An Evolutionary Approach for Ontology Driven Image Interpretation

Germain Forestier, Sébastien Derivaux, Cédric Wemmert and Pierre Gançarski

LSIIT - CNRS - University Louis Pasteur - UMR 7005
Pôle API, Bd Sébastien Brant - 67412 Illkirch, France
Email: {forestier,derivaux,wemmert,gancarski}@lsiit.u-strasbg.fr

Abstract. Image mining and interpretation is a quite complex process. In this article, we propose to model expert knowledge on objects present in an image through an ontology. This ontology will be used to drive a segmentation process by an evolutionary approach. This method uses a genetic algorithm to find segmentation parameters which allow to identify in the image the objects described by the expert in the ontology. The fitness function of the genetic algorithm uses the ontology to evaluate the segmentation. This approach does not needs examples and enables to reduce the *semantic gap* between automatic interpretation of images and expert knowledge.

1 Introduction

Automatic interpretation of images becomes a more and more complex data mining process. For example, in the field of remote sensing, the rapid evolution in terms of spatial resolution (image size) and spectral resolution (number of bands) increases the complexity of available images. Automatic analysis methods are needed to avoid a manual processing which is often costly. The most promising and the most studied approach is the *object oriented* approach which consists in identifying objects composed of several connected pixels and having an interest for the domain expert, by using a segmentation algorithm.

There exists many algorithms of segmentation like the watershed transform [12] or region growing [7]. These algorithms need a complex parametrization like the selection of thresholds or weights which are usually meaningless for the user. Thus, a difficult task for the user is to find the link between his knowledge about the objects present in the image and the appropriate parameters for the segmentation allowing to create and identifying these objects.

The use of genetic algorithm [4] is a solution to find an optimal (at least near-optimal) parameters set. They can be used to optimize a set of parameters if an evaluation function of these parameters is available. The existing methods of segmentation optimization with genetic approaches [9,1,11,3] are based on evaluation function where examples of segmented objects provided by the expert are needed. If examples are not available, it is possible to use some unsupervised criteria [1,3], which evaluate the intrinsic quality of a segmentation (e.g. region

This is the author's version of an article published at EvoIASP 2008. The final authenticated version is available online at: https://doi.org/10.1007/978-3-540-78761-7_30

homogeneity). Nevertheless these unsupervised criteria are often insufficient to produce a good quality segmentation, especially for the analysis of complex images.

In this article, we propose to use domain knowledge to evaluate the quality of a segmentation. Indeed, with the oriented object approach the expert is able to express his knowledge about objects of the image. It allows a natural and intuitive description of the objects potentially present in an image. Thus, an ontology (i.e. a knowledge base) can be used to define the different objects (i.e. concepts) and their characteristics. Then, the coherence of a segmentation can be evaluated thanks to the concepts defined in the ontology. This approach does not needs examples and uses the knowledge defined in the ontology.

The outline of this paper is the following. In Section 2, we introduce the used segmentation algorithm and a description of the needed parameters. In Section 3, we present the formalization of the knowledge through an ontology. In Section 4, we study the proposed evolutionary approach used to find the set of parameters for segmentation thanks to an evaluation using the ontology. Finally, we present experimentations on the interpretation of images for Earth observation.

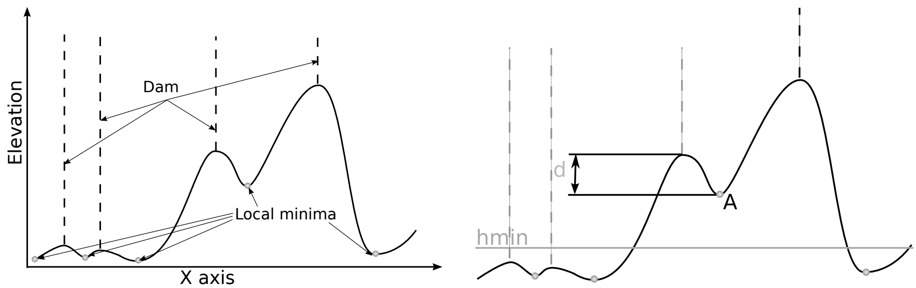
2 Image segmentation

The watershed segmentation is a well-known segmentation method which considers the image to be processed as a topographic surface. In the immersion paradigm from Vincent and Soille [12], this surface is flooded from its minima thus generating different growing catchment basins. Dams are built to avoid merging water from two different catchment basins. A example of a cut of an elevation image and its minima is presented in figure 2 (a).

