VEGA-CONSTELLATION TOOLS TO ANALYZE HYPERSONSPECTRAL IMAGES

V. Savorskiy\textsuperscript{a,b}, E. Loupian\textsuperscript{a}, I. Balashov\textsuperscript{a}, A. Kashnitskii\textsuperscript{a}, A. Konstantinova\textsuperscript{a}, V. Tolpin\textsuperscript{b}, I. Uvarov\textsuperscript{a}, O. Kuznetsov\textsuperscript{a}, S. Maklakov\textsuperscript{b}, O. Panova\textsuperscript{b}, E. Savchenko\textsuperscript{b}

\textsuperscript{a} Space Research Institute RAS, 117997, 84/32 Profsoyuznaya Str, Moscow, Russia - evgeny@d902.iki.rssi.ru
\textsuperscript{b} Kotel’nikov Institute of Radioengineering and Electronics of Russian Academy of Sciences, Fryazino Department, Vvedenskogo sq., 1, Fryazino, Moscow region, Russia - savor@inbox.ru

KEY WORDS: Services, Remote, Infrastructure, Distributed, System, Internet/Web

ABSTRACT:
Creating high-performance means to manage massive hyperspectral data (HSD) arrays is an actual challenge when it is implemented to deal with disparate information resources. Aiming to solve this problem the present work develops tools to work with HSD in a distributed information infrastructure, i.e. primarily to use those tools in remote access mode. The main feature of presented approach is in the development of remotely accessed services, which allow users both to conduct search and retrieval procedures on HSD sets and to provide target users with tools to analyze and to process HSD in remote mode. These services were implemented within VEGA-Constellation family information systems that were extended by adding tools oriented to support the studies of certain classes of natural objects by exploring their HSD. Particular developed tools provide capabilities to conduct analysis of such objects as vegetation canopies (forest and agriculture), open soils, forest fires, and areas of thermal anomalies. Developed software tools were successfully tested on Hyperion data sets.

1. INTRODUCTION
Recently a remarkable progress has been observed towards greater open access to hyperspectral data (HSD). So, for example, NASA space data centers and NASA contractors increased access to data of hyperspectrometers Hyperion (EO-1, 2003) and HICO (HICO, 2014) mounted onboard satellites NASA EO-1 and International Space Station (ISS), respectively. In 2013, the Russian satellite “Resurs-P” with domestic hyperspectrometer HSA produced by S.A. Zverev Krasnogorskiy Zavod was launched (Arkhipov et al., 2014), whose data also became available for wide research community. This new situation calls for creating data provider services to manage HSD, including not only data ordering capabilities but also a set of analytic tools for remote analysis and processing HSD on data provider facilities.

One of the main obstacles to widening hyperspectrometer application is the fact that satellite hyperspectrometers cannot provide regular and sufficiently frequent coverage of Earth areas on global or even on regional scale. This is why currently hyperspectrometer applications are interesting mainly for scientific researchers. In this regard, the possibility of joint analysis of HSD with multispectral data (MSD) is of great interest. It is necessary primarily for a clear and deeper understanding of various objects and phenomena that we can now permanently monitor using a number of satellite Earth observing systems. Therefore, the introduction of capabilities to work simultaneously with hyperspectral and multispectral data into Earth observation (EO) information systems (IS), whose services and information resources are designed to solve scientific and applied tasks, is actual and relevant.

Assimilation of HSD in EO IS implies not merely adding HSD sets but also upgrading existing functionality to deal with unique HSD features. At least, such software (SW) tools as spectral analysis, HSD classification, and hyperspectral indices exploration should be added. One should keep in mind that these tools have to manage massive and rapidly growing HSD arrays in distributed infrastructure. A common approach to deal with the problem was elaborated in (Loupian et al., 2012). The main conclusion of this work is that efficient usage of EO data (fast data processing, reliable and comprehensive results, capabilities to analyze multiyear global data sets, etc.) can be achieved if we provide users not only with EO data sets and products but also with remotely accessible SW tools to manage and analyze these data sets via Web interfaces. This approach to managing HSD is implemented in the present work.

