New Developments and Challenges in Remote Sensing, Z. Bochenek (ed.)
©2007 Millpress, Rotterdam, ISBN 978-90-5966-053-3

A toolbox for multi-temporal analysis of satellite imagery

A. Pinheiro[1], P. Gonçalves[2], H. Carrão[1,3] & M. Caetano[1]
[1]*Portuguese Geographic Institute (IGP), Remote Sensing Unit, Lisbon, Portugal; andre.pinheiro@igeo.pt, hugo.carrao@igeo.pt, mario.caetano@igeo.pt*
[2]*INRIA Rhône-Alpes, St. Ismier, France; paulo.goncalves@inria.fr*
[3]*Statistics and Information Management Institute (ISEGI), New University of Lisbon, Lisbon, Portugal.*

Keywords: time series analysis, MODIS, land cover characterization, Support Vector Machines

ABSTRACT: In this paper we present a MATLAB toolbox that was developed for the analysis and classification of medium spatial resolution satellite imagery. Its main functionality is the time series analysis of satellite imagery for the land cover characterization. This multi-temporal assessment is of extremely high importance, since different land cover classes exhibit specific spectral reflectances as function of time and their exploration should significantly improve classification scores obtained from single date measurements.
For the classification procedure of single date or multi-temporal imagery data we implemented the Support Vector Machine (SVM) algorithm. In addition to the multi-temporal analysis and classification tasks of satellite images the presented toolbox also includes procedures for sample collection of land cover classes and flag quality analysis of image pixels.

This toolbox is the first version of an ongoing development effort that currently exploits MODIS images in TIF format. For further development, we plan to implement MERIS images.

1 INTRODUCTION

Land cover cartography has an important role in our days. The needs to map the urban expansion or to establish environment policies are two application examples that explore this kind of cartography. Traditionally, land cover maps used to be produced by visual interpretation of aerial photos, but this is a hard and time-consuming process. Nowadays, it is of extremely high importance that land cover cartography production could be fast, recent and accurate (good correspondence between the maps and reality). In this sense, during last decades, satellite imagery has been exploited to automatically extract land cover information from specific spectral reflectances of Earth's surface. The reflectances recorded on an image are related to the available bands of the sensor and to the date of capture, and the distinction between different land cover classes is dependent of such measurements. Thus, mapping of land cover often requires processing of satellite images collected at different time periods and at many spectral wavelengths (Maxwell *et al*. 2002). The main goal of exploring images time series is to register

different states of the vegetation growth that are useful to characterize the land cover classes spectral "behavior" along time.

Recently launched Earth Observation (EO) sensors, such as the *MEdium Resolution Imaging Spectrometer* (MERIS) and the M*oderate Resolution Imaging Spectroradiometer* (MODIS), exhibit enhanced spectral and temporal resolutions, as well as superior standards of calibration, georeferencing and atmospheric correction, and detailed per pixel data quality information. The automatic exploitation of these improved temporal and spectral features will, as stated before, provide a better-quality interpretation of the images and a more time-accurate cartography. However, high dimensional remote sensing imagery provides a challenge to the current classification techniques. Thus, the development of new methods to analyze high dimensional satellite imagery data became extremely necessary (Haertel and Landgrebe, 1999; Landgrebe, 2002; Shah *et al.* 2003). A new supervised classification system based on the statistical learning theory (Vapnik, 1998), defined as Support Vector Machine (SVM), has recently been applied to the problem of high dimensional remote sensing data classification (e.g., Carrão *et al*. 2006; Gonçalves *et al.* 2005; Pal and Mather, 2005; Marçal *et al.*, 2005; Mercier and Lennon, 2003; Huang *et al.* 2002; Zhu and Blumberg, 2002). The results show that SVM can obtain a better classification performance and has a more generalization capacity than classifiers that aim to minimize the training error rate alone.