The segmentation result is defined by the locations of the dams (i.e. the watershed lines). In this approach, an image gradient is most often taken as the topographic surface, since object edges (i.e. watershed lines) are very probably located at pixels with high gradient values (high heterogeneity areas). To build the gradient image, each pixel is replaced by the difference between the maximal value and the minimal value of a 3×3 windows centered on the pixel. The final elevation image is obtained by combining the elevation of the different spectral bands thanks to an Euclidean norm.

The watershed has the advantage to be a completely unsupervised method without parameters. Nevertheless, it produces generally an over-segmented result, which means an image where each object (e.g. a house) is represented by several regions (e.g. the two sides of its roof). To resolve this problem, many methods can be used independently or simultaneously.

When the gradient image is computed, a threshold of the gradient [5] can be made. Every pixel having an inferior value to the threshold is set to zero. Thus, the small variations within the homogeneous region are deleted. On figure 2 (b), the line *hmin* represents a threshold, and the value beside its are considers as null. Another method consists in using the depth of the basins [8]. Let m_r be the elevation of a local minimum of the basin r and d_r be the minimal elevation when



(a) Immersion of a elevation map.

(b) Illustration of over-segmentation reduction method for the watershed transform.

Fig. 1. Example of a watershed segmentation (a) and effect of over-segmentation reduction methods (b).

it will be separated of an another basin by a watershed. Every local minimum where $d_r - m_r < d$, with d a given threshold, will not be considered during the basin immersion step. On the figure (b), the local minimum A will not be taken into account during the immersion because its dynamic, the difference between m_r and d_r, is too small. Finally, it is also possible to use a region merging technique [5]. Two regions can be separated by an heterogeneous area (implying a generation of a frontier by the watershed) but they can be spectrally similar (using the average value). To solve this problem, it is possible to use a filter which merge adjacent regions having an euclidean distance between their means lower than a threshold ft.

These different techniques can be used simultaneously to reduce over segmentation caused by the watershed and need the selection of three parameters (the level $hmin$, the threshold d and the fusion threshold ft). The optimal values of these parameters are difficult to find because the value for a given parameter depends heavily of the values selected for the other ones. Moreover, there exists a lot of local optima which increase the difficulty to find the best solution.

3 Geographical objects ontology

We propose here a model allowing the representation of geographic objects through an ontology and a matching process which allows to compare a region build during a segmentation and the different concepts defined in the ontology. The matching process has been fully described in [2], we remain here the principal functionalities.

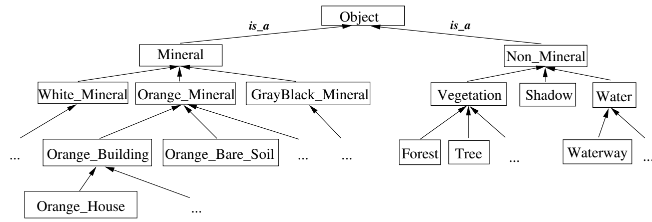


Fig. 2. Excerpt of the ontology.

3.1 Ontology description

The proposed ontology is composed of a hierarchy of concepts (an extract is given on figure 2). In this hierarchy each node corresponds to a concept. Each concept has a label (e.g. *house*) and is defined by its attributes. Each attribute is associated to an interval of accepted values (e.g. $[50; 60]$) and a weight (in $[0; 1]$) representing its importance to recognize the geographical object corresponding to this concept (1 meaning that the attribute is very relevant). The values of these concepts have been filled by geographers experts thanks to their knowledge about the morphology of urban objects and machine learning tools [10].

3.2 Region matching

A matching mechanism of the region allows to evaluate the similarity between a region built during a segmentation and the concepts defined in the hierarchy of the ontology. The region matching consists in verifying the validity of feature values of a region (spectral response, size, elongation, ...) according to the properties and the constraints defined in the concepts of the ontology. The measure of matching computes the similarity between the characteristics $v_1^r \dots v_n^r$ of a region r and the attribute $a_1^k \dots a_n^k$ of a concept k is composed of a local component (dealing with the inner properties of the concept) and a global component (evaluating the pertinence in the hierarchy of concepts).

