an active role in training the system, to create relationships between primitive features and definitions. If the user is not an expert, he/she can just use the definitions prepared by other experts for retrieving images from the database.

The expert can train the system by pointing and clicking via a graphical user interface on 'positive' and 'negative' image-structure examples (and therefore the corresponding clusters), in two steps:
– first of all to identify specific broad 'cover-type' definitions, related to broad domains of possible user interest (e.g. geology, forestry, …)
– therafter to create from the aggregation of the above definitions, and the possible use of additional 'training' pixels, more precise definitions with semantic meaning (i.e. 'concepts', like wood, water, grass, urban area, etc.).

A simple Bayesian network links primitive feature clusters and definitions and these associations can be stored and made available to users for subsequent interactive sessions.

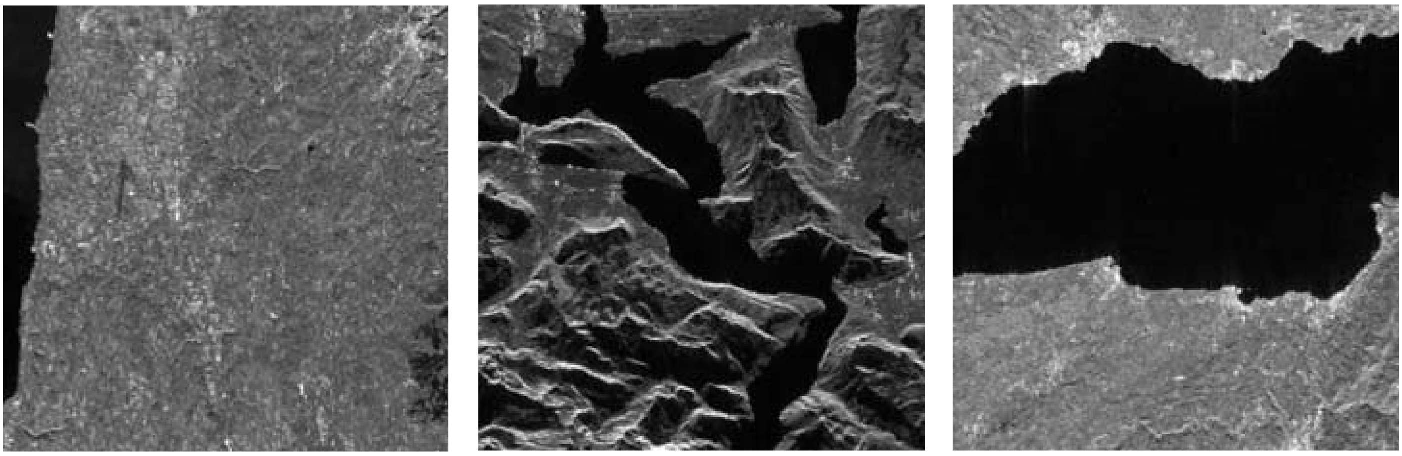
With this approach, we are modelling and learning about the user's interests and actions. We are implementing machine-learning methods to answer the question: What is the user trying to do? We are exploring ways to design an information mining and interpretation system that adapts to the user's particular interests and incorporates contextual information to determine the user's intentions and degree of satisfaction with the results. It should provide a breakthrough by establishing a new pattern for user-EO system (archive) interaction, and a quantum leap with respect to the more traditional feature-extraction systems. The aim with KIM is to help users uncover the most relevant image information content, by providing an 'eye' with which to delve into multi-sensor and multi-temporal image data archives.

Examples of application scenarios

The KIM system provides a wide variety of 'mining' tools, including semantic querying by image content and image example, and interactive classification and learning of image content. As an example, we can take an archive of Landsat TM and Space Radar Laboratory X-band SAR SRL/X-SAR images covering the whole of Switzerland, one of which is shown in Figure 4. The Landsat and X-band SAR scenes have been partitioned into sub-scenes of 2048 x 2048 pixels, with all data geo-coded in a pre-processing step. The user has available a catalogue list of semantically valid land-cover structures, e.g. lakes, forests, alpine valleys, cities, etc. A study of the dynamics of inhabited areas, for example, requires the detection of built-up areas as a first step. The KIM system can search for all SAR images likely to contain settlements larger than a specified threshold, returning the set shown in Figure 5.

