"FUSION FOR CLASSES IN DIFFICULTY" FOR ACCURATE AND SPEED TROPICAL RAINFORESTS CLASSIFICATION

Robin Pouteau, Benoit Stoll

"FUSION FOR CLASSES IN DIFFICULTY" FOR ACCURATE AND SPEED TROPICAL RAINFORESTS CLASSIFICATION

R. Pouteau and B. Stoll

South Pacific Geosciences (GePaSud) Laboratory
University of French Polynesia, BP 6570, 98702 Faa'a, French Polynesia
{robin.pouteau, benoit.stoll}@upf.pf

ABSTRACT

The accuracy of rainforests classification is generally improved by the input of multisensory data since complex vegetation type identification benefits from complementary information. However, in some cases, multisource fusion can also deteriorate accuracy when irrelevant sources are added. Thus, we introduce a fusion method for classes "in difficulty". Our method outperforms the classical global approach consisting in performing fusion for all classes. Moreover, the fusion processing time can significantly decrease when several classes are put aside. This operational method can be used effectively to enhance accuracy and processing speed when analyzing the wealth of information available from remote sensing products.

***Index Terms*—** Vegetation mapping, image classification, support vector machines, multisensory imagery, data fusion

1. INTRODUCTION

Today, an increasing number of sensors of greater diversity are available to the remote sensing community. Such a variety of spectral, spatial and temporal resolutions has very useful complementary properties and can therefore outperform conventional single-source approaches [1]-[9].

A range of fusion algorithms and schemes have been proposed and compared over the past two decades which highlights that multisource fusion is a key research topic. To our knowledge, the first attempt is [1] where fusion of visible Landsat MSS bands with infrared Landsat MSS bands is performed to map 11 classes in an agricultural landscape of New South Wales (Australia) using a probabilistic scheme that employs a global membership function and the Dempster's orthogonal sum combination rule. Optical Landsat MSS data and ancillary data (elevation, slope and aspect) are fused in [2] to map 10 classes in a montane forest of Colorado (USA) using the minimum Euclidean distance, the maximum likelihood classifier (MLC) and the minimum Mahalanobis distance. The first comparative fusion of both optical and SAR data is probably [3] which used optical Daedalus 1268 ATM data with "PLC-band, fully polarimetric NASA/JPL SAR sensor" data to map 6 classes in an agricultural landscape of Feltwell (UK) using structured-Neural Networks (NN), fully connected-NN and the *k*-nearest neighbors. Then, [4] fused optical Landsat TM data with ERS-1 SAR data to map 12 classes in an anthropogenic area of Lisbon (Portugal) using MLC, NN, the majority voting and the logarithmic opinion pool. In [5], optical Landsat TM data and ERS-1 SAR data are fused to map 16 classes in an agro-forest landscape of Gothenburg (Sweden) using MLC, NN and the sequential maximum *a posteriori*. Next, [6] fused optical Landsat TM data with ancillary data (elevation, slope and distance to water body) to map 14 classes in an agricultural landscape of Oklahoma (USA) using both C4.5 algorithm and support vector machines (SVM). Two optical IKONOS images are merged in [7] to map 6 urban units in Reykjavik (Iceland) using NN and a fuzzy decision rule. The ability of MLC, DT, "boosted-DT" and SVM to fuse optical Landsat-5 TM and SPOT-5 data with Envisat ASAR and ERS-2 data is compared in [8] over an agricultural landscape of Bonn (Germany) where 8 classes occur. More recently, [9] fused optical SPOT-2 data with ALOS/PALSAR data to map 6 classes in an urban/peri-urban area of Hochiminh (Vietnam) using MLC and SVM.