The main aim of the work was to develop tools to work with HSD in a distributed information infrastructure, i.e. primarily in the remote users access mode via Web interfaces. They should allow users to not only do search-and-retrieval operations on HSD but also provide capabilities to analyze and configure HSD processing procedures in remote mode.

Currently there is an abundance of EO IS (see e.g. references in (Loupian et al., 2012)) providing Web access to catalogues and archives of EO data. However, implementations of SW tools to work with EO data, i.e. to analyze and process them by user requests, in remote mode via common Web interfaces of an EO IS are scarce. Such type of IS was implemented in VEGA-Constellation IS family by research and development team of the Space Research Institute of the Russian Academy of Sciences (IKI RAS) (Loupian et al., 2011; Bartalev et al., 2012; Uvarov et al., 2014; Savorskiy et al., 2014; VEGA, 2015). Design and development of VEGA-Constellation information systems is based on original GEOSMIS technology also elaborated by the IKI RAS team.
2. GEOSMIS TECHNOLOGY ARCHITECTURE

GEOSMIS technology was designed by IKI RAS development team (Tolpin et al., 2011; Savorskiy et al., 2012). GEOSMIS architecture consists of two functional levels whose functional interactions are shown in Figure 1 (Tolpin et al., 2011):

- Presentation level includes Web and GIS interfaces;
- Application level includes:
  - Web services (map services, metadata services, and data management services);
  - System API, enabling access to data, metadata, and various resources and interfaces.

Presentation level is a subsystem that supports interaction with user. This level allows users to work with the system via Web interface or via specialized applications, e.g. GIS. As typical GIS is provided with all basic tools to work with map data, its interactions with GEOSMIS should run on data service level. This is the reason why GEOSMIS is oriented on the development of unified data access interfaces. Presentation level is shown in Figure 2 (Tolpin et al., 2011).

Work of such interfaces is enabled by a kernel consisting of two object modules:
- Map object module is implemented as smisMap;
- Metadata object module is implemented as smisMeta.

Map object module displays map information retrieved by data services. In addition, the module allows to send spatial queries and to export displayed data to print service or even to other systems. Metadata object module supports catalog data search and retrieval of data necessary for information displaying on rendering service.

Application level is a subsystem that should enable interaction between interfaces and data. This level consists of CGI services, core services, system API, and plugins as shown in Figure 3 (Tolpin et al., 2011; Savorskiy et al., 2012).

Figure 1. GEOSMIS architecture.

Figure 2. Presentation level architecture.

CGI services receive data via HTTP protocols. In order to provide Application level functionality, three types of CGI services are envisaged:
- Map service;
- Metadata service;
- Data service.

All three types of Web services are realized in frame of one common strategy i.e. are integrated in one common service. Those common services enable support of map data receiving (via GetMap request that is generated in interface by smisMap module), queries on metadata receiving (via GetMetadata request that is generated in interface by smisMeta module), requests on data addition and modification, e.g. addition of polygons and information about them. So, CGI interfaces are capable to provide direct interaction with metadata retrieving services and map information displaying. It is important that they can interact with any amount of spatially distributed services.

System API is a separated part of Application level. It is composed of a set of Perl modules that are located on servers and can get data and work with them on server side. All the other Application level services work with data via those modules. Each module takes responsibility for a definite data type or system part. This allows flexible system configuration, error fixing, and fast software development.