An extensive research of the existing commercial products for satellite images analysis confirmed that there are not specific tools implemented for land cover classification that take advantage of temporal and spectral features of new sensors. Also, the inspected softwares do not include the SVM algorithm as standard classifier to deal with high dimensional satellite data. Some of the applications and toolboxes explored in the framework of this specific research area are listed below, but none of them exploit entirely the high dimensional satellite data currently available for land cover characterization:

- Rosario, S. *et al* (2006) – "A toolbox to perform land cover classifications with a big amount of features". It uses MATLAB standard supervised and no supervised classifiers and permits the usage of dynamic classes;
- Jönsson *et al*. (2006) – Toolbox that can be used in MATLAB software, the main goal of which is to filter time series signals, http://www.nateko.lu.se/remotesensing/;
- Cruz *et al*. (2004) – "A MATLAB Toolbox for Hyperspectral Image Analysis". It has several tools to deal with satellite images and to do image classification.

In this paper we present a toolbox developed in MATLAB software for the analysis and classification of medium spatial resolution satellite imagery. Specifically, we implemented an algebraic model to fit inter-annual reflectance time series of satellite images that can be used as an additional sustaining feature for land cover characterization (Gonçalves *et al.* 2006). Additionally, we also implemented several standard functions for satellite image analysis and digital processing, as well as the SVM supervised learning approach for automatic land cover classification. Presently, this toolbox deals with time series of MODIS medium resolution images.

2 TOOLBOX'S OVERVIEW

This toolbox was developed in the software MATLAB 7.0.1. We used the GUIDE graphical tools to build a user-friendly toolbox to be used by non-experts on satellite image analysis and land cover classification. MATLAB is an efficient software that is able to manage large datasets of satellite images as well other auxiliary information. We decided to use the MATLAB software because of its optimal performances in raster calculation. In this paper we present the toolbox characteristics and functionality.

The SVM classifier was implemented through the Spider object orientated environment for machine learning in MATLAB. The Spider description and tutorial can be found freely available at: www.kyb.tuebingen.mpg.de/bs/people/spider/tutorial.html.

The user can work with MODIS imagery from any geographic area in the world, but always respecting a specific cartographic projection that must be defined within the working project. This property is very important, mainly due to two reasons: (1) when importing land cover samples by specifying the geographical coordinates, they must match obviously the image cartographic projection; (2) when performing time series analysis of satellite imagery, the pixel area in one date must correspond geographically to the same pixel area in the other dates.

The toolbox can be divided into three main groups:

- Visualization and manipulation of satellite images: this item includes the spectral bands and vegetation indexes visualization options; mask selection based on the MODIS quality factor information; and selection of multi-temporal images list;
- Sample collection and time series analysis: this group proposes to the user two different modes for land cover samples collection to be used in the supervised classification; the information of each sample, in each band and at each date can be explored for data pre-processing and for classification tasks;
- Classification: this function is presently in an experimental phase, regarding the classifier's algorithm implementation; the classification parameters can be defined here.

3 WORKING WITH THE TOOLBOX

Figure 1 shows the main window of the toolbox. When the user is working with the toolbox, this window will be always active. All available functions and utilities are in this window, which we divided in several blocks, as shown in Figure 1, and are described below:

1 – Start a new project or import an existing one (Figure 2);
2 – Open the images and visualization options (Figure 3);
3 – Get pixel information (Figure 4);
4 – Box with some information to optimize the usage of application;

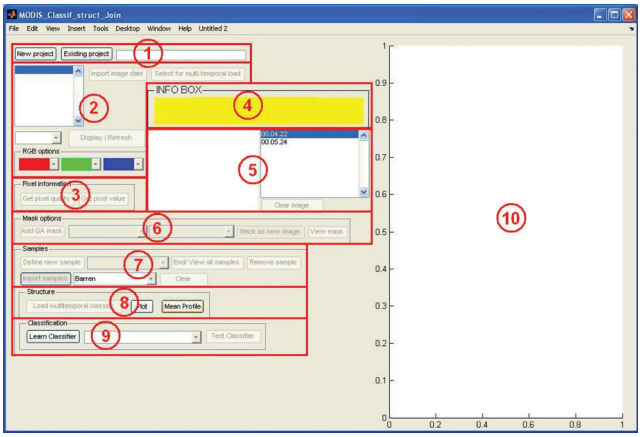

Figure 1. Toolbox's main window.

