PREPERATION OF SCENARIOS FOR THE PERFORMANCE OPTIMIZATION OF A CONTENT-BASED REMOTE SENSING IMAGE MINING SYSTEM

Gottfried Schwarz and Mihai Datcu

German Aerospace Center (DLR), Remote Sensing Technology Institute, Oberpfaffenhofen, Germany

ABSTRACT

Recent development in the design of modern satellite ground segments include systems and tools for automated content analysis allowing users to conduct systematic semantic searches within satellite image data archives. The need for such tools becomes more and more pressing as future space-borne imaging sensors will deliver enormous quantities of data that cannot be studied manually. For instance, typical examples from a European perspective are described in [1] and [2]. Within this framework, the European Space Agency (ESA) has started to fund the Earth Observation Librarian (EOLib) project to set up the next generation of image information mining systems [3]. Here we report on the preparation of scenarios that are needed for training and to verify and optimize the performance of such systems.

***Index Terms*—** Remote sensing, image mining, scenarios

1. INTRODUCTION

Image mining systems for satellite ground systems have to support routine queries by system operators as well as interactive user queries based on combinations of image examples, extracted image features, differences between features, image metadata, differences between metadata, semantic context annotations, label combinations, ontologies, and queries formulated in SQL, or via a query builder. For further details and the current state-of-the-art, see the references contained in [4]. This wide range of functionalities calls for the preparation of typical scenarios where a system developer can verify all processing steps and monitor the system performance. Running a scenario within a realistic environment can even result in improvements of the technical system layout. Therefore, the application range and the maturity of the final system will depend on the proper selection of scenarios during early project phases.

2. GENERAL SCENARIOS

Extensive user consultations, discussions with ground segment experts within [3], and tests with existing data and tools have led to a general scenario concept for our project that consists of four main classes of scenarios:

- Routine surveys
- Rapid mapping tasks
- Search for unexpected events
- Detection of gradual changes

Routine surveys are typical data analysis tasks being conducted by ground segment operators in order to explore the content of image archives, or to run regular searches for known targets. On the other hand, rapid mapping applications are typical for disaster monitoring or simple change detection where pre- and post-event images have to be compared for damage assessment, image content variations, or differences in semantic labeling. While disasters may range from flooding to a volcano eruption or a technical accident, a detected change could be a new construction site in a city, or farming activities in the countryside. Therefore, the features to be compared may be rather diverse. Similarly, when we work with extended image time series, we can search for unexpected and isolated events, or for gradual changes. Typical cases are environmental phenomena (e.g., vegetation effects); thus, analyses of time series data and efficient technical implementations have become an attractive field of research during the last years [5].

3. DETAILED SCENARIOS

In order to provide a more detailed description of concrete scenarios we refer to the TerraSAR-X instrument. Since the beginning of 2008 this SAR imager has delivered a large quantity of remote sensing data. The development of TerraSAR-X was a private public partnership project. As a consequence, there is an industrial enterprise (Astrium Geo-Information Services) that offers commercial data services to its customers (www.astrium-geo.com), and a public institution (the German Aerospace Center) providing data access to scientific users (https://terrasar-x.dlr.de).

We assume that typical user demands and application cases that evolved during these many years now provide a stable basis for the selection of representative scenarios. When we look at the services being offered to commercial customers, we find the main topics listed in Table 1 (taken from the Astrium Geo-Information Services website).

Delivery of standard image products: - High resolution spotlight images - Spotlight images - Stripmap images - ScanSAR images
Change detection maps
Topographic mapping
Land use / Land cover mapping
Digital surface models / Digital terrain models
Surface movement monitoring
Emergency response information
Reference data following military standards (STANAG)

Table 1: Commercially offered services

In contrast, when we look at the typical topics of the regular TerraSAR-X Science Team Meetings, one can extract more published details about specific application cases of the user community and we can compile the list of topics given in Tables 2a and 2b. The main topics and the detailed subjects listed below were taken from the program of the 5^{th} TSX Science Team Meeting of June 10-12, 2013.

Polar applications and glaciers - Geophysical parameters, validation - Mapping, elevation, ice sheets, mass budgets and land cover classification - Ice motion, dynamics, kinematics, subsidence and land cover dynamics - Ice loss, deformation, volume changes, calving - Landform dynamics
Forest applications - Backscatter - Biomass - Damages
Topographic applications - Mapping, pattern discovery - Digital surface and digital elevation models - Ground control - Façade structures
InSAR applications (earthquakes, volcanoes, infrastructure, urban areas) - Tomography and object reconstruction - Deformations, movements, subsidence - Persistent scatterers and reflectors - Geodynamics, landslides, slopes - Geohazard supersites - Source modeling - Anthropogenic events

Table 2a: Scientifically relevant topics (Part a)