The previously mentioned fusion algorithms comparative studies mainly focus on simple anthropogenic structures. According to the results of these comparisons, the best contemporary fusion algorithm is arguably SVM. Although the classification scheme proposed by [8] is quite simple to implement and adapted to classification of different nature data, its main drawback is that fusion is performed for all classes globally. Indeed, for some classes, multisource fusion can also deteriorate accuracy found in monosource when a non-relevant source is used [1]. This paper aims to assess, over a structurally complex model, an extension of this method: the "fusion for classes in difficulty".

2. MATERIAL AND METHODS

2.1. Study site and ground data collection

This study focuses on tropical rainforests which is a subject of great interest to scientists around the world. In witness

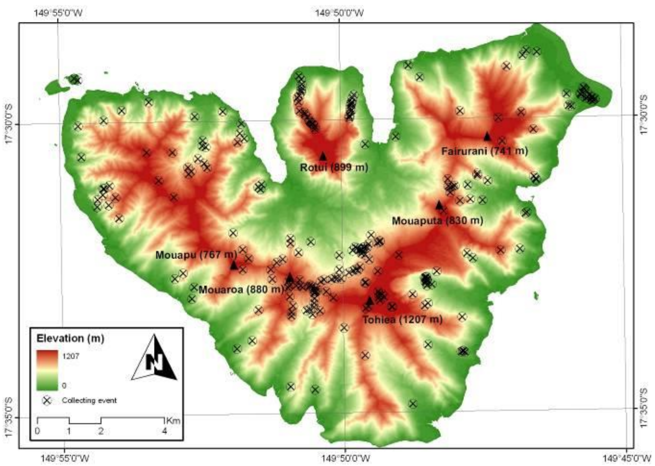


Figure 1. Presentation of the Moorea island and localization of the 255 collecting points.

thereof, United Nations General Assembly declares 2011 as the International Year of Forests. Here, we argue that multisource image fusion is critical for classifying complex structures since each complementary source can contribute to the classification success. Optical, infrared, SAR, digital elevation model (DEM) and multitemporal data can therefore be useful for species identification according to their physico-chemical, anatomical, structural, ecological and phenological properties respectively.

The present study is conducted in French Polynesia (South Pacific) and more precisely on the island of Moorea (140 km² with a highest summit reaching 1,207 m) (Figure 1) where 17 vegetation types occur.

Fifteen 125 m² circular regions of interest are selected and geolocalized with a handheld Trimble® GeoXH™ GPS for each vegetation type (*ca.* 1/2,000 of the island area is sampled in total). Half of this area is used for classification training and half for validation. Balance data sets are systematically used to avoid under- or over-representation problems [10].

2.1 Remotely sensed data

The experiment is carried out on three data types, namely an optical source, a SAR source and a digital elevation model (DEM), projected in the WGS 84 – UTM 6 South coordinate system (Figure 2).

2.2.1. Optical spectral data

For the first image type, we selected a four-bands and 0.60 m-resolution Quickbird scene from November 9, 2006. It is orthorectified using the cubic convolution approximation technique, more suitable than nearest neighbor and bilinear interpolation techniques [11]. The near infrared band is useful for vegetation studies [12] and very high spatial resolution is critical for plant species discrimination [13], [14] using texture metrics for example.

2.2.2. Optical textural data

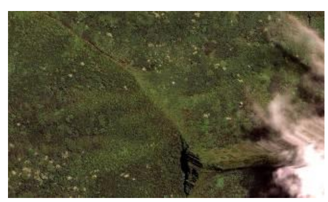
Eight gray-level co-occurrence matrix (GLCM) texture metrics are extracted from these data: mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation [15]. They are calculated on the 4 bands and in 3x3, 9x9 and 15x15 pixels window sizes. To prevent classification from the Hughes phenomenon, we select the most relevant band and window size by calculating the mean Jeffries-Matusita separability for each combination. Separability may be an adapted metric when using SVM since they do not aim to describe classes as conventional approaches but to separate them [16]. The winning combination is the texture calculated on the green band in a 15x15 pixels window. Since GLCM texture metrics have a strongly different nature from the spectral information, they are considered as a separated source as in [6].