5 – Visualization of several information (pixel information, coordinates, reflectances, samples, etc);
6 – Mask option based on pixel quality flag (Figure 5);
7 – Collection of samples (Figure 6);
8 – Multi-temporal data base production (Figure 7);
9 – Classification (auxiliary tools in Figure 8 and Figure 9; classification result in Figure 11);
10 – Images and samples visualization.

Block 1 which is the beginning of the toolbox is shown in Figure 2. The user can initiate a new project or import an existing one (with the samples and multi-temporal images already defined).

When starting a new project, the first step is to import and visualize the satellite images (Figure 3). In the menu below the list of the image will appear the bands included in each image file. It is possible to visualize each band individually, two different vegetation indexes (Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI)) and all possible RGB composites. With this toolbox, the user can only read images in TIF format (double, 8 bits and 16 bits).

In the same block, the user can define the list of images to be included in the multi-temporal analysis. Selecting an image from the initial image list, the **select for multitemporal load** button will be activated. If the user clicks on this button, the selected image will be placed in the multitemporal image list (Block 5 of Figure 1). The chronological order of the selection is very important because the temporal analysis is based on the input order.

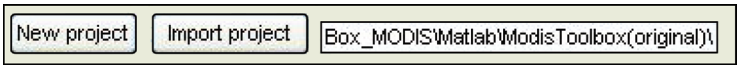

Figure 2. Block1: Start a project.

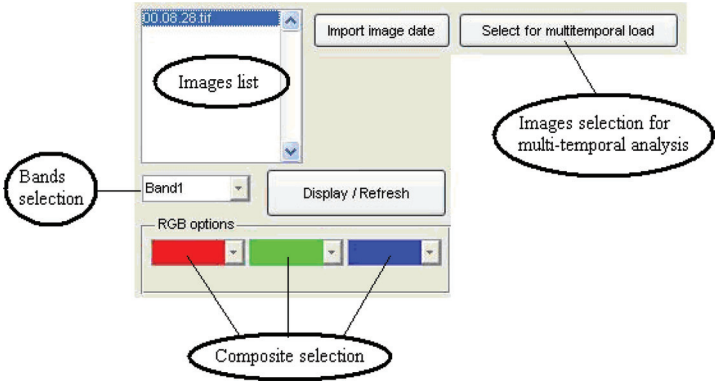

Figure 3. Block2: Managing satellite images.

Figure 4 shows the tool where the user can extract some information from a selected pixel. This tool provides two kinds of information: a quality factor that is sensor dependent (this is the example from MODIS); and reflectance values from the existing bands. When the user clicks on one of these buttons it will appear the information from a previous selected pixel. To deactivate this function, the user must click the right mouse button. Some information to guide the user will appear in the **information box** of Block 4 of Figure 1. It is recommended that the user follows the displayed instruction because they were thought to optimize the performance of the toolbox.

The mask options (Figure 5) is a pixel selection with a given criteria, which are the quality factor of MODIS images (left list in Figure 4) that can be seen individually in the Block 3 of Figure 1. The flag quality allows the user to create masks that can be used to eliminate pixels with radiometric problems or simply to monitor the geographical location of some specific

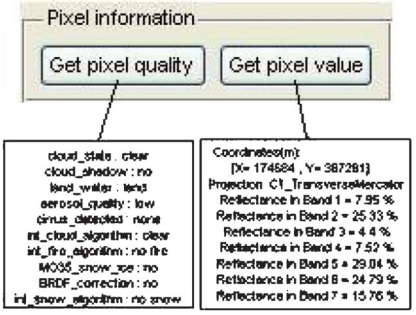

Figure 4. Block 3: Pixel information.

