

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Remote Sensed Image Processing on Grids for Training in Earth Observation

Dana Petcu<sup>1</sup>, Daniela Zaharie<sup>1</sup>, Marian Neagul<sup>1</sup>, Silviu Panica<sup>1</sup>,  
Marc Frincu<sup>1</sup>, Dorian Gorgan<sup>2</sup>, Teodor Stefanut<sup>2</sup> and Victor Bacu<sup>2</sup>

<sup>1</sup>West University of Timisoara

<sup>2</sup>Technical University of Cluj Napoca  
Romania

## 1. Introduction

Remote sensing involves techniques that use sensors to detect and record signals emanating from target of interest not in direct contact with the sensors. Remote sensing systems integrate cameras, scanners, radiometers, radar and other devices, and deal with the collection, processing, and distribution of large amounts of data. They often require massive computing resources to generate the data of interest for their users.

Nowadays, remote sensing is mainly applied to satellite imagery. Satellites have proven in the last two decades their powerful capabilities to allow the Earth observation on a global scale. This observation is currently used in strategic planning and management of natural resources. The applications based on satellite data are often encountered in at least six disciplines: (1) agriculture, forestry and range resources in vegetation type, vigor and stress, biomass, soil conditions, or forest fire detection; (2) land use and mapping for classification, cartographic mapping, urban areas, or transportation networks; (3) geology for rock types, delineation, landforms, or regional structures detection; (4) water resources for water boundaries, surface, depth, volume, floods, snow areal, sediments, or irrigated fields detection; (5) oceanography and marine resources for marine organisms, turbidity patterns, or shoreline changes detection; (6) environment for surface mining, water pollution, air pollution, natural disasters, or defoliation detection.

Current applications involving satellite data needs huge computational power and storage capacities. Grid computing technologies that have evolved in the last decade promise to make feasible the creation of an environment, for these kinds of applications, which can to handle hundreds of distributed databases, heterogeneous computing resources, and simultaneous users. Grid-based experimental platforms were developed already at this century's beginning with a strong support from NASA and ESA.

In this context, the chapter presents for the beginners an overview of the technological challenges and user requirements in remote sensed image processing, as well as the solutions provided by the Grid-based platforms built in the last decade. Section 2 starts with a short description of the basic principles of the satellite imagery, the technical problems and state of the art in solving them. It points also the fact that the training activities in Earth observation are not following the intensity of the research activities and there is a clear gap between the request for specialists and the labor market offer.

Source: Image Processing, Book edited by: Yung-Sheng Chen,  
ISBN 978-953-307-026-1, pp. 572, December 2009, INTECH, Croatia, downloaded from SCIYO.COM

For initiated readers, Section 3 of this chapter presents a complex case study: the solutions provided by the recent developed platform, namely GiSHEO, in what concerns the image processing services, workflow-based service composition, and user interaction combined with e-learning facilities.

For experts, Section 4 presents the results obtained by applying the services offered by the GiSHEO platform in order to assist archaeologists in identifying marks corresponding to buried archaeological remains.

## **2. Problems in satellite imagery and current solutions**

The following section shortly presents the basic principles in satellite imagery and the existing solutions for fast response to its high request for resources of various kinds: computational, storage, or human.

### **2.1 Remote sensed image processing – basic principles**

Remote sensing data measures reflected or emitted radiation from surfaces in different parts of the electromagnetic spectrum like visible, ultraviolet, reflected infrared, thermal infrared, microwave and so on.

Multiband or multispectral data consist of sets of radiation data that individually cover intervals of continuous wavelengths within some finite parts of the electromagnetic spectrum. Each interval makes up a band or channel. The data are used to produce images of Earth's surface and atmosphere or to serve as inputs to complex analysis programs.

An image is produced by radiation from ground areas that are samples for a larger region. The radiation varies depending on the reflectance, absorption or emittance properties of the various ground objects. The sampling area varies from a square meter to a squared kilometer depending on the sensor position and accuracy. Each radiation measure is associated with a gray level tone when is displayed on a computer output device. Usually a sampled area corresponds to a pixel on a display.

The multiband data collected by one sensor have differences from one band to another. The constant band to band response for a given feature or class of materials is interpreted as its spectral signature (a plot of wavelengths versus an intensity function like reflectance). If three bands are each assigned to one of the primary colors, blue, green, and red, a color composite is obtained.

Nowadays, hyperspectral sensors are making accurate and precise measurements of individual materials using spectrometers operating from space. The resulting data set produces a detailed spectral signature.

The most simple image processing operations that are performed on satellite data, named transforms in (Mather, 2004), are those allowing the generation of a new image from two or more bands of a multispectral or multi-temporal image. It is expected that the final image has properties that make it more suited to a particular purpose than the original one or ones. For example, the numerical difference between two images collected by the same sensor on different days may provide information about changes that have occurred between the two dates, while the ratio of the near-infrared and red bands of a single-date image set is widely used as a vegetation index that correlates with the difficulty to measure variables such as vegetation vigor, biomass, and leaf area index (see details in (Mather, 2004)). We should note that change detection is the most common used for satellite imagery, being important, for example, in meteorological studies and disaster management.

The more complex image processing concepts and methods involved in satellite imagery are dealing with spectral transforms, including various vegetation indices, principal components and contrast enhancement, independent component analysis, vertex component analysis, convolution and Fourier filtering, multiresolution image pyramids and scale-space techniques such as wavelets, image spatial decomposition, image radiometric and geometric calibration, spatial decomposition, thematic classification using traditional statistical approaches, neural networks, or fuzzy classification methods, image modeling, two-dimensional time series modeling, image fusion for better classification or segmentation, or multi-image fusion.

Several recent books are aiming to present in details the digital image processing procedures and methodologies commonly used in remote sensing. Books like (Jong & Meer, 2005; Schowengerd, 2007; Chen, 2007; Chen, 2008) are covering most of the above described topics. Other books provide an introduction view to a level meaningful to the non-specialist digital image analyst, as (Richards & Jia, 2006) does, or to the level of graduate students in the physical or engineering sciences taking a first course in remote sensing, as (Schott, 2007). The book (Borengasser et al, 2008) describes case studies for the use of hyperspectral remote sensing in agriculture, forestry, environmental monitoring, and geology. Topics for agriculture, forestry, and environmental monitoring applications include detecting crop disease, analysing crop growth analysis, classifying water quality, mapping submerged aquatic vegetation, and estimating hardwood chlorophyll content. For geology applications, topics include detecting hydrocarbons and identifying and mapping hydrothermal alteration.

## 2.2 Grid computing for remote sensed image processing

Remote sensing is a major technological and scientific tool for monitoring planetary surfaces and atmospheres. Practical applications focusing on environmental and natural resource management need large input data sets and fast response.

To address the computational requirements introduced by time-critical applications, the research efforts have been directed towards the incorporation of high-performance computing (HPC) models in remote sensing missions. Satellite image geo-rectification and classification are the first candidates for parallel processing. The book (Plaza & Chang, 2008) serves as one of the first available references specifically focused on describing recent advances in the field of HPC applied to remote sensing problems.

Satellite image processing is not only computational-intensive, but also storage-intensive; therefore special technologies are required for both data storage and data processing. Hundreds of gigabytes of raw sensor data are generated per day and these data must be fast processed to produce the data required by the final users. Moreover, satellite image processing also involves different types of image processing techniques and algorithms. For each type of image processing an analysis is needed in order to point out several requirements as determining a suitable processing type, data movement issues and workflow management. Furthermore, satellite image processing applications require not only the processing of large volumes of data, but also various types of resources, and it is not reasonable to assume the availability of all resources on a single system. In this context, the Grid based technologies promised to make feasible the creation of a computational environment handling not only heterogeneous computing resources, but also hundreds of distributed databases, and simultaneous users.

There are at least three reasons for using Grid computing for satellite image processing: (a) the required computing performance is not available locally, the solution being the remote computing; (b) the required computing performance is not available in one location, the solution being cooperative computing; (c) the required computing services are only available in specialized centres, the solution being application specific computing.

An early paper (Lee et al., 1996) describes a metacomputing application that integrates specialized resources, high-speed networks, parallel computers, and virtual reality display technology to process satellite imagery; the inputs of the near-real-time cloud detection code are two-dimensional infrared and visible light images from satellite sensors.

Later on, realizing the potential of the Grid computing for the satellite imagery, several projects were launched at the beginning of this century to make the Grid usage idea a reality. Within the European DataGrid project an experiment aiming to demonstrate the use of Grid technology for remote sensing applications has been carried out; the results can be found for example in the paper (Nico et al., 2003). In the same period, (Aloisio & Cafaro, 2003) presented an overview of SARA Digital Puglia, a remote sensing environment that shows how Grid technologies and HPC can be efficiently used to build dynamic Earth observation systems for the management of space mission data and for their on-demand processing and delivering to final users.

Since 2005, the GEOGrid project (Sekiguchi et al., 2008) is primarily aiming at providing an e-Science infrastructure for worldwide Earth sciences community; it is designed to integrate the wide varieties of existing data sets including satellite imagery, geological data, and ground sensed data, virtually, again enabled by Grid technology, and is accessible as a set of services. Later on, D4Science (Tsangaris et al., 2009) studied the data management of satellite images on Grid infrastructures.

The testing phase has finished with the study delivered by the European DEGREE project (DEGREE consortium, 2008) about the challenges that the Earth Sciences are imposing on Grid infrastructure, as well as several case studies in which Grid are useful.

The second stage is the one of the production environments. (Cafaro et al, 2008) describes the standard design of a current Grid computing production environment devoted to remote sensing. For example, a special service was developed in the frame of the European BEinGRID project (Portela et al., 2008) to process data gathered from satellite sensors and to generate an multi-year global aerosol information; through the use of Grid technologies the service generates data in near real time. The platform called Grid Processing On Demand, shortly G-POD (Fusco et al., 2008), aims to offer a Grid-based platform for remote processing of the satellite images provided by European Space Agency (ESA) and it offers several satellite image processing services for environmental management. G-POD has proved its usefulness of the concept for real applications like flood area detection. The platform for satellite imagery search and retrieval, called Ground European Network for Earth Science Interoperations - Digital Repositories, shortly GENESI-DR (GENESI-DR consortium, 2008), offers to an end-user an interface for digital data discovery and retrieval; raw data are processed using G-POD facilities. The Landsat Grid Prototype LGP is using Grid computing to generate single, cloud and shadow scenes from the composite of multiple input scenes, the data for which may be physically distributed; the system ingests multiple satellite scenes, calibrates the intensities, applies cloud and shadow masks, calculates surface reflectance, registers the images with respect to their geographic location, and forms a single composite scene (Gasster et al., 2008). Ongoing EGEE-3 and SEE-Grid-SCI European



Grid-based e-infrastructures projects are currently building environmental applications based on satellite data including also some of the ones provided by GENESI-DR.

