IMAGE INFORMATION MINING - EXPLORATION OF EARTH OBSERVATION ARCHIVES

MIHAI DATCU
German Aerospace Center - DLR
Remote Sensing Technology Institute - IMF
Oberpfaffenhofen, D-82234 Wessling
Tel: +49 8153 28 1388
Fax: +49 8153 28 1444
Email: mihai.datcu@dlr.de

KLAUS SEIDEL
Remote Sensing Group
Computer Vision Lab ETH
CH 8092 Zürich - SWITZERLAND
Tel: +41 1 632 5284
Fax: +41 1 632 1251
Email: seidel@vision.ee.ethz.ch

Abstract
The new generation of high resolution imaging satellites acquires huge amounts of data which are stored in large archives. The state-of-the-art systems for data access allow only queries by geographical location, time of acquisition or type of sensor. This information is often less important than the content of the scene, i.e. structures, objects or scattering properties. Meanwhile, many new applications of remote sensing data are closer to computer vision and require the knowledge of complicated spatial and structural relationships among image objects.

We are creating an intelligent satellite information mining system, a next generation architecture to help users to gather rapidly information during courses of actions, a tool to add value and to manage the huge amount 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.

Research topics are within the frame of Baysian learning, content-based querying, data modelling, adaptation to user conjecture.

Keywords: remote sensing, information mining, archiving
1. Introduction

Information mining opens new perspectives and a huge potential for information extraction from large volumes of heterogeneous images and the correlation of this information with the goals of applications.

The image archives are heterogeneous, huge data repositories, they are high complexity sources of valuable information, e.g. the Earth Observation data archives contain millions of optical, radar and other types of images and data. The exploration of their content is not an easy task. Among the promising methods proposed in the last years are the methods of data and information mining. However, accessing the image information content involves highly complex problems arising primarily from the huge volume of data, the rich information content, and the subjectivity of the user interpretation. The present article makes an analysis of the Image Information Mining methods seen as an information transmission problem: the source of information is an image archive, the receiver is the community of users. Data and information mining are exploratory processes focusing on the techniques for analyzing and combining raw data and detecting patterns and regularities within the data set. The success of the exploratory information search depends on the capacity to capture and describe the full complexity of the data. Thus we use a concept integrating multiple methods: information theory, stochastic modelling, Bayesian inference, machine learning. Information theory deals with encoding data in order to transmit it correctly and efficiently. The theory of stochastic processes and machine learning deal with estimating models of data and predicting future observations. There is a relationship between these fields: the most compact encoding of the data is by the probabilistic model that describes it best, thus there is a fundamental link between information and probabilistic models. This link is basic to the implementation of optimal algorithms for information extraction, detecting causalities, and for the design of information systems implementing image information mining functions. The article presents and analyzes several methods for mining the information content of large image repositories, and demonstrates image mining functions, like, search by example, search by data model, exploration in the scale space and image complexity, knowledge acquisition, and adapting to the user conjecture.

2. From content based image retrieval to mining the image information

The continuous expansion of multimedia into all sectors of activity faces us with a double explosion:

- the number of image data sets
- the data size and information variability of each image

E.g. with a digital camera we can acquire 10 Gb of images during a 3 weeks holiday, a satellite sensor can acquire 100 Gb per day.

It has been known since many years that classical image file text annotation is prohibitive for large data bases. The last decade is marked by important research efforts in the development of Content based Image Retrieval (CBIR) concepts and systems [11]. Images in an archive are searched by their visual similarity with respect to color, texture or shape characteristics. While image size and information content is continuously growing CBIR was not any more satisfactory and Region Based Information Retrieval (RBIR) was developed [11]. Each image is segmented and individual objects are indexed by primitive attributes like color, texture and shape. Thus, RBIR is a solution to deal with the variability of image content.

However, both CBIR and RBIR have been computer centered approaches, i.e. the concepts could only marginally or not at all adapt to the user needs. Consequently, the image retrieval systems have been equipped with relevance feedback functions [1].
The systems are designed to search images similar to the user conjecture. The algorithms are based on analyses of the probabilities of an image relative to the search target. A feedback mechanism which takes this part into account is introduced.

Another interesting approach was developed based on a learning algorithm to select and combine feature grouping and to allow users to give positive and negative examples. The method refines the user interaction and enhances the quality of the queries [8].

Both previously mentioned concepts are first steps to include the user in the search loop, they are information mining concepts. Also, these are methods in the trend of designing human centered systems.

3. Images and image information

Compared with Data Mining the field of Image Information Mining reaches much higher complexity resulting from:

- the huge volume of data (Tb to Pb)
- the variability and heterogeneity of the image data (diversity of sensors, time or conditions of acquisition, etc)
- the image content, its meaning is many times subjective, depending on the users interest
- the large range of user interest, semantics and contextual (semiotic) understanding.