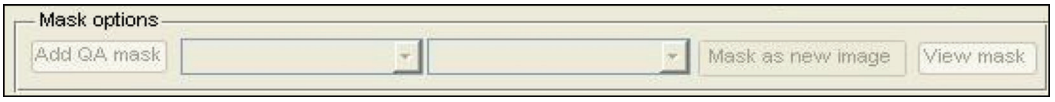

Figure 5. Block 6: Mask options.

occurrences. For example, the user can enhance pixels without clouds, pixels in land or water surfaces, etc. This information can be helpful in the samples selection.

The selection of land cover samples for supervised classification can be done by two different modes:

- On screen selection: one of the nine default classes (listed in Figure 6). The collection of these samples is made directly on the image, selecting each pixel for the default land cover class;
- Dynamic selection: the user can be defined any class that is not listed in the toolbox, by importing from an EXCEL file with the coordinates reference.

The Block 8 (Figure 7) of this toolbox is very important and is essential to the multitemporal classification task. To activate the **load multitemporal classes** button, the user must have already selected the samples and the time series images. Clicking on the **load multitemporal classes** button, the toolbox will create a MATLAB database with the value of reflectance in each band and at each date for the samples areas.

This multitemporal information by pixel can always be visualized by clicking on the **Plot** button. The **Plot** tool (Figure 8) allows the user to analyze the temporal profiles of spectral reflectance of any class in any band. This profile is defined with the existing dates and is useful to see the behavior of each class for those dates.
The user has several plot options, namely:

- The profiles can correspond to all the samples from each class, or from a sub group of samples;
- The user can choose several line types to different classes or samples;
- Instead of all samples profile, it is possible to analyze the mean class profile. However, this option is only available after using the **Mean Profile** tool (Figure 9).

The Mean Profile tool is also essential for the classification task. The main goal of this tool is to pre-process the temporal reflectances of selected samples.

Two important issues are available in this tool:

- Median calculation of spectral reflectances for each band and for each class (to remove spectral noise derived by clouds);

Figure 6. Block 7: Samples selection.

Figure 7. Block 8: Multi-temporal structure and graphical analysis.

- Interpolation of the reflectances for a given number of days not available in the imported image.

In this reflectances pre-processing, it is recommended to include all the classes and all the bands before the classification. Note that, just in case of all the reflectances are filtered and interpolated with the same parameters, it can be performed a correct multitemporal classification.

The final task implemented in this toolbox is to perform a land cover classification (Figure 10). The classification is, presently, in an experimental phase and is based in the input images and only in the selected samples. For these propose, the user has two types of classifier, and for each of these classifiers there are two options (Linear and Gaussian):

- K-Nearest neighborhood (KNN);
- Support Vector Machine (SVM)

The user can select the classes and the bands to include in the classification. Presently, the classification involves just the sample areas and not the entire image as it is intended to be implemented in the future. For the classification task, the samples are randomly divided in two subgroups (train and test). The algorithm "learns" with the multi-temporal information from the train samples and then performs a classification with the test samples. This task is performed with the cross validation method. In Figure 11, the classification result for nine classes is shown. The classes involved in the process are listed in the legend with the respective color. The axes of the graphical result represent:

- Horizontal – number of pixels per class;
- Vertical – Classes. The representation of these classes is related with the legend (by order and by color).

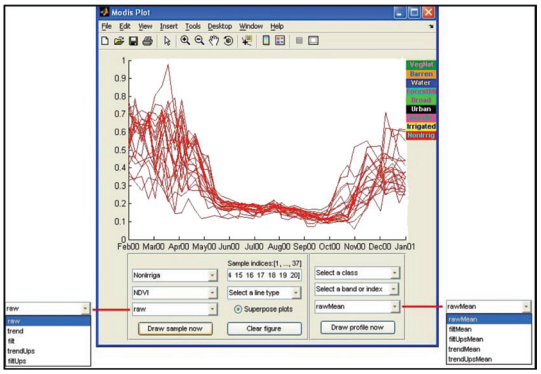

Figure 8. Plot tool: Temporal profiles visualization.