The Committee on Earth Observation Satellites (CEOS), an international coordinating body for spaceborne missions for the study of the Earth, maintains a working group on information systems and services with the responsibility to promote the development of interoperable systems for the management of Earth observation data internationally. The WGISS Grid Task team is coordinating efforts of ESA, NOAA and NASA projects.

A new trend is to make use of service-oriented architectures. A Service Grid reflects the recent evolution towards a Grid system architecture based on Web services concepts and technologies. The Service Grids' potential for remote sensing has already been pointed out at the beginning of this evolution, for example in (Fox et al., 2005). Wrappers are used to encapsulate proprietary image processing tools as services and furthermore allowing their usage in more complex applications. This is the road taken in the last years also by small scale research platforms like MedioGrid (Petcu et al, 2008) or Grid-HIS (Carvajal-Jimenez et al, 2004) trying to support national requests for remote sensing applications.

### **2.3 Training in Earth Observation**

The rapid evolution of the remote sensing technologies is not followed at the same developing rate by the training and high education in this field. Currently there is only a few number of resources involved in educational activities in Earth Observation. One of the most complex is EduSpace (ESA, 2007).

Recognizing the gap between research or production activities and the training or education ones, a partnership strategy for Earth Observation Education and Training was established in 1999 for an effective coordination and partnership mechanism among CEOS agencies and institutions offering education and training around the world. The key objective is to facilitate activities that substantially enhance international education and training in Earth observation techniques, data analysis, interpretation, use and application. In this context, the CEOS Working Group of Education, Training and Capacity Building is collecting an index of free Earth observation educational materials (CEOS, 2009).

### **3. Grid-based platform for remote sensed image processing – GisHEO**

We have developed recently a platform, namely GisHEO (On Demand Grid Services for Training and High Education in Earth Observation (GisHEO Consortium, 2008)) addressing the issue of specialized services for training in Earth observation. Special solutions were proposed for data management, image processing service deployment, workflow-based service composition, and user interaction. A particular attention is given to the basic services for image processing that are reusing free image processing tools, like GDAL or GIMP.

Our aim is to set up and develop a reliable resource for knowledge dissemination, high education and training in Earth observation. In order to answer to the on-demand high computing and high throughput requirements we are using the latest Grid technologies. A special features of the platform is the connection with the GENESI-DR catalog mentioned in the previous section.

Contrary to the existing platforms providing tutorials and training materials, GisHEO intends to be a living platform where experimentation and extensibility are the key words.

The platform design concepts were shortly presented in (Panica et al., 2009) and the details about the e-learning component can be found in (Gorgan et al., 2009). In this section we

shortly present the architecture and technologies that are used and then we are going in details related to the basic image processing services and interfaces.

3.1 Platform architecture

While the Grid is usually employed to respond to the researcher requirements to consume resources for computational-intensive or data-intensive tasks, we aim to use it for near-real time applications for short-time data-intensive tasks. The data sets that are used for each application are rather big (at least of several tens of GBs), and the tasks are specific for image processing (most of them very simple). In this particular case a scheme of instantiating a service where the data are located is required in order to obtain a response in near-real time. Grid services are a quite convenient solution in this case: a fabric service is available at the server of the platform that serves the user interface and this service instantiates the processing service where the pointed data reside.

Figure 1 presents the conceptual view of the implemented architecture.

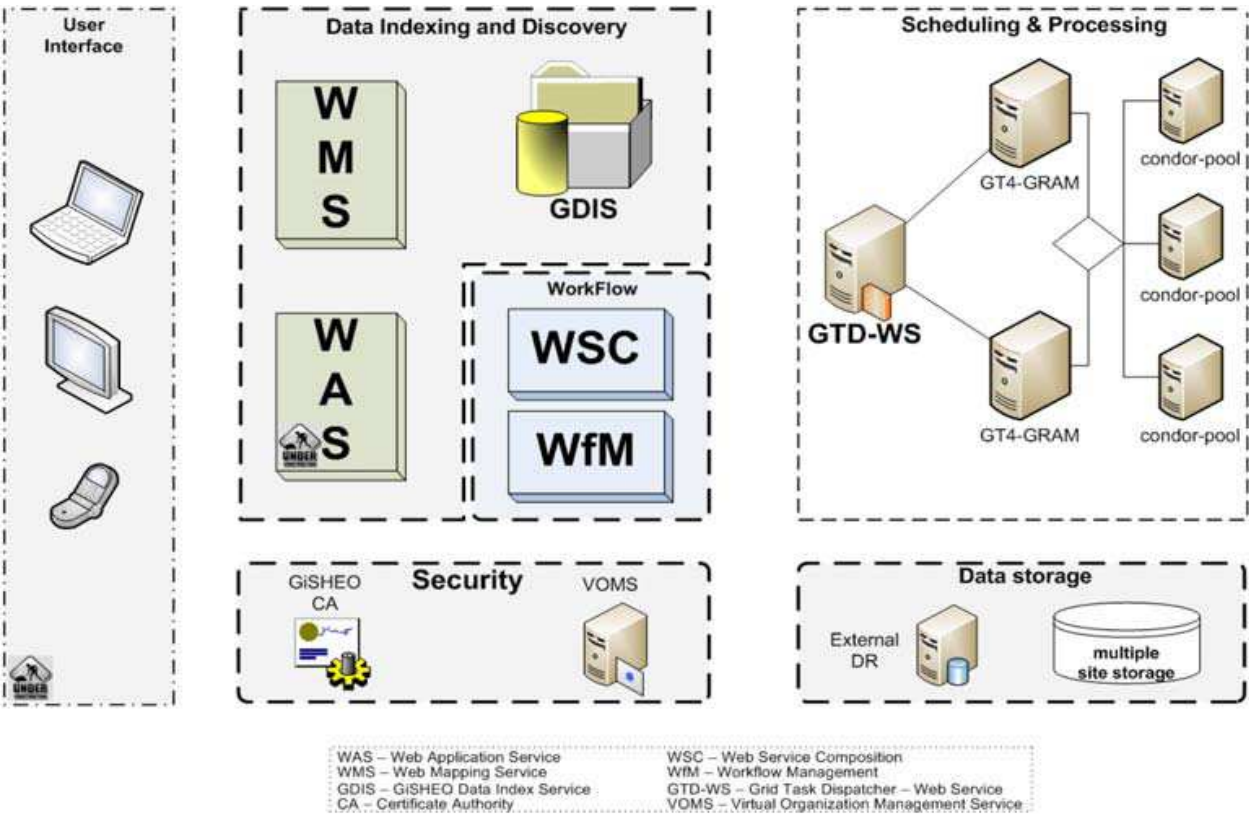


Fig. 1. GiSHEO architecture

The WMS is the standard Web Mapping Service ensuring the access to the distributed database and supporting WMS/TMS & VML. WAS is the acronym for Web Application Service that is invoked by user interface at run-time and allows workflows description. GDIS is a data index service – more precisely a Web service providing information about the available data to its clients. It intermediates access to data repositories, stores the processing results, ensures role based access control to the data, retrieves data from various information sources, queries external data sources and has a simple interface that is usable by various data consumers. The platform has distributed data repositories. It uses PostGIS for storing

raster extent information and in some cases vector data. Moreover the data search is based on PostGIS spatial operators.

The physical platform is based on four clusters that are geographically distributed at four academic institutions. Due to the low security restriction between the four institutions, data distribution between the clusters is done using Apache Hadoop Distributed File System. The data transfer from and to external databases is done using GridFTP – this is for example the case of the connection with GENESI-DR database.

The Workflow Service Composition and Workflow Manager are the engines behind WAS and are connected with the tasks manager. Each basic image processing operation is viewed as a task. Several tasks can be linked together to form a workflow in an order that is decided at client side (either the teacher, or the student interface). The workflow engine is based on an ECA (Event-Condition-Action) approach since it offers much more dynamism and adaptability to changes in workflow tasks and resource-states than other classic workflow engines. In order to respond to the special requirements of the platform a rule-based language has been developed.

The GTD-WS (Grid Task Dispatcher Web Service) is a service-enabled interface for easy interoperability with the Grid environment. EUGridPMA signed certificate are required to access the full facilities of the platform.

A particular component of WAS is eGLE, the eLearning environment. It uses templates to allow teachers specialized in Earth observation to develop new lessons that uses Earth observation data.

### 3.2 Basic services for image processing

We divide remote sensing processing operations into two types: basic and complex. Basic operations represent basic image processing algorithms that can be applied on a satellite image (histogram equalization, thresholding etc.). Complex operations are represented by the complex image processing algorithms (i.e. topographic effect regression) or by a composition of two or more basic operations. In Grid terms this operations must be exposed using some Grid-related technologies in order to interact with other grid components. Two related technologies can be used here: Grid services and Web services.

Web services (WS) are Internet application programming interfaces that can be accessed remotely (over a network) and executed on a remote system that hosts the requested services. Grid services can be seen as an extended version of Web services.

In our platform the Web services serve as an interface for the processing algorithms. This interfaces can be accessed remotely (normally using an user interface like a Web portal) and allowing the execution on a computational Grid of different types of processing techniques (distributed or parallel) depending on each algorithm in part.

In the framework of GiSHEO project we have developed a number of basic services that are useful in Earth observation e-learning process. In the following, some of them are presented together with visual examples.

The service for grayscale conversion (Figure 2) receives as input a coloured satellite image in any desired format (TIFF, PNG, BMP etc.) and transforms the triplets of values corresponding to each pixel in a value in the range 0-255. The service for histogram equalization (Figure 3) is used to increase the global contrast of an image; the histogram of the resulting image will be flat (pixels with the same value will not be separated into new values, however, so the histogram may not be perfectly flat). The service for quantization



(Figure 4) is used for reducing the number of colours used in a image; our implementation uses a multilevel quantization.



Fig. 2. Grayscale conversion: input parameter – an image file; output - the converted grayscale image file

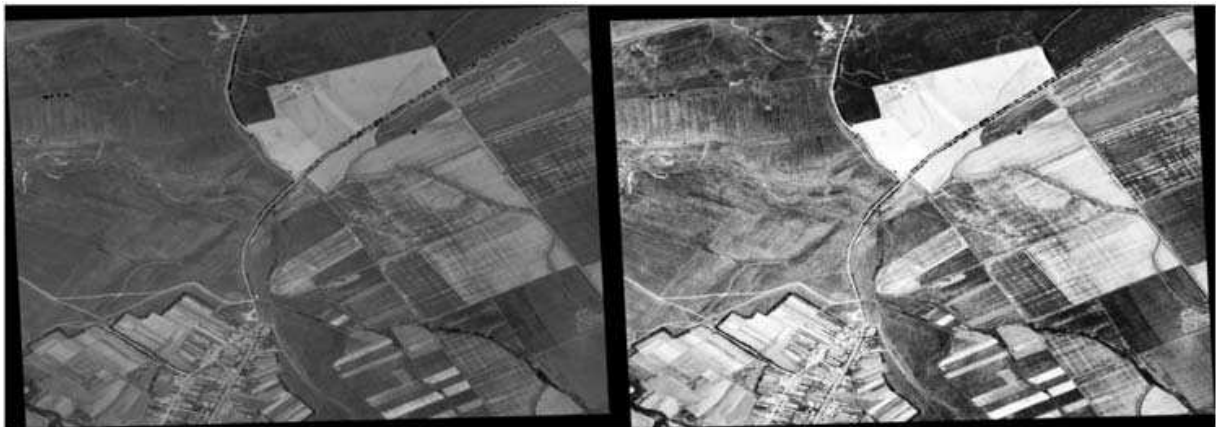


Fig. 3. Histogram equalization: input parameter – an image file; output – the transformed image



Fig. 4. Quantization: input parameters – image and number of gray levels; output – image

The service for thresholding (Figure 5) refers to a processing technique for image segmentation; in our implementation the user must chose a threshold (T) and this is used to compute the entire image. The service for blending images (Figure 6) is used for blending

two images; different types of blending are supported. The service for image embossing (Figure 7) implements an algorithm for embossing an image.