The above approach formed the basis for flexible and robust development of analytical SW tools running in remote access mode with distributed Earth observation data sets. This is proved by GEOSMIS technology implementation (Tolpin et al., 2011) in various monitoring systems operating in Russia, including:

- Integrated interface to work with the data of informational system of remote monitoring of the Federal Forestry Agency of Russia (Bartalev et al., 2010; Efremov et al., 2011);
- Web interface to work with data of industry-wide monitoring system of Federal Agency for Fisheries of Russia (Solodilov et al., 2011);
- Integrated data catalog of the Scientific Center for Earth Operative Monitoring of Russian Space Agency. The interface provides the capability to analyze data from various satellite systems (Bourtsev et al., 2011a);
- Joint access system of the European, Siberian and Far Eastern centers for receiving and processing satellite data “Planeta” of Rosgidromet (Bourtsev et al., 2009; Bourtsev et al., 2011b).
3. HSD TOOLS IMPLEMENTING TASKS

VEGA-Constellation is a long-term project of IKI RAS aiming at creating information systems or services based on a unified GEOSMIS technology. The different information systems within the VEGA-Constellation are described in detail in (Loupian et al., 2011; Bartalev et al., 2012; Uvarov et al., 2014; Savrasov et al., 2014). Its interface is shown in Figure 4.

In order to work in information systems of VEGA family with HSD, it was necessary to solve the following tasks:

- create a technology for obtaining HSD from various sources including remote ones;
- create a technology for HSD backup which should ensure rapid search, selection and display of HSD from long-term and operational archives;
- create basic interfaces that should provide capabilities to analyze HSD, including complex analysis of HSD and their products, together with the other information products stored in information systems of VEGA-Constellation family.

4. HIGH PERFORMANCE ALGORITHMS FOR HYPERSPECTRAL DATA MASSIVE MANAGEMENT

The data files from Hyperion USGS archive (Hyperion, 2014) were used as the main source of HSD during the work phase of design and implementation of basic HSD management and analytic software for usage in VEGA environment. For these purposes the IKI RAS team developed its original technology of data exchanges. This technology is described in detail in (Loupian et al., 2012a). It should be mentioned that using the same approach one can realize HSD uploading from another, different from USGS, data sources or providers.

The data were downloaded in the form of zip-compressed files containing 242 GeoTIff images, one per each spectral channel of Hyperion (in the band range from 355.59 to 2577.8 nm), and an ASCII text file with metadata in MTL format. Each file contains a 16-bit (int16) image in UTM projection. Each pixel represents radiation flux intensity.

After downloading, all data are automatically included into a 242-channel GeoTIff file. This procedure is completed by inclusion in the file of a 5-level (5-scale) pyramid. Scale pyramid is formed in order to provide scaling of data in order to significantly speed up access to data sets. This is particularly important for viewing many images in on-line remote mode. After testing, the best compression DEFLATE algorithm was chosen. It allows reducing the storage capacity by 10% compared to the original archive volume. Original spatial resolution, projection and values of the basic scale image remain unchanged.

So, we are able to display, analyze and process data in subsequent procedures without losing their original quality.

Along with the data file, the text annotation in a format suitable for mastering a specialized database is automatically generated. This specialized database is organized by basic software for management of satellite data archives, also developed by IKI RAS (Balashov et al., 2013). Depending on the scene angle, the special product identifiers specific to daytime and nighttime data are received. Such information is supplied to archives built using FDB (File Data Base) technology developed by ISR RAS (Efremov et al., 2004) as well. Hyperion data archive is organized in the same way as the majority of VEGA-Constellation archives, developed by IKI RAS. Namely, high spatial resolution data sets (Landsat, SPOT, Canopus-B, etc) are stored in the same way (Loupian et al., 2011; Bartalev et al., 2012; Uvarov et al., 2014). Currently IKI RAS HSD archives contain more than 11,000 scenes (mostly of the territory of Northern Eurasia). Their total volume exceeds 3.5 TB.

