Original Article

Formalization of Image Processing Chains for the Dynamics Forest Cover using Landsat Satellite Multi-Sensor and Multi-Temporal

Arisetra Razafinimaro^{1,2,3}, Aimé Richard Hajalalaina^{1,2,3}, Zojaona Tantely Reziky⁴, Eric Delaitre⁵

¹Department of informatics, School of Management and Technological Innovation,

Fianarantsoa University, Madagascar

²Laboratory of Computer Science and Mathematics Applied to Development, Fianarantsoa University

³Informatics-Geomatics Mathematics and Applications, Fianarantsoa University, Madagascar

⁴Higher Polytechnic School of Antananarivo, Antananarivo University, Madagascar

⁵Research Institute for Development, Montpellier, France

Received Date: 18 December 2020 Revised Date: 28 January 2021 Accepted Date: 31 January 2021

Abstract - The Landsat series are among the most widely used satellites in the forestry field, providing multispectral, multi-sensor (RBV, MSS, TM, ETM+, OLI, and TIRS) and multi-temporal images. Methods have been developed by remote sensors for processing these images to monitor forest dynamics but are still not yet capitalizable through distributed computing. The problem that arises is how to make the experimental protocol for satellite image processing for monitoring forest dynamics interoperable while enhancing Landsat multi-sensor images? To answer this question, the notion of work context is introduced to formalize this experimental protocol in the form of a chain of processing chains so that they can be reused, shared, and exchanged, mutualized. interoperable. This formalization is illustrated by the construction of the knowledge base of the working context and the modeling of abstract and concrete chains for the processing of chronological series of Landsat images for the monitoring of forest cover dynamics. An implementation using a WPS (Web Processing Service) server and OTB (Orfeo ToolBox) is implemented for the execution of the concrete chains designed through the distributed computing system.

Keywords – Formalization, Processing chains Image processing, Landsat; Forest covert

I. INTRODUCTION

The world's forest area is shrinking enormously due to heavy deforestation on our planet earth. In addition, several researchers have measured the evolution of forest cover via the solution of remote sensing because satellite image processing is considered a tool that leads to easily estimates the evolution of a phenomenon [1-4]. Landsat satellites are among the best solution and have provided pleasant information for the environmental field. These satellites produce very large space-time images, which have been archived since 1972 and up to the present day. These possibilities make it possible to carry out broad studies following years. The remote detectors valorize these data by putting protocols for the analysis of environmental phenomena while complex treatments are dedicated for these Spatio-temporal data. This context encourages remote sensors to work closely with computer scientists to model and automate in order to simplify these rigorous treatments.

More Scientifics works have been devoted to the study of forest cover using Landsat satellite images for the production of forest surface situation maps, the estimation of rates of change, the implementation of algorithms and processing methods, and the rationale for implementing natural resource management strategies [2],[3],[4],[5],[6],[7]. In this, their results revealed the reliability of the information obtained using Landsat satellite image types. However, the formalization of the processing chains of Landsat satellite images remains to be resolved to make the experiments interoperable in order to facilitate the capitalization of knowledge and the reuse of experiences in the same field. This work is associated with the formalization of SPOT image processing chains carried out by [8].

Resource and processing organizations are recommended for the study of forest cover dynamics using multi-dates and multi-sensor Landsat satellite images to ensure interoperability. We propose to adopt the notion of working context in order to highlight the targeted actors, the organizations at the data level as well as to favor the processing organizations planned for multi-date, multi-source, and multi-sensor data. First, we will talk about the notion of the work context. Then we will propose models taking into account all the criteria recommended by the working context using the modules of the OTB library and publish these models using WPS Server followed OGC normes [9-12].

II. WORK CONTEXT

Setting up a work context requires experimentation in the field study in order to highlight all the necessary components. The compatibility of each composition ensures the validity of a workflow [13]. Therefore, the intervention of experts promotes the validation of the system with the help of organizations. In addition, the construction of a work context conceives the organization of data, the organization of human resources, and the organization of processing. The aim is often to reference future users based on predefined roles and

rights in the human resources organization as well as to reference resources (data/treatments). Various metamodel formalisms lead to the modeling of a processing chain such as UML (Unified Modeling Language), SPEM (Software & Systems Process Engineering Metamodel), BPMN (Business Process Definition MetaModel), BPEL (Business Process Execution Language), and SWM (Simple Workflow Metamodel). Scientific research in the formalization of processing chains has been carried out on SWM. Its formalisms are more adapted to the formalization of processing chains. We will use SWM formalism. Fig. 1 illustrates the components of a work context metamodel.



Fig. 1 Work context [13]

Like all metamodel formalisms, SWM has its particularity in the term of graphic language, which is proposed by Lin et al. in 2011. Fig. 2 shows us the formalism of SWM.

Traitement	Donnée	Port	Lien
Traitement atomique	Donnée	0	Donnée Donnée [Condition]

Fig. 2 SWM Formalism

III.PROPOSAL OF MODEL ORGANIZATIONS

A. Organization of human resources

The formalization of Landsat satellite image processing chains for forest cover requires indispensable human resources to ensure their interaction. However, we identify actors such as the Landsat Image Provider, the satellite image processing expert, the Processing Chain Developer, and the User who will interact directly with the resources.

• The Landsat Image Provider, the one who shares and distributes the satellite data needed at the time of processing. In our case, we have been working with Landsat satellite images throughout the processing, so we assume NASA as our provider.

• The expert in satellite image processing, he designates the protocols for implementing the experimental processing and builds the work context by determining the resource hierarchies (data and processing). Then, he can locate resources in order to add them to the work context by creating and filling in the associated descriptions for each of them [14].

• The developer of the processing chain ensures the operation of the processing, especially at the concretization of the manipulated chain, by checking the types and number of parameters to be inserted. He must understand the working context in order to adjust the needs and its developments.

• User: Once all the systems and data are available, the user serves all the resources by launching his treatment and verifies the results obtained, of which he occupies the validation of result his classification. Fig. 3 illustrates all the players involved in this formalization of the processing chains.



Fig. 3 Organization of Human Resource

B. Organization of the data

The variation in the characteristics of Landsat satellite images represents the evolution made by NASA since 1972 and up to the present. For forest cover, the use of multi-date images increases the quality of the result by giving a strong observation on the evolution of the forest surface. This interest recommends that we take advantage of Landsat multi-sensor images in order to determine the evolution of each study. Moreover, the result of the previous step allows us to trigger the next step during the execution of the processing chain. Thus, we admit that these manipulations of the different data require such an organization. For these reasons, we define four categories of data:

• Raw data for forest cover, represented by the category Raw Data Forestry Cover, includes data delivered by suppliers who have not undergone any prior processing. These include data from Landsat satellite sensors.

• Preprocessed data for forest cover, represented by the category Preprocessed Forest Cover Data, includes data that has been preprocessed but is not usable.

• Intermediate forest cover data is represented by the category Intermediate Forest-Cover Data, which includes data that has undergone intermediate processing.

• The final data for forest cover, represented by the category Final Data Forestry Cover, includes data that can be used and exploited by the end-users for the analysis and interpretation of forest cover dynamics. Fig. 4 illustrates the proposal for the organization of the data.



Fig. 5 Organization of raw Landsat images

C. Organization of treatments

The main objective of this organization is focused on the automation of processing. This automation requires the choice of method to be chosen in order to avoid manual processing, which depends on the characteristics and sensors of the images to be used. This characteristic leads to the differentiation of the processing to be applied. Moreover, all these factors recommend an organization of processing. Thus, in order to facilitate the processing, we propose to divide this organization into three categories separately:

• Pre-processing of forest cover data, represented by the category Pre-processing forest Cover, groups together the

processing of raw data delivered by the Landsat satellite data distributor;

• Intermediate processing of forest cover data, represented by the category IntermediateForestryCover Processing, includes the processing of data whose outputs are not yet usable or analyzable by users;

• Final treatments of the forest cover data, represented by the category Final TreatmentsForestryCover, group those treatments whose output data are directly usable by users. Fig. 6 illustrates the organization of the processing for forest cover dynamics using Landsat satellite images.



Fig. 6 Organization of forest cover treatments using Landsat images

By combining the existing data and processing, we can plot the resource graphs in Fig. 7.



Fig. 7 Resource plot of intermediate treatments for forest cover dynamics using Landsat

IV. ABSTRACT MODELS FOR FOREST COVER DYNAMICS

The organizations in the work context lead us to draw up abstract models that illustrate the type of processing corresponding to the characteristic of Landsat satellite sensors for the study of forest cover dynamics. These models are grouped into three phases such as pre-processing, intermediate processing, and final processing.

A. Abstract pre-processing models

The pre-processing phase takes into account the characteristic of each sensor as well as the possible combination of sensors in order to achieve the use of multi-dates and multi-source images. Thus, we design four abstract chains for pre-processing. This first abstract pre-processing chain is specific for Landsat satellite images of the first series and the second series [15]. In this, we consider the comparison of two images for which this string is valid for the same sensor combination, such as Landsat 1 (RBV-RBV or MSS-MSS), Landsat 2 (RBV-RBV or MSS-MSS), Landsat 3 (RBV-RBV or MSS-MSS), and Landsat 4 (MSS-MSS, TM-TM).this proposal can be applied to both images. Fig. 8 illustrates the steps of this first chain.



Fig. 8 Abstract Model Of Landsat Image Preprocessing First And Second Series From The Same Sensor And Satellite

The second abstract pre-processing chain is specific for Landsat first series and second series satellite images, where the two images come from different sensors such as Landsat 1 (RBV- MSS), Landsat 1(RBV)-Landsat 2(RBV), etc. According to the theory, the use of images from different sensors requires an overlay, one of which is chosen as a reference image. Fig. 9 shows us the second proposition of an abstract pre-processing chain.



fig. 9 Abstract model of Landsat image preprocessing first and second series of different sensors

This third abstract pre-processing chain is intended for sensors that have the panchromatic band, such as Landsat 7 (ETMplus- ETMplus) and Landsat 8 (OLI-OLI). This chain consists of the fusion with the panchromatic band to produce a multispectral image at 15m resolution. In addition, it is possible to use this panchromatic band with Landsat's first and second series satellite images. Fig. 10 shows the structure of this band.



Fig. 10 Abstract model of Landsat image preprocessing with a merge

The last abstract pre-processing string proposal is used for Landsat satellite images that have the panchromatic band that these two images need from different sensors as much as Landsat 7(ETM+)-Landsat 8(OLI). In this, we will have to choose one of these two sensors as a reference image to use for the second overlay. Fig. 11 illustrates the chain of this pre-processing.



Fig. 11 Abstract model of Landsat image preprocessing through merging and overlaying with the reference image

B. Abstract models for the intermediary

The intermediate processing phase focuses on the classification of our images. Generally, we design two kinds of abstract chains for this phase. The first one is designed for the images of the RGB sensor. This sensor has only three spectral bands only. Then, the second one is considered for the other sensors, which have the Red and Near-Infrared band to pass to the calculation of NDVI Vegetation index than the filtering or the forest thresholding [16].Fig. 12 illustrates the chain that went directly to the preparation of the classification of our image after the pre-processing.



Fig. 13 illustrates the intermediate processing chain that implements drill thresholding after the NDVI calculation. This chain is compatible with all Landsat sensors except the Landsat 8 RGB sensor and TIRS.



Fig. 13 Abstract chain of the intermediate treatment passed through the thresholding process

C. The abstract model of final treatments

The third phase consists of providing a result ready to be analyzed directly. The tasks are focused on cleaning the classification results and then comparing the two images in the detection of change [17]. Thus, the use of the two images of different dates is necessary to obtain an accurate result.





Fig. 14 Abstract Chain of Landsat Final treatments for Forest Cover

D. Complete abstract model using two different images ETM+ sensor dates

This abstract model is obtained by combining three (3) abstract models (Fig. 10, 13, 14), which are used for each image. In fact, these three abstract models are constituted by the pre-processing phase, the intermediate phase, and then the final phase. Fig. 15 illustrates the combination of these phases to have a complete model allowing the study of forest dynamics using two ETM+ sensor images.



Fig. 15 Abstract model of a forest dynamics processing chain using 2 images of different dates from ETM+.

V. CONCRETE MODEL FOR FOREST COVER DYNAMICS

Tools are possible to use for the realization, but our choice of tools is particularly based on OTB for the implementation of our chains because it has extensions to the other programming language such as Python. Intuitively, this can facilitate its implementation to the automation of processing chains. In this, we use the OTB version 6.4.0. This version is freely downloadable from the official OTB website [18],[19].

A. Concrete models of treatment chains for pretreatment

Fig. 16 shows a concrete chain for Landsat 1 (RBV-RBV or MSS-MSS), Landsat 2 (RBV-RBV or MSS-MSS), Landsat 3 (RBV-RBV or MSS-MSS), and Landsat 4 (MSS-MSS, TM-TM), using the abstract chain Fig. 8 with the Optical calibration and orthorectification modules.



Fig. 16 Concrete channel using OTB for preprocessing Landsat 1,2,3,4,5 of the same sensor

Fig. 17 shows us a concrete chain for the first series and the second series of Landsat images while both images are captured by different sensors. This concrete chain has implemented the Optical Calibration, Orthorectification, Superimpose modules of the OTB. In addition, the equivalence of this chain is Fig. 9 in the case of an abstract model.



Fig. 17 Concrete chain for pre-processing Landsat images of first and second series, from different sensors

Fig. 18 illustrates the concrete chain planned for sensors that have the panchromatic band, such as Landsat 7 (ETMplus-ETMplus) or Landsat 8 (OLI-OLI). Moreover, it can be used both for the fusion of multispectral first series Landsat images (radiometrically and geometrically corrected) with the panchromatic band of the ETMplus or OLI sensor. The OpticalCalibration, Orthorectification, Superimpose, Pansharpening module of the OTB are required to build this model.



Fig. 18. Concrete chain for pre-treatment based on panchromatic band fusion

Fig. 19 shows us the concrete pre-processing chain planned for the use of the two scenes that have the panchromatic band. In this, the possible combination is Landsat 7(ETM+)-Landsat 8(OLI). Thus, we will have to choose one of the two images as a reference image. The OpticalCalibration, Orthorectification, Overlay, Pansharpening module of the OTB is necessary for its implementation.



Fig. 19 Concrete pre-processing chain using the second superposition

B. Concrete models of treatment chains for the intermediate

As in the abstract models, we have two different types of the processing chain for the intermediate phase, such as the specific intermediate phase for the GVR sensor and the intermediate phase passing in the vegetation index calculation. The concretization of intermediate processing of the RBV sensor uses. The Compute image statistics, train image classifier, and image classifier module of the OTB library. Fig. 20 shows the concrete model of the intermediate phase for the GVR sensor images.



Fig. 20 Concrete chain of intermediate processing for the RGB sensor

Fig. 21 illustrates our proposal on the concrete model of the processing chain for the intermediate phase passing through the forest threshold. This model uses the Radiometric Index, Band Math, Compute Image Statistics, Train Image Classifier, Image Classifier module of the OTB.



Fig. 21 Concrete chain of the intermediate treatment passed through the thresholding process

C. Concrete model of final treatments

The concretization of the third phase in our proposal uses two OTB modules, such as Classification Map Regularization, and Multivariate Alternation Dection on which our chains produce the change between the two images. It is important to note that all Landsat sensors have the same processing in this phase. Fig. 22 illustrates this model.



Fig. 22 Concrete chain of Landsat final treatments for forest cover

D. Complete concrete model using two images of different dates from the ETM+ sensor

Fig. 23 shows a global view of the concretization of the processing chains using two images of different dates from the ETM+ sensor and the OTB library to determine the forest cover dynamics.



Fig. 23 A concrete model of a processing chain for forest dynamics using two ETM+ sensor images and the OTB

VI. DISCUSSIONS

The results of the formalization illustrated the specification of Landsat data, the specification of human resources, and the specification of treatments on the study of forest cover. These formalizations also help us to highlight some major technical aspects of the forest estate while respecting the parameters required for processing. Thus, having identified the advantages and strengths of this formalization, we will first summarize the results obtained by showing the interests within.

On formalization, we have organized the images from the different Landsat sensors in relation to their types. In addition, we have also set up the organization of tasks such pre-processing, intermediate processing, and final as processing for the study of the evolution of the forest surface. In the pre-processing, we have four different chains taking into account the characteristic of each sensor as well as the prediction for the comparison of the two images. Then, we have two chains for the intermediate treatment, of which this part uses the results of the pre-treatments that we have proposed. Finally, we have only one type of chain for the final processing. This formalization allowed us to organize the human resources while respecting all the responsibilities. Consequently, this organization has easily led us to the automation of the processing chain, which makes our work simply because the presence of experts is not mandatory when launching the treatments. A simple user can carry out the treatment by specifying the data in entries, and then he only verifies the sequence of tasks. In addition, this formalization will solve the user's problem of reusing tasks. The latter sublimates the main interest in the formalization of a processing chain compared to the manual processing chain.

Reuse: this property corresponds to the fundamental interest of formalization with task automation. Once a processing chain is formalized, the execution, taking into account different new data, is always performed automatically. This indication makes the chain reusable by allowing the possibility of repeating the process several times without the assistance of experts. Thus, non-expert users can carry out their treatments by reusing the chains and their own data.

Sharing of experience: For the policy of sharing scientific experience, this formalization exactly values the sharing of experience between researchers in Landsat satellite image processing for the study of forest surface evolution. We believe that this work will contribute to the launching of scientific research in the field of remote sensing and the environment. We strongly wish for the maximum diffusion of these chains by allowing everyone to freely use the processing chains that we have proposed.

Processing method: these processing chains are proposals that we have elaborated on according to our knowledge and our experiments, as well as the collaboration with experts in the field of satellite image processing. Indeed, some methods are absolutely simple and fast, but we have approached this type to have precise results and to avoid the interaction of experts as we have seen in the manual treatments. This method is oriented to the use of the OTB modules of which is based on a supervised classification algorithm of maximum likelihood. The final objective is to compare two images of different dates, from the same sensor or not, in order to determine the evolution of the forest cover. Validation of processing results: we know that the training areas used for classification ensure the quality of the processing results that we can measure from the kappa index obtained. If the value of this index is inferior to the acceptable value, then we must improve the training area. The experts take care of the quality of the treatment result when setting up these treatment chains. Field validation is also necessary to enhance the result during the treatments.

Automation: this type requires programming skills in order to realize the need. In this, we cannot neglect the sensor characteristics of the satellite images used, and then all the necessary parameters depend on these characteristics. These characteristics can be resided in the image metadata or from knowledge and experience. Thus, the certainty of the parameters (mandatory or optional) ensures task chaining. On the other hand, the use of the python language facilitates this kind of work in image processing.

VII. CONCLUSION

This article consists of the formalization of processing chains for the study of forest cover dynamics using multidates and multi-sensor Landsat satellite images. The aim of this work is to promote the reuse of the processing chain without the presence of the expert at the time of processing, the capitalization of knowledge, and the sharing of experiments between researchers in the same field. These researchers interact directly with interoperability at the level of scientific research, particularly in the forestry field.

Moreover, this formalization absolutely allows us to set up organizations at the level of the human resources involved, the heterogeneous data, and the treatments according to the notion of the work context. This notion also leads to the proposal of abstract and concrete chains respecting the variation in the characteristics of the images obtained for each Landsat sensor. The implementation of these chains and the automation of the processing are done using the OTB library, the Zoo-project web service, and the Python language.

From the perspectives of this work, we now have readyto-use processing chains for forest cover. This allows us to introduce ourselves to the use of such Grid Computing and Cloud Computing architectures in order to share, mutualize resources in an optimal way and also to use resources on the internet. In addition, we can also consider developing a platform that allows us to automatically process Landsat satellite images from any type of input image sensor. In this, this platform will analyze the metadata of the input images in order to know the type of sensor and the adapted chain.

REFERENCES

- DIAFT et MECNT; Direction des Inventaires et Aménagement Forestières, Ministère d'Environnement Conservation de Nature et Tourisme, Protocole méthodologique de l'évaluation du couvert forestier national de référence en République Démocratique du Congo(2015) 1-34.
- [2] J.Guellec,Possibilités d'utilisation d'images Landsat améliorées à l'echelle de 1/200.000,pour la connaissance des forêts:Revue Bois et Forêts des Tropiques(193)(1980) 41-56.
- [3] Leslie Bouetou-Kadilamio, Suspense Averti Ifo, Stoffenne Binsangou, Changement de Couverture Forestière Dans le département de la Likouala (République Du Congo) durant la période de 1986 à 2015: European Journal scientific. (13) (2017) 322-343.
- [4] Marine Boulogne. Vulnérabilité des paysages forestiers dans le parc de Ranomafana (Madagascar):dynamiques environnementales et trajectoires agroforestières. Environnement et Société. Université Grenoble Alpes;(2016)
- [5] Cyrille Chatelain, Mitia R. Hanitrarivo, Bric F., L. Rakotozafy, Ralph Bolliger, lacopo Luino, Patrick Ranarison, Laurent Gautier, Cartographie de la couverture forestière du massif deBeanka, Région Melaky, Ouest de Madagascar: Malagasy nature (7) (2013) 85-103.
- [6] Rakotomalala F.A,Rabenandrasana J.C.,Andriambahiny J.E.4,Rajaonson R.,Andriamalala F.,Burren C.,Rakotoarijaona J.R,Parany B.L.E, Vaudry,R.,Rakotoniaina S.,Ranaivosoa R.,Rahagalala P.,Randrianary T.,Grinand C., Estimation de la déforestation des forêts humides à Madagascar utilisant une classification multidate d'images Landsat entre 2005,2010 et 2013:Revue Française de photogrammétrie et de la télédétection (211-212)(2015) 11-23
- [7] Aimé Richard Hajalalaina, Manuel Grizonnet, Eric Delaître, Solofo Rakotondraompiana, Dominique Hervé, discrimination des zones humides en foret malgache, proposition d'une methodologie multiresolution et multisource utilisant orfeo toolbox:Revue Francaise de Photogrammetrie et de Télédétection(201)(2013) 37-48.
- [8] Aimé Richard Hajalalaina, Dominique Hérvé, Eric Delaitre, Thérèse Libourel; Modeling process chain of SPOT images for resources uncertainty to monitor change in forest cover: Spatiale Accuracy (2016) 38-45
- [9] Christopher Michaelis, Daniel P. Ames Evaluation and Implementation of the OGC Web Processing Service for Use in

Client-Side GIS: Geoinformatica (13)(2009) 109-120

- [10] Fitzke, J., Greve, K., Müller, M., and Poth, A., Deegree–ein Open-Source-Projekt zum Aufbau von Geodaten infrastrukturen auf der Basis aktueller OGC-und ISO- Standards: GeoBIT/GIS, (9)(2003) 10-16.
- [11] M. Dumas, M.-C. Fauvet, Intergiciel et Construction d'Application Reparties (2008) 77-101.
- [12] Marc Gilgen Open GIS Consortium, Apercu et perspectives de l'OpenGIS dans le domaine du Web Mapping: EPFL-SIRS (2001);
- [13] Yan Lin; Méthodologie et composants pour la mise en œuvre de workflows scientifiques. Thèse de doctorat, Ecole Doctorale Information, Structure et Système, Université de Montpellier, France, 2(2011) 1-192
- [14] Aimé Richard Hajalalaina; Aide à la formalisation des chaînes de traitement de données spatiales Capitalisation des connaissances sur les thématiques forêt et climat à Madagascar; (2015).
- [15] B. Lounis, A. Belhadj AÏSSA, Processus de Correction Radiométrique Relative PCRR Appliqué Aux Images Landsat TM Multi-Dates, International Conference: Science of Electronic, Technologies of Information and Telecomunications (2005)(27-31).
- [16] Boulfroy, E., J. Khaldoune, F. Grenon, R. Fournier et B. Talbot. Conservation des îlots de fraîcheur urbains - Description de la méthode suivie pour identifier et localiser îlots de fraîcheur et de chaleur (méthode en 9 niveaux). CERFO et Université de Sherbrooke. Rapport 2012-11 (2013) 40.
- [17] Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Digital change detection methods in ecosystem monitoring: A review. International Journal of Remote Sensing, (25)(2004) 1565-1596.
- [18] Orfeo ToolBox, The ORFEO ToolBox Software Guide UpdatedforOTB-6.4.0.www.orfeotoolbox.org/packages/ OTBSoftwareGuide.pdf (2018).
- [19] Inglada J. et Christophe E., The Orfeo Toolbox remote sensing image processing software. Dans: IEEE International Géoscience and Remote Sensing Symposium, Le Cap, Afrique du Sud,(4)(2009)733– 736.
- [20] Surendiran, R., and Alagarsamy,K., 2010. Skin Detection Based Cryptography in Steganography (SDBCS). International Journal of Computer Science and Information Technologies (IJCSIT), 1(4) 221-225.