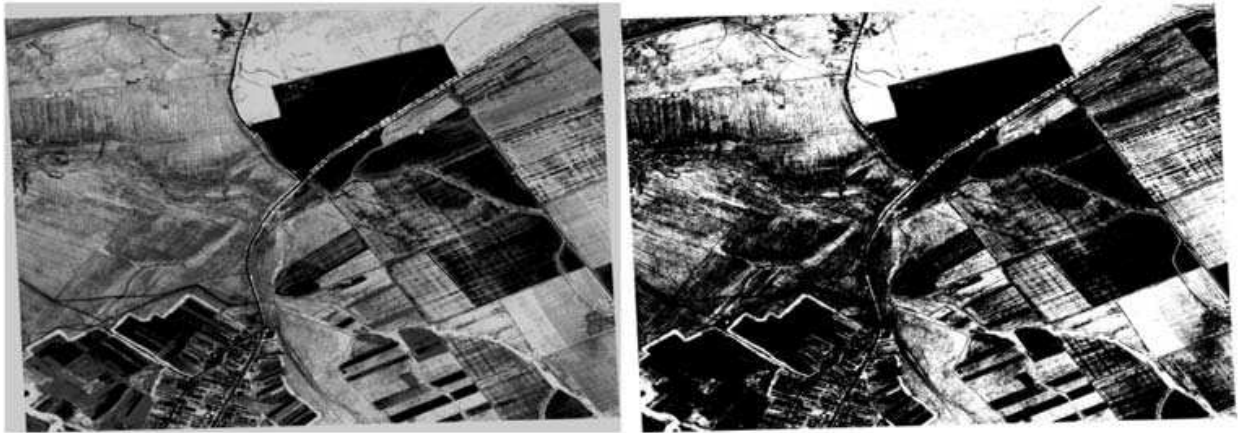


Fig. 5. Thresholding algorithm with  $T=125$  (input parameters - image and threshold  $T$ ; the image can be TIFF, PNG, BMP etc; output - image)

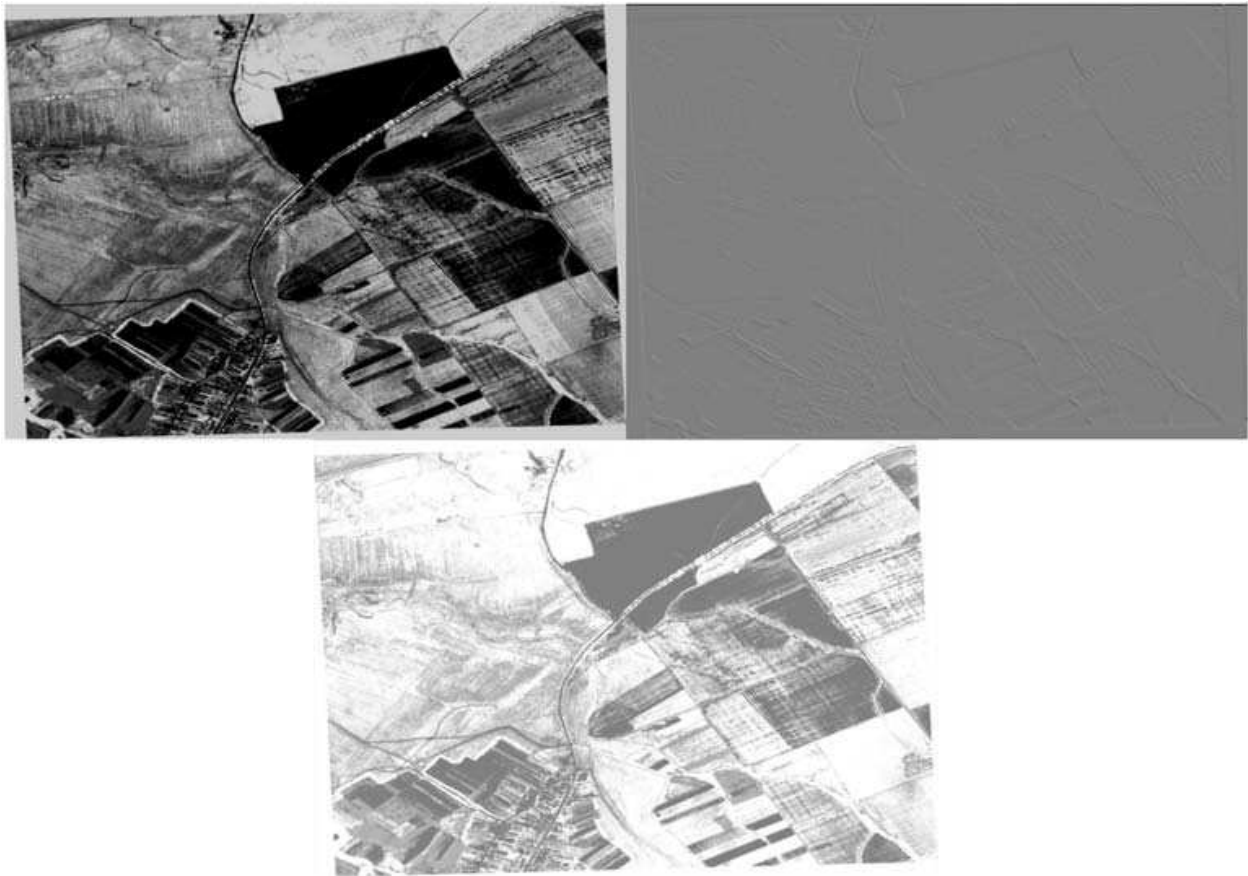


Fig. 6. Image blend: input parameters - layer | image1 | image2 | blend type (numerical value); output - blended image

The service for image transformation using a binary decision tree (Figure 8) is used for a quick image transformation using a binary decision tree to detect areas with water, clouds, forest, non-forest and scrub. The service for layers overlay is used for overlaying different layers; it has several images at inputs and produces one image. The service for vegetation

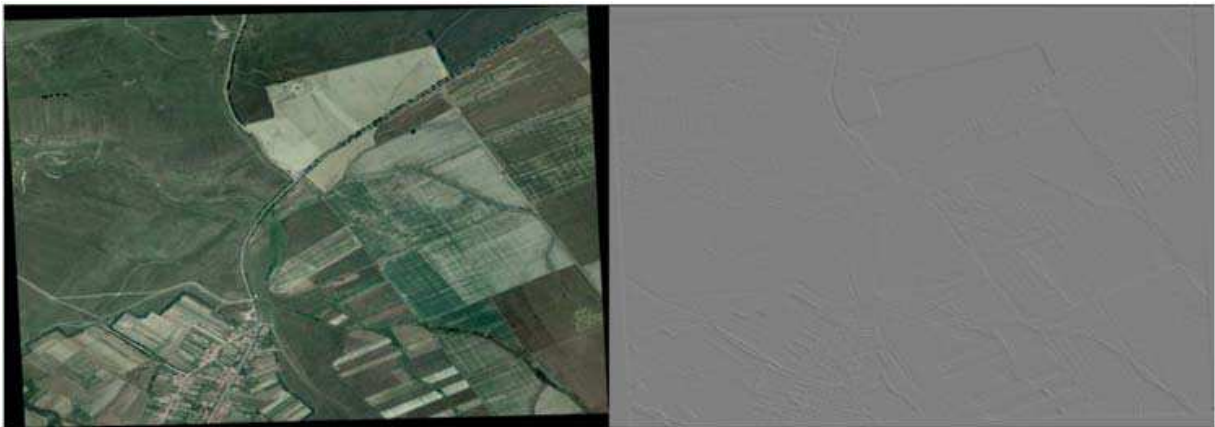


Fig. 7. Emboss service: input parameters – image | emboss level | grey level; output - image

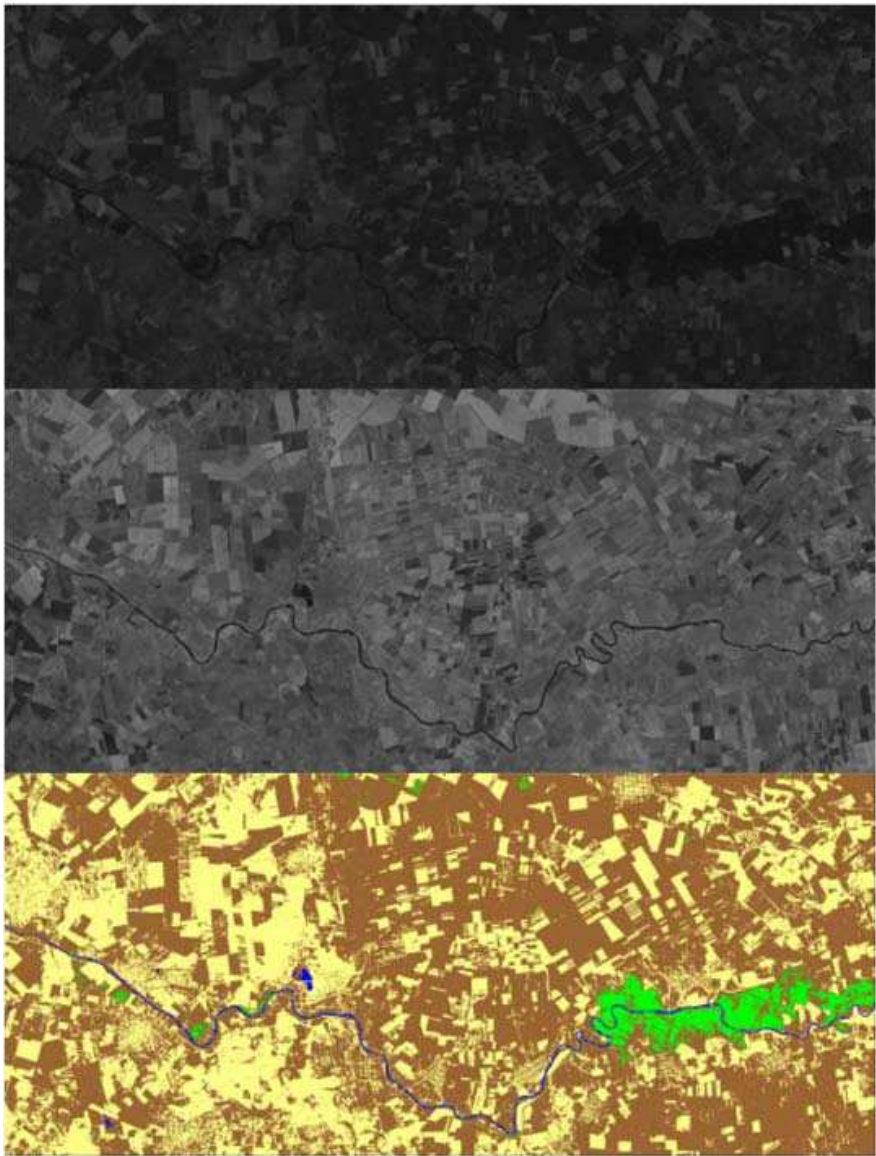


Fig. 8. Image transform: inputs - red-band image | infrared-band image | w | c | nf | f | s, where w-water value, c-clouds value, nf-nonforest value, f-forest value, s-scrub value; output – 5-color image resulted after applying a binary decision tree algorithm