Figure 6 shows the result of a search for the cover structure 'glacier'. Each image has associated with it a classification map identifying the structure of interest and a table recording all

Figure 4. Lake Constance on the Swiss-German border (Landsat TM)



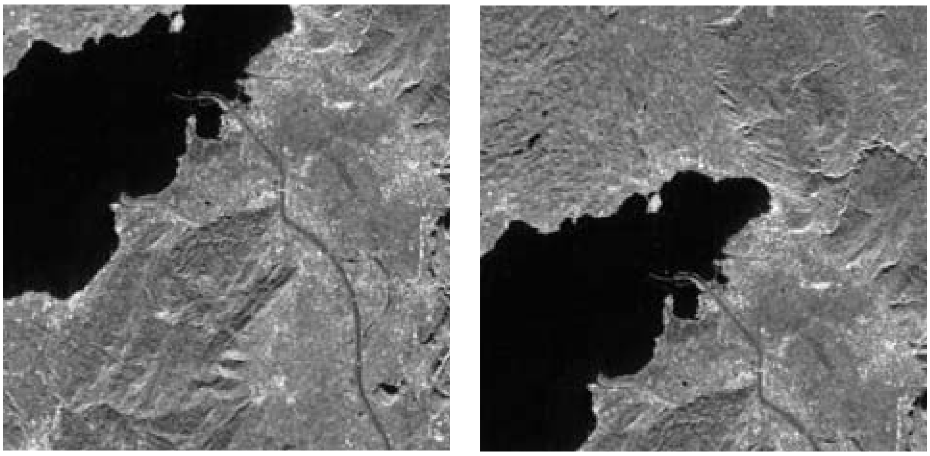
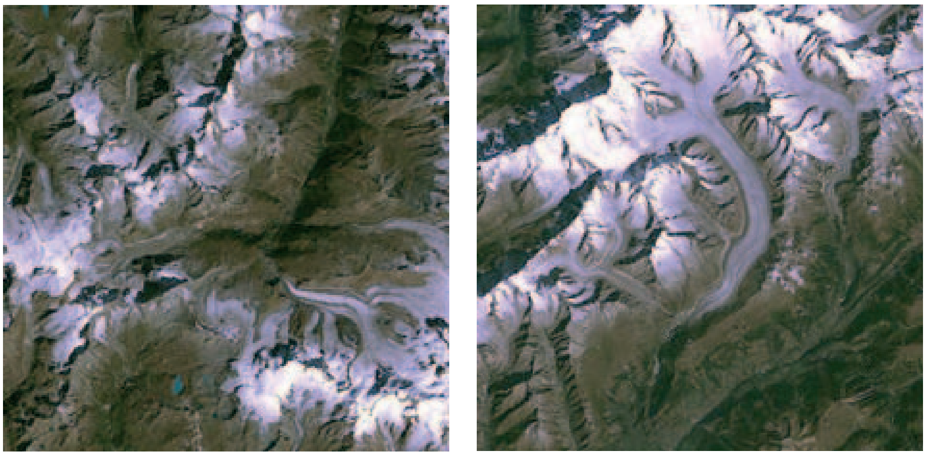


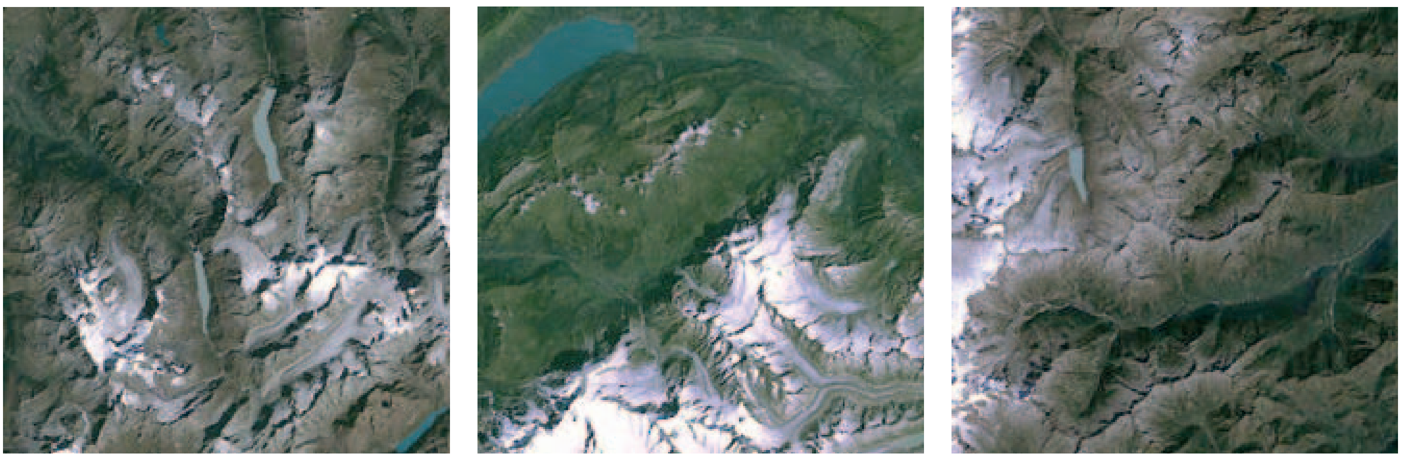
Figure 5. X-SAR images containing built-up areas