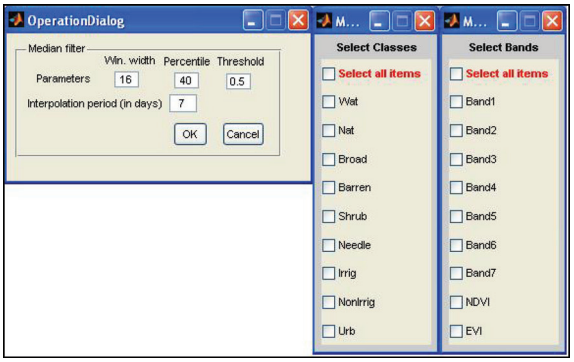

Figure 9. Mean Profile: Signal processing.

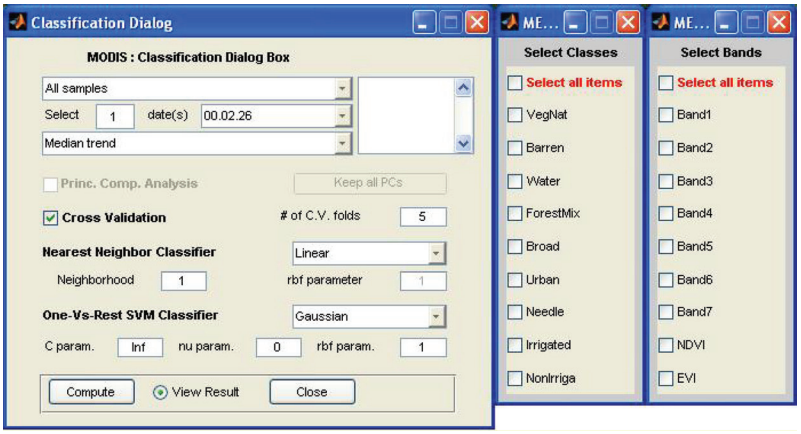

Figure 10. Classification menu.

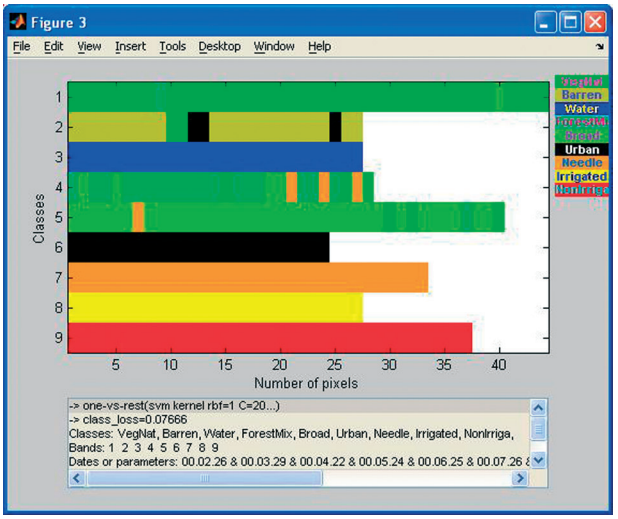

Figure 11. Classification results.

In the graphic, the user can analyze the performance of the classification. For example, in the 9th class (non-irrigated field) there are no classification errors because all the samples belonging to this class were classified as non-irrigated field. An example of classification errors is given in the 2nd class (barren soil). Some of its original pixels were classified as urban areas or as mixed forest.

In the menu below this graphic, there is some useful information, namely: classifier parameters, class loss result, classes, bands and dates involved in the classification.

4 CONCLUSION

In this paper we describe in detail a toolbox developed in MATLAB software, still in an experimental phase, which main goal is to take advantage of time series of MODIS images for land cover classification. This toolbox is user-friendly and easy to learn. It takes little practical time to start working with it and to manage all its capabilities. Only the SVM and KNN classification algorithms are implemented, but other classifiers may be easily adopted and integrated in this toolbox. The SVM performed good classification results using high dimensional satellite image datasets, and so we think that it was a valid decision to integrate this learning approach in the toolbox. Further developments rely on the classification of the entire satellite image and not just in the sample areas. Moreover, we think that this toolbox can be adopted for further and different applications, for example, to process and analyze MERIS medium resolution satellite images.