index computation is used for computing the normalized difference vegetation index (NDVI) that shows whether an area contains a live green vegetation or not; it also supports the calculation of enhanced vegetation index (EVI); the input parameters are the red-band image and near infrared-band image (default MODIS-EVI values are used  $L=1$ ,  $C1 = 6$ ,  $C2 = 7.5$ , and  $G$  as gain factor = 2.5).

Basic services presented above can be used as single services or composed ones (see more about Web service composition in the next subsection). Figure 9 gives an example of a resulted image in a composed execution. The input image is a standard aerial image. The applied services are the following ones: gray-scale conversion, histogram equalization, quantization and thresholding.

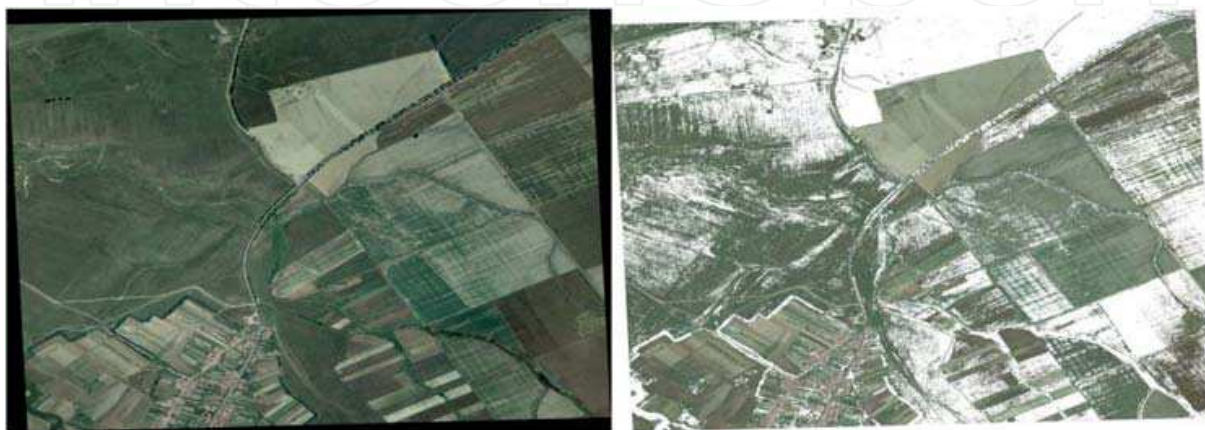


Fig. 9. Shape detection using a composition of several services (grayscale conversion, histogram equalization, quantization and thresholding)

### 3.3 Workflows and user interfaces

Processing large satellite data requires both a computational and storage effort. Usually operations on them are done in order to gain some insight on features which are not visible in the original image such as features visible at different bandwidths, changes over various time periods in ground features (floods, fires, deforestation, desertification, ice coverage), artificial or natural ground formations which are not easily distinguishable (ancient roads, fortifications, impact craters). These operations usually require two major steps before obtaining the final image(s). The first stage implies extracting relevant information from the satellite images such as geo-coding information, layers or bands. This information can be later used in the resulting image on the surface of the planet and in the actual processing step. This subsection will deal mostly with the latter step and with relevant examples.

Processing images in order to obtain relevant information usually requires several steps each of them consisting of a simple operation such as: obtaining the negative, gray level conversion, histogram equalization, quantization, thresholding, band extraction, embossing, equalization, layers subtraction etc. As it will be detailed later in this section, choosing the operations and their order depends on the desired result. This selection and ordering can be made either by the user or can be automatically generated starting from the desired output and following the chain of image combinations which lead to it. In this latter case it is up to the user to choose the solution which best fits its initial image characteristics (image type, codification, stored information). To support the user actions, we have developed a workflow language (Frincu et al, 2009) together with a set of tools for users not familiar with

programming which can be used both for visually creating a workflow (Figure 10) and for automatically generate a solution given a user defined goal. The application is then responsible for converting the visual workflow into a specific language developed by us which can then be executed. After defining the workflow the user can then select a region containing one or more images on which the workflow is to be applied. The communication is asynchronously and the user can directly observe how each of the images is gradually converted by the workflow.

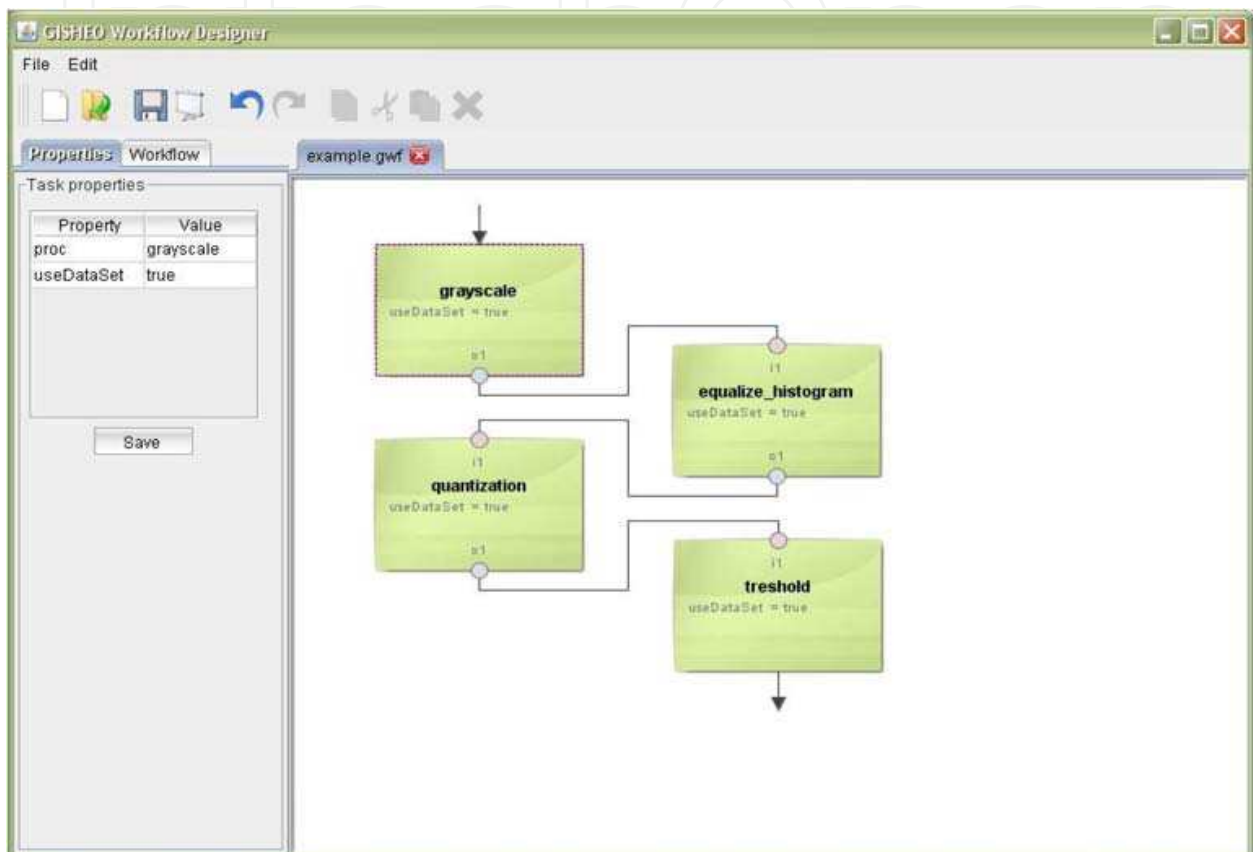


Fig. 10. Visually designing a GiSHEO workflow

In the frame of the GiSHEO project we are mostly interested in applying image transformations for historical and geographical use (see the case study from the next section). In this direction we have identified several workflows made up of basic image operations which allow the users to better observe features relevant for the two previously mentioned fields. For example in history and in particular in archaeology there is a constant need for identifying ancient sites of human settlements, fortresses or roads. Different flows of image transformations (e.g. gray level conversion, histogram equalization, quantization and thresholding) could enhance the image in order to make the marks induced by archaeological remains easier to be identified by visual inspection.

Identifying the required operations is followed by actually designing the workflow which usually implies binding them together in a certain order. Designing the workflow can be achieved in two ways either by using a visual tool such as the one presented in Figure 10, showing how a sequence of operations can be created, or by directly using the workflow language developed especially for this task which will be briefly presented in the following paragraph.



As previously mentioned visually designing the workflow is only the first step as it needs to be transformed into a language understood by the workflow engine (named in GisHEO OSyRIS - Orchestration System using a Rule based Inference Solution). The workflow language named SILK (Simple Language for workFlows) is rule based and can also be used directly by users preferring to define workflows without using the visual designer. A sequence of basic image processing operations can be defined using SILK as follows:

```
# Initial activation task
A0=[o1:output="imageId", "instances"="1"];
# The following tasks belong to the processing workflow
A=[i1:input, o1:output, "processing"="image grayscale(image)", "isFirst"="true"];
B=[i1:input, o1:output, "processing"="image equalize-histogram(image)"];
C=[i1:input, o1:output, "processing"="image quantization(image)"];
D=[i1:input, o1:output, "processing"="image threshold(image)", "isLast"="true"];
# Compute grayscale from the initial image
A0[a=o1] -> A[i1=a];
# Apply histogramequalization to the grayscale image
A[a=o1] -> B[i1=a];
# Apply quantization to the equalized image
B[a=o1] -> C[i1=a];
# Apply thresholding to the quantized image
C[a=o1] -> D[i1=a];
```

After transforming the visual workflow in the SILK language the workflow is executed using a workflow engine and the result is sent back to the user which is able to view it inside its web portal. In Figure 11 it can be seen how a selection of four images from a geocoded map is displayed after being processed. The sequence of operations corresponds to the sequence previously described and exemplified in Figure 10.