5. TECHNOLOGY OF HSD ANALYTIC TOOLS

Specialized capabilities that operate within the Web mapping interfaces based on GEOSMIS technology were created (Kashnitskii et al., 2015) in order to work with HSD. These capabilities allow integration of functions that can work with EO data in a variety of information systems based on GEOSMIS technology. Created capabilities allow developing analytic tools to deal with data of individual spectral channels, build color synthesis of arbitrarily selected channels, conduct joint analysis of HSD and data from other satellite sensors, to deal with hyperspectral indices, etc. (see Section 6.8 for descriptions of these capabilities).

The technology enabling HSD analysis in remote access mode should possess specialized SW components (aggregated in 3 subsystems as presented in Figure 5) in order to implement the following typical milestones for satellite data processing (Kashnitskii et al., 2015):

- prepare source data for processing, selection of required bands and spatial extent;
- run processing procedures;
- create a technology for obtaining HSD from various sources including remote ones;
- create a technology for HSD backup which should ensure rapid search, selection and display of HSD.

This contribution has been peer-reviewed.

doi:10.5194/isprsarchives-XLI-B4-235-2016

237
Web interface
Queue lookup and running the task
Preparing source data
Source data providing service
Task started
Processing
Task completed
Tasks database
Tasks queue
Web interface

Figure 6. Life cycle of processing request.

The IKI RAS team designs all components, which are created as part of GEOSMIS technology, as multipurpose ones for application in various information systems and remote sensing data centers. In accordance with the described principal schematic, a variety of data processing procedures is supported by common life cycle of processing request which is shown in Figure 6. Via Web mapping interface (running on IC) user can make preferable job and task settings. Since the performance of individual processing procedures can take a long time, the interface supports control of processing execution. User task assignment is recorded in ADB. After it, user is informed that the task is queued. Each processing stage with all its parameters is recorded in ADB. If user waits for the job completion, Web interface periodically interrogates ADB task manager TEC and informs user on the status of the job running.

6. VEGA DATA SELECTION POSSIBILITIES

Selection tools are of great importance for HSD handling due to the necessity to search and retrieve HSD sets associated with disparately located and rarely observed objects. Therefore, search and detection of target objects by standard means (Savorskiy et al., 2014), i.e. without any auxiliary information, is very difficult time-consuming procedure leading to inefficient use of system services. To solve this problem, many efforts were devoted to enhancing the selection capabilities of VEGA system with different kinds of auxiliary information available in VEGA databases. VEGA Web interfaces allow using MSD obtained in simultaneous observations, e.g. Landsat, SPOT, Canoptus-B data from IKI archives, and active wild fire maps from IKI databases as auxiliary information when selecting HSD sets (Savorskiy et al., 2014).

Along with studying mobile flares from wild forest fires, VEGA services added a capability to detect and explore static, or immobile flares that often can be attributed to anthropogenic activities. The procedure for selecting HSD subsets that describes thermal anomalies of anthropogenic origin is based on VEGA information product called Permanent Flares. These objects are registered in special VEGA DB. They are retrieved from multispectral or hyperspectral observations as fire flames. However, unlike wild fires, e.g. forest or prairie ones, permanent flares do not change their location within few days.

VEGA services allow visualizing the Permanent Flares objects and use this capability for the identification and localization of permanent flares depicted as tiny blue circles in Figure 7. Figure 8 shows an image of one of such permanent flare sites (© Google Planet Earth) registered as Permanent Flares object in VEGA DB. Analysis of the location indicates that this object is apparently a gas-flaring torch on the territory of a petroleum-producing enterprise.

Figure 7. VEGA visualizing and subselection services: “Permanent Flares” objects (Hyperion observations, 20.05.13, West Siberia). Target gas flare is located in red circle.1

Figure 9 shows a hyperspectral image of the Permanent Flares object presented in Figure 8. Spectral features confirm that there is a possible gas flare. Note that image in Figure 9 is presented in default visualization mode, however, even in this case one can detect particular hyperspectral features associated with fires (see Figure 10.b). Substantial improvement in recognition of targeted objects can be achieved by application of hyperspectral index technique (see Section 8.2 and Figures 14).
7. GENERAL-PURPOSE ANALYTIC TOOLS

