CLASSIFICATION OF FOREST STRUCTURE USING VERY HIGH RESOLUTION PLEIADES IMAGE TEXTURE

Beguet B.[1,2], Chehata N.[1,3], Boukir S.[1], Guyon D.[2]

[1.] G&E Laboratory (EA 4592), IPB / University of Bordeaux,
1 Allée F. Daguin, 33607 Pessac Cedex, France.
[2.] INRA, UMR1391 ISPA, 33140 Villenave d'Ornon, France.
[3.] LISAH UMR 144, IRD, El Menzah 4, Tunis, Tunisia.

ABSTRACT

The potential of very high resolution Pléiades image texture for forest structure mapping was assessed on maritime pine stands in south-western France. A preliminary step showed that multi-linear regressions allow a reliable prediction of forest variables (such as crown diameter or tree height) from a set of features automatically selected among a huge number of Haralick texture features with various spatial parameterizations. In a second step, to assess Pléiades image texture contribution for classification, Random Forests (RF) classification was performed to discriminate four forest structure classes from recent reforestation to mature stand. Two texture feature selection strategies are compared: (1) the previous regression-based modelling using in situ tree measurements (2) the RF-variable importance using a visual photo-interpretation. Both methods produced comparable classification accuracies. The results highlight the contribution of processes automation and the need for using both Pléiades image resolutions (panchromatic and multispectral) to derive the best performing texture features.

Index Terms**—**Forest, Texture, Feature Selection, Classification, Pléiades

1. INTRODUCTION

Very High resolution (VHR) images allows a detailed description of forest structure (tree distribution and size) at the stand level, by exploiting the strong relationship between tree spatial structure and image texture when the pixel size is smaller than the tree dimension. This expectation meets the strong need for spatial forest resources inventory at the stand level and its changes due to forest management, land use or catastrophic events. Our objective is to assess the potential of Pléiades images to map the forest structure. The approach was conducted in two complementary steps:

- The first concerns the estimation of the main forest structure variables (crown diameter, stem diameter, height, density or tree spacing) from the image texture. On these bases, different forest structure classes could be considered and a pixel-based image classification could be processed in order to produce the finest possible spatial information on forest structure.
- The second aims to classify these forest structure classes when no ground measurements are available. This issue is explored by proposing an alternative methodology.

2. MATERIAL

The study covers a 80-km² site located in south-western France (Tagon–Marcheprime, between Bordeaux and the Bay of Arcachon) within the largest European maritime pine (*Pinus pinaster* Ait.) forest. It was conducted using two Pléiades image sets in panchromatic (Pan, resampled spatial resolution= 0.50m) and multispectral bands in blue, green, red and near infrared (MS, 2m-), acquired on 26 June and 8 August 2012.

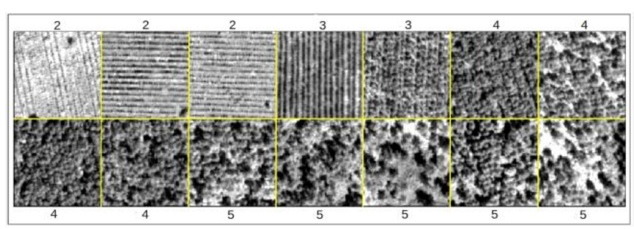

Figure 1: Forest image samples corresponding to ground measurements (ordered by structure class). Panchromatic Pléiades image, 26 June 2012. Class 2 = crown diameter < 1m, tree rows are visible ; Class 3 = crown diameter is between 1m and 2.5m, tree rows are visible ; Class 4 = crown diameter is between 2.5m and 4.5m, tree rows are no more visible; Class 5 = crown diameter > 4.5m.

Some forest structure variables were measured during summer 2012 on a large number of maritime pine stands (n=111) at the end of the tree pine growth time: tree height (varying from 1.7m to 26.2m), crown diameter (0.77m - 10.69m), diameter at breast height (0.02m - 0.56m), density (150tree/ha - 6729tree/ha) and tree spacing (1.31m - 8.77m). The trees sampled in each stand were included in a 20m × 20m square plot representative of the forest structure in the 80m × 80m area that encloses it.

For each sampled stand we considered an image sample covering a square area centered on a field measurement plot with a slightly larger width (60 meters, i.e. 120 pixels in Pan and 30 pixels in MS). Some image samples are given in figure 1 as a non-exhaustive illustration of the variability of forest structure and image texture.