Fig. 11. Processed selection of four images inside the GisHEO web interface

After obtaining the result users can then either choose another image selection or change the workflow.

As it can be noticed from the previous paragraph the user interaction module is composed of several parts including the visual workflow designer which can be used independently for creating the workflows and the web portal which allows users to select a previously defined workflows, to choose a region comprised of several images and to apply the

workflow on them. The users could still use and navigate the map in search for other potential targeted images while the processing is still running. The visual designer and the engine API together with the automatic workflow generator are available on demand from the project repository.

The recent developed gProcess Platform (Radu et al, 2007), incorporated in GiSHEO platform, provides a flexible diagrammatical description solution for image processing workflows in the Earth Observation field. Abstract workflows (or Process Description Graphs (PDG)) and instantiated workflows (or Instantiated Process Description Graphs (IPDG)) are the workflows that can be defined using the gProcess Platform. Both workflows representation is based on DAGs. The PDG is a pattern definition of the workflow because it contains only the conceptual description of the workflows. On the other hand, the IPDG representation is linked to specific input data or resources. Based on this only the IPDGs can be executed in the Grid infrastructure. For describing workflows, we are using different types of nodes. The input data or data resources are used to specify satellite images or data values (int, float, string data types). The data resources are inputs for operators, sub-graphs and services. Operators are algorithms implemented to run over the Grid. The difference between operators and services is related to the way in which these nodes are executed in the Grid environment. Sub-graphs are used to integrate graphs that may have been developed by others and they can be used to define a hierarchical representation of the algorithm. We adopted an XML based representation as a persistent storage solution for the processing graphs. For every graph node, we defined an XML tag and a series of attributes that defines the nodes.

As a future development of the gProcess Platform, we intend to add the possibility to define in workflows some control structures like *for* or *if* statements. This will involve the ability of creating more complex image processing algorithm for execution over the Grid.

The gProcess architecture (Figure 12) is based on the client-server model. The server side enables the access to the Grid infrastructure using a set of services (EditorWS, ManagerWS, ExecutorWS and ViewerWS). The User Oriented Application Level and Application Level are encapsulated in the client side. The Client Java API accomplishes the access to the server side; this layer creates a transparent invocation level to the server side services. User Oriented Application Level exposes a set of user interaction components (EditorIC, ManagerIC, ViewerIC). The complex functionality is developed by using the Application Level, which combines the editor, manager and viewer functionality.

Each of the Web services exposed by the gProcess Platform is managing different functionality. The EditorWS provides information that is used to describe workflows, like the list of operators, the available sub-graphs or services, what satellite images can be used, etc. The EditorIC component that supports the user's editing operations for the workflow development uses this Web service. The interactive workflow design, the visualization of the workflow at different levels (by expanding or contracting sub-graphs), or user editing tools are developed using this interaction component.

Different information related to the existing workflows (PDGs or IPDGs) are exposed by the ManagerWS. Another functionality of this Web service is related to the interaction with the gProcess repository to integrate new operators, services, sub-graphs, to include new workflows or to access already designed workflows. The ManagerIC component can be used to instantiate workflows to different data resources (satellite images), to manage the model resources (operators, services, sub-graphs). The operator integration and monitoring user interface are implemented in Flex.

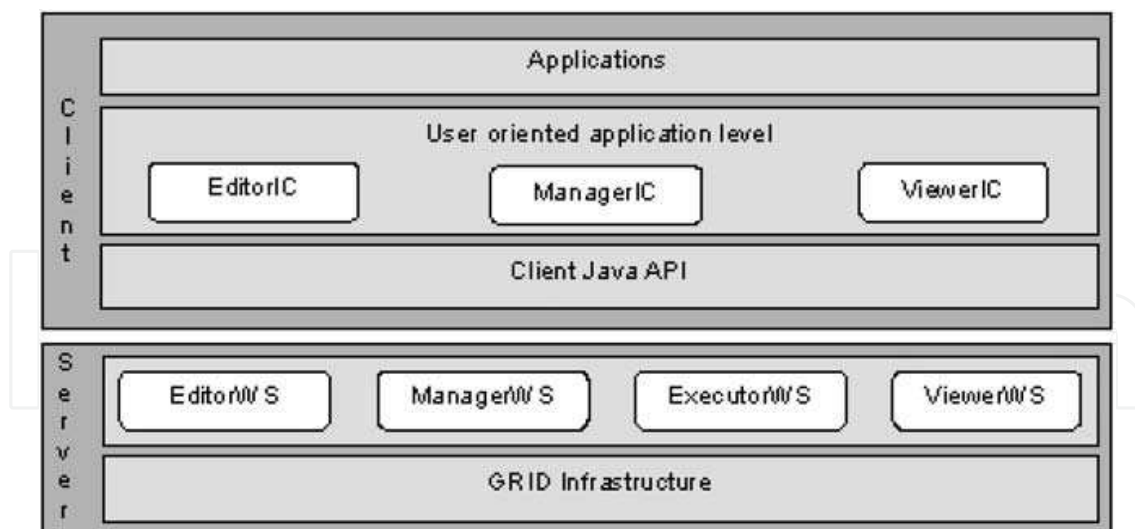


Fig. 12. gProcess platform based architecture

The instantiated workflows (IPDGs) can be executed over the Grid by using the ExecutorWS. Another important functionality of this Web service is the monitoring support for the executing workflows. An internal data structure that maps the workflow definition is created, and it is used to submit the operators to execution.

**GProcess Operators Integration.** The atomic component that implements a specific functionality is integrated in the gProcess platform like an operator. Currently a basic set of image processing operators are included in the gProcess operator's repository. The gProcess Platform supports the integration of user-developed operators. The operators must be written in Java and they have to extend a base class. The workflow description (the IPDG file) contains only the id from the repository. At execution time, the data from the database is retrieved and used to submit the operator to execution. An operator can have multiple input data resources. In the database, we store an invocation pattern that is used at the execution time. For example for the add operator, the pattern can be the following: *OperationExec [Add-?,?,?]*. The *OperationExec* represents the class that is executed. The *Add* represents the operator name. The last "?" specify the output and the other "?" characters are specifying the inputs. At execution time, this pattern is replaced with the execution command, for example: *OperationExec [Add-omania\_b2.tif,romania\_b3.tif,add\_result.tif]*.

In the graphical user interface (Figure 13), in order to integrate a new operator to the repository, the user must complete the following steps. The general information section must be completed with the operator name and also with a short description of the operator. In the input section, the user can select the input type for each data resource that is needed by that operator. In the same manner in the output section, the user can select the output type. In the upload tab, the user must select the executable class for the operator and also the dependencies that are required at execution time. After completing these steps, the user can define workflows by using the newly added operators.

**Workflow example - EVI algorithm.** The Enhanced Vegetation Index is used to enhance the vegetation signal by reducing the atmosphere influences. The input data for the algorithm are satellite images. Based on a set of processing steps the algorithm highlights the vegetation areas. The input resources are the NIR, Red and Blue spectral bands. The basic formula is:

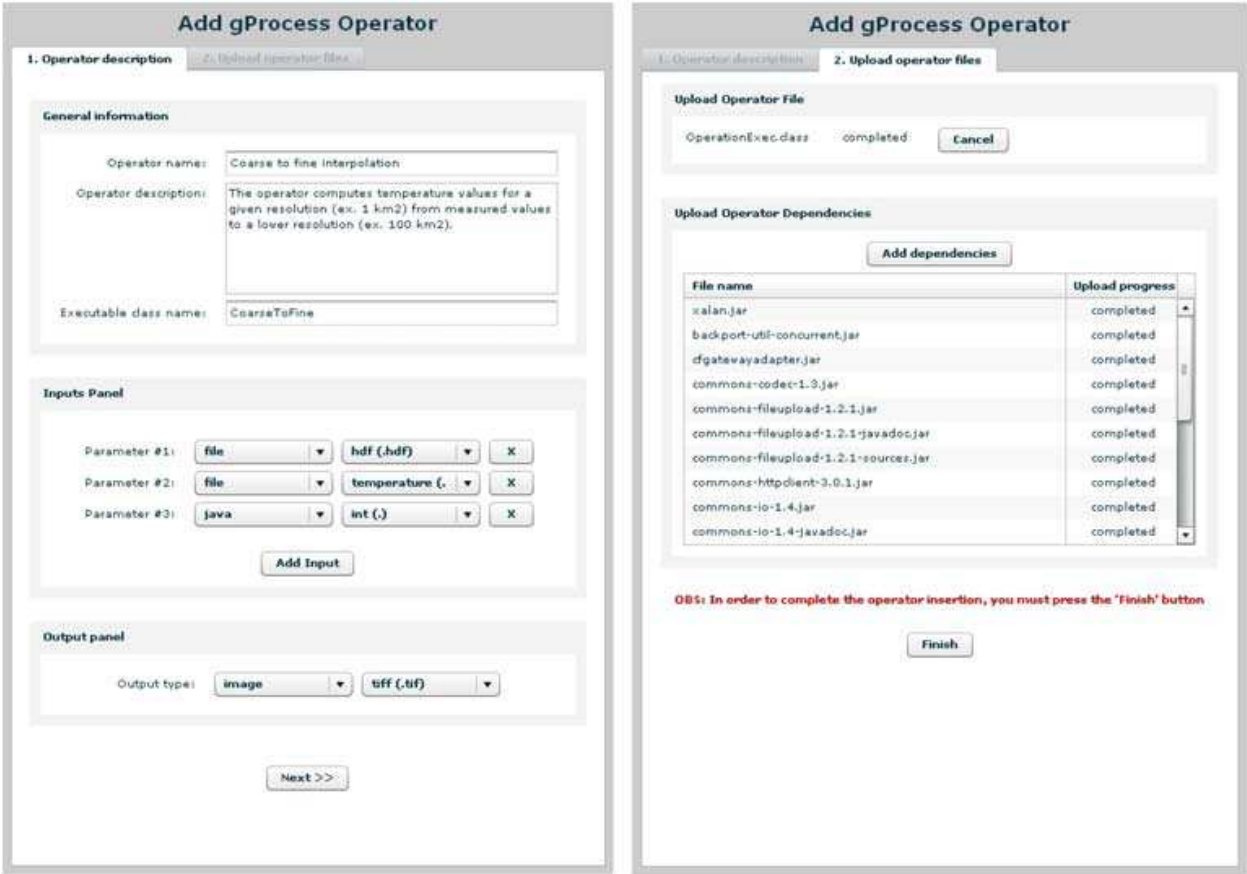


Fig. 13. gProcess interface for integration of new operators

$$\frac{2.5 * (NIR - RED)}{NIR + 6 * RED + 7.5 * BLUE + 1}$$

In order to design the workflow for this algorithm we have to identify the atomic components, and to rewrite the formula using the available operators. Since we have only binary operators for addition, subtraction, etc. the new formula is the following:

$$\frac{2.5 * (NIR - RED)}{[NIR + (6 * RED)] + [(7.5 * BLUE) + 1]}$$

In Figure 14 we exemplify the EVI graphical representation of the workflow. A sample from the XML definition of the workflow is the following:

```
<?xml version="1.0" encoding="UTF-8"?>
<Workflow>
<Nodes>
  <Resource id="1" name="B4" description="NIR spectral band">
    <LocalResource path="romania_B4.tif" />
    <PostConditions>
      <Output idTypeDB="3" />
    </PostConditions>
  </Resource>
  ...