The *degree of validity* $Valid(v_i^r, a_i^k)$ evaluates the validity between the extracted characteristics v_i of a region r and the boundaries of accepted values of the attribute a_i of the concept k.

$$Valid(v_i^r, a_i^k) = \begin{cases} 1 & if \;\; v_i^r \in [min(a_i^k); max(a_i^k)] \\ \frac{v_i^r}{min(a_i^k)} & if \;\; v_i^r < min(a_i^k) \\ \frac{max(a_i^k)}{v_i^r} & if \;\; v_i^r > max(a_i^k) \end{cases}$$

The measure of *local similarity* $Sim(r, k)$ compares all the common characteristics of the region r with the attributes of the concept k. The value λ_i^k is the weight of a_i^k and represents the importance of a_i^k to identify k.

$$Sim(r,k) = \frac{\sum_{i=1}^{n} \lambda_i^k Valid(v_i^r, a_i^k)}{\sum_{i=1}^{n} \lambda_i^k}$$

The *matching score* $Score(r,k)$ evaluates the relevance of the matching between the region r and the concept k within the hierarchy of concepts. The matching score is a linear combination of local similarity measure obtained with the concepts k_j of the path starting from the root of the ontology to the studied concept k_m. This calculation integrates the proof of the concepts β_i to advantage lower concept in the hierarchy.

$$Score(r,k_m) = \frac{\sum_{j=1}^{m} \beta_j Sim(r,k_j)}{\sum_{j=1}^{m} \beta_j}$$

With such a matching process, each region produced by the segmentation can have a score which represents its suitability to the concepts formalized in the ontology.

4 Genetic algorithm

In this section, we are interested in using an evolutionary approach to find the parameters of the segmentation algorithm by using the knowledge contained in the ontology. We start by describing the genetic algorithm and we detail the chosen evaluation function.

4.1 Description

A genetic algorithm is an optimization method. Given an evaluation function $\mathbb{F}(g)$ where g is taken in a space $\mathbb{G}$, the genetic algorithm searches the value of g where $\mathbb{F}(g)$ is maximized. Genetic algorithms are known to be effective even if $\mathbb{F}(g)$ contains many local minima. In order to consider this optimization as a learning process, it is required that the optimization performed on a learning set could be generalized to other (unlearned) datasets.

Here we consider g (the genotype in the genetic framework) as a vector containing the parameters to be tuned automatically in the watershed segmentation process. All these parameters are normalized in $[0;1]$, so here $\mathbb{G} = [0;1]^3$ as we consider 3 parameters to optimize: *hmin*, d and ft.

A genetic algorithm requires an initial population defined as a set of genotypes to perform the evolutionary process. In this process, the population evolves to obtain better and better genotypes, i.e. solutions of the optimization problem under consideration. In order to build the initial population, each genotype is randomly chosen in the space $\mathbb{G}$ except one which uses default parameters. By this way, we ensure that the final solution is as good as the default one. In our case, the default set of parameters is $\{0,\ 0,\ 0\}$, thus disabling the various over-segmentation reduction methods described previously.

Once the initial population has been defined, the algorithm relies on the following steps which represent the transition between two generations:

1. assessment of genotypes in the population.
2. selection of genotypes for crossover weighted by their score rank, as discussed in the following subsections.
3. crossover: two genotypes (p_1 and p_2) breed by combining their parameters (or genes in the genetic framework) to give a child e. For each parameter g_i, $g_i(e)$ is computed as the value $\alpha \times g_i(p_1) + (1-\alpha) \times g_i(p_2)$ where α is a random value between 0 and 1. We apply an elitist procedure to keep in the next generation the best solution of the current generation.
4. mutation: each parameter may be replaced by a random value with a probability $\mathbb{P}_m$. Thus we avoid the genetic algorithm to be trapped in a local minimum. As indicated previously, the best genotype of a generation is kept unchanged.

We use a mutation rate of $\mathbb{P}_m = 1\%$ and a number of generations of 15, as experiments shown that more generation do not increase the results.

4.2 Choice of the evaluation function

A critical point of genetic algorithm methods is how the quality of potential solutions (i.e. genotypes) is estimated, through evaluation criterion. We use here the ontology knowledge to drive the evolutionary process and find the set of parameters which allows to maximize the discovery of objects within an image. Thus, we propose to use as evaluation function, the percentage of the image which is identified by the ontology. Let $\mathcal{R}^g$ be the set of regions of a segmentation obtained with the parameters g and $\mathcal{R}^g_o$ be the set of regions identified by the ontology ($\mathcal{R}^g_o \subseteq \mathcal{R}^g$). The percentage of the surface of the image recognized by the ontology is defined as :

$$\mathbb{F}(g) = \frac{\sum_{r \in \mathcal{R}^g_o} Area(r)}{\sum_{r \in \mathcal{R}^g} Area(r)}$$

with $Area(r)$ a function returning the surface in pixels of the region r. The surface of the identified regions has been preferred to their number to evaluate the result. Indeed, a segmentation algorithm can produce many small regions which do not cover any type of concept in the ontology and thus that can not be identified by the ontology. These small regions can perturb a criteria based on the number of regions. The criterion based on the surface of the regions allows to quantify the quality of the segmentation according to the knowledge present in the ontology. The increasing of this criterion indicates that the regions built by the segmentation correspond more and more to the description of the geographical objects described in the ontology. By maximizing this criterion we build a segmentation matching with the expert knowledge about geographical objects.

