Intelligent Geo-Search System: A semantic aware method for retrieving remote sensing images

Dimitris Sykas[1,3] , Alkyoni Baglatzi[2], George Vafiadis[3]

[1] Laboratory of Remote Sensing, National Technical University of Athens
H. Polytechniou Str. 9, 15780 Zografos Campus, Greece
dimsyk@gmail.com

[2] Laboratory of Cartography, National Technical University of Athens
H. Polytechniou Str. 9, 15780 Zografos Campus, Greece
baglatzi@mail.ntua.gr

[3] MapTronic Limited
2 Woodberry Grove, London N12 0DR, UK, dsykas@maptronic.co.uk, gvafiadis@maptronic.co.uk

Abstract: Technological advances in remote sensing increased the availability of satellite images with different spatiotemporal and spectral characteristics. The problem of obtaining data has been transformed into the problem of retrieving the most appropriate data for each user's needs. The biggest challenge is to bridge the gap between the low level semantics (detectable objects) of the images and the semantic information immanent in them with the view to designing intelligent geo-search systems. One key concern is the balance between the amount of knowledge that is formalized and the performance of the search system.

In the current work we designed a methodology for semantically aware image retrieval systems, targeted at expert users and certain type of satellite data. By knowing the constraints of the data, the expertise of the users (i.e. that they won't try to search for a car in a Landsat image [900 m^2 pixel area]), and the domain knowledge to be formalized complexity reduction and increase of performance and accuracy of results was achieved. The main objective was to build a data and domain specific search system. Search was based on tags and spectral indexes, which composed the metadata of each image within the system.

The proposed Intelligent Geo-Search System (IGSS) consists of 3 steps: (a) knowledge formalization (ontology design and tag formulation), (b) spectral library building, (c) remote sensing image metadata formation. In the first phase, knowledge about the geographic objects, processes and indexes is formalized as an ontology. In the second phase, a spectral library is built consisting of the signatures of the objects contained within the ontology. The third step is the tagging of the remote sensing images using an endmember extraction algorithm and a labelling algorithm. During this step indexes are calculated (NDVI, NDMI, BRI and FSI). Statistics for each index (minimum and maximum values, mean, standard deviation and distribution type) are stored along with the tags.

In that way, queries can be formulated that enable both object detection (e.g. burned areas and forest type identification) and phenomena/hazards quantification (e.g. increased risk for forest fire), making domain oriented questions answered more robust. For example IGSS can handle the complex queries: “Give me all the Landsat ETM images of Greece where 30% of the burned areas have a BRI value greater than 0.2” or “Give me images with a fir forest that had a burned area greater than 5 km^2 and the FSI value is less than 0.3.”

In this work, we propose a method for bridging the gap between raw data and the semantic information in them. As a use case we applied our methodology on the forest fire domain by using Landsat ETM images. An ontology was developed for these data and a sample of 30 images was used to test queries. The search result showed that the IGSS returned successfully the relevant images and ignored the ones which didn't match the search criteria.

Keywords: tagging, semantic search, endmembers, ontologies, image retrieval