Decision Trees to Detect Changes in Remote Sensing Image Time Series

Thales Korting, Leila Fonseca, Gilberto Câmara

National's Institute for Space Research – Image Processing Division
{tkorting, leila, gilberto}@dpi.inpe.br

Abstract. This work provides a description of a method to detect land changes, using a time series of remote sensing images. The time series is interpreted as a sequence of calculated features. A visual representation of the sequence is a curve, which behaves in time. The method uses a supervised classification algorithm of decision trees to infer about changes. Three main steps are necessary to detect changes in these images. The first is defining the change typology, performed by the analyst. The second is setting up a reference set. The last step includes mining change signatures.

1 INTRODUCTION

The Earth is constantly changing, and such changes are resultant from two main driving forces, namely nature and man. Natural forces apply changes gradually, however human made changes are abrupt. Therefore, sudden changes do not consider the impacts in the whole scenario. As stated by Boriah et al. (2008), converting natural land-cover into human-dominated cover types is a global challenge with many unknown environmental effects.

One example of land change and its effects over nature is modern agriculture. Modern land-use practices, while increasing the short-term supplies of material goods, may undermine many ecosystem services in the long run, on both regional and global scales. Worldwide changes to forests, farmlands, waterways, and air are driven by the need to provide food, water, and shelter to more than six billion people (Foley et al. 2005). It is urgent that all changes be looked into, mainly the ones that cause bad impact in the ecosystem.

Satellite observations offer new opportunities for understanding how the Earth is changing, for discovering what reasons cause these changes, and for predicting future changes. Remotely sensed data, combined with information from ecosystem models, offers an unprecedented opportunity for predicting and understanding the behaviour of the Earth's ecosystem (Tan et al., 2001). Such pool of information allows revealing complex and important patterns on applications of environmental monitoring and analysis of land-cover dynamics.

The development of effective methods for multitemporal data analysis is one of the most important and challenging issues for the RS community (Bruzzone et al., 2003). However, it is difficult to deal with Earth Science patterns, because of their spatio-temporal nature. In accord to Mota et al., (2009), identifying changes means discovering how the objects gain or lose their identity, how their properties change, as well as which changes happen simultaneously.

Remote sensing satellites provide a continuous and consistent set of information about the Earth's land and oceans, providing rich data that helps us to follow changes in our planet. An image series is a set of images taken from the same scene at different acquisition times. Therefore it is fundamental to note the spatial resolution, atmospheric conditions of every image, and registration constraints. The variation of the features (or attributes) extracted from the images defines a trajectory. Figure 1 shows an example of trajectory represented as a curve in the attribute-time space. The trajectory portion that represents a changing pattern in a specific interval defines the object *change signature*.

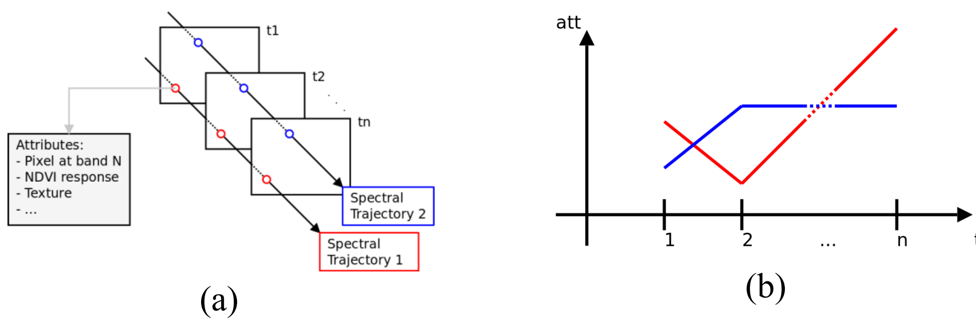

Figure 1: a) Features (attributes) defining trajectories. b) Plotting trajectories through time.

Trajectories can tell the history of certain events found in an image time series. The recovery of such history is useful for understanding of land evolution. Techniques based on data mining and information mining are feasible to deal with large amount of data. Heas and Datcu (2005) define information mining as the nontrivial process of analyzing data in the perspective of discovering implicit but potentially useful information.