5 Experimentations

The proposed method have been evaluated on an image of Strasbourg taken by the satellite Quickbird. The size of the image is 900 x 900 pixels and the spectral resolution is 8 bits for each of the four band. The figure 3 presents the image to interpret. The figure 4 (a) presents an extract of the image and the figure 4 (b) presents the segmentation of this extract by the watershed without parametrization. The figures 4 (c) (d) (e) and (f) present extracts of segmentation with the parameters found during the genetic evolution. We observe an amelioration of the construction of the objects, the image being better identified by the ontology with the number of generation. To validate these results we have evaluated the quality of the segmentations obtained with ground truth of three thematic classes, *house*, *vegetation* and *road*. These evaluations are done on geographic objects built and labeled manually by an expert. Three quality indexes have been used to evaluate the quality of the segmentations.

The first index used is the *recall*. It consists in considering the identification of the ontology as a classification of the image. The pixels of the objects identified by the ontology are then compared to the pixels of the objects provided by the expert:

$$recall = \frac{number\ of\ well\ labelled\ pixels}{number\ of\ expert\ pixels}$$

It takes its values in $[0;1]$, the more it is near to 1, the more the image is well identified.

The second index used is the index of *Janssen* defined in [6]. It evaluates the concordance between the expert objects and the corresponding regions in the segmentation. For each expert object i and each corresponding region j having the biggest intersection with the object i, this index is defined as :

$$Janssen_{(i,j)} = \sqrt{\frac{Area(i \cap j)}{Area(i)} \times \frac{Area(i \cap j)}{Area(j)}}$$

It takes its value in $]0;1]$, 1 meaning a perfect correspondence between the expert objects and the regions of the segmentation.

The third and last index is the index of *Feitosa* defined in [3]. It also evaluates the correspondence between the expert object and the corresponding regions in the segmentation. With the same notation, it is defined as :

$$Feitosa_{(i,j)} = \frac{Area(i \setminus (i \cap j)) + Area(j \setminus (i \cap j))}{Area(i)}$$

It takes its value in $[0, (Area(i) - 1) + (Area(j) - 1)]$, the nearer from 0 it is, the more the regions correspond to the expert objects.

We have compared our evaluation criterion to these three criteria. For each criterion a mean is computed on the set of objects provided by the expert. The goal of this analysis is to check that maximizing our criterion leads to a real amelioration of the segmentation. During the evolutionary process we have evaluated each individual according to the different criteria. Thus, 200 possible parametrizations have been evaluated. The different set of parameters have been ordered according to our criterion of evaluation. The figure 5 shows the curves

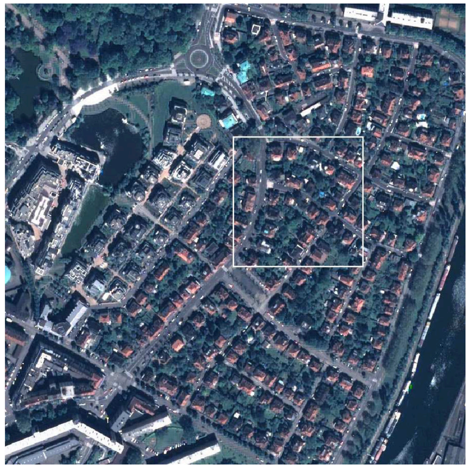


Fig. 3. Quickbird image of Strasbourg (France). The square area used to illustrate the segmentation is outlined in white.

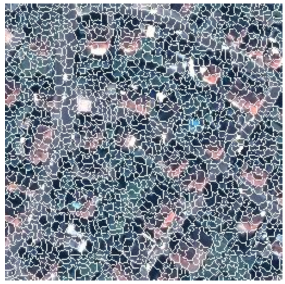
for the three indexes. We can notice that for the three cases, the two curves seem to have the same behaviour and are highly correlated. These results show that optimizing our criterion is relevant and allows to compute segmentation of quality without forcing the expert to provide examples.

Finally, the table 1 presents values of these indexes for the segmentations with the parameters found during the generations 1, 3, 5 et 11 of the genetic evolution. After the 11^{th} generation, the quality of the results does not increase significantly. The genetic algorithm has found the limit of the ontology recognition (approximatively 52% as shown on table 1). This limit can be explained in two ways : first many pixels in the image do not belong to the concepts given in the ontology (noise from the sensors, shadows, cars, etc.), and second the concepts as defined in the ontology do not match all objects in the image.