2.2.3. SAR data

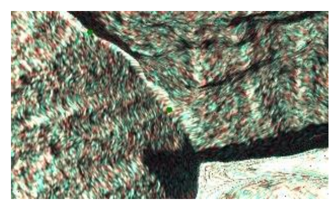
For the second image type, two 2.75 m-resolution StripMap TerraSAR-X ©DLR (2010) acquisitions were programmed over Moorea on April 30, 2010 in VV-VH polarizations and on August 28, 2010 in HH-HV polarizations. The scenes are geometrically corrected using a 5 m-resolution DEM. Speckle noise is reduced by using the enhanced Frost filter in a 7x7 pixels window, showing the best mean Jeffries-Matusita separability after several tests.

2.2.4. DEM-extracted data

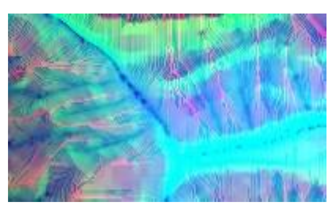
The third data set is the DEM produced from a



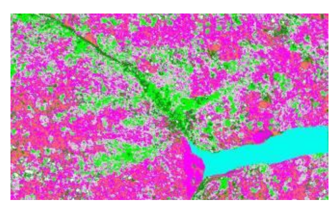
Quickbird
R: Red, G: Green, B: Blue



TerraSAR-X ©DLR (2010)
R: Xvv, G: Xhv, B: Xhh



DEM
R: WI, G: Slope, B: Elevation



Fusion for classes in difficulty
Colour correspondence in table 2

Figure 2. Composite illustrations of the source images and classification with the "fusion for classes in difficulty" method on a subscene centred on the mont Tamarutoofa.

Table 1. Accuracies (%) achieved by SVM using different sources (OA refers to the overall accuracy based on the mean of accuracy of each class).

Sources	OA	Kappa
1. Optical spectral data	62.7	60.5
2. Optical textural data	50.3	47.3
3. SAR data	36.7	32.7
4. DEM-extracted data	64.2	62.0
All with method in [8]	**67.0**	**64.9**
All with "fusion for classes in difficulty"	**78.2**	**76.9**

photogrammetric restitution at a scale of 1/5000 based on aerial photography from 1997 at a scale of 1/15000. With a 5 m-resolution, it enables extraction of topographical variables typically impacting plant distribution in montane ecosystems: elevation, slope, aspect, windwardness and a wetness index (WI) [17]. The latter was used as an index of water drainage with low WI values representing convex positions like mountain crests and with high WI values representing concave positions like coves or hillslope bases. It is a function of the slope angle β (in radians) and the specific catchment area (*As*) expressed as m² per unit width orthogonal to the flow direction (1).

$$Wetness\ index = \ln\left(\frac{As}{\tan\beta}\right) \qquad (1)$$

2.3. Fusion of support vector machines

Machine learning algorithms such as SVM have the advantage to be non-parametric and to be able to weight heterogeneous sources according to their relevance [2]. SVM is introduced by [18] and extensively described by [19]-[21]. It is arguably one of the most successful algorithms for multisource fusion [6], [8], [9]. SVM consists in projecting vectors into a high dimension feature space by means of a kernel function then fitting an optimal hyperplane that separates classes using an optimization function.

We compare the following two fusion schemes based on SVM (Figure 2):

- method in [8]: a single SVM is trained on each source separately and a rule image is generated for each class from each source. Then an additional SVM is trained on all rule images to perform the fusion;

- "fusion for classes in difficulty": this fusion method starts from the observation that, for some classes, multisource fusion can also deteriorate accuracy found in monosource when a non-relevant source is added [1]. Thus, the general principle is the same but fusion is performed only when no single source is able to classify satisfactorily a class or a set of classes, *i.e.* when the condition (2) is fulfilled:

$$\min(PA_{source\ i};UA_{source\ i}) < \min(PA_{fusion};UA_{fusion}) \qquad (2)$$

wherein $i \in [1,4]$, PA is the producer accuracies and UA the user accuracies. Should the opposite occurs, the class is not considered as "in difficulty", the spatial distribution of the considered class is the one found in the most accurate single-source classification (the source having the best min(PA;UA)) and the class is expelled from fusion. If a pixel belongs to several classes with this process, the class with the best min(PA;UA) wins.