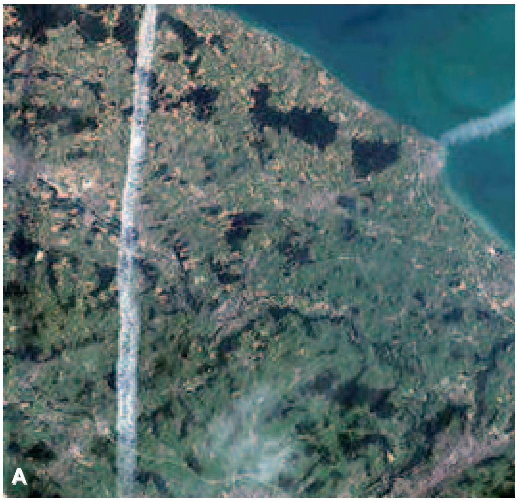
of the ‘structures’ that the users have searched for in this image. The latter is a record of the ‘image content’ seen from the perspectives of users with different interests or backgrounds. The record is created during the interactive learning step.

The interactive learning function is a valuable mining tool for exploring the unknown content of large image archives. A Graphical User Interface (GUI) enables the user to select, by clicking on the image, those structures of greatest interest, which then appear in red on a gray-scale visualisation of the relief

Figure 6. Result of a search for cover type glaciers







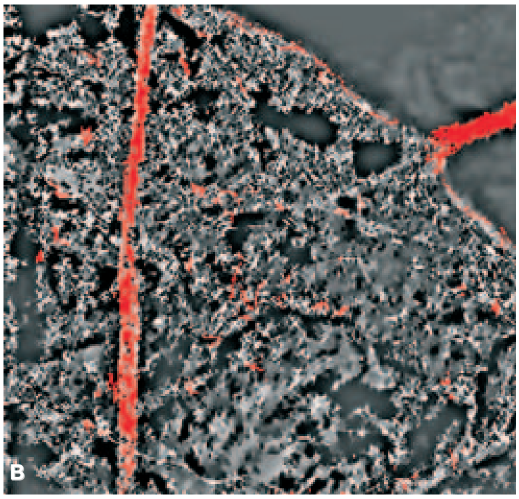


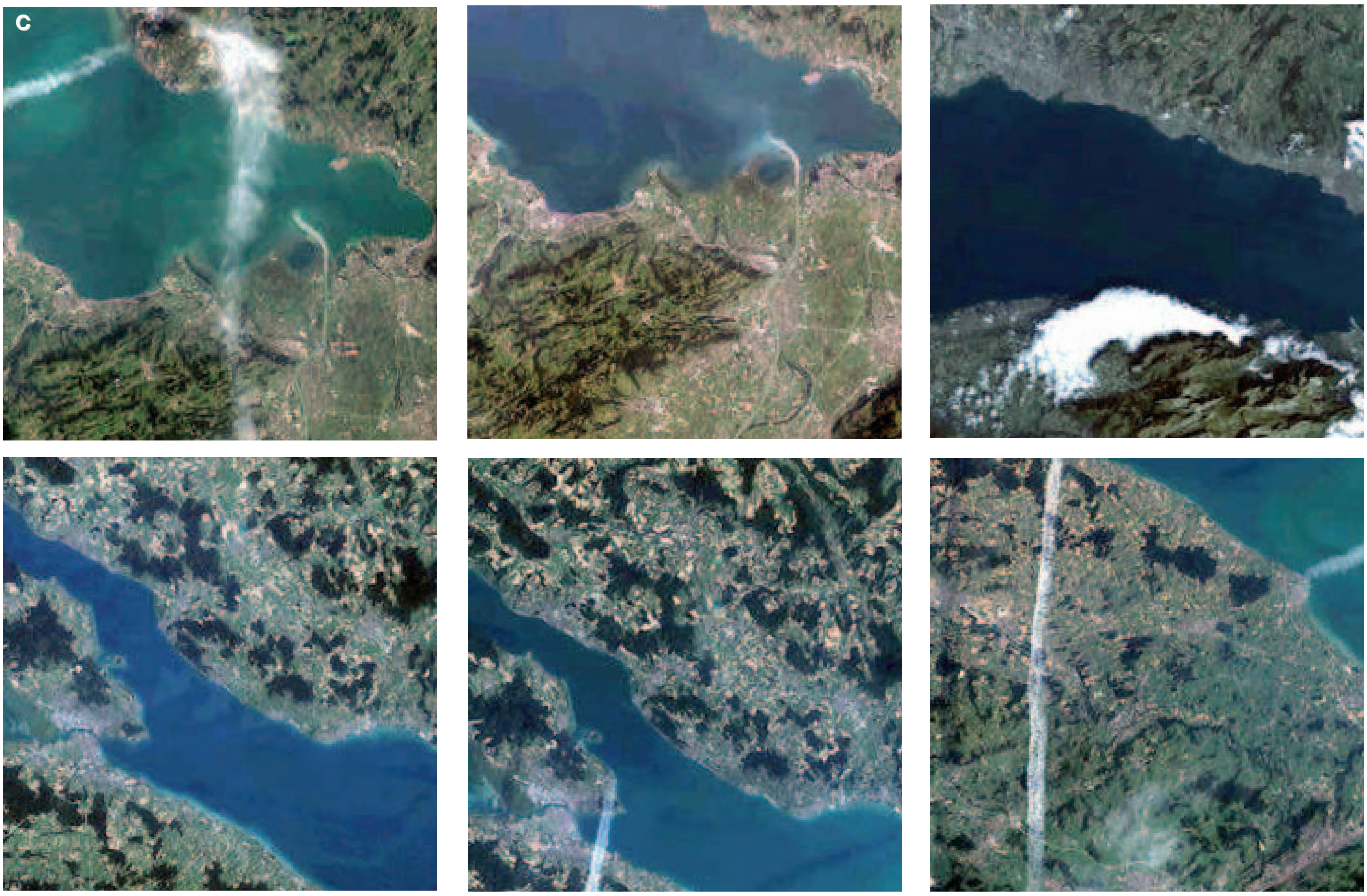
Figure 7