Table 1. Results of the evaluation of the method on expert objects for 4 generations.

generation	Ontology	Recall	Janssen	Feitosa
1^{st}	29.64 %	24.98 %	0.32	25.19
3^{rd}	33.92 %	27.83 %	0.35	17.55
5^{th}	37.56 %	31.74 %	0.42	6.63
11^{th}	51.91 %	49.72 %	0.48	7.10

(a) raw image

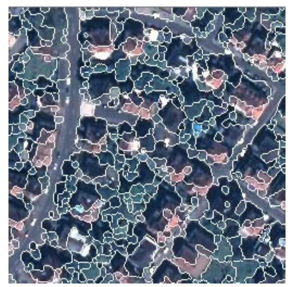


(b) 5% recognized (without parameters)

(c) 29% recognized (1st generation)

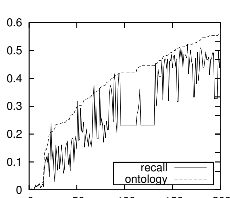
(d) 33% recognized (3rd generation)

(e) 37% recognized (5th generation)

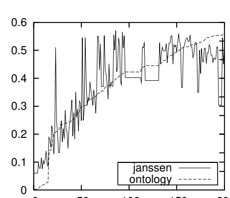


(f) 51% recognized (11th generation)

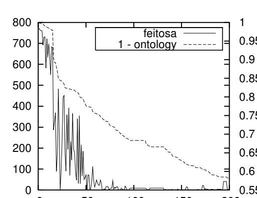
Fig. 4. Extracts of segmentations obtained at different generations during a genetic evolution. The outline of the regions is drawn in white.



(a) Recall



(b) Janssen



(c) Feitosa

Fig. 5. Evolution of evaluation functions for 200 individuals ordered by the criterion based on the ontology.

6 Conclusion

In this article, we have presented how an evolution algorithm could fill the semantic gap between meaningless parameters for a segmentation algorithm and knowledge of a domain expert. Results show the relevance of this approach. In the future, we want to introduce contextual knowledge like the position of the objects between each others. The evolutionary algorithm will be more complex and will check the constraints defined by this contextual knowledge.

References

1. Bhanu, B., Lee, S., Das, S.: Adaptive image segmentation using genetic and hybrid search methods. IEEE Transactions on Aerospace and Electronic Systems 31(4), 1268–1291 (1995)
2. Durand, N., Derivaux, S., Forestier, G., Wemmert, C., Gancarski, P., Boussaid D, O., Puissant, A.: Ontology-based object recognition for remote sensing image interpretation. In: IEEE International Conference on Tools with Artificial Intelligence. pp. 472–479. Patras, Greece (october 2007)
3. Feitosa, R.Q., Costa, G.A., Cazes, T.B., B., F.: A genetic approach for the automatic adaptation of segmentation parameters. In: International Conference on Object-based Image Analysis (2006)
4. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional (January 1989)
5. Haris, K., Efstradiadis, S.N., Maglaveras, N., Katsaggelos, A.K.: Hybrid image segmentation using watersheds and fast region merging. IEEE Transaction On Image Processing 7(12), 1684–1699 (1998)
6. Janssen, L., Molenaar, M.: Terrain objects, their dynamics and their monitoring by the integration of gis and remote sensing 33, 749–758 (1995)
7. Mueller, M., Segl, K., Kaufmann, H.: Edge- and region-based segmentation technique for the extraction of large, man-madeobjects in high-resolution satellite imagery. Pattern Recognition 37(8), 1619–1628 (2004)
8. Najman, L., Schmitt, M.: Geodesic saliency of watershed contours and hierarchical segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(12), 1163–1173 (1996)
9. Pignalberi, G., Cucchiara, R., Cinque, L., Levialdi, S.: Tuning range image segmentation by genetic algorithm. EURASIP Journal on Applied Signal Processing 2003(8), 780–790 (2003)
10. Sheeren, D., Puissant, A., Weber, C., Gancarski, P., Wemmert, C.: Deriving classification rules from multiple sensed urban data with data mining. In: 1rst Workshop of the EARSel Special Interest Group Urban Remote Sensing, Berlin (march 2006)
11. Song, A., Ciesielski, V.: Fast texture segmentation using genetic programming. In: IEEE Congress on Evolutionary Computation. vol. 3, pp. 2126–2133 (2003)
12. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Pattern Analysis and Machine Intelligence 13(6), 583–598 (1991)