HAL
open science

Classification of MODIS Time Series with Dense Bag-of-Temporal-SIFT-Words: Application to Cropland Mapping in the Brazilian Amazon

Adeline Bailly, Damien Arvor, Laetitia Chapel, Romain Tavenard

To cite this version:

Adeline Bailly, Damien Arvor, Laetitia Chapel, Romain Tavenard. Classification of MODIS Time Series with Dense Bag-of-Temporal-SIFT-Words: Application to Cropland Mapping in the Brazilian Amazon. IEEE International Geoscience and Remote Sensing Symposium, Jul 2016, Beijing, China. 10.1109/IGARSS.2016.7729594 . halshs-01343211

HAL Id: halshs-01343211

https://shs.hal.science/halshs-01343211

Submitted on 7 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CLASSIFICATION OF MODIS TIME SERIES WITH DENSE BAG-OF-TEMPORAL-SIFT-WORDS: APPLICATION TO CROPLAND MAPPING IN THE BRAZILIAN AMAZON

Adeline Bailly[1,2], *Damien Arvor*[1], *Laetitia Chapel*[2], *Romain Tavenard*[1]

[1]Université de Rennes 2, LETG-Rennes COSTEL, Rennes, France

[2]Université de Bretagne-Sud, IRISA, Vannes, France

ABSTRACT

Mapping croplands is a challenging problem in a context of climate change and evolving agricultural calendars. Classification based on MODIS vegetation index time series is performed in order to map crop types in the Brazilian state of Mato Grosso. We used the recently developed Dense Bag-of-Temporal-SIFT-Words algorithm, which is able to capture temporal locality of the data. It allows the accurate detection of around 70% of the agricultural areas. It leads to better classification rates than a baseline algorithm, discriminating more accurately classes with similar profiles.

***Index Terms*—** Time Series Classification, MODIS, Vegetation Index, D-BoTSW

1. INTRODUCTION

Mapping croplands is of primary interest to estimate cultivated areas and agricultural production or to qualify the agricultural practices and their potential impacts on the environment. This assessment is especially true for tropical areas. Indeed, these regions which concentrate major global natural resources are expected to supply a large proportion of increasing production demand at global scale. Since the early 2000s, the high temporal resolution of the MODIS sensor onboard the TERRA and AQUA satellites allows the monitoring of phenological cycles in order to classify vegetation types. Such approach appears to be especially relevant in tropical areas where vegetation index time series enable one to limit the constraints related to high cloud cover rates. In the Brazilian Amazon, many studies based on MODIS vegetation index time series were carried out to map croplands and crop types. Different approaches were successfully proposed to process MODIS time series including wavelet analysis [1], feature extraction [2] usually associated to a classifier, *e.g.* decision tree classifiers [3, 4] or maximum likelihood [5]. However, in a context of climate change and crop expansion to other Brazilian states (Maranho, Piaui, Tocantins and Bahia), agricultural calendars are expected to evolve so that it is necessary to implement new methodologies to process time series. Such methodologies should 1) be independent from dates to only focus on time series profiles and 2) consider the entire time series not to loose information during feature extraction.

In this paper, we investigate the use of the recently proposed Dense Bag-of-Temporal-SIFT-Words algorithm for the classification of Brazilian Amazon crops using MODIS time series. This method exhibits the aforementioned desirable properties and has been shown to significantly outperform most state-of-the-art time series classification algorithms [6], while being computationally efficient.

2. STUDY AREA AND DATASET

2.1. Study area

The study area is located in the state of Mato Grosso, in the southern Brazilian Amazon (Fig. 1). This state has suffered dramatic land use changes since the 1970s due to the rapid progress of an agricultural frontier. This frontier was the result of efficient public policies to encourage people to settle the large and plane areas of native savannas and forests. It led to the establishment of a powerful agricultural sector dedicated to the production of commodities such as soybean, maize or cotton in large-scale farms. Until the mid-2000s, the agricultural sector was especially criticized for its severe impacts on forests since Mato Grosso has long been known for its high deforestation rates. Nonetheless, deforestation decreased for 10 years and the historical extensive agriculture model is being replaced by intensive agriculture practices [7]. For example, double cropping systems have been generalized in Mato Grosso (from 35 to 62% of the net cropped area [8]) and especially concern the cultivation of soybean followed by maize, millet or cotton. In this case, soybean is sown between late September and late November and cultivated until early January and late March. When soybean is harvested, maize and cotton is sown to be harvested between late May (maize) and mid-July (cotton). Although the average agricultural calendar is well known in this region, slight differences may occur at regional scale due to rainfall variability [9] or logistics issues. Indeed, for largest farms (up to 35,000 ha), the sowing