Methods (e.g., for change detection) - Target analysis and temporal filtering - Distribution analysis, statistics, segmentation and vegetation parameters - Data fusion and merging - Operating modes and decompositions - Information mining, content understanding and ontologies - Big data - Calibration aspects and signal path delays - Virtual labs and education
Hydrology applications - Soil moisture - Soil texture - Seasonal changes - Water surface detection
Ocean applications - Maritime monitoring - Sea state and ice - Wind fields, wind direction - Mapping of species
Land cover - Methodological approaches incl. interferometry - Classification, signatures, textures, prospection - Mapping of settlements, habitats and forest - Crop and snow pack monitoring - Damage assessment - Occupancy of parking spaces

Table 2b: Scientifically relevant topics (Part b)

The application areas listed in Tables 1 and 2 have to be seen in conjunction with the actually commanded data takes of the instrument and its operating modes. Table 3 illustrates - as of end of May, 2013 - how many images contained in the TerraSAR-X level 1b image product catalogue of the German Aerospace Center (www.eoweb.dlr.de) have been recorded with the different operating modes of the TerraSAR-X instrument.

TerraSAR-X operating mode	*Number of scenes*
Scan SAR	10,354
Stripmap	53,165
High resolution spotlight	15,100
Spotlight	8,024

Table 3: TerraSAR-X scenes per operating mode

When we look into the geographical distribution of the entire catalogue of image takes then it becomes apparent that the large-scale stripmap mode is needed to prepare a worldwide digital elevation model, while the high resolution images are concentrated within specific areas of the globe. This means that stripmap image scenarios have to cover any type of surface feature combination, while high resolution images are needed for specific application fields and deserve extra care during the preparation of typical scenarios. This is also understandable from a user perspective: a user capitalizing on urban developments in India will have distinct goals that differ widely from polar research activities looking into ice age parameters in Antarctica. This also applies to analyses studying local vegetation characteristics and farming practices in different countries.

In the following, we present some examples of typical images pertaining to our scenarios of high resolution spotlight images (mostly HH polarization images).

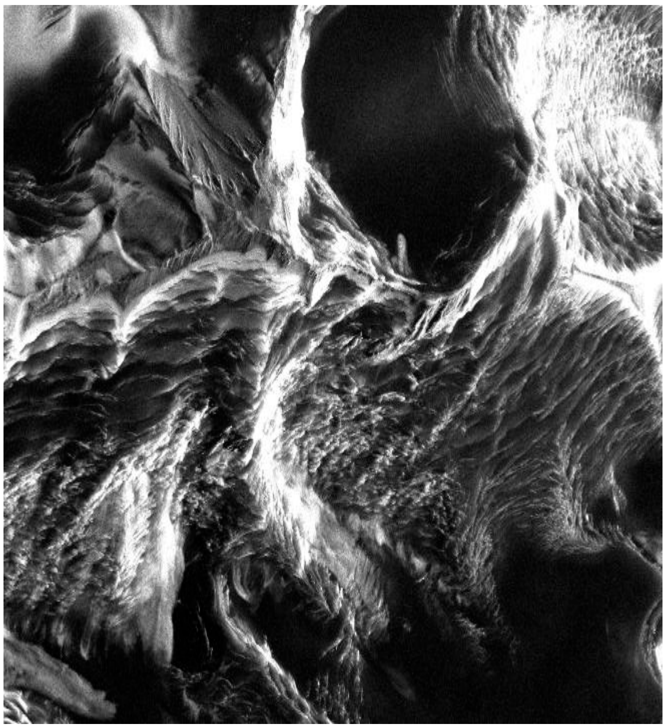

Fig. 1: Mountains in Antarctica: high mountains and slopes (Latitude -83.9° / longitude 169.0°; Jan. 29, 2013)

4. IMPACT ON SYSTEM COMPONENTS

Each scenario has a defined workflow, starting with the selection of images to be analyzed and the operating mode of the selected instrument, the data ingestion into the data mining system, the semantic annotation of objects appearing in the images, the inclusion of external information (e.g., cartographic information), and the provision of visual information for further knowledge discovery via visual data mining (e.g., a display of image content after dimensionality reduction). Thus, each scenario entails a sequence of processing steps that have to be tested and validated. For a number of processing steps - such as feature extraction - alternative algorithms exist; the most appropriate algorithm has to be selected based on the instrument imaging mode and the actual image recording parameters.