ACKNOWLEDGEMENTS

This study was carried out in the framework of the project ''LANDEO – User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data'' (PDCTE/MGS/49969/2003) funded by ''Programa Dinamizador das Ciências e Tecnologias para o Espaço'' from ''Fundação para a Ciência e Tecnologia'', and from the ''Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data'' from European Space Agency (ESA).

REFERENCES

Arzuaga-Cruz, E. Jimenez-Rodriguez, L.O. Velez-Reyes, M. Kaeli, D. Rodriguez-Diaz, E. Velazquez-Santana, H.T. Castrodad-Carrau, A. Santos-Campis, L.E. Santiago, C. (2004), ''A MATLAB Toolbox for Hyperspectral Image Analysis'', Geoscience and Remote Sensing Symposium, 2004. IGARSS '04.

Carrão, H., P. Gonçalves, and M. Caetano, 2006, MERIS Based Land Cover Characterization: A Comparative Study. Proceedings of the ASPRS 2006 Annual Conference – Prospecting for Geospatial Information Integration, Reno, USA, 1–5 May, 2006.

Gonçalves, P., Hugo Carrão, Andre Pinheiro and Mário Caetano (2005). Land cover classification with Support Vector Machine applied to MODIS imagery. *Proceedings of the 25th EARSeL Symposium*, June 6–11, 2005, Porto, Portugal.

Gonçalves, P., Hugo Carrão, and Mário Caetano (2006). Parametric model for inter-annual reflectance time series. *Proceedings of the 26th EARSeL Symposium*, May 29th – 2nd June, 2006, Warsaw, Poland (accepted for publication).

Haertel, Victor and David Landgrebe (1999). On the Classification of Classes with Nearly Equal Spectral Responses in Remote Sensing Hyperspectral Image Data. IEEE *Transactions on Geoscience and Remote Sensing*, l(37): 2374–2386.

Huang, C., Davis, L.S. and Townshend, J.R.G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4): 725–749.

Jönsson, P., Eklundh, L., ''TIMESAT'', an on going work: http://www.nateko.lu.se/remotesensing/ (consulted last time in April, 2006).

Landgrebe, David (2002). Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem, (Invited), *Special Issue of the IEEE Signal Processing Magazine*, 19(1): 17–28.

Marçal, A.R.S., Borges, J.S., Gomes, J.A., and Costa, J.F. Pinto Da (2005). Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing*, 26(7): 1347–1362.

Maxwell, S.K., Hoffer, R.M., and Chapman, P. L. (2002). AVHRR channel selection for land cover classification. *International Journal of Remote Sensing*, 23(23): 5061–5073.

Mercier, G. and M. Lennon (2003). Support vector machines for hyperspectral image classification with spectral based kernels. Proceedings of IEEE Intl. Geoscience & Remote Sensing Symposium, IGARSS 2003 Conference, 21–25 July, Toulouse, France.

Pal, M., and Mather, P.M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5): 1007–1011.

Rosario, S., ''Hyperspectral image analysis toolbox'' http://www.censsis.neu.edu/software/hyperspectral/html/hyper_help/hyper_notes.html(consulted last time in 2005).

Shah, C.A., Watanachaturaporn, P., Arora, M.K., and Varshney, P.K. (2003). Some Recent Results on Hyperspectral Image Classification. In IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data. October 27–28, 2003, NASA Goddard Spaceflight center, Greenbelt, MD.

The spider tutorial: www.kyb.tuebingen.mpg.de/bs/people/spider/tutorial.html (consulted last time in May, 2006).

Vapnik, V. (1998). *Statistical Learning Theory.* Wiley.

Zhu, G., and Blumberg, D.G. (2002). Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel. *Remote Sensing of Environment*, 80: 233–240.