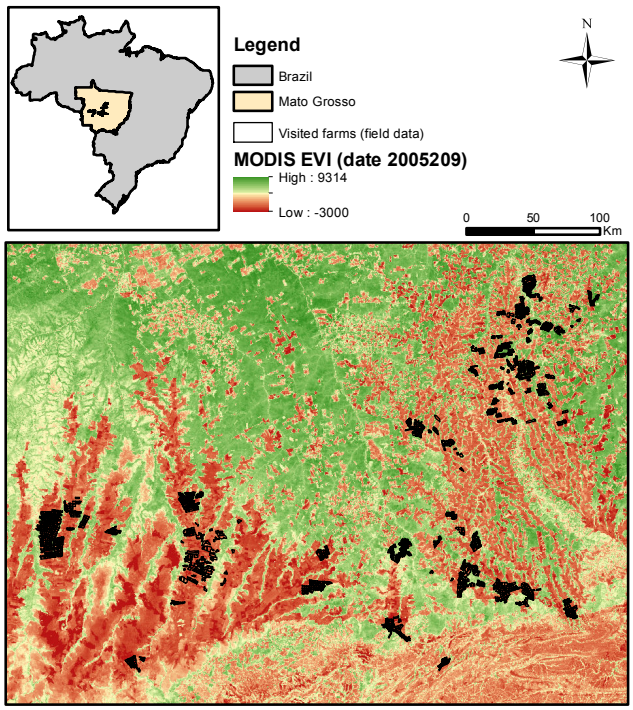

Fig. 1: Map of the study area with field data location.

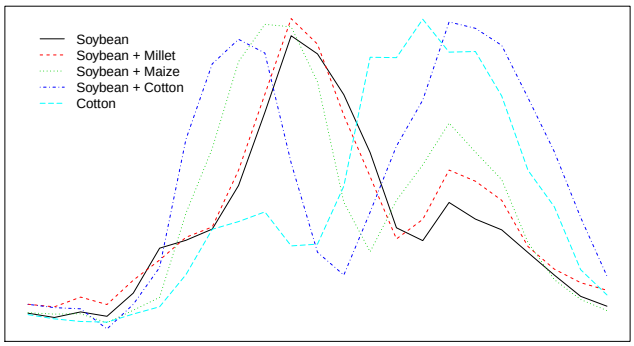

Fig. 2: Mean EVI time series of the 5 crop classes.

or harvesting period can last up to 3 months. This is important to emphasize since it implies a large temporal variability in vegetation index time series for pixels of a same crop class thus affecting any attempt to classify these classes.

2.2. Dataset

We acquired two years of MODIS vegetation index time series corresponding to the cropping periods 2005-2006 and 2006-2007. The dataset is thus made of 46 MODIS images from July 2005 (DOY 209) to July 2007 (DOY193), i.e. 23 MODIS images per year. This study period was chosen to match with field data in order to enable statistical validation of the method. The vegetation index time series refer to Enhanced Vegetation Index (EVI) of the MOD13Q1 product (tile h12v10). The EVI is usually preferred to the Normalized Difference Vegetation Index (NDVI) because it saturates less than NDVI and it is less sensitive to atmospheric conditions [10]. When missing data occurs (*e.g.* due to meteorological conditions), we performed linear extrapolation using closest available data points before and after the missing data. The field data used for validation is the same as in [5]. It is based on an extensive field campaign carried out in 2007 to collect validation data about crop type, crop yield and sowing and harvesting dates for two harvests (2006 and 2007). Seventy-six farms were visited and mapped in 13 municipalities from the two main agricultural regions in Mato Grosso, i.e. along the BR163 road and the Chapada dos Parecis. Five crop classes were identified: 1) soybean, 2) soybean + millet, 3) soybean + maize, 4) soybean + cotton and 5) cotton. Classes 1 and 5 refer to single cropping practices whereas classes 2, 3 and 4 refer to double cropping practices. For each class, 500 pixels were randomly selected each year. Thus, we built two datasets (each of them composed of 2500 pixels described by time series of length 23) on two different years in order to assess the method's robustness by training and validating on different years. Fig. 2 shows per-class average EVI time series; note that there exists intra-class temporal variability.

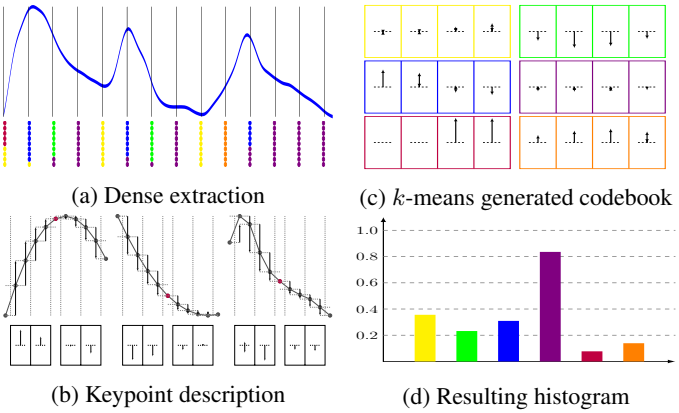

Fig. 3: D-BoTSW overview (from [6]): (a) A time series and its dense-extracted keypoints. (b) Keypoint description is based on the time series filtered at the scale at which the keypoint is extracted. Descriptors are quantized into words. (c) Codewords obtained *via* k-means, the color is associated with the dots under each keypoint in (a). (d) Histograms of word occurrences are finally given to a classifier.