```

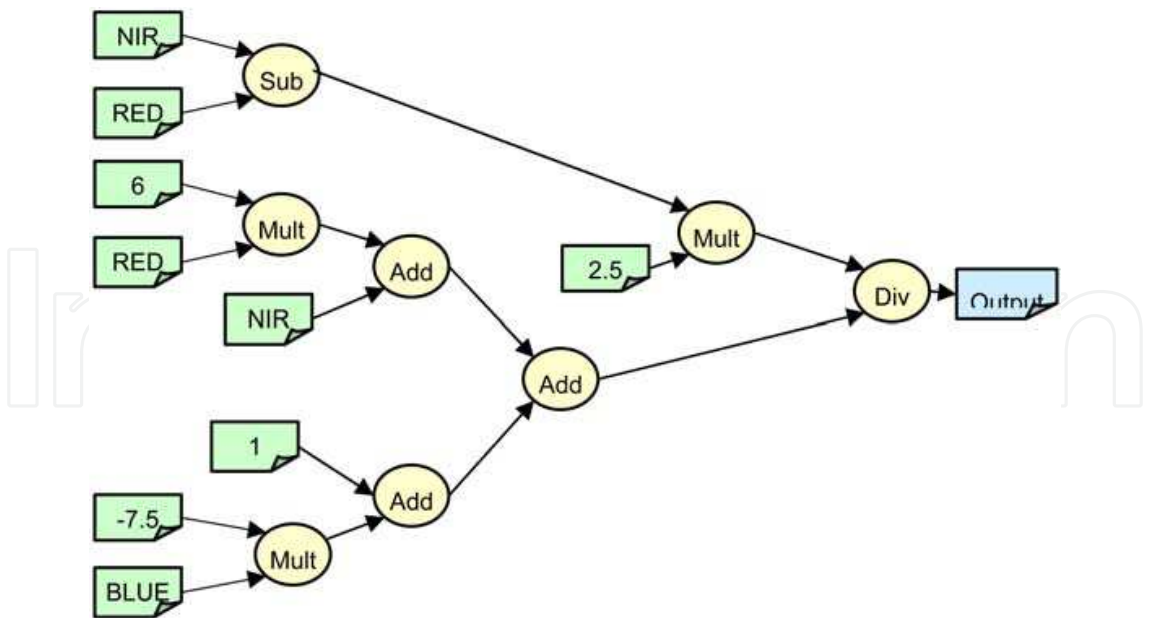


Fig. 14. EVI workflow

```
<Operator id="3" name="Sub" description="" idDB="2">
  <Preconditions>
    <Input id="1"/>
    <Input id="2"/>
  </Preconditions>
</Operator>
...
<Operator id="15" name="Div" description="" idDB="4">
  <Preconditions>
    <Input id="3"/>
    <Input id="14"/>
  </Preconditions>
</Operator>
</Nodes>
<Groups>
</Groups>
</Workflow>
```

The processing result is presented in Figure 15 for a particular example.

**Monitoring interface.** The ExecutorWS is the component that, based on a user selected IPDG, submits the jobs to execution. The monitoring component from this Web service updates the database with the current execution status for each node from the workflow description. The monitoring interface (Figure 16) displays this monitoring information, node name, start time, end time, execution status (waiting, submitted, running or completed). If the job is completed then the result can be visualized (if the result is an image), or it can be downloaded.

3.4 E-learning components

The aim of the eGLE application is to provide the non-technical specialists in Earth observation with an environment that will allow them to search and retrieve information





average computer users, but the functionalities implemented must allow the launching of complex Grid operations with minimum restrictions.

There are three main steps that a teacher must complete in order to create a Grid based lesson and they are described in what follows.

**Step 1. Acquire the information needed for the lesson.** In a transparent manner, without having any knowledge over the location of the data or the protocol needed to access it (HTTP, FTP, GFTP etc.), the teacher is able to browse and search for information based on keywords, time intervals or latitude-longitude defined areas. This modality of information search and retrieval will be available at first only for public, non-secured repositories, as the secured access problem is a more complex issue that requires some specific user knowledge and actions. Another type of information that the teacher can include into the lesson are the results of his own computations executed over the Grid. Through the platform visual tools included in the eGLE interface, the teacher can describe his own PDGs, create iPDGs, and launch them in execution, monitor the execution progress and access the results without possessing any technical information related to Grid.

**Step 2. Organize and display the lesson content.** Once the information needed for the lesson is acquired, the teacher should be able to setup the lesson structure, to organize logically the information and to define the desired display settings (e.g. text size and color). As the amount of data included into the lesson can be very large (satellite images, videos, files with measured values etc.) or may be accessible only at runtime (the custom computations launched by students) the offline lesson development using dedicated desktop applications is not an option. The eGLE Platform provides the teacher with all the functionalities needed to create the visual appearance of the lesson through the usage of visual containers like *tools*, *patterns* and *templates*.

*Tools* are visual and functional elements specialized on a certain content type (image, video, text, graph etc.) and represent the atomic parts (smallest division) of the lesson. They are developed by the programmers and integrated into the eGLE platform in order to be used by teachers through the environment interface. The tools are concerned with data retrieval and display mechanisms and provide only an API that can be used to customize their settings (specify the data to be accessed and displayed – image, video etc., modify their visual appearance – width, height, text color, text size etc) according with the content they are specialized on.

*Patterns* represent visual containers and logical information organizers. They can be created directly by the teachers at authoring time (Fig. 17, Select Pattern) through a wizard like interface that will allow them to customize their visual appearance (ex. number of columns). On each column of a pattern can be integrated a different tool that will be chosen on the second step of the wizard (Fig. 17, Select Tools). Once a pattern is created, it can be reused by the teacher (or by other teachers) with the same visual settings but with different tools included.

*Templates* are visual containers and patterns collections that define the general layout and settings at lesson global level. The visual attributes defined at template level will provide unitary visual formatting throughout the entire lesson, and can be overridden at pattern or tool level if necessary.

**Step 3. Data binding and user interaction description.** After creating a pattern and selecting the desired tools to be integrated into the lesson, the teacher has the opportunity to specify the instantiated data that will be displayed in each tool by moving to the step 3 of the wizard (Fig. 17, Edit content). At this point each tool will display visual user interface





#### 4. Case study: remote sensing processing services for research in archaeology provided by GiSHEO

Remote sensing techniques proved to be useful in non-intrusive investigation of archaeological sites by providing information on buried archaeological remains (Gallo et al., 2009; Lasaponara & Masini, 2007). The presence of different remains in the ground can generate different marks identifiable in high resolution panchromatic and/or multispectral images: crop marks, soil marks, shadow marks and damp marks (Gallo et al, 2009). The crop marks are determined by the particularities of vegetation above different kind of remains. The crop above wall foundations is negatively influenced because of lack of water in soil while the crop above buried pits or ditches is positively influenced because of soil nutrients. The soil marks consist of changes of the soil colour or texture. Both crop and soil marks can be identified from panchromatic images and different spectral bands (for instance crop marks are easier to be identified by using red and near infrared bands). The main steps in extracting marks related to archaeological remains are described in the following.

**Step 1. Data preparation.** This step can consist in rectifying the distortions in the image and/or in fusing data corresponding to panchromatic of different spectral bands (Lasaponara & Masini, 2007). Particularly useful in identifying crop marks are the vegetation indices and other thermal parameters which can be computed from different spectral bands (e.g. red and near infra red bands).

**Step 2. Applying image enhancement techniques.** In order to emphasize the marks of buried remains, the images are processed by applying basic techniques. Since a good contrast is very important in making different marks easily to identify by the human eye, the contrast enhancing techniques (contrast stretching or histogram equalization) are frequently used in processing archaeological images (Aqduş et al., 2008). Other techniques frequently used to detect crop and soil marks is edge detection, edge thresholding and edge thinning (Lasaponara & Masini, 2007). In the case of multispectral images an important processing step is that of computing the principal components which help in identifying changes in surface variability.

**Step 3. Extracting knowledge from processed images.** Automatic identification of archaeological sites from digital images is a difficult task, since the small anomalies induced by the buried remains are usually hidden by stronger marks corresponding to the structures currently existing on the ground (roads, constructions, trees, rocks etc). Therefore the final identification and interpretation of the marks should be made by the expert by visually inspecting the enhanced image and by corroborating his observations with additional information (e.g. historical maps, current roads network etc).

The case study we conducted aimed in experimenting with different workflows of enhancement operations applied to high-resolution panchromatic images. The images correspond to several sites in Timis county, Romania where on ground research proved the existence of different remains, e.g. clusters of pits, tombs, roman fortifications etc. Different sequences of operations have been applied on a set of images selected by an archaeologist who also interpreted and validated the results.

The results obtained for three images are presented in Figs. 18, 19 and 20, respectively. In all cases we were looking for soil marks. The results in Fig. 18 were obtained by first converting the initial image in gray scale and then applying contrast stretching (Fig. 18a) or histogram equalization (Fig. 18b). For the image presented in Fig. 19a besides the contrast enhancing