3. RESULTS

Regarding fusion methods, accuracies are improved when multiple sources are used for classification (Table 1). The experimental results clearly show the positive impact of complementary multisensory imagery for forest classification, especially with "fusion for classes in difficulty". With the latter, 7 classes are considered as "in difficulty" and the 10 other classes are classified from a single adapted source (Table 2). For example, DEM-extracted data are adapted to classify the coastal and high-

Table 2. Vegetation types occurring on Moorea and source from which the classification is based (F refers to "fusion for classes in difficulty"; source # refers to table 1).

Classes (colour in figure 2)	**Source**	**PA**	**UA**
Plantations			
Pinus caribaea	1	72.1	67.1
Falcataria moluccana (coral)	1	79.0	46.1
Cocos nucifera	F	90.4	90.4
Coastal vegetation			
Typha domingensis	4	100	100
Low- to mid-elevation mesic to moist vegetation			
Metrosideros collina (dark green)	F	83.1	86.3
Casuarina equisetifolia	1	90.1	91.0
Dicranopteris linearis	1	98.2	93.0
Leucaena leucocephala	F	80.3	79.2
Syzygium cumini	F	73.9	60.7
Miscanthus floridulus	1	86.0	88.4
Low- to mid-elevation moist to wet vegetation			
Neonauclea forsteri (magenta)	F	61.7	41.3
Aleurites moluccana (white)	2	92.5	86.5
Inocarpus fagifer	F	50.0	39.4
High-elevation vegetation			
Montane cloud forest (cyan)	4	95.3	96.4
Summit shrubland	4	98.9	94.8
Ubiquitous			
Hibiscus tiliaceus (green)	3	24.6	47.3
Spathodea campanulata	F	55.2	70.0

elevation vegetation whereas fusion performs worse. Indeed the addition of non-relevant sources acts as a bias in the classification process. "Fusion for classes in difficulty" main advantage is to not be too global and to select classes which can benefit from multiple information sources.

The fusion processing time has a quadratic relation with the number of classes considered for SVM fusion. As a result, by limiting the number of classes in the fusion, processing time can be significantly reduced. For example, with Q=7 (for the "fusion for classes in difficulty"), fusion processing time is *ca.* a sixth of the computational time with Q=17 (for the method in [8]).

4. CONCLUSION

Two fusion schemes based on SVM were compared for classification of optical, SAR and ancillary data on a structurally complex tropical rainforest. We introduced an operational method consisting in fusing these data only for classes being "in difficulty". Our method outperformed the classical global approach in term of accuracy and reduced by a factor 6 the fusion processing time in our study case.

AKNOLEDGEMENT

The authors are grateful to the Government of French Polynesia, its Urbanism Department for providing the DEM data and its Research Department for funding this project. We are also grateful to the German Aerospace Center (DLR), German Remote Sensing Data Center (DFD) and their TerraSAR-X Science program for providing the SAR Data. We also would like to thank Marie Fourdrigniez (BioConsulting), Jean-Yves Meyer (Research Department of the Government of French Polynesia), Ravahere Taputuarai (MaNature) for their collaboration and the Moorea Biocode Project for logistic support.

REFERENCES

[1] T. Lee, J.A. Richards, and P.H. Swain, "Probabilistic and evidential approaches for multisource data analysis," *IEEE Trans. Geosci. Remote Sens*., vol. 25, no. 3, pp. 283-293, 1987.