In this scenario, the objective of this work is to describe a method to identify land use and cover changing patterns using data mining. A classification algorithm based on decision trees do not consider the feature's scale but find the best separability indices using multiple amplitudes. Given the changing patterns, these algorithms identify change signatures, build the classification model, and classify full image time series according to the model.

2 DECISION TREES

A decision tree is a set of conditions applied to intervals of features values. Observations satisfying the condition at each junction go to the left branch, and the others go to the right. Terminal nodes, or leaves of the tree, suit to the classes. We show a simple decision tree classifier in Figure 2, using two features to distinguish three classes.

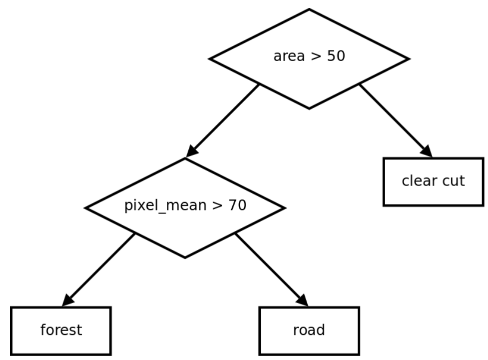

Figure 2: Example of a decision tree: two features are tested, reaching three classes.

Hastie et al. (2009) express the problem of forming a decision tree recursively, in a divide-and-conquer method. In the first stage, one feature is selected to split the data into a proper interval. It splits the training set into subsets, and goes recursively for each branch, using the remaining instances. When some branch only owns instances of a single class, it becomes a leaf.

Every node matches a noncategorical feature, and every connection means a feature interval. A leaf in the tree means the categorical value expected for all instances described by the path between the root and the leaf.

According to Quinlan (1993), three basic assumptions are the base of decision trees. Given a set T of training samples, and their matching classes $C_1, C_2, \ldots\ C_k$, there are three possibilities:

- T contains one or more samples, all belonging to a single class C_j. In this case, the decision tree for T is comprised by a leaf identifying the class C_j;
- T contains no samples. Further information is necessary to infer the class for the objects in T. The algorithm can use the overall majority class, for instance;
- T contains samples of different classes. The objective is to refine T into single-class subsets. Based on a single feature, which has one or more mutually exclusive outcomes, the algorithm chooses one test. Decision

tree for T consists on a node identifying the test, and one branch for each valid outcome. The same scheme is applied recursively to each subset of training cases.

But how to define the best feature to split the data? According to Witten and Frank (2005), if we had a measure of the purity of each node, we could choose the feature that produces the purest daughter nodes. Based on the information theory, *entropy* plays a central role as a measure of choice, information, and confusion (Shannon, 2001). The entropy H, with two possibilities with probabilities p and $q = 1 - p$, is calculated as

$$\mathrm{H}(p, q) = -(p \log p + q \log q).$$

Extending this equation to more than two probabilities, we have:

$$\mathrm{H}(p_1, p_2, \dots p_N) = -p_1 \log p_1 - p_2 \log p_2 \dots - p_N \log p_N$$

Given this measure, one can calculate the *information value* (info) for the selected features, converting the instances of classes to a set of probabilities:

$$\mathrm{info}(v_1, v_2, \dots v_N) = \mathrm{H}\left(\frac{v_1}{\mathrm{D}}, \frac{v_2}{\mathrm{D}}, \dots \frac{v_N}{\mathrm{D}}\right)$$

$$\mathrm{D} = \sum_{i=1}^{N} v_i$$

The value gain measures the advantage of using a certain feature in despite to another. The info contained in all elements inside each sample class minus the info for the number of instances that go to every branch results the gain. The feature's amplitude and statistical distribution do not influence the computation of gain. This grants an independence from data standards, and therefore becomes more flexible for classification. In multitemporal analysis, in which one can have many features, it makes this algorithm a satisfactory choice for classifying change signatures.

3 METHOD

Three main steps compose the method for detecting changes in image time series.

3.1 Defining the change typology

Through visual interpretation of the temporal series the analyst recognizes places in which changes have occurred. This task is performed by defining labels of the found changes.

There are different ways to perform visual interpretation. One is by showing the time series in a key-framing scheme, where the images are displayed one after the other. It allows the user to see when and where abrupt changes took place. Another way is to display small image crops around the interest region, side by side. The change typology can also be noticed by plotting through time certain features from the imagery, such as vegetation indices. By noticing the trajectory of certain characteristics in the feature space, it is possible to make inferences.

3.2 Building a reference Set

Since the analyst knows the typology of change, the description of change signatures is performed by selecting samples, which are representative regions over the images. This is carried out by picking regions up where the changes are apparent. Such regions, obtained from segmentation, and its features are associated to one of the defined labels.

Labelled samples are used to build a decision tree. The classifier divides the data into similar trajectories. Every group of trajectories leads to a changing pattern, therefore it is called *change signature*, the key point of our method. Signatures will be used to classify the complete data set and to make inferences about the features, which of them were the most (or less) important in the patterns discovery.

Another possible and interesting inference is the relationship between changes and objects features. Nevertheless, our method can aid the analyst to infer that certain features are more responsible for different types of changes, including natural (e.g., clouds or seas) and human made (e.g., urban growth). When interpreting a decision tree, the analyst follows the most important features and thresholds defined by the algorithm.

3.3 Mining change signatures

Given the change signatures detected by the decision tree, the algorithm classifies all segmented regions in the images. The result is a map highlighting the different changes. Therefore the analyst notices when, what and where changes happened, for example *deforestation*, *urban* and *forest growth*, or *no change*. The same model shall be tested into different data sets, to evaluate the possibilities of extending one typology of change into different areas where similar context of evolution is expected.

Large datasets are commonly used to detect changes in remote sensing data. Even applying techniques to reduce the dimensionality of data and to select best subset of features, the remaining data still present high volume. We aim to use decision trees to fight against this constraint, since it is neither affected by the volume of data, nor by the different scales of the features.

4 CONCLUSIONS

This work presented a method to detect changes in image time series. The method uses the supervised classification algorithm of decision trees, which partitions the feature space. The analyst defines expected typology of change on the time series with a sample set. The algorithm is used to infer the best splits on the feature space to classify the entire time series.

Since the features contain information of time, they can be used to infer the periods of change. And as the data contains geographical descriptors, they are also used to answer where the changes have occurred.

Future works include creating and describing new features from the series, developing different techniques for time series visualization, and extending the classification algorithm to carry out nonrigid thresholds.

5 REFERENCES

Boriah, S., V. Kumar, M. Steinbach, C. Potter, S. Klooster (2008). "Land cover change detection: a case study". ACM SIGKDD international conference on knowledge discovery and data mining. 14: 8570-865.

Bruzzone, L., P. Smits, J. Tilton (2003). "Foreword special issue on analysis of multitemporal remote sensing images". IEEE Transactions on Geoscience and Remote Sensing. 41: 2419-2422.

Foley, J. et al (2005), "Global consequences of land use". Science, New York. 309: 570-574.

Hastie, T., R. Tibshirani, J. Friedman (2001). "The Elements of Statistical Learning - Data Mining, Inference and Prediction". Springer, New York.

Heas, P., M. Datcu (2005). "Modeling trajectory of dynamic clusters in image time-series for spatio-temporal reasoning". IEEE Transactions on Geoscience and Remote Sensing. 43: 1635-1647.

Mota, J., Câmara, G., M. Escada, O. Bittencourt, L. Fonseca, L. Vinhas (2009). "Case-Based Reasoning for Eliciting the Evolution of Geospatial Objects". Conference on Spatial Information Theory.

Quinlan, J. (1993). "Programs for Machine Learning". Morgan Kaufmann, San Mateo. 302.

Shannon, C. (2001). "A Mathematical Theory of Communication". Mobile Computing and Communications Review, SIGMOBILE Mob 5.

Tan, P., M. Steinbach, V. Kumar, C. Potter, S. Klooster, A. Torregrosa (2001). "Clustering earth science data: Goals, issues and results". Fourth KDD Workshop on Mining Scientific Datasets 25.

Witten, I. and E. Frank (2002). "Data Mining: Practical Machine Learning Tools and Techniques". Morgan Kaufmann, San Francisco. 558.