In general, by image we understand picture thus relating it to the (human) visual perception and understanding. A picture is characterized by its primitive features such as color, texture, shape at different scales. Its perception and understanding is in form of symbols and semantics in a certain semiotic context [12].

However, the concept of image is beyond the pictorial understanding. Images are multidimensional signals, like computer tomography, hyperspectral images or results of simulations. They are communicated to users via 2-dimensional visual projections. These images can contain quantitative, objective information, as acquired by an instrument.

In Fig. 1 an example is presented for the visualization of a data set of a Digital Terrain Model (DEM) in comparison with a color rendered satellite image of the same Alpine region. The visual information in the DEM image is not easy to read. The information of terrain elevation is contained in the image samples. The color image, however, shows the complexity of pictorial information.

In the perspective of image information mining both types of images, pictorial and multidimensional signals give rise to the same problem. Their understanding depends on the accuracy of:

- information content modelling
- modelling the users understanding.

Thus, image information mining can be seen as a communication task. The source of information is the large heterogeneous image archive. The receiver is the community of users. The accuracy of communication, i.e. the success of finding the information needed as exploration results, depends on the accuracy of the previously assumed levels of modelling.
4. Information mining: concept

We developed a theoretical concept for image information representation and adaptation for the user conjecture [2,3,4,6,7]. A quasi-complete description of the image content is obtained by utilization of a library of models. The feature extraction is equivalent to the splitting the image content into different information channels. An unsupervised clustering is done for each information channel as an information encoding and data reduction operation. Then, during the operation of the system, an interactive learning process allows the user to create links, i.e. to discover conditions between the low-level signal description and the target of the user.

Figure 1. Quantitative versus pictorial information
Top: Visualization of a digital Elevation Model (DEM) data set of Davos, Switzerland. The information on terrain height is contained in the pixel intensity, the information is quantitative and is not rich in visual meaning.
Bottom: Satellite image (Landsat TM) of the same area. The information is pictorial, aggregation of colors, textures and geometrical objects at different scales makes it possible to understand the scenery of an alpine ski resort.
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, color and image texture carries information about the structure of object surfaces. However, in the case of modelling high complexity signals, a large number of sources coexists within the same system, thus multiple candidates models are needed to describe the information sources in the image. Also, to reduce complexity, to capture the class structure, and discover causalities and to provide computational advantages, the models are likely to be analyzed hierarchically. The hierarchical information representation is further presented and depicted in Fig. 2:

- **Image data**: the information is contained in the samples of the raw data. It is the lowest level of information representation.
- **Image features**: 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 non-stationary sources, accuracy also depends on adaptation to local characteristics. For a quasi-complete characterization the image content, information is extracted in form of parameters characterizing: color or spectral properties, texture as interactions among spatially distributed samples, the geometrical attributes of image objects.
- **Meta features**: estimation of the image features, 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 cooccurrence matrix are not comparable with parameters of Markov random fields. The meta features have semantic value.
- **Cluster model**: the signal features have n-dimensional representations. Due to observation noise or model approximations the feature space is not occupied homogeneously. Thus, another level of information abstraction is the type of feature grouping, i.e. the cluster models, and the associated parameters. The obtained clusters represent information only for each category of the features.
- **Semantic representation**: 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. Augmentation of data with meaning requires a higher level of abstraction. The extracted information, represented in the form of classes is fused in a supervised learning process. Prior
information in the form of training data sets or expert knowledge is used to create semantic categories by associations to different information classes. Thus, the observations are labelled and the contextual meaning is defined.

In order to implement the hierarchical representation of the image information content, the data are pre-processed. First, the image features are extracted for different image scales. Then the image features are clustered, and further a signal content index is created using the cluster description, the scale information and the type of stochastic model assumed for the image parameters. A Bayesian learning algorithm 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, and model types. The index of each image pixel is encoded by the spatial correspondence of the class information. 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. The hierarchical information, meta-information, associations and semantic labels are stored and managed by a Data Base Management System.

4.1. Mining driven by primitive signal features

Mining driven by primitive signal features, such as spectral signatures or structural patterns, is enabled by the exploration of the links between the cluster and image data levels. Examples of spectral and textural signature mining are depicted in Fig. 3. The spectral mining is an example of physical, quantitative model exploration. For the Landsat-TM images used in the example only 6 spectral bands were selected.

![Figure 3. Example of image information content extraction in a Landsat TM image of Switzerland.
Left: Spectral image content, in red, obtained by the correlation of a specified cluster model with the pixel position in the image.
Right: Texture image content obtained in similar manner, however, the textural information characterizes structures, thus the resulting classification has connected areas. The information is indexed enabling the discovery of all images with similar spectral or textural properties.](image-url)
4.2. Mining information theoretical measures

In the exploration of large image archives with rich information content it is important to group the data according to various objective information measures. This helps the users to orient within the search process.