[2] J.A. Benediktsson, P.H. Swain, and O.K. Ersoy, "Neural network approaches versus statistical methods in classification of multisource remote sensing data," *IEEE Trans. Geosci. Remote Sens*., vol. 28, no. 4, pp. 540-552, 1990.

[3] S.B. Serpico and F. Roli, "Classification of multisensory remote-sensing images by structured neural networks," *IEEE Trans. Geosci. Remote Sens*., vol. 33, no. 3, pp. 562-578, 1995.

[4] J.A. Benediktsson and I. Kanellopoulos, "Classification of multisource and hypersperctral data based on decision fusion," *IEEE Trans. Geosci. Remote Sens*., vol. 37, no. 3, pp. 1367-1377, 1999.

[5] D.B. Michelson, B.M. Liljeberg, and P. Pilesjo, "Comparison of algorithms for classifying Swedish landcover using Landsat TM and ERS-1 SAR data," *Remote Sens. Environ.*, vol. 71, no. 1, pp. 1-15, 2000.

[6] X. Song, G. Fan, and M. Rao, "Automatic CRP mapping using nonparametric machine learning approaches," *IEEE Trans. Geosci. Remote Sens*., vol. 43, no. 4, pp. 888-897, 2005.

[7] M. Fauvel, J. Chanussot, and J.A. Benediktsson, "Decision fusion for the classification of urban remote sensing images," *IEEE Trans. Geosci. Remote Sens*., vol. 44, no. 10, pp. 2828-2838, 2006.

[8] B. Waske and J.A. Benediktsson, "Fusion of support vector machines for classification of multisensory data," *IEEE Trans. Geosci. Remote Sens*., vol. 45, no. 12, pp. 3858-3866, 2007.

[9] H.T. Chu and L. Ge, "Synergistic use of multi-temporal ALOS/PALSAR with SPOT multispectral satellite imagery for land cover mapping in the Ho Chi Minh city area, Vietnam," *Proc. IGARSS*, pp. 1465-1468, 2010.

[10] B. Waske, J.A. Benediktsson, and J.R. Sveinsson, "Classifying remote sensing data with support vector machines and imbalanced training data," In *Multiple Classifiers Systems*, Heidelberg, Springer Berlin, pp. 375-384, 2009.

[11] F. Arif, M. Akbar, and A.-M. Wu, "Projection method for geometric modeling of high resolution satellite images applying different approximations," *PSIVT*, pp. 421-432, 2006.

[12] J.H. Boureau, *Manuel d'interprétation des photographies aériennes infrarouges. Application aux milieux forestiers et naturels*. Nogent-sur-Vernisson, France: Institut Forestier National, 2008.

[13] Y. Xie, Z. Sha, and M. Yu, "Remote sensing imagery in vegetation mapping: a review," *J. Plant Ecology*, vol. 1, no. 1, pp. 9-23, 2008.

[14] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger, "Remote sensing for biodiversity science and conservation," *TRENDS in Ecology and Evolution*, vol. 18, no. 6, pp. 306-314, 2003.

[15] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Systems Man Cybernet*., vol. 3, pp. 610-621, 1973.

[16] G.M. Foody and A Mathur, "The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM," *Remote Sens. Environ.*, vol. 103, no. 1, pp. 179-189, 2006.

[17] P.E. Gessler, O.A. Chadwick, F. Chamran, L. Althouse, and K. Holmes, "Modeling soil-landscape and ecosystem properties using terrain attributes," *Soil Sci. Soc. Am. J*., vol. 64, pp. 2046-2056, 2000.

[18] V. Vapnik, *Statistical learning theory*. New York: Wiley, 1998.

[19] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition," Data Min. Knowl. Disc., vol. 2, no. 2, pp. 121-167, Jun. 1998.

[20] B. Schölkopf and A. Smola, *Learning with kernels*. Cambridge, MA: MIT Press, 2002.

[21] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Technical report, Department of Computer Science & Information Engineering, National Taiwan Univ., Taiwan, 2009.