3. METHODS

We used texture features derived from the co-occurrence matrices (GLCM): the eight most common Haralick's texture features [1], and also mean and variance (first order statistics). They are commonly used for forest studies [2,3,4]. An automatic modeling procedure based on multiple linear regression of each forest variable by the texture features [3] is proposed and evaluated based on image samples corresponding to the in-situ measurements. The approach consists in selecting as predictors the most efficient combination of texture features calculated in Pan (and then resampled to MS resolution) and MS bands: best R^2 and RMSE while ensuring a minimal collinearity and a balanced contribution importance [3]. This allowed us to explore a wide range of parameters of GLCM (window size, displacement and orientation) and thus to optimize them automatically.

Five forest structure classes, representative of crown dimension and spatial distribution are defined. The first class, corresponding to clear cuts and stands with trees younger than 3 years is not treated in this study (as no measurements were done on these stands, where the pine trees are less tall than non-forest species). Classes 2 to 5 are defined in figure 1.

For pixel-level classification, two feature selection methods were compared: the previous modelling-based approach using 111 stands and a Random Forest (RF) ranking using variable importance [5,6]. The latter, which could be proposed when no ground measurements are available, is based on 40 plots among 111 and their forest structure classes are recognized by visual photo-interpretation of the Pan image. Then, all texture features with various parametrizations are put into the RF classifier, the best parameterizations are selected by RF variable importance [4].

Finally, the classification is processed by Random Forests [7,8] using 100 trees for the two sets of selected features. Two independent sets of image samples are used for learning (n=80) and for test (n=80). They are also independent of the samples used for feature selection and are selected by photointerpretation of the Pan image. The classification performances of both feature selection methods are compared.

		Classification					
		2	3	4	5	Rowtotal	Commission Error %
Reference	2	14416	3174	233	177	18000	24,86
	3	1392	13115	3396	97	18000	27,14
	4	791	864	14821	1524	18000	17,66
	5	5	1014	2538	14443	18000	19,76
	Column total	16604	18167	20988	16241	72000	
	Omission Error %	13,18	27,81	28,38	11,07		
	O.A =	78,88					
	Kappa =	71,84					

Table 1: Classification confusion matrix obtained with the texture features selected by RF variable importance.

		Classification					
		2	3	4	5	Rowtotal	Commission Error %
Reference	2	13899	3587	379	135	18000	22.78
	3	2132	12495	2716	657	18000	30.58
	4	900	1200	12201	3699	18000	32.21
	5	0	444	2191	15365	18000	14.63
	Column total	16931	17726	17487	19856	72000	
	Omission Error %	17.9	29.51	30.22	22.61		
	O.A =	74.94					
	Kappa =	66.59					

Table 2: Classification confusion matrix obtained with the texture features selected by the regression-based modelling of forest variables.

4. RESULTS AND DISCUSSION

The modelling results show that the best performing regression models combine panchromatic and multispectral (different spatial and spectral resolutions) texture features. The prediction accuracies (RMSE) obtained using the five best performing texture features as predictors of forest variables are similar for both dates: ~1.1m for crown diameter, ~3m for tree height, ~0.9m for tree spacing, ~586tree/ha for density and ~0.066m for diameter at breast height. These results are very satisfactory. Therefore we can expect to discriminate several forest structure classes from image classification.

Both feature selection methods used for classification provide a subset of 50 texture features to which the four original spectral bands are added. The classification results are presented in Table 1 and Table 2. Only results from the June image are presented (they are similar to August image). The overall classification accuracy reaches 78.8% using the RF importance feature selection, improving the results by almost 4% with comparison to the modelling based feature selection. However, due to the complexity of forest structure and photo-interpretation labelling, these results are considered as similar. This indicates the potential of our methodology to identify forest structure classes when no ground measurements are available.

We notice that the larger error rates (both in commission and omission) occur in the intermediate classes (3 and 4) for both feature selection methods. These confusions can be observed in Figure 2. They highlight the complexity of forest structure classification, as the class borders are fuzzy since forest variable distributions are continuous. Figure 2 also shows that most of the classification errors appear on stand borders which is inherent to a spatial descriptor such as texture. A regularization step should enhance the classification quality and it will be investigated in order to produce an operational forest structure map.

The introduction of the class 1 (clear cuts and stands with trees younger than 3 years), which seems very different from others both in radiometry and texture, should increase the classification accuracy, and it will be done in further work.