One important characteristic is the scale at which relevant information is concentrated. We used an multiscale stochastic process for automatic scale detection and segmentation [9,10]. An example is shown in Fig. 4. The exploration of image archives by scale is a process which is implicitly using a priori knowledge assumed by the user: the ratio of the image resolution and size of objects he is searching for.

The complexity of the images is another information theoretical measure used to rank images. The complexity is defined as the Kullback-Leiber divergence between the cluster level and the image data level. The complexity depends on the quality and type of model used. In Fig. 5 examples of ranking images are presented according to their spectral and textural complexity.
4.3. Mining by interactive learning

Interactive learning is the process of discovering the links between the user interest (target), the image content in terms of describing models and the images containing the assumed structure [3,7]. In a first step, interactive learning uses a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures. In a second step, also using a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures. In a second step, also using a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures. In a second step, also using a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures. In a second step, also using a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures. In a second step, also using a Bayesian network to create the links between the concept and cluster levels. During interactive learning the image data (quicklooks) are used to give examples and to index the spatial position of the target structures.
approach, a probabilistic search over the image space is performed. At this stage the links between the concept level, clusters and image data levels are created. The learning process uses positive and negative examples, both from the user and machine. It is a man-machine dialog.

In Fig. 6 an example is presented for the exploration of different models (texture at various scales and spectral signatures) to discover different semantic objects in the data.
Figure 6. Interactive training

Top: Interactive training using fusion of spectral and textural information at the finest image scale. The target semantics is “meadow”.

Bottom: On the same image, interactive training using fusion of texture information estimated for scales 1:2 and 1:3, the target semantics is “mountain”.

Interactive learning is an information mining which enables adaption to the user conjecture. It is a pure exploratory function based on learning, fusion, and classification processes, using the pre-extracted image primitive attributes, and allowing an open, very large semantic space. The user defined target is generalized over the entire image archive, thus allowing further exploration.
The results of the probabilistic search are depicted in Fig. 7 for the cases indicated in Fig. 6.

Figure 7. Probabilistic search
Top: the result of probabilistic search for images containing “meadow”.
Bottom: the result of probabilistic search for images containing “mountains”. Both query results correspond to the interactive training as defined in Fig. 6.
4.4. Semantic Content Based Image Retrieval

Following automatic processing at data ingestion or in a semi-automatic manner using an interactive learning process, the system can create links between the concept level and the image data and cluster levels.

The user is enabled to specify semantic queries at concept level and the system returns all images with the specified content and a classification on individual images. An example is given in Fig. 8.

Figure 8. Semantic CBIR on Synthetic Aperture Radar X-SAR SRL images of Switzerland. Top: Result of semantic query - discovering settlements. The images have been automatically analyzed at ingestion in the archive, and a catalogue entry was created for all images containing built up areas. Bottom: Each image has attached the result of the classification, the regions marked in red correspond to villages and cities, thus the result of the query is the list of images, augmented with the expected semantic image content.

In the case of Earth Observation the geographical location is also used as meta-information allowing finding the location of the intensity images as indicated in Fig. 9.

4.5. Knowledge driven image information mining and user conjecture

During the interactive learning and probabilistic search the database management system (DBMS) holds a record of:

- the user semantic
• the combination of models able to explain the user’s target
• the classification of the target structure in each individual image
• a set of statistical and information theoretical measures of goodness of the learning process.

This information and associations represent a body of knowledge, either discovered or learned from the various system users. This information is further being used for other mining tasks. This acquired and learned information and knowledge is itself object of mining, e.g. grouping of semantic levels, relevance feedback, joint grouping between the semantic space and the statistical or information theoretical measures of goodness of the learning process.

5. Information mining: system architecture

The system is implemented in a server-client architecture as presented in Fig. 10.

This concept was implemented and successfully demonstrated with an on-line experimental system, see http://isis.dlr.de/mining. The novel mining functions presently provided by the system are further presented.

6. Conclusions

We based and developed a new concept for image information mining. We regard the mining process as a communication task, from a user centered perspective. The hierarchy of information representation, in conjunction with the quasi-complete image content description, enables implementation of a large variety of mining functions. The concept was demonstrated for a variety of Earth Observation data. Further work is being performed for the development of intelligent satellite ground segment systems, and value adding tools. However its potential is broader, other fields of applications are possible, such as medical imagery, biometrics, etc.
The proposed concept is far away from being fully exploited. Presently ongoing theoretical development is profundating the problematic of image complexity. In the case of high heterogeneity observations the complexity and the course of dimensionality are two key issues which can hinder the interpretation. Therefore, as an alternative solution to the “interpretation”, we propose an exploratory methodology approached from a information theoretical perspective in a Bayesian frame.

