Automatic Learning of Structural Knowledge from Geographic Information for updating Land Cover Maps

Meriam Bayoudh[1,2], Emmanuel Roux[1], Richard Nock[3], Gilles Richard[4]

1 ESPACE-DEV, UMR228 IRD/UMII/UR/UAG, Institut de Recherche pour le Développement, Centre IRD de Guyane, Route de Montabo BP165; 97323 Cayenne Cedex
firstname.name@ird.fr
2 Université des Antilles et de la Guyane, Cayenne, French Guiana, France
3 CEREGMIA, UAG, Schoelcher, Martinique, France, richard.nock@martinique.univ-ag.fr
4 IRIT, University of Toulouse, France, richard@irit.fr

The availability of remote sensing images increases due to the multiplication of the earth observation satellites, the improvement of the image spatial and temporal resolutions and the changes in data distribution policies. Such data quantities make the image processing and interpretation a new challenge for engineers and researchers. Therefore, new approaches have to be developed in order to automatically updating land cover/use maps that provide useful information to decision makers. The great majority of the methods devoted to satellite image supervised classification consider pixel information within the image regions associated to different classes, in order to learn class spectral signatures. Structural aspects within the pixel neighborhood are essentially considered by computing textural indexes within the same regions. Very few methods are based on structural knowledge discovery at a higher semantic level, like the one constituted by regions themselves, referred to as "objects" in the following.

Our work consists in predicting the membership of the land cover/use classes by first inducing classification rules from existing land use/cover maps and from various complementary geographic information layers. More specifically, we propose to explicitly extract structural (relational) knowledge on the objects, in space and/or time.

We propose the application of the Inductive Logic Programming (ILP). ILP is a learning approach which aims to induce rules from a set of examples and from background knowledge encoded in logic programs. The rules cover a maximum of positive examples and invalidate a maximum of negative examples. In our case, positive and negative examples are defined by the sets of objects which, respectively, belong and do not belong to the class of interest. Objects are described by attributes characterizing intrinsic features (class, area, perimeter, compactness, fractal dimension, crossed by a road/river) and relational features (inclusion, adjacency, north/south, east/west). For quantitative variables, each of the object values is re-coded by indicating whether it is below or above the 10th, 20th, ... and 90th percentiles of the empirical distribution. As land cover/use typologies usually consist in several classes, the one-vs-rest induction method is applied, providing as many classifiers as there are classes.

The methodology was applied to update a land cover maps of the French Guiana littoral. We exploited land cover maps provided by the French Forestry Commission (Office National des Forêts – ONF). Such maps are based on a 39-class typology and are related to years 2001, 2005 and 2008. The hydrological and road networks were exploited as complementary geographical information.

343 classification rules were induced. A ten-fold cross validation of each classifier provided accuracy rate between 91.1% and 100.0% and an overall multi-model accuracy equal to 98.6%. However, these values greatly overestimate the overall system performances. In fact, the sensibility values vary from 0 to 100%, 31% and 72% of the classifiers exhibiting sensibility values above 80% and 50%, respectively. As required, induced rules exhibit structural knowledge as shown with the following example:

```
class_y(Object_A, Mangrove) :- adjacent(Object_A, Object_B),
    class_y-7(Object_B, Flooded or swampy forest),     class_y-3(Object_A, Mangrove).
```

Such rule means that an object *A* belongs to the class `Mangrove` at year y if it belonged to the same class three years before and is adjacent to an object *B* which belonged to the class `Flooded or swampy forest` 7 years before. The following example shows how numerical information coding has permitted to learn value intervals:

```
class08(Object_A, Scattered built) :- class_y-3(Object_A, Scattered built),
    fractal_dimension(Object_A, ≤321), fractal_dimension(Object_A, >243).
```

Such rule means that an object *A* belongs to the class `Scattered built` if it belonged to the same class three years before and its fractal dimension is comprised between 243 and 321.

The application of the ILP to discover structural knowledge from maps provided simple and intelligible classification rules the predictive power of which appears very promising. Induced rules could be exploited to automatically classify new objects and/or help operators using object-based classification procedures.