In addition to tools implementing selection capabilities of VEGA system, IKI RAS team developed VEGA Web interfaces that incorporate a set of general-purpose analytic tools. They are designed for exploration of a variety of Earth objects in contrast to more specialized ones oriented for investigation of a definite type, or class, of natural objects (Section 8). These tools can also handle both HSD and MSD. Their functionalities were realized in the following remotely accessible user services (Savorskiy et al., 2014; Kashnitskii et al., 2015):

- Web mapping (cartographic) interfaces for user access to HSD resources combined with means of MSD access;
- space data visualization procedures which were enhanced to work both with HSD and MSD;
- data classification Web tools which were upgraded to work with HSD (Figure 10),
- HSD spectral analyzer.

VEGA capabilities were tested by applying the VEGA Web tools for classification of landscapes that contained vegetation cover and thermal anomalies. For example, results presented in Figure 10 show a determination of a Open Flame class on a typical Siberian forest landscape. One can easily select fire fronts (red ring in the center), both visually and by numeric analysis (as belonging to Open Flame class). One of the main conclusions of the test analysis is an independent confirmation of the necessity to develop dimensionality reduction algorithms in order to increase reliability of class differentiation due to decreasing a required volume of teaching data samples for reaching necessary representative statistics. This is one of the main objectives for introduction of hyperspectral indices (HSI) for using in HSD analysis (Section 8.2).

Spectral features of HSD can be analyzed on-line by VEGA Spectral Analyzer (VSA) instrument. VSA is a Web interface tool that was developed by the IKI RAS team for retrieving the parameters of spectral profiles of interactively selected points of hyperspectral image attributed to explored Earth objects. It is one of the principal analytic tools designed as Web services for HSD applications. As an example, Figure 11 shows the results of examining a forest fire. This is the same forest fire as was used in classification presented in Figure 10. So, these services are complementary to each other and available via one common Web interface.
Analysis of HSD thermal anomalies associated with open flame on a forest territory demonstrated their principal spectral features. They allow determining both thermal anomaly location and type. One of such features is a presence of spectral intensity maximum near 2.2 μm that is comparable with spectral intensities near 1.1 μm. In absence of flame on a forest plot, the spectral intensities near 2.2 μm is substantially lower than spectral intensities near 1.1 μm. This is illustrated in Figure 11, where one can see an increase of registered spectral intensities in spectral band over 2.2 μm in presence of open flame.

8. SPECIAL ANALYTIC TOOLS

In addition to general-purpose analytic tools, VEGA incorporates a set of special tools designed for investigation of particular classes of natural objects. Two of such tools are presented in Sections 8.1 and 8.2. They are spectral portraits system and hyperspectral index system (HSI) for investigation of vegetation canopy.

8.1 Spectral portrait system

The spectral portraits system accumulates knowledge about spectral properties of different surface types. A library of spectral portraits is being created to enable comparability of spectral images of individual plots of surface with reference standards to make decisions on their belonging to determined classes. The library is populated with reference spectral portraits, each produced by averaging a set of spectral profiles of the same definite surface type. It allows us to reduce the effect of uncertainties caused by spatial/temporal variability of surface parameters and provide representative statistics.

Spectral profiles obtained from HSD are distinguished by a high degree of details and appear in the shape of continuous curves. A large number of hyperspectral channels makes it difficult to use the entire data array for classification. Therefore, reference spectral portraits, which are of particular interest for research purposes, have to contain only the most informative parts of the spectrum for later use in classification tasks. Using reference spectral portraits implies the following user scenario:

1. Conduct visual detection of a group of characteristic points in a satellite hyperspectral image that can be assumed to belong to a particular surface type.
2. Start Web interface depicting spectral profiles related to the selected set of points of the hyperspectral image.

