Image Information Mining
for
Earth Observation
at
ESA

Sergio D’Elia
Head of Research and Technology Development Section
EO Ground Segment Department
Sergio.DElia@esa.int
**EO Data Exploitation Chain**

- **Distributors**
- **Value Adders**
- **Service Providers**

**Data Providers**

- **Archives (PBytes)**

- **EO Data / Products**

- **Knowledge**

- **Information / Service**

- **Non-EO Data**

- **Users (KBytes)**

Support / Automation possible?
Expand use of EO data by supporting / automating the identification in images of user relevant information (Features (*))

- Assist and facilitate image analysis and interpretation
  Information Discovery

- Permit image selection also on the basis of information content
  (in addition to the classical search arguments: mission, area, time, ...)
  Content Based Image Selection

- Permit direct access to the image features
  Feature Access

Explore possibilities offered also for archives by

- Feature Extraction Algorithms
- Probabilistic Information Mining

(*) Objects or homogeneous areas adhering to specific models (cloud, fire, forest, ...
Specialised algorithms permit the identification of features by image post processing (with possible revision by expert)

Advantages
- Best tuned for the specific feature
  (a set of conditions uniquely marking the feature must be identified)

Disadvantages
- Reprocessing of the whole image set required for
  - new versions of the algorithms
  - new algorithms capable to detect new features
- Features defined a priori for all users
First step: automatic ingestion of images
   - Identify at various scales image Primitive Features (characteristics of pixels/surrounding areas: radiometry, texture, geometry, ...)
   - Cluster Primitive Features’ data (for storage minimisation)

Second step: user interaction with the system
   - The user trains the system by providing positive and negative examples (by clicking with left or right mouse button on image pixels)
   - The system links dynamically, via a probabilistic network, the training commands to weighted combinations of Primitive Features and returns best matching images
   - The user loops until satisfied
   - The user assigns a semantic definition (for future use or for use by others)
**Advantages**

- New training permits to detect new features without re-ingesting all images (if feature compatible with the ingestion parameters)
- A user can define features closer to his expectation and perception (a feature can be slightly different or have different meanings for different users: a forest for a geologist, an environmentalist, a forester, a urban planner, …)

**Disadvantages**

- Requires storage proportional to archive size and feature resolution (currently it is possible to handle small archives for medium scale features or only a few images at pixel resolution)
- Probabilistic (confidence vs. feature extraction being assessed)
IIM Research Topics

**Probabilistic IM**
- Primitive Features (automatic extraction and multi-dimensional clustering)
  - Learning Systems / Training
  - Probabilistic Search
    - Ingestion Automation
    - Incremental Clustering
    - Semantic Aggregation
    - Multi Data-type Handling

**Workflow**
- Fast Databases, Scalability
- Classification, Ontology
- Knowledge Discovery
- Mediation

**Prototypes**
- Surveys of Algorithms
- Impact Crater Discovery
- Feature Extraction

**Further Research**
- Co- / absolute-registration
- Space-time (geometry, time series)
- Data / Information Fusion
- KD, Mediation (advanced)
# IIM Research Pattern at ESA

**Image Information Mining - Time Series**

<table>
<thead>
<tr>
<th>Year</th>
<th>IIM-TS</th>
<th>UGEI</th>
<th>UGEIP</th>
<th>MIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Universal Geometry Engine for EO Images (GSTP) - Registration**

<table>
<thead>
<tr>
<th></th>
<th>UGEI</th>
<th>UGEIP</th>
<th>MIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal Geometry Engine for EO Images - Preparation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple Image Registration Tools</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Knowledge Centred Earth Observation - IIM Services**

<table>
<thead>
<tr>
<th></th>
<th>KEO</th>
<th>MIMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MERIS Information Mining Service (EOP)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Knowledge Enabled Services**

**Impact Crater Discovery (GSP) - Feature Extraction**

<table>
<thead>
<tr>
<th></th>
<th>ICDY</th>
<th>SURF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Crater Discovery (GSP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey / classification of advanced features manipulation techniques and tools for EO</td>
<td></td>
<td>Survey of algorithms for automatic recognition of impact craters (GSP)</td>
</tr>
</tbody>
</table>

**KIM Validation (EOP) - Probabilistic IM**

<table>
<thead>
<tr>
<th></th>
<th>KIMV</th>
<th>KES</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIM Validation (EOP)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Knowledge Enabled Services**

**Knowledge driven information mining in remote sensing image archives**

<table>
<thead>
<tr>
<th></th>
<th>KIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
</tr>
</tbody>
</table>
IIM Research Results

Probabilistic IM
- Primitive Features (automatic extraction and multi-dimensional clustering)
- Learning Systems / Training
- Probabilistic Search
- Ingestion Automation
- Incremental Clustering
- Semantic Aggregation
- Multi Data-type Handling

System / Interface
- Workflow
- Fast Databases, Scalability
- Classification, Ontology
- Knowledge Discovery
- Mediation

Feature Extraction
- Prototypes
- Survey of Algorithms
- Impact Crater Discovery

Further Research
- Impact Crater Discovery

Information Discovery
- CB Image Selection

IIM Services for archives
- (= above + basic Feature Access)

KEO
- IIM Services for archives

KD, Mediation (advanced)
- Space-time (geometry, time series)
- Data / Information Fusion

KIMV, MIMS
- Information Discovery
KIM Architecture

Multi-sensor sequence of images

Data ingestion

Primitive features extraction

Data acquisition, pre-processing, archiving

Image archive

Inventory

Browsing engine

Query engine

Classification

Knowledge acquisition

Interactive learning

Information fusion and interpretation

User
KIM User Training Interface

Give positive and negative examples
Available from November 2004 (with > 30 EO related services)

- Easies service publication, identification, fruition and chaining
- Supports on-line and off-line services
- Services remain at Service Providers’ sites (local control)
- Based on standards, implements four commands (web services)
  - Request for Quotation, Order
  - Search, Present
- ESA IIM projects directed to provide SSE compatible services (KES-B, MIMS, KEO)
The Image Information Mining Coordination Group (IIMCG) coordinates EO related IIM activities in Europe.

Founded by agencies and research institutes (ASI, CNES, CNR, DLR, EC, ESA, ETHZ, EUSC).

The group operates on a voluntary basis to:
- Promote European research and development on IIM
- Interface with European and National programmes
- Foster cooperation within members
- Provide a reference point for external entities
- Facilitate / promote standards and use of products

The group organises periodical workshops / conferences on IIM.
Useful Links

- ESA EO GS research and technology development projects
  http://earth.esa.int/rtd

- IIMCG
  http://earth.esa.int/rtd/IIMCG

- IIMCG workshops / conferences
  http://earth.esa.int/rtd/IIMCG/events.html

- SSE Introduction Workshop
  http://earth.esa.int/rtd/Events/SSE_2004

- SSE
  http://services.eoportal.eo

- KIM (Image Information Mining Prototype)
  http://www.acsys.it:8080/kim/index.html