Knowledge Discovery Architecture Concept for the Payload Ground Segments

Espinoza-Molina, Daniela - *German Aerospace Center, IMF*
Mihai, Datcu - *German Aerospace Center, IMF*

In recent years the ability to store large quantities of Earth Observation (EO) satellite images has greatly surpassed the ability to access and meaningfully extract information from it. The state-of-the-art of operational systems for Remote Sensing data access (in particular for images) allows queries by geographical location, time of acquisition or type of sensor. Nevertheless, this information is often less relevant than the content of the scene (e.g. specific scattering properties, structures, objects, etc.). Moreover, the continuous increase in the size of the archives and in the variety and complexity of EO sensors require new methodologies and tools - based on a shared knowledge - for information mining and management, in support of emerging applications (e.g.: change detection, global monitoring, disaster and risk management, image time series, etc.). In addition, the current Payload Ground Segments (PGS) are mainly designed for Long Term Data Preservation (LTDP), in this article we propose an alternative solution for enhancing the access to the data content. Our solution presents a knowledge discovery architecture concept, whose intention is to implement a communication channel between the PGS (EO data sources) and the end-user who receives the content of the data sources coded in an understandable format associated with semantics and ready for the exploitation. This architecture concept encapsulates several techniques such as image content exploration based on signal processing analysis, knowledge discovery based on information modeling, and queries of the image archive based on data mining methods. Our new concept is developed in a modular system composed of the following components 1) the data model generation implementing methods for extracting relevant descriptors (low-level features) of the sources (EO images), analyzing their metadata in order to complement the information, and combining with vector data sources coming from Geographical Information Systems (GIS). 2) A database management system, where the database structure supports the knowledge management, feature computation, and visualization tools because of the modules for analysis, indexing, training and retrieval are resolved into the database. 3) Data mining and knowledge discovery tools allowing the end-user to perform advanced queries and to assign semantic annotations to the image content. The low-level features are complemented with semantic annotations giving meaning to the image information. The semantic description is based on semi-supervised learning methods for spatial-temporal and contextual pattern discovery. 4) Scene understanding counting on annotation tools for helping the user to create scenarios using EO images as for example change detection analysis, etc. 5) Visual data mining providing Human-Machine Interfaces for browsing the archive using 2D or 3D representation. The visualization techniques perform an interactive loop in order to optimize the visual interaction with huge volumes of data of heterogeneous nature and the end-user. This system produces information about EO product contents which is usually hidden in raster data (image), time series, metadata, and GIS objects. The system extends established EO data center access interfaces (catalogue search and ordering, data browse and download) by search functions at higher semantic level, allowing finding single products for specific applications based on image content interpretation. This concept extends the classical PGS functions by knowledge discovery and data mining functions.