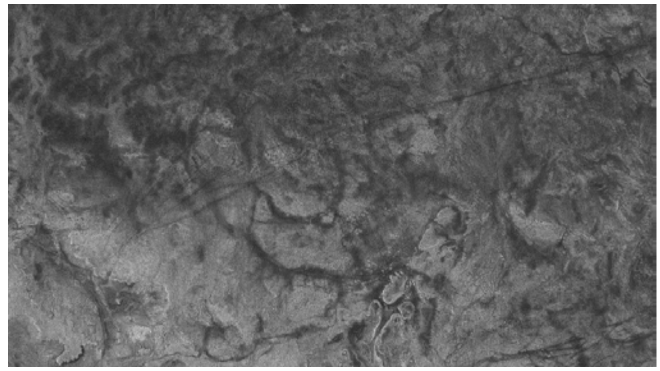

Fig. 2: Desert in Libya: backscatter from sand and rocks (Latitude 21.3° / longitude 23.0°, Nov. 20, 2007)

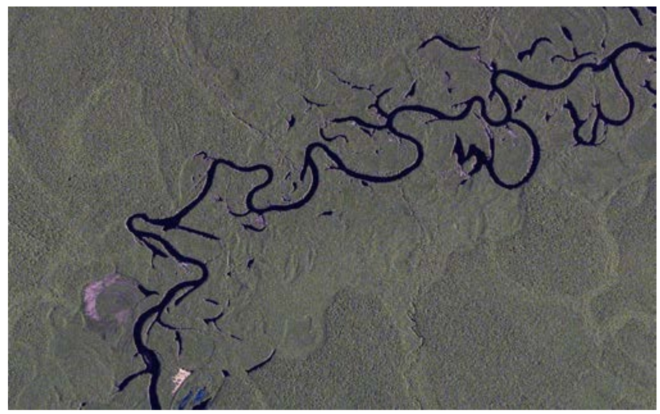

Fig. 3: Rainforest in Brazil: uniform SAR calibration target (Latitude -1.0° / longitude -64.0°; Jan. 01, 2010)

Fig. 4: Eastern China: settlements and agriculture (Latitude 34.05° / longitude 117.55; Mar. 31, 2013)

Fig. 5: Oil pier in Kuwait: industrial facilities
(Latitude 29.35°/ longitude 47.76°; May 09, 2013)

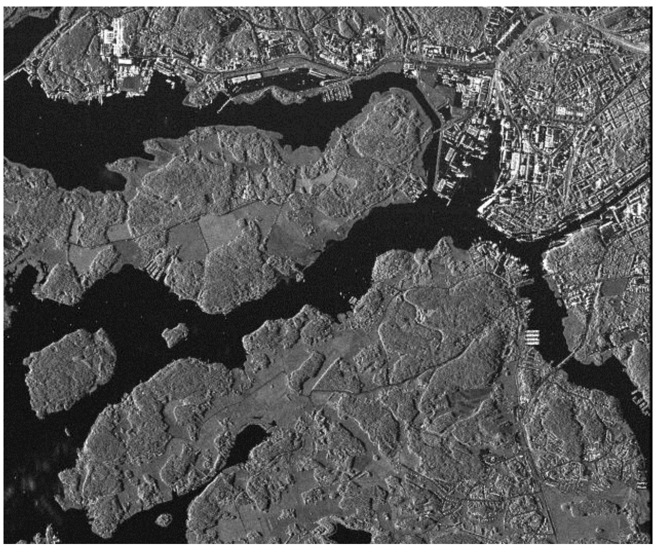

Fig. 6: Harbor of Turku, Finland: forest, water and city
(Latitude 60.42°, longitude 22.20°; Aug. 18, 2010)

5. EXPERIENCES GAINED

During test runs with images taken over all continents and many vegetation zones, it turned out that many land use parameters differ between countries and regions. Typical differences are the texture of built-up areas, the type and characteristics of vegetation, the size and shape of agricultural parcels, and the characteristics of the regional road network. As a consequence, the numerical features extracted from these country specific objects will differ, albeit their semantic labels may be the same. This had led to the conception of regional data collections with homogeneous parameters, and a semantic image retrieval that takes these local properties into account.

6. OUTLOOK

Currently, we expand our database with more and more images. Image mining tests will show us what kind of semantic annotation will be most useful for world-wide image content retrieval. First consolidated results about image content differences in different countries and regions of the world have already been presented in [6].

Fig. 7: Canberra, Australia: regular street network
(Latitude -35.28° / longitude 149.12°; Apr. 10, 2010)

7. REFERENCES

[1] R. Torres and more than 20 co-authors, "GMES Sentinel-1 mission", *Remote Sensing of Environment,* Vol. 120, pp. 9-24, 2012

[2] M. Drusch and 14 co-authors, "Sentinel-2 ESA's Optical High-Resolution Mission for GMES Operational Services", *Remote Sensing of Environment*, Vol. 120, pp. 25-36, 2012

[3] http://deepenandlearn.esa.int/tiki-index.php?page=EOLIB+Project

[4] M.N. Klaric, G.J. Scott, C.-R. Shyu, "Multi-Index Multi-Object Content-Based Retrieval", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 50, pp. 4036-4049, 2012

[5] R.S. Lunetta, C.D. Elvidge, "*Remote sensing change detection: environmental monitoring methods and applications*", Taylor & Francis, London, 1999

[6] C. Dumitru, "Earth Observation Image Semantics and Interoperability Aspects in Support of Content Retrieval and Data Mining", Big Data from Space, 2013, Frascati, Italy