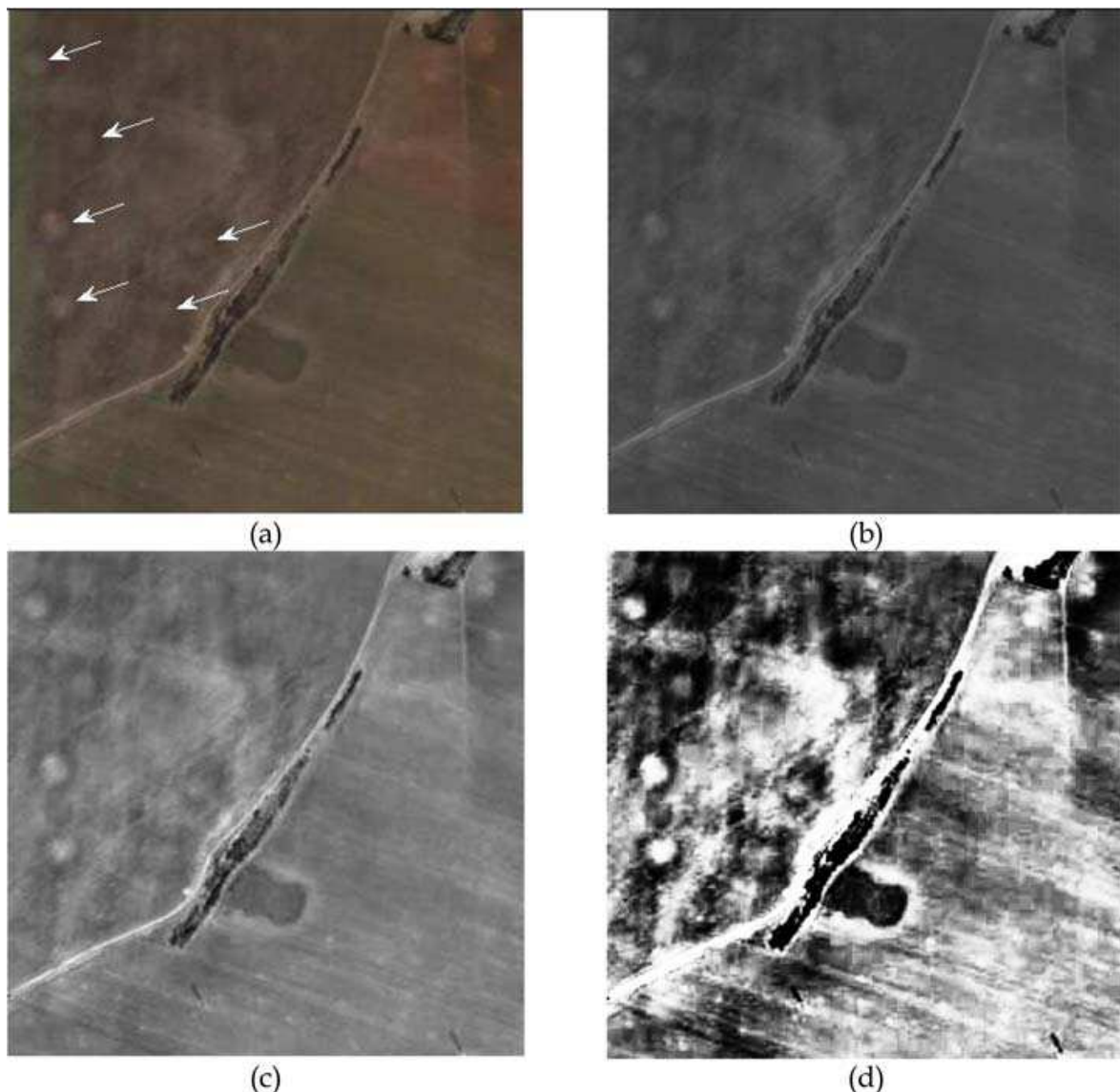


Fig. 18. Soil marks enhancement (pits cluster and ancient square delimitation) (a) Original image with position of pits marked by arrows (b) Gray level image (c) Gray level image enhanced by contrast stretching (c) Gray level image enhanced by histogram equalization.

operations (Figs. 19b and 19c) a sequence of several operations has been also applied (Fig 19d). The workflow used in this case consisted in a sequence of six operations: gray scale conversion, histogram equalization, edge detection (Sobel filter), thresholding, inversion and erosion.

In the case of the image presented in Fig. 20a besides contrast enhancement by histogram equalization (Fig. 20b) we also applied an emboss filter to the gray scale image (Fig. 15c) followed by histogram equalization (Fig. 20d).

These examples just illustrate the possibility of enhancing panchromatic images by applying flows of basic image processing operations. Thus tools allowing the construction of workflows of simple operations or just the selection of predefined workflows could be useful in training of students in landscape archaeology. Further work will address to construct the support for extracting archaeological marks from multispectral images.



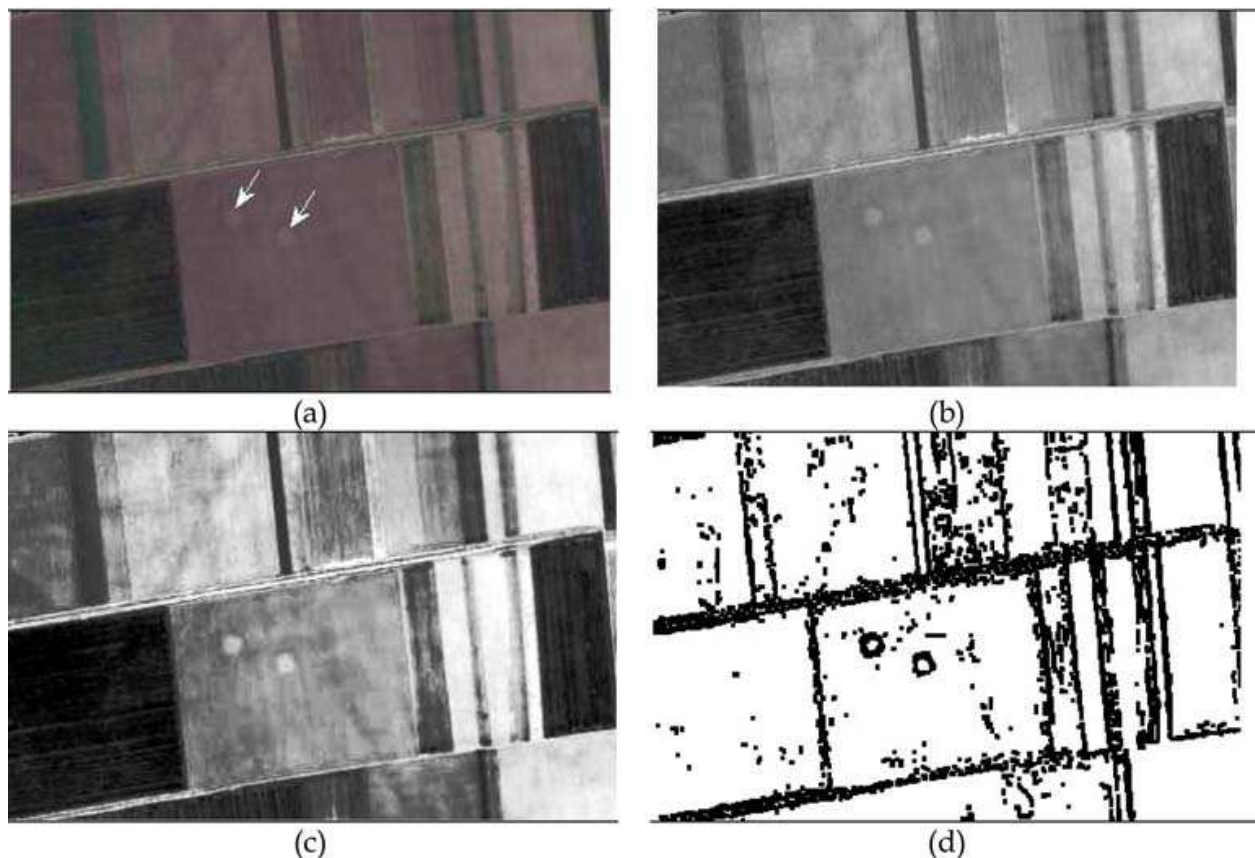


Fig. 19. Soil marks enhancement (burial mounds) (a) Original image with positions of burial mounds marked by arrows; (b) Gray scale image enhanced by contrast stretching; (c) Gray scale image enhanced by histogram equalization; (d) Result of a sequence of operations on the gray scale image: histogram equalization, edge detection, thresholding, inversion, erosion.

## 7. Conclusion

In order to respond to the need of training and high education platforms for Earth Observation, a Grid-based platform for satellite imagery has been recently developed and its services are exposed in this paper. The development is far from being complete. Complex services are envisioned to be constructed in the near future and intensive tests and comparisons with other approaches are planned to be performed in the next year.

## 8. References

- Agdus, S.,A.; Drummond, J. & Hanson, W.S. (2008). Discovering archaeological crop marks: a hyperspectral approach. *The Intern. Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, XXXVII (B5), Beijing(2008), pp. 361-366,ISSN1682-1750
- Aloisio, G. & Cafaro, M. (2003). A dynamic Earth observation system. *Parallel Computing* 29-10, pp. 1357-1362, ISSN 0167-8191
- Borengasser, M.; Hungate, W.S. & Watkins R. (2008). *Hyperspectral Remote Sensing: Principles and Applications*, CRC Press Taylor & Francis, ISBN 978-1-56670-654-4, Boca Raton

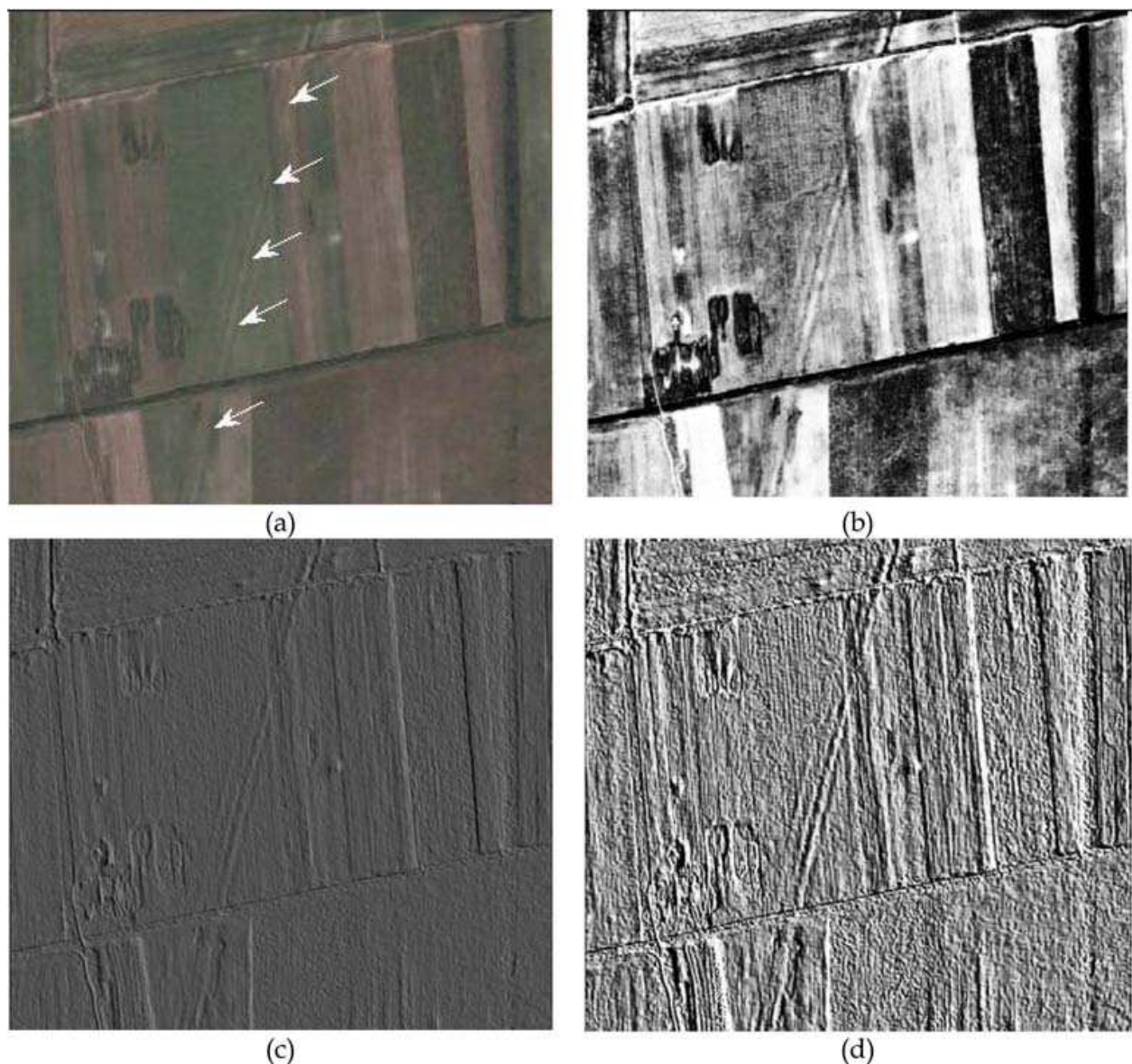


Fig. 20. Soil marks enhancement (roman fortifications) (a) Original image with linear fortification marked by arrows; (b) Gray scale image enhanced by histogram equalization; (c) Embossed gray scale image; (d) Embossed gray scale image enhanced by histogram equalization.