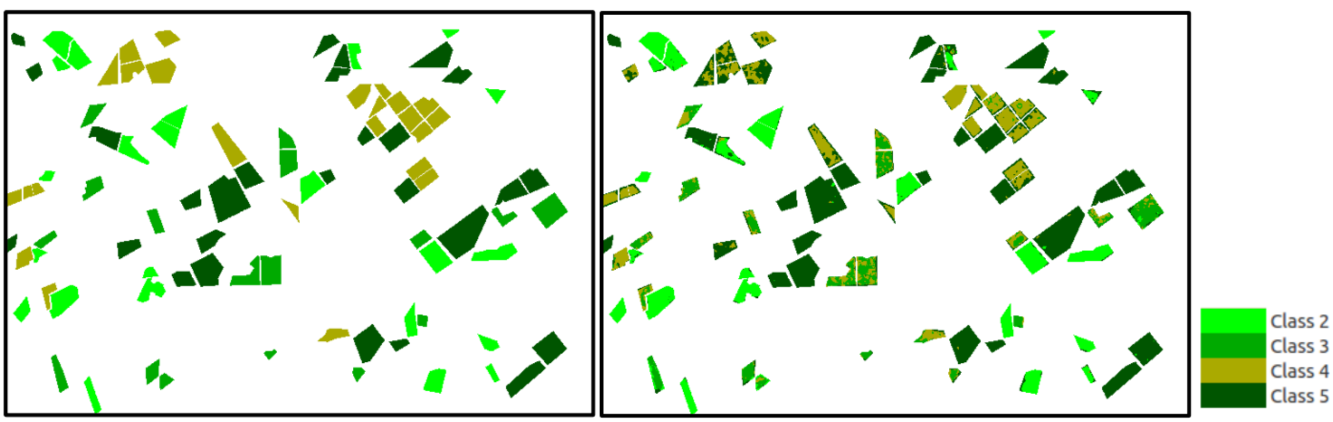

Figure 2: Reference map (left) and classification map (right) obtained with the texture features selected by RF ranking of forest variables on a part of the study area.

5. CONCLUSION

In conclusion, the results show the potential of Pléiades image texture to estimate the main forest structure variables and to map forest structure.

The proposed methodologies allow us to optimize the use of VHR Pléiades images, calculating texture features with various spatial parameterizations on both types of spatial and spectral resolution (50cm in Panchromatic, 2m in multispectral mode) before finding their best combinations for both forest variable estimation and forest structure classification.

The texture features representative of forest variables allow us to classify forest structure with a satisfying accuracy into four classes from recent reforestation (>3 years) to mature stand. When no ground measurements are available, the proposed methodology appears to be a satisfying alternative to drive a forest structure classification task.

The robustness of the proposed methodologies will be assessed using a third Pléiades image, acquired in winter (20 February 2013, i.e. with a lower sun and a dry understory vegetation). In addition, their application to multi-annual images would allow us to assess their ability to detect and map strong forest changes such as forest cuts, urban sprawl or storm damage.

6. ACKNOWLEDGEMENTS

This research was funded with grants from Conseil Régional d'Aquitaine and CNES (Centre National d'Etudes Spatiales). The Pléiades images were provided by the Orfeo thematic program set up by CNES. The work is based on the use of OrféoToolBox [9].The authors are very grateful to Bernard Issenhut (INRA, Unité expérimentale Forêt Pierroton) for the measurements in the field and the good quality of the provided forest data. Further thanks to Christian Germain from MS Lab (Bordeaux), Jordi Inglada from CESBIO (Toulouse) and Jean-Pierre Wigneron from INRA (Bordeaux) for discussions.

7. REFERENCES

[1] R.M. Haralick, K. Shanmugan and I. Dinstein, " Texture features for image classification", IEEE Transactions Systems, Man and Cybernetics, pp. 601-621, 1973.

[2] F. Kayitakire, C. Hamel and P. Defourny, " Retrieving forest structure variables based on image texture analysis and ikonos-2 imagery", Remote Sensing of Environment 102, pp. 390-401, 2006.

[3] B. Beguet, N. Chehata, S. Boukir and D. Guyon, " Retrieving forest structure variables from very high resolution satellite images using an automatic method", ISPRS Annals of Photogrammetry, Remote Sensing and the Spatial Information Sciences I-7, pp. 1-6, 2012.

[4] M. A. Castillo, M. Ricker, and B. H. J. De Jong, Estimation of tropical forest structure from spot-5 satellite images.International Journal of Remote Sensing, vol. 31-10, pp. 2767–2782, 2010.

[5] R. Genuer, J.M. Poggi, and C. Tuleau-Malot, " Variable selection using random forests", Pattern Recognition Letters vol. 31, pp. 2225-2236, 2010.

[6] Dash M., Liu H., Feature Selection for Classification, Intelligent Data Analysis, vol. 1, p. 131-156, 1997.

[7] L. Breiman, " Random forests", Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.

[8] M. Pal, Random forest classifier for remote sensing classification.
International Journal of Remote Sensing, vol. 26(1), pp. 217–222, 2005.

[9] J. Inglada and E. Christophe, The Orfeo toolbox remote sensing image processing software. IEEE International Geoscience and Remote Sensing Symposium, vol. 4, pp. IV-733–IV-736, 2009.