Another direction is the analysis of cluster models from the perspective of an “objective” semantic approach, aiming at the elaboration of methods to understand the nature of the feature space.

A direction of application of the developed methodology is the mining of temporal series of images, considering the integration of spatio-temporal signal analysis.

Even the concept of learning the user conjecture was at some extent demonstrated. Difficult problems are under further research, such as developing image grammars and representation of image content in different contextual environments. This is a semantic problem which can arise between different users when they define or describe the same structures differently, requiring the primitive attributes, features, domains, values, or causalities to be translated.

A number of challenges, mainly in the design of multidimensional DBMS, man-machine interfaces, distributed information systems, will probably be investigated soon.

ACKNOWLEDGEMENT

The project has been supported by the Swiss Federal Institute of Technology (ETH) Research Foundation Advanced Query and Retrieval Techniques for Remote Sensing Image Archives (Grant: RSIA 0-20255-96). The author would like to thank Michael Schröder and Hubert Rehrauer for converting the concept into algorithms and setting up the Multi-Mission Demonstrator (MMDEMO).

REFERENCES


Where is the information?

Where is the information? - The information is in models. Information about an unknown variable $x$ can be obtained if we are able to model the incertitude of the variable. If the probabilistic model is $p(x)$, the information is

$$i(x) = \log \frac{1}{p(x)}$$

In the case of mining information in large Earth Observation (EO) archives, we need to build hypotheses about the possible models describing the data. We base our model on the prior knowledge of the image formation process, thus being able to describe the causalities which link the physical parameters of the scene to the image content and further to the user understanding in terms of symbols and semantics he is using.

In the figure below the generic image formation process upgraded with the user models is presented.

Traditionally, in interpretation of EO data two operations are performed:

1. image understanding which has as goals the classification and/or recognition of 2-dimensional structures in the data
2. scene understanding with the extraction of scene parameters from the 2-dimensional image observation.

In the case of image information mining the problem to face is more complex. The large amount of data is implying a broad diversity of the image structures and therefore a much larger class of models is needed.

In our approach we propose a quasi-complete image content description for optical data with the use of

1. spectral signatures
2. texture
3. multiscale structural analyses

4. geometrical features

However, other sensor data require different types of models, e.g. synthetic aperture radar data needs specific modelling of the coherent phenomena.
A Bayesian Approach for Image Information Mining.

A Bayesian stochastic model is the tuple

\[ \{ p(y|x), p(x) \} \]

where \( p(y|x) \) is the parametric likelihood model and \( p(x) \) is a prior distribution defined in the parameter space.

The Bayesian Theorem

\[ p(x|y) = \frac{p(y|x)p(x)}{p(y)} \]

shows how the information in \( x \) is updated by the information contained in the data \( y \).

A Bayesian hierarchical model assumes the prior \( p(x) \) is decomposed in conditional distributions

\[ \{ p(x|\theta_1), p(\theta_1|\theta_2), \ldots, p(\theta_\mu) \} \]

where \( p(\theta_\mu) \) is the marginal distribution.

From this perspective information mining in image archives is defined as a Bayesian learning of a hierarchical model.

\[ \{ p(I|\omega), p(\omega|L) \} \]

which explains the dependencies between the image data \( I \), the image content extracted with different models \( \{\omega\} \) and the semantic labels \( L \) defined by the user.

The inference is carried on as a Bayesian network, as presented in the figure below.

Three mining functions can be implemented:

1. the search of all images containing a target semantic label

\[ p(\omega_j|I_k) \quad p(L_i|\omega_j) \]
The three mining functions are exploiting the hierarchical causalities modeled by a Bayesian network.
Insert 3:

Swiss National Archive for Remote Sensing Images

In recent years most of the available satellite images over the territory of Switzerland have been stored in the Swiss National Remote Sensing Image Archive. This archive was initiated and is maintained by the National Point of Contact (NPOC) of Switzerland in cooperation with the Swiss Federal Office of Topography and the Remote Sensing Labs (RSL) at the University of Zürich (www.npoc.ch).

The archive is installed at the Centro Swizzero di Calcolo Scientifico (CSCS, Swiss High Performance Computer Center) in Manno which belongs to the Swiss Federal Institute of Techonology (ETH).

As a result of an ongoing research activity at the Computer Vision Lab (ETH) a Multi-Mission Demonstrator (MMDEMO) has been set up in order to show the advantages as soon as more sophisticated catalogue systems become available. A variety of datasets of the national archive are used in the demonstrator for content-based querying and retrieval techniques:

http://www.vision.ee.ethz.ch/~rsia/