Cafaro, M.; Epicoco, I.; Quarta, G.; Fiore, S. & Aloisio, G. (2008). Design and implementation of a Grid computing environment for remote sensing. In: *High Performance Computing in Remote Sensing*, 281-305, Chapman&Hall/CRC, ISBN 978-1-58488-662-4

Carvajal-Jimenez, C.L.; Lugo-Beauchamp, W. & Rivera, W. (2004). Grid-HSI: using Grid computing to enable hyper-spectral imaging analysis. In: *Procs.3rd Intern. Conf. Communications, Internet&Information Technology*, Hamza. M.H. (Ed.), pp. 583-588, ISBN 0-88986-445-4, St. Thomas, USA, Nov. 2004, IASTED/ACTA Press, Calgary

CEOS (2009). Working group on education, training and capacity building - WGEdU Overview. Available at <http://oislabs.eumetsat.org/CEOS/webapps/>

- Chen C.H. (Ed.) (2007). *Signal and Image Processing for Remote Sensing*. CRC Press, Taylor & Francis Group, ISBN 978-0-8493-5091-7, Boca Raton
- Chen C.H. (Ed.) (2008). *Image Processing for Remote Sensing*. CRC Press, Taylor & Francis Group, ISBN 978-1-4200-6664-7, Boca Raton
- DEGREE consortium (2008). Earth Science White Paper on Grids. Available at [http://www.eu-degree.eu/DEGREE/internal-section/wp6/DEGREE-D6.1.2\\_v2.8.pdf](http://www.eu-degree.eu/DEGREE/internal-section/wp6/DEGREE-D6.1.2_v2.8.pdf)
- ESA (2007). EDuspace The European Earth observation Web site for secondary schools.. Learning with Earth Observation. Europe from Space. Disaster Monitoring. Global Change. Available at <http://www.eduspace.esa.int/>
- Frincu, M.E.; Panica, S.; Neagul, M. & Petcu, D. (2009). Gisheo: On demand Grid service based platform for EO data processing, *Procs.HiperGrid'09*, ISSN: 2066-4451, Romania, May 2009, Volume 2, pp. 415–422. Politehnica Press, Bucharest.
- Fox, G.; Pallickara, S.; Aydin G. & Pierce M. (2005). Messaging in Web service Grid with applications to geographical information systems. In: *Grid Computing: A New Frontier of High Performance Computing*. Grandinetti L. (Ed.), pp. 305–331, ISBN 978-0-444-51999-3, Elsevier, Amsterdam.
- Fusco, L.; Cossu, R. & Retscher, C. (2008). Open Grid services for Envisat and Earth observation applications. In: *High Performance Computing in Remote Sensing*, Plaza A. & Chang C.(Eds.),237–280,Chapman&Hall/CRC,ISBN978-1-58488-662-4,Boca Raton
- Gallo, D.; Ciminale, M.; Becker H. & Masini N. (2009). Remote sensing techniques for reconstructing a vast Neolithic settlement in Southern Italy. *Journal of Archaeological Science* 36, Issue 1 (January 2009), pp. 43-50, ISSN 0305-4403
- Gasster, S.D.; Lee, C. A. & Palko, J.W. (2008). Remote sensing Grids: Architecture and implementation. In: *High Performance Computing in Remote Sensing*, Plaza A. & Chang C.(Eds.),203–236,Chapman&Hall/CRC,ISBN978-1-58488-662-4,Boca Raton
- GENESI-DR Consortium (2008). Ground European network for Earth Science interoperations – Digital repositories. Available at <http://genesi-dr.eu>
- GiSHEO Consortium (2009). On-demand Grid services for high education and training in Earth observation. Available at <http://gisheo.info.uvt.ro>.
- Gorgan D.; Stefanut T. & Bacu V. (2009). Grid based training environment for Earth observation. In: *Advances in Grid and Pervasive Computing*, LNCS 5529, Abdennadher, N. & Petcu, D. (Eds.), pp. 98-109, ISBN 978-3-642-01670-7, Springer, Berlin.
- Jong, Steven M. de; Meer, Freek D. van der (Eds.) (2005). *Remote sensing image analysis:: including the spatial domain*. Series: Remote Sensing and Digital Image Processing, Vol. 5, Kluwer Academic Publishers, ISBN 1-4020-2559-9, Dordrecht
- Lasaponara, R. & Masini, N. (2007). Detection of archaeological crop marks by using satellite Quickbird multispectral imagery. *Journal of Archaeological Science*, Vol. 34, Issue 2 (February 2007), pp. 214-221, ISSN 0305-4403
- Lee, C.A.; Kesselman, C. & Schwab, S. (1996). Near-realtime satellite image processing: metacomputing in CC++. *IEEE Computer Graphics & Applications*, Vol. 16, Issue 4 (July 1996), pp. 79-84, ISSN 0272-1716
- Mather, P. M. (2004). *Computer Processing of Remotely-Sensed Images: an Introduction*. John Wiley & Sons Ltd, ISBN 978-0470849194, Chichester, UK
- Masini, N.; Rizzo, E.; Lasaponara, R. & Orefici, G. (2008). Integrated remote sensing techniques for the detection of buried archaeological adobe structures: preliminary



- results in Cahuachi (Peru). *Advances in Geoscience*, 19 (2008), pp. 75-82, ISSN 1680-7340
- Nico, G.; Fusco, L. & Linford, J. (2003). Grid technology for the storage and processing of remote sensing data: description of an application. *Procs. SPIE: Fuhisada H. et al. (Eds.)* vol. 4881, no. 677 (April 2003), pp. 677-685, Agi Pelagia, Greece, Sept. 2002.
- Panica, S.; Neagul, M.; Petcu, D.; Stefanut, T. & Gorgan, D. (2009). Designing a Grid-based training platform for Earth observation. In: *Procs. SYNASC 08*, pp. 394-397, ISBN: 978-0-7695-3523-4, Timisoara, Sept. 2008, IEEE Computer Press, Los Alamitos
- Petcu, D.; Gorgan, D.; Pop, F.; Tudor, D. & Zaharie, D. (2008). Satellite image processing on a Grid-based platform. *International Scientific Journal of Computing*, vol. 7, issue 2 (August 2008), pp. 51-58, ISSN 1727-6209
- Plaza A. & Chang C. (Eds.) (2008). *High Performance Computing in Remote Sensing*. Chapman & Hall/CRC, Taylor & Francis Group. ISBN 978-1-58488-662-4, Boca Raton
- Portela, O.; Tabasco, A.; Brito, F. & Goncalves, P. (2008). A Grid enabled infrastructure for Earth observation. *Geophysical Research Abstracts*, Vol. 10 (2008), ISSN 1607-7962
- Radu, A., Bacu, V. and Gorgan, D.: Diagrammatic description of satellite image processing workflow. *Proc. SYNASC07*. pp. 341-348, ISBN 0-7695-3078-8, Timisoara, Sept. 2007, IEEE Computer Press, Los Alomitos
- Richards, J. A. & Jia, X. (2006). *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, ISBN: 978-354064-860-4, New York.
- Schott J. R. (2007). *Remote Sensing: The Image Chain Approach*. 2<sup>nd</sup> edition, Oxford University Press, ISBN 978-019517-817-3,, Oxford, USA
- Schowengerdt R.A. (2007). *Remote Sensing: Models and Methods for Image Processing*, 3<sup>rd</sup> Edition, Academic Press, Elsevier Inc., ISBN 978-0-12-369407-2, Burligton, USA
- Sekiguchi, S.; Tanaka, Y.; Kojima, I.; Yamamoto, N.; Yokoyama, S.; Tanimura, Y.; Nakamura, R.; Iwao, K. & Tsuchida, S. (2008). Design principles and IT overviews of the GEOGrid. *IEEE Systems Journal*, Vol.2, No.3 (Sept.2008), 374-389, ISSN 1932-8184
- Tsangaris, M.; Kakalettris, G.; Kllapi, H.; Papanikos, G.; Pentaris, F.; Polydoras, P.; Sitaridi, E.; Stoumpos, V. & Ioannidis, Y. (2009). Dataflow processing and optimization on Grid and cloud infrastructures, *Bulletin of the Technical Committee on Data Engineering*, Vol. 32, No. 1 (March 2009), pp. 67-74, IEEE Computer Society



## **Image Processing**

Edited by Yung-Sheng Chen

ISBN 978-953-307-026-1

Hard cover, 516 pages

**Publisher** InTech

**Published online** 01, December, 2009

**Published in print edition** December, 2009

There are six sections in this book. The first section presents basic image processing techniques, such as image acquisition, storage, retrieval, transformation, filtering, and parallel computing. Then, some applications, such as road sign recognition, air quality monitoring, remote sensed image analysis, and diagnosis of industrial parts are considered. Subsequently, the application of image processing for the special eye examination and a newly three-dimensional digital camera are introduced. On the other hand, the section of medical imaging will show the applications of nuclear imaging, ultrasound imaging, and biology. The section of neural fuzzy presents the topics of image recognition, self-learning, image restoration, as well as evolutionary. The final section will show how to implement the hardware design based on the SoC or FPGA to accelerate image processing.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Dana Petcu, Daniela Zaharie, Marian Neagul, Silviu Panica, Marc Frincu, Dorian Gorgan, Teodor Stefanut and Victor Bacu (2009). Remote Sensed Image Processing on Grids for Training in Earth Observation, Image Processing, Yung-Sheng Chen (Ed.), ISBN: 978-953-307-026-1, InTech, Available from: <http://www.intechopen.com/books/image-processing/remote-sensed-image-processing-on-grids-for-training-in-earth-observation>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821



© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen