Raster-vector cooperation algorithm for GIS

Application to ecological units delineation

Enguerran Grandchamp
Département Mathématiques et Informatique
GRIMAAG, Université des Antilles et de la Guyane
Pointe-à-Pitre, Guadeloupe France
e-mail: egrandch@univ-ag.fr

***Abstract* - During the last decade, Geographic Information Systems (GIS) considerably evolved: the amount of stored data explodes and the tools used to treat them have been improved.**
During the same period the arrival of very high spatial resolution satellite images (less than 1 meter) gives an inexhaustible source of information actually largely unexploited.
We generally use images (satellite or airborne) in GIS by stacking vector information to visualize thematic maps but without exploiting the images as an information source.
We propose in this paper a way to use satellite images as an information source in order to (i) produce new information layers (ii) validate some fusion operation made with the other information layers.
We present an iterative algorithm to select and fuse layers (with eventually a relaxation). We use an homogeneity criterion based on local texture analysis to validate the selection.
After running this algorithm, new information layers are produced using the images and depending on the previous fusions and selections.
We apply this algorithm to search Ecological Units within forest in order to guide the research of observation samples for the underlying ecosystems.

Keywords - ***Remote sensing; GIS; Information Fusion; Texture Analysis; environment survey***

I. INTRODUCTION

Many information have been collected over the earth surface and Geographic Information Systems (GIS) have been fulfilled by different ways (automatically, manually …) and by different actors (biologists, autorities …) depending on the application (geography, country planning, resources preservation …).
The very high resolution satellite images from QuickBird (0.6m) and IKONOS (1m) allow extracting information from satellite image scenes with a fair accuracy [1, 2, 3, 4].
In this study we propose to combine both information sources in order to produce useful maps representing several information at the same time. To select the interesting information among the different layers we propose an iterative optimization process. This leads to theoretical units which are validate by a image analysis. We use an homogeneity criterion based on the visual aspect of the different units. This criterion is computed from the satellite images with a color and texture analysis.

In a second step, when we obtain final units we compute some features from the color and texture and produce new information layers.

As an example this allows studying the increase or decrease of forest surfaces. With classical remote sensing analysis (segmentation and classification based on multispectral analysis) it is possible to detect land cover changes by comparing multi-date classifications. But it doesn't allow a thin follow-up of the different ecosystems inside the global forest. Moreover, some ecosystems could disappear while the forest has globally increased.
The most reliable approach to realize such a task is based on ground truth and allows precisely defining and localizing each ecosystem, but it is time expensive. It has been made for the studied forest with coarsed sampling using small squares manually investigates.
By combining the information layers we can split the forest into theoretical ecosystems (Ecological Units) and validate them with the image analysis.

Section 2 defines the notion of unit. Section 3 presents the different steps of the algorithm by detailing fusion and homogeneity criterion. Section 4 and 5 present an application of the algorithm for ecological units (EU) delineation. Then section 6 gives the conclusions and perspectives of this study.

II. UNIT DEFINITION

In the scope of this article we only deal with layers regrouping areas having the same value for a defined feature. Network integration (rivers, roads …) is not discussed here.

Each layer is composed of several classes. Each class is composed of several disjointed areas.

We name these classes *Units* if the rest of this paper.

When we combine two layers, the intersection of the units leads to new smaller units having several common features.

The number of units obtained after crossing layers can quickly increase. Specially if there are some localization errors. Indeed, a bad localization leads to undesired units which disturb the treatment of the other units and a correction process must be defined.

Figure 1. illustrates the notion of Unit within a layer (a, b), the crossing of two layers (c) and errors linked to geo-localization (d). We remark a blue line along the green unit due to a lack of precision in the localization.

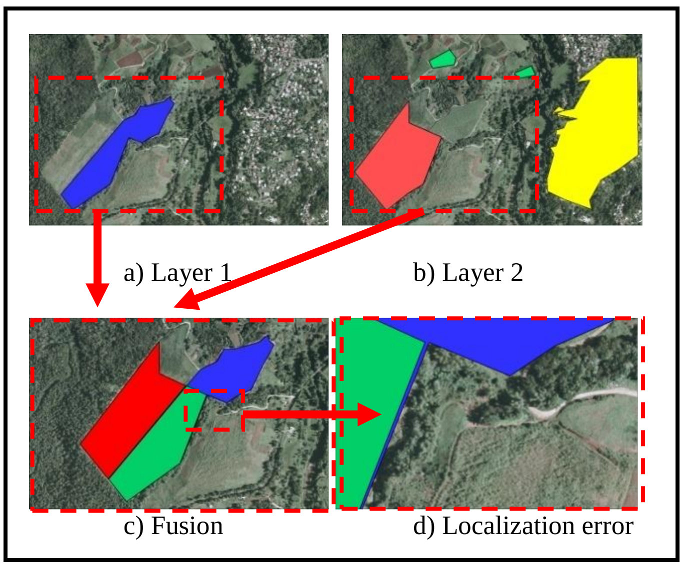


Figure 1. Unit definition

The preprocessing of the layers to correct localization errors is not detailed in this paper.

A theoretical unit is then an homogenous area regarding the features obtained from layers.

In an image analysis point of view, a unit must be homogenous regarding a criterion based on texture and color components.

Theoretical units are not necessary homogenous according to image analysis.

The algorithm described in the following section is specially design to determine the layers to use in order to obtain both homogeneities.

III. ALGORITHM

A. *Principle*

The objective is then to obtain a selection of layers leading to visually homogenous units. The use of all layers is not useful because it leads to a lot of small units.

Nevertheless, the more layers we combine to obtain the units, the more representative they are. So we have to make a tradeoff between size and representativeness.

The proposed algorithm is iterative and its main steps are: *(i)* a selection process extracts a sub-set of information layers from the SIG *(ii)* a fusion process computes the units *(iii)* an evaluation process computes the homogeneity criterion and decides if the algorithm stops or not *(iv)* after having selected the final layers, a classification process is used to produce new units (from non homogeneous units) or information layers (from homogeneous units) using image analysis.

Figure 2 illustrates the algorithm.

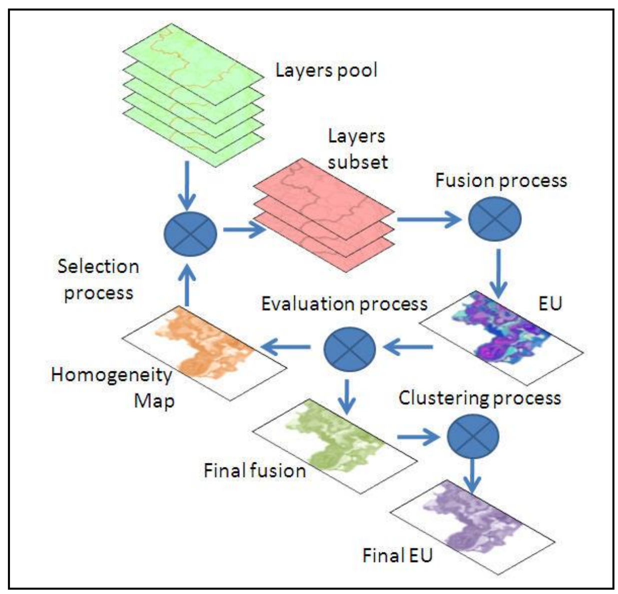


Figure 2. Algorithm

B. *Homogeneity criterion*

Validation and selection of layers are based on an homogeneity criterion computed by analyzing color and textures of satellite images.

This computation is done separately on all units.

Color [8] and textures features are: mean and standard deviation over each color band, co-occurrence matrices [9], Gabor filters [10], Laws filters [11], Hue moments [12] and fractal dimension [13]. We obtain 25 normalized features integrating geometrical, statistical, frequential and fractal approaches.

The homogeneity coefficient associated with a unit is the inverse of the standard deviation of the features. As a unit can be composed of several disjoined areas, the coefficient is computed on each one separately. Then a global coefficient is computed for the whole unit. The higher the coefficient is, the more homogeneous the unit is.

C. *Layers selection*

The first selection is composed of every layers. This leads to a lot of small unit having a maximal homogeneity.

Indeed, there is no other layer to split a unit in more homogenous ones. Then the homogeneity coefficient can only decrease but to reach wider units.

This is a combinatorial step and an exploration of every combination is generally not possible if we manage more than 10 layers.

The selection evolves by removing, adding or replacing a layer.

The selection criterion is based on the previous coefficient computed independently on each layer.

D. *Layers fusion*

The layers fusion consists on computing polygons intersections. This step is time consuming and localization errors have to be treated before to fuse in order to ensure the validity of the resulting units. More details about this step can be found in future publications.

E. *Evaluation*

The evaluation step consists on computing the homogeneity coefficient, the number of units and their localization and surface. According to this features we decide to stop or not the process.

F. *Classification*

After having selected the layers we solve local homogeneity problems by applying a clustering on non homogeneous units.

We use K-Means [14] and SVM [15] clustering algorithm.

When a unit is not homogenous we try to localize sub units by using a clustering on the color and texture features.

The resulting clusters lead to new units as soon as most of the cluster samples are regrouped (we compute the spatial repartition of the corresponding points).

Then a vectorization of the clusters is done to obtain a new information layer. We use a snake function to estimate the polygon representing the new unit borders. Finally a name is given by the experts to each area.

IV. Application to Ecological Units delineation

Nowadays the satellite images resolution allow extracting vegetation with a fair accuracy [1, 2, 3, 4] and studying their border with a one tree precision [5, 6, 7].

However, when we analyze the whole forest, the surface can globally increase whereas some ecosystems disappear. Indeed, the wide surface of the forest and the complex structure of the canopy don't allow splitting the forest into different ecosystem only using image analysis.

An ecosystem is relatively complex to define and localize but experts observe that same ecosystems evolved in close environmental conditions. Another remarks is that visual aspect of identical ecosystems is close, the validation using the proposed homogeneity criterion is then well adapted.

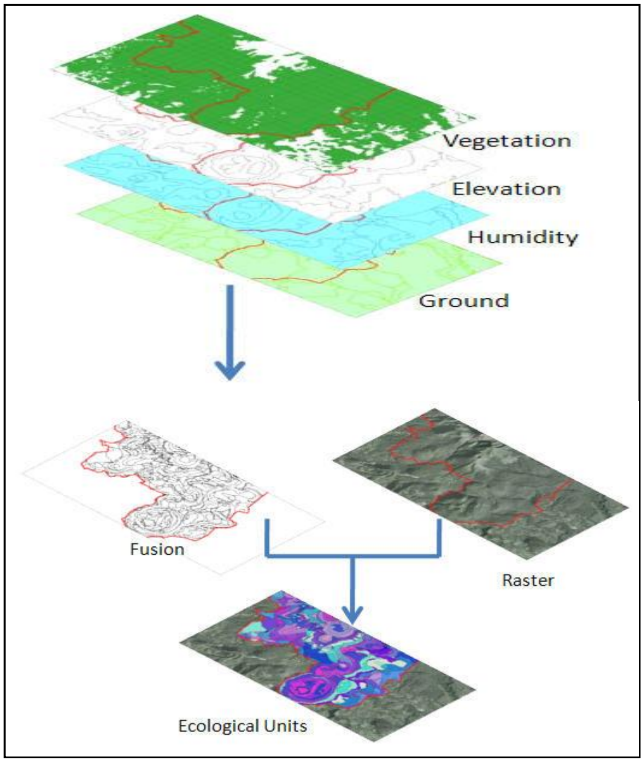


Figure 3. Ecological Units (EU)

So we use the previous algorithm to access the Ecological Units (EU) which represent potential ecosystem. Reducing the number of units (and then increasing their surface) is essential for biologists in order to reduce the number of survey samples.

We use 26 information layers such as humidity, ground, unrefined vegetation, temperature, slope … We use very high resolution images from IKONOS and QuickBird satellites. Color and textures features are computed on 5x5 windows and the features vector is normalized.

V. Results

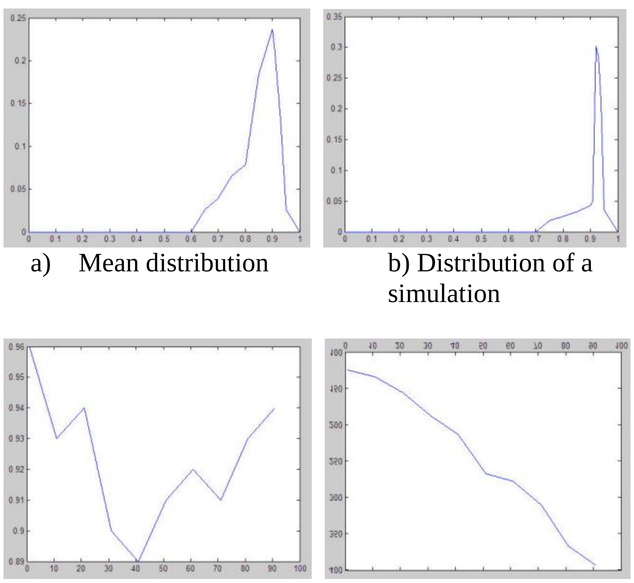
A. *On synthesis data*

Before applying the algorithm on real data we validate the approach on synthesis data. We define a virtual set of ecosystems, a virtual set of layers and a synthesis image composed of different textures with the following rules: an ecosystem is represented by a unique texture but two different ecosystems can be represented by the same texture.

A first test set was built in order to have a correspondence between the ecosystems and the Ecological Units. This set validates the selection, fusion and evaluation processes. The algorithm reaches the sub set of layer corresponding to the ecosystems in 100% of the 100 simulations.

A second test set introduces errors in the layers definition. There is no more subset of layers leading to the ecosystems. This leads to non homogenous ecological units. The best resulting sub set is then improved using the clustering process to extract units from non homogenous ones.

Figure 4. a) presents the distribution of the mean homogeneity coefficient computed at the end of the algorithm for 100 simulations. Figure 4. b) presents the distribution of the homogeneity coefficient for one simulation. Homogeneity coefficient are relatively high (upper than 0.9). Figure 4. c) shows the evolution of the coeeficient during the simulation. As expected is globally decrease during the simulation. Figure 4. d) shows for the same simulation the mean surface of a units which increase as expected.



a) Mean distribution

b) Distribution of a simulation

c) Evolution during simulation

c) Unit surface evolution

Figure 4 Homogeneity coefficient

B. On real case data

We use the National Parc of Guadeloupe for the real cases experimentation. Some of the ecological units have been explored by experts to validate the results.

Results are visually satisfactory (homogenous ecological units) but other ground validations are required to statistically validate the approach.

Figure 5. presents an extract of the ecological units obtained at the end of the algorithm (before clustering). We can see 3 EU visually close but splited by the information layers combination. Homogeneity coefficients are respectively: 0.93, 0.7 and 0.95.

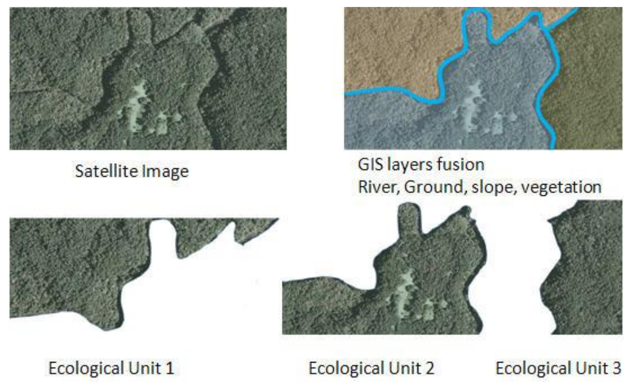


Figure 5. Resulting Ecological Units

EU number 2 is not homogeneous and the clustering process must be applied to eventually detect sub units. In order to illustrate the clustering process (based on color and texture features) we apply this step on each EU.

Figure 6. presents clustering results using k-Means algorithm. The number of clusters is set to 2 for EU 1 and 3 because the unit is homogenous (one for the ecosystem and one for the shadow). Resulting clusters are representative of the classes.

In the EU number 2 we set the number of classes to 3 (one for the main ecosystem, one for the ecosystem responsible of the non homogeneity and one for the shadow).

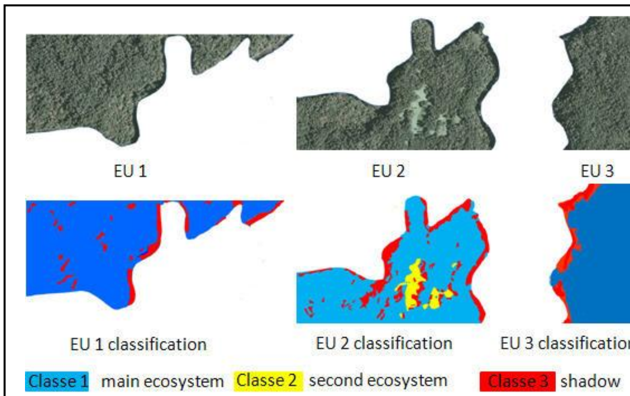


Figure 6. Clustering results

The second ecosystem is clearly localized (yellow on figure 6.).

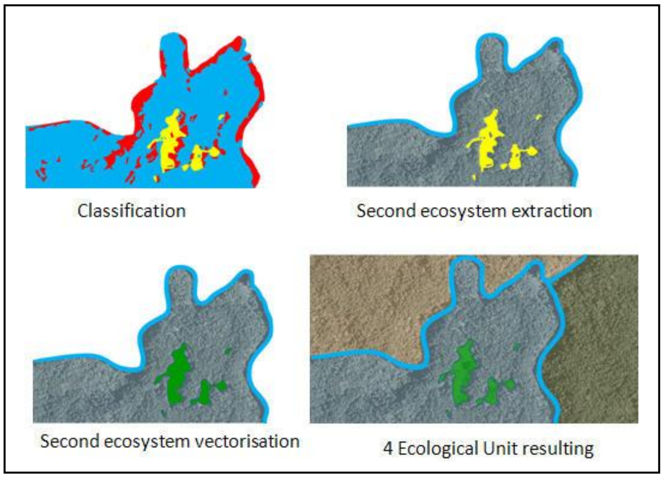


Figure 7. New ecological units

Figure 7. shows the vectorisation process results. This step allows addying new information layers and new units to the GIS. The name of the unit is given by the experts and in this case is linked to a ground slippage.

VI. CONCLUSION

GIS offer a powerfull framework to efficiently apply algorithm developped for remote sensing and image analysis. By combining remote sensing and vectorial information layers we can automatically restrict the area to apply the algorithm. Moreover we use the satellite images as real information sources and produce new information layers.

The problem is combinatorial and we can't explore every combination to select the best layer subset. So we propose a local iterative algorithm in order to reach a solution in an acceptable time.

We validate the approach using both synthesis and real data. We can improved the algorithm by developping a method to automatically set the different threshold (homogeneity, spatial dispertion …).

A ground truth will also be useful to statistically validate the approach.

REFERENCES

[1] Chen Junying; Tian Qingjiu, Vegetation classification model based on high-resolution satellite imagery, Proceedings of SPIE, Remote sensing of the environment, 19-23 August, 2005, Guiyang City, China

[2] [Wulder, Michael A.; White, Joanne C.; Fournier, Richard A.; Luther, Joan E. And Magnussen, Steen. Spatially Explicit Large Area Biomass Estimation: Three pproaches Using Forest Inventory and Remotely Sensed Imagery in a GIS. Sensors 2008, 8, 529-560.

[3] Qian Yu, Peng Gong, Nick Clinton, Greg Biging, Maggi Kelly, and Dave Schirokauer, Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery, Photogrammetric Engineering & Remote Sensing Vol. 72, No. 7, July 2006, pp. 799–81

[4] Kasper Johansen, Nicholas C Coops, Sarah E Gergel, Yulia Stange, Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification, Remote Sensing of Environment, Vol. 110, No. 1. (14 September 2007), pp. 29-44

[5] A. Gougeon and al**,** Individual Tree Crown Image Analysis – A Step Towards Precision Forestry. Natural Resources Canada, 2001.

[6] A. Gougeon and al**,** Individual Tree Crown Image Analysis – A Step Towards Precision Forestry. Natural Resources Canada, 2001.

[7] M. Abadi, Phd Thesis, Couleur et texture pour la représentation et la classification d'images satellite multi-résolutions, 2008

[8] ANGWINE, S. J., AND HORNE, R. E. N. The colour Image Processing Handbook (oelectronics, Imaging and Sensing). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.

[9] Gotlieb, C. C., and Kreyszig, H. E. Texture descriptors based on cooccurrence matrices. Computer Vision Graphic Image Process 51, 1 (1991), 70–86.

[10] Manjunath, B. S., and Ma, W. Y. Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Anal. Mach. Intell. 18, 8 (1996), 837–842.

[11] Laws, K. I. Rapid texture identification. In SPIE Image Processing for Missile Guidance 238 (1980), 376–380.

[12] Hu, M.-K. Visual pattern recognition by moment invariants. Information Theory, IEEE Transactions on 8, 2 (1962), 179–187.

[13] Mandelbrot, B. B. Fractals: Form, Chance and Dimension. W. H. Freeman and Co., 1977.

[14] Scarpa, G. Haindl, M., and Zerubia, J. Hierarchical finite state modeling for texture segmentation with application to forest classification. Research Report 6066, INRIA, France, 2006

[15] E. Grandchamp, GIS information layer selection directed by remote sensing for ecological unit delineation, IGARSS 2009