(A) Jet condensation trails in a Landsat image
(B) The user selects the 'interesting' pixels by pointing and clicking
(C) The result of a search for similar features after the selection performed in Figure 7B

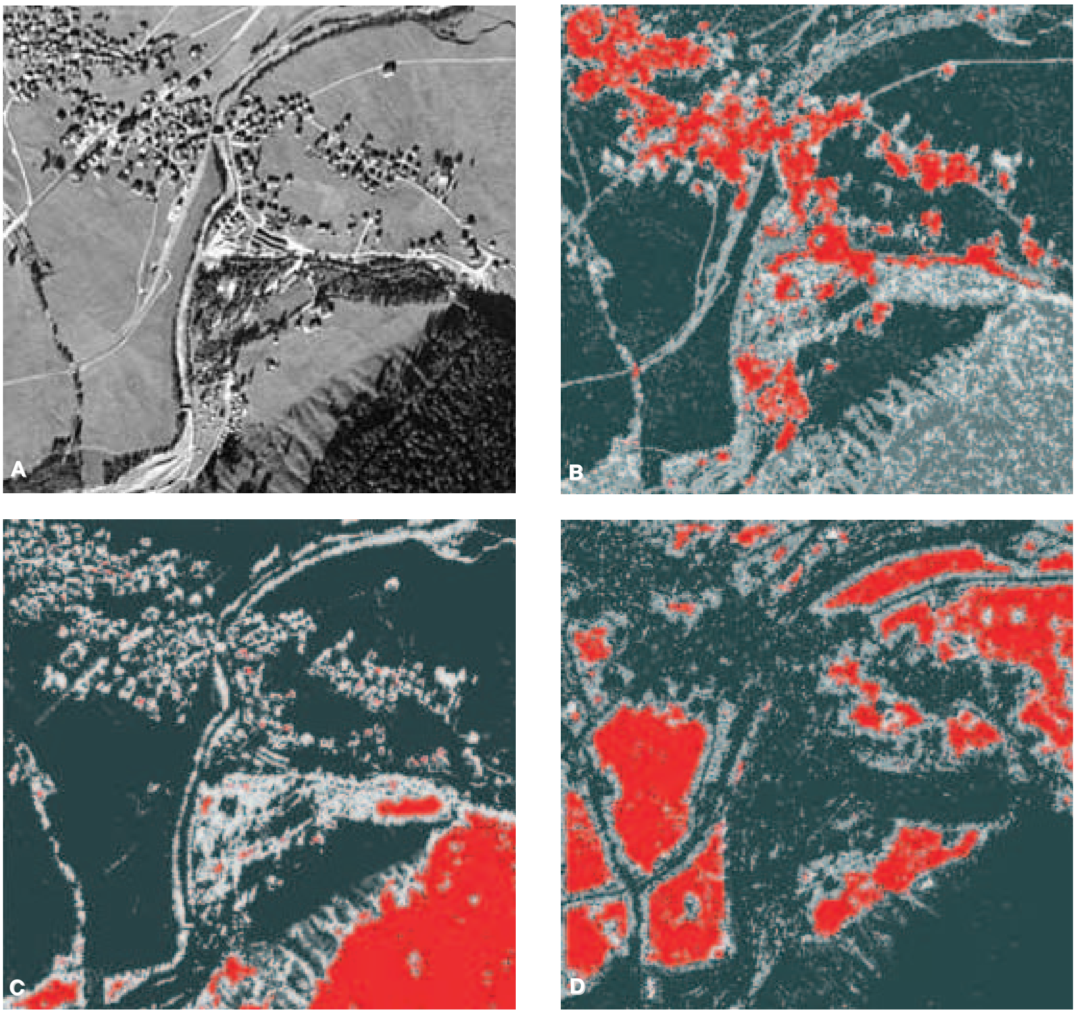
according to the Bayesian learning of the structure recognition (Figs. 7A and B). The scenario in which Figures 7A,B have been produced assumes that the user is interested in finding/studying the condensation trails produced by jet aircraft. The classification is based on the fusion of the image information that best explains the selected structures, i.e. spectral and textural image parameters. KIM thereby provides the user with those images most likely to contain similar structures (Fig. 7C), ranked according to their relevance.