3. Conduct visual analysis of spectral profiles for the group of points in order to estimate the representativeness of the samples set and spectral separability of object types.
4. Assign the averaged profile of the group of characteristic points to be standard spectral portrait of particular surface type, name the spectral portrait, and save it in “Spectral portrait” DB.
5. In order to check class affiliation of any new surface, compare its spectral profile with previously defined standard spectral portraits.

Storage of spectral portraits is accomplished by MS MySQL DB. VEGA table structure allows to store unlimited number of portraits for every user if they have unique portrait names. Each portrait may include a series of spectral profiles that are based on various characteristics, such as spectral brightness or spectral reflectivity.

Implementation of the spectral portrait system is illustrated in Figure 12. It shows spectral portraits of some plant species obtained under the described technology (Figure 12 (a)). Figure 12 (b) presents the map of vegetation cover prepared using the spectral portrait data.

8.2 Hyperspectral index system

VEGA system also includes a toolkit called Image Algebra. It allows user to perform arithmetic, logic and various mathematical transformations on any data, including hyperspectral data. The values in the individual channels of satellite images can be converted by a formula given by user directly in the interface. The result is bitmaps created from existing raster layers using arithmetic and logical expressions involving integer and floating point numbers, normalization procedures, and mathematical functions.

With Image Algebra, VEGA cartographic Web interface is capable to calculate different spectral indices "on the fly" and to test experimentally new thematic products. For example, the user can build difference indices using absorption and emission spectral bands for the study of mineralogical and chemical composition of the underlying surface. It should be noted that such operations can be carried out on all data sets available in the archives of VEGA family systems. It means that users can combine not only HSD channels of one scene, but also any channels of multi-temporal scenes of any instrument stored in VEGA archives.
In addition to create, store and edit new hyperspectral indices, the VEGA system provides users with preset hyperspectral indices for such target areas as vegetated (forest and agriculture) plots, open soil surfaces, forest fires locations, areas of thermal anomalies. For each of them VEGA offers a specific set of indices for estimating parameters of the study objects with great accuracy and reliability. Experiments on the use of different indices for Hyperion data showed good results in the study of characteristics of these objects.

A general view of VEGA Web interface for work with hyperspectral indices is shown in Figure 13. It demonstrates a Transformed Chlorophyll Absorption Ratio Index (TCARI) (Haboudane et al., 2002) map of an agricultural area near the southern part of the Volga River. One can see a significant Figure 5. The principal scheme of GEOSMIS technology implementation and non-irrigated ones in late summer (dry season for this area).

Effectiveness of preset hyperspectral index applications is illustrated by VEGA presentation of an Aerosol Free Vegetation Index (AFVI1600) (Karnei et al., 2001) map of the vicinity of a gas flare (Figure 16). Notice drastic enhancement in visualization capabilities of the AFVI1600 map (Figure 14.b) in the presence of smoke in comparison to NDVI (Rouse J.W. et al., 1974) (Figure 14.a) or RGB composites (Figure 9). AFVI1600 mapping is also a very productive way to get information on forest fire environment required to produce reliable forecasts of fire dynamics, since, unlike NDVI map, this map provides information on wood stock under smoke layer (Figure 15).

CONCLUSIONS

The paper presents a variety of remotely accessed services that enable both search and retrieval procedures on HSD sets and their processing and analysis. These services were developed by the IKI RAS team within VEGA-Constellation information systems on the basis of GEOSMIS technological solutions. In particular, the developed tools enable spectral analysis of HSD sets in remote access mode. They also enable flexible applications of hyperspectral indices for exploration of natural resources. Efficiency of the proposed approach is confirmed and illustrated by Hyperion (EOS-1) data analysis results. The tools proved to be especially useful for HSD analysis in presence of clouds, aerosols or smoke.

ACKNOWLEDGEMENTS

The work is supported by RFBR grant №13-07-00513, 13-07-12116 ofi_m, 16-07-00956. Landsat and Hyperion data were received from the U.S. Geological Survey.
REFERENCES