3. METHOD

Dense Bag-of-Temporal-SIFT-Words (D-BoTSW) consists in describing local[1] patches of a given time series and then gath-

[1] In the following, "local" will refer to temporal locality, not spatial one.

Parameter	Possible values	Optimal value
# of blocks n_b	$\{2, 4, 6, 8\}$	8
block size a	$\{2, 4, 6, 8\}$	2
# of clusters k	$\{2^5, 2^6, \cdots, 2^{10}\}$	1024
C_{SVM}	$\{0.1, 1, 10\}$	1

Table 1: D-BoTSW parameter sets. Please refer to [12] for a detailed description of these parameters.

ering local representations into a global one using a Bag-of-Words (BoW) approach. An overview of this process is presented in Fig. 3. We choose the widely-used local descriptor SIFT [11] as the local feature in our setup. It consists in describing gradients around interest points in images and has been adapted for time series classification [12]. SIFT features are extracted every τ_{step} time instants at all SIFT scales, leading to a highly redundant representation that enables to fully capture all local behaviors in the time series. At that point, each time series is hence described by a set of local SIFT features. This set of features is transformed into a BoW (*i.e.* a k-dimensional histogram) through k-means quantization. Histograms representing time series are finally given to a standard classifier (here we use a linear SVM) in order to learn how to discriminate classes. The resulting description of the time series is then independent from the dates (only local temporal information is kept in SIFT features), while retaining information from the whole time series.

4. EXPERIMENTS

We compare D-BoTSW with a baseline classifier that takes raw data as input. Those classifiers are often shown to be accurate with low computational complexity, but are not able to take into account the temporal distortions within the time series of the same class. Since the classification step of D-BoTSW is performed using a linear SVM, our baseline is chosen to be a linear SVM taking EVI time series as 23-dimensional input. We use the publicly available D-BoTSW source code[2] and the R package kernlab [13] for running the experiments. For both methods, parameters are learned through cross-validation on the training set (see Table 1 for an overview of the range of possible values for the D-BoTSW parameters) and the performances are then evaluated on the test set. Parameters a and n_b that represent the size and the number of block used for keypoint description respectively, vary in a set of small values due to the small length of the time series. Parameter k that corresponds to the number of words generated by the k-means, needs to be small enough to avoid over-fitting and large enough to precisely represent the different kinds of local neighborhoods.

Both methods allow the accurate identification of the crops of around 70% of the test set (Table 2). D-BoTSW exhibits the best performances, as it is able to tackle temporal distortion within each class. Classes soybean, soybean + millet and classes soybean + millet and soybean + maize have similar EVI profiles (as shown in Fig. 2) and are then particularly difficult to differentiate (Table 3).

[2] https://github.com/a-bailly/dbotsw

Method	Cohen's Kappa	Accuracy
D-BoTSW	**0.645**	**0.716**
Linear SVM	0.624	0.699

Table 2: Cohen's Kappa and classification accuracy for D-BoTSW and Linear SVM.

D-BoTSW and linear SVM performs similarly to differentiate soybean from soybean + millet. It is due to the fact that vegetation starts growing after the soybean harvest until the end of the rainy season, leading to a new (but smaller) increase, close to the EVI profiles obtained in presence of millet. D-BoTSW allows for better identification of soybean + maize and soybean + cotton. Indeed, soybean + maize has a similar profile than soybean + millet and D-BoTSW is able to catch more accurately the small magnitude difference present at the end of the time series. The same happens with class soybean + cotton which is better discriminated from class soybean + maize; as well as class cotton from class soybean + cotton with D-BoTSW.

5. CONCLUSION

MODIS vegetation index time series are used to identify crop types in the state of Mato Grosso, Brazil. Results confirm the ability of the recently proposed Dense Bag-of-Temporal-SIFT-Words time series classification algorithm to differentiate between the different crop types, with better performances on classes with similar EVI profiles than a baseline algorithm. Future works include the classification of crops from other states in Brazil, in which the use of D-BoTSW is particularly relevant as time series are even more likely to present time shifts due to climate variability. Moreover, further development should be carried out in order to explicitly incorporate time information in the D-BoTSW representation without sacrificing the good invariance properties of D-BoTSW.

6. ACKNOWLEDGMENTS

This work has been partly funded by ANR project ASTERIX (ANR-13-JS02-0005-01) and Région Bretagne.

7. REFERENCES

[1] Gillian L. Galford, John F. Mustard, Jerry Melillo, Aline Gendrin, Carlos C. Cerri, and Carlos E.P. Cerri,

Actual classes \ Predicted classes	soybean	soybean + millet	soybean + maize	soybean + cotton	cotton	
soybean	**297**	118	26	25	34	0.594
soybean + millet	136	**271**	83	8	2	0.542
soybean + maize	35	82	**357**	22	4	0.714
soybean + cotton	1	5	46	**425**	23	0.850
cotton	9	5	10	37	**439**	0.878

Actual classes \ Predicted classes	soybean	soybean + millet	soybean + maize	soybean + cotton	cotton	
soybean	**327**	114	20	3	36	0.654
soybean + millet	143	**298**	51	3	5	0.596
soybean + maize	39	152	**286**	18	5	0.572
soybean + cotton	5	14	69	**388**	24	0.776
cotton	13	14	6	18	**449**	0.898

Table 3: Confusion Matrices obtained with D-BoTSW (left) and a linear SVM (right).

"Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 576–587, Feb. 2008.

[2] Anibal Gusso, Damien Arvor, Jorge Ricardo Ducati, Mauricio Roberto Veronez, and Luiz Gonzaga da Silveira, "Assessing the MODIS Crop Detection Algorithm for Soybean Crop Area Mapping and Expansion in the Mato Grosso State, Brazil," *The Scientific World Journal*, vol. 2014, pp. 1–9, 2014.

[3] Douglas C. Morton, Ruth S. DeFries, Yosio E. Shimabukuro, Liana O. Anderson, Egidio Arai, Fernando del Bon Espirito-Santo, Ramon Freitas, and Jeff Morisette, "Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon," *Proceedings of the National Academy of Sciences*, vol. 103, no. 39, pp. 14637–14641, Sept. 2006.

[4] J. Christopher Brown, Jude H. Kastens, Alexandre Camargo Coutinho, Daniel de Castro Victoria, and Christopher R. Bishop, "Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data," *Remote Sensing of Environment*, vol. 130, pp. 39–50, Mar. 2013.

[5] Damien Arvor, Milton Jonathan, Margareth Simes Penello Meirelles, Vincent Dubreuil, and Laurent Durieux, "Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil," *International Journal of Remote Sensing*, vol. 32, no. 22, pp. 7847–7871, Nov. 2011.

[6] Adeline Bailly, Simon Malinowski, Romain Tavenard, Thomas Guyet, and Laetitia Chapel, "Dense Bag-of-Temporal-SIFT-Words for Time Series Classification," Tech. Rep., Jan. 2016.

[7] Marcia N. Macedo, Ruth S. DeFries, Douglas C. Morton, Claudia M. Stickler, Gillian L. Galford, and Yosio E. Shimabukuro, "Decoupling of deforestation and soy production in the southern Amazon during the late 2000s," *Proceedings of the National Academy of Sciences*, vol. 109, no. 4, pp. 1341–1346, Jan. 2012.

[8] Damien Arvor, Margareth Meirelles, Vincent Dubreuil, Agns Bgu, and Yosio E. Shimabukuro, "Analyzing the agricultural transition in mato grosso, brazil, using satellite-derived indices," *Applied Geography*, vol. 32, no. 2, pp. 702 – 713, 2012.

[9] Damien Arvor, Vincent Dubreuil, Josyane Ronchail, Margareth Simes, and Beatriz M. Funatsu, "Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil): Spatial patterns of rainfall regimes in Mato Grosso," *International Journal of Climatology*, vol. 34, no. 8, pp. 2622–2633, June 2014.

[10] Alfredo Huete, Kamel Didan, Tomoaki Miura, E. Patricia Rodriguez, Xiang Gao, and Laerte G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sensing of Environment*, vol. 83, no. 1-2, pp. 195–213, 2002.

[11] David G. Lowe, "Object Recognition from Local Scale-Invariant Features," in *International Conference on Computer Vision*, 1999, pp. 1150–1157.

[12] Adeline Bailly, Simon Malinowski, Romain Tavenard, Thomas Guyet, and Laetitia Chapel, "Bag-of-Temporal-SIFT-Words for Time Series Classification," *ECML-PKDD Workshop on Advanced Analytics and Learning on Temporal Data*, 2015.

[13] Alexandros Karatzoglou, Alex Smola, Kurt Hornik, and Achim Zeileis, "kernlab – an S4 package for kernel methods in R," *Journal of Statistical Software*, vol. 11, no. 9, pp. 1–20, 2004.