

ASQuLD: an Advanced Semantic Query System for Large Satellite Database

Alan Beccati^{1,2}, Mario Cavicchi¹, Marco Bascetta¹, Roberto Fabrizi¹ and Stefano Natali¹
¹ Meteorological and Environmental Earth Observation, Ferrara, Italy
² University of Ferrara, Computer Science Department, Ferrara, Italy

Introduction

Many satellite sensors have observed the Earth-atmosphere system for decades, collecting hundreds of terabytes of compressed data. Most of the existing databases are not accessible on-line, and queries are performed only on existing catalogues that host few information (geographic, temporal and, in few cases, cloud cover) for each BD Element (image).

The future scenario is even more complex, with multi-hyperspectral sensors/constellations planned to generate many gigabytes per day.

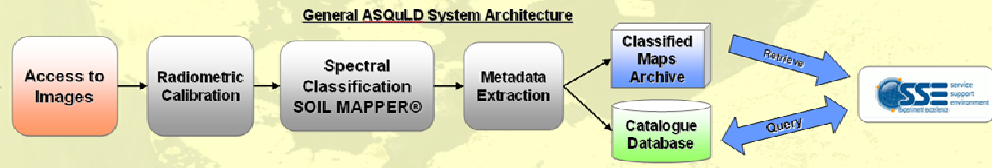
The proposed approach, named Advanced Semantic Query system for Large Satellite Database (ASQuLD), aims at improving the entire databases exploitation permitting effective queries based on Land Cover semantic parameters.

ASQuLD has been already implemented for the entire ATSR-2 / AATSR database (13years of data, 20TB) and is planned to be implemented for a test database of AVNIR-2 data (3 years of data over Europe) in the framework of the ESA projects CARD and SPA, but the same approach can be applied the entire range of optical sensors (from mid to very high resolution).

System Architecture

Main modules:

- > Access to satellite images
- > Data processing:
 - Radiometric Calibration
 - Classification with SOIL MAPPER®
 - Metadata extraction
- > Catalogue creation
- > Classification Map storage
- > Access (advanced queries)



Focus on the Preliminary classification maps – SOIL MAPPER®

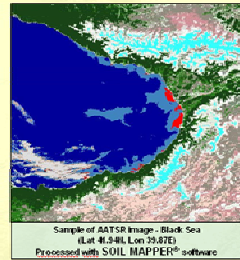
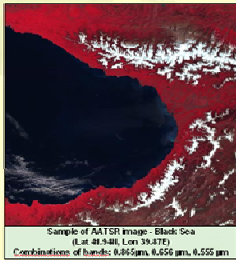
SOIL MAPPER® is an automatic unsupervised, multi-sensor, spectral classifier of remotely-sensed imagery that, starting from radiometrically calibrated TOA reflectance / temperature values, generates Spectral Classification maps.

Currently supported sensors:

MODIS, AVHRR, AATSR, MERIS, Landsat 5/7, ASTER, SPOT4-5, IRS 1-C/4, IRS-P6, IKONOS, ALOS/AVNIR-2, QuickBird.

Spectral categories detected by SOIL MAPPER®:

- Vegetation
- Snow / Ice
- Water / Shadows
- Bare soil / Built-up
- Clouds
- Outliers



Access stored data – ATSR2-AATSR Classification Map Data Catalogue

The system allows accessing in real time the entire preliminary classification maps database, through geographic, temporal and semantic queries.

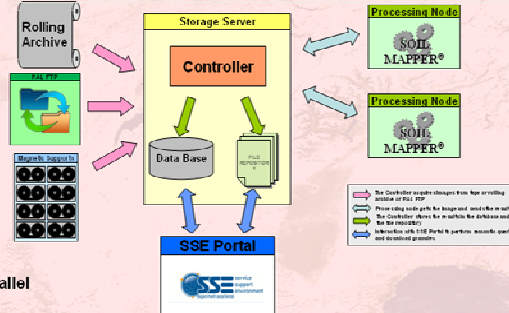
The semantic queries are performed through an SSE service web interface. The user can select a range of dates, an area of interest (AOI) and one rule for each macro-class of spectral category (in terms of percentage threshold - greater than, smaller than - of pixels of the granule belonging to the specific class).

The screenshot shows a 'Search Page' with a world map and various search filters. On the right, there is a 'Present Page' with a 'Class Report' table. The table lists various spectral classes and their corresponding pixel counts and percentages.

Class	Pixel Count	Percentage
Vegetation	6,000,000	100.00%
Bare soil / Built-up	1,000,000	16.67%
Water / Shadows	1,000,000	16.67%
Clouds	1,000,000	16.67%
Outliers	1,000,000	16.67%

ASQuLD (A)ATSR Implementation

- The system is composed by two node types:
- Each **Processing Node** is devoted to automated processing of ATSR-2/AATSR data to build classified maps and related metadata.
 - The **Storage Server** node is devoted to storage and coordination functions; in particular it provides:
 - > system storage space
 - > system main database
 - > images acquisition function from:
 - Magnetic and optical media
 - Rutherford Appleton Laboratory (RAL) FTP archive
 - European Space Agency (ESA) rolling archive
 - > access to result images and metadata
 - > processing coordination for processing nodes



(A)ATSR Data Coverage

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1995												
1996												
1997												
1998												
1999												
2000												
2001												
2002												
2003												
2004												
2005												
2006												
2007												
2008												
2009												
2010												

The system comprises two processing nodes (for parallel and fail safe processing) and one storage node.

References

- [1] Classification Application-services and Reference Datasets – CARD Project, <http://earth.esa.int/rd/Projects/CARD/>
- [2] MEEO S.r.l., SOIL MAPPER® System and Products Description, Issue 6.0, 01/09/2009, <http://www.meeo.it>
- [3] Baraldi, V., Puzolo, P., Blonda, L., Bruzzone, and C. Tarantino "Automatic Spectral Rule-Based Preliminary Mapping of Calibrated Landsat TM and ETM+ Images", IEEE Trans. Geosci. And Remote Sensing, vol. 44, no. 9, pp. 2563-2586, Sept. 2006.
- [4] SPA Project Web Site <http://earth.esa.int/rd/Projects/SPA/>
- [5] CARD Project Web Site <http://earth.esa.int/rd/Projects/CARD/>
- [6] SSE Web Site <http://services.eoportal.org/>

Conclusions

The proposed ASQuLD system provides a general and flexible architecture that permits to perform advanced (semantic) queries over large image databases. Users will avoid wasting time downloading or browsing previews of tens of images to find by trial and error the right images to be processed, thereby saving time by exploring the entire database in real time. Its general approach permits to apply such a system to existing and past missions (like Landsat that generated more than 500TB of data) as well as to future missions (like the sentinel constellation). Its modularity permits a complete customization of the system on customer's needs (e.g. data access interfaces different than SSE).