The last examples, in Figure 8, are based on an aerial photograph of a small Swiss village and show how the user can explore the image, marking by pointing and clicking the features of interest. In this case, the red areas associated with 'built-up areas' (Fig. 8B), 'forests' (Fig. 8C), 'meadows' (Fig. 8D) and 'roads' (Fig. 8E), have been generated by a previous user clicking on the image, the KIM interactive learning module being able, in real time, to generate a set of supervised image classification maps. The interactively induced image classification is generalised over the entire image archive.

Conclusions

The technologies for knowledge-driven image information mining are reaching a sufficient level





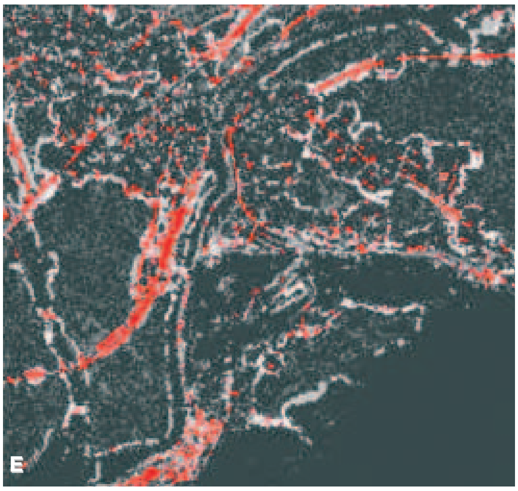


Figure 8.
(A) Aerial photograph of a Swiss village
(B) Selection of built-up area
(C) Selection of forest
(D) Selection of meadows
(E) Selection of roads

of maturity for their integration into commercial products, as has been demonstrated here for a variety of remote-sensing applications. This opens new perspectives and offers huge potential for correlating the information extracted from remote-sensing images with the goals of specific applications.

These technologies shift the focus from data to information, meeting user needs, promoting scientific investigations, and supporting the growth of the value-adding industry, service providers and market, by permitting the provision of new services based on information and knowledge. They will also profoundly affect developments in fields like space exploration, industrial processes, exploitation of resources, media, etc.

The KIM prototype has demonstrated that:
– the results of advanced and very highly complex algorithms for feature extraction can be made available to a large and diverse user community
– the users, who can access the image information content based on their specific background knowledge, can interactively store the meta-information and knowledge
– a new paradigm for the interaction with and exploitation of EO archives can be implemented, paving the way for much easier access to and much wider use of EO data and services.

Acknowledgements
The authors are very grateful to J-L. Valero and H. Touron from the European Union Satellite Centre for their support and comments provided during the preliminary evaluation of the KIM prototype.