TOWARDS EFFICIENT SATELLITE IMAGE TIME SERIES ANALYSIS: COMBINATION OF DYNAMIC TIME WARPING AND QUASI-FLAT ZONES

Jonathan Weber[†]*, François Petitjean*[‡] *and Pierre Gançarski*[‡]

[†] Université de Lorraine, LORIA, Vandœuvre-lès-Nancy, F-54506, France
[‡] University of Strasbourg, LSIIT, UMR 7005, Illkirch, F-67412, France

ABSTRACT

Satellite Image Time Series (SITS, for short) are useful resources for Earth monitoring. Upcoming satellites will provide a global coverage of the Earth's surface with a short revisit time (five days); a huge amount of data to analyze will be produced. In order to be able to analyze efficiently and accurately these images, new methods have to be designed. In this article, we propose to combine a spatio-temporal segmentation pre-processing method – quasi-flat zones, which have been recently extended to video analysis – and the distortion power of DTW to simplify the representation of the SITS, in order to reduce both the time and the memory consumption. Experiments carried out on a series of 46 images show that the memory consumption can be reduced by an order of magnitude without reducing the relevance of the analysis.

***Index Terms*—** Remote Sensing, Crops, Time series analysis, Image segmentation, Image classification.

I. INTRODUCTION

SATELLITE Image Time Series (SITS, for short) are useful resources for Earth monitoring. They currently have either high temporal resolution (SPOT-VEGETATION, MODIS) or high spatial resolution (LANDSAT). The ESA's SENTINEL program will soon provide both high temporal and high spatial resolution SITS. Current satellites as the Taiwanese FORMOSAT-2 already provide such data but without a global coverage of Earth's surface and with only four spectral bands.

These new satellites, such as SENTINEL-2 (global cover every five days with 10 m to 60 m resolution and 13 spectral bands), will produce an important mass of data. As a consequence, new methods for SITS analysis, both efficient and accurate, have to be developed. Dealing with such a data deluge is actually a common scientific challenge [1].

Our research group focuses on the comparison of radiometric time series. The similarity measure is the key tool of many classification algorithms. In this way, having a similarity measure between radiometric time series makes it possible to easily provide a temporal analysis of the sensed scene. We have recently introduced the Dynamic Time Warping (DTW) similarity measure to the remote sensing community [2]–[4], which makes it possible to consistently compare radiometric series with different lengths and sampling. This ability allows us to simply remove cloud-covered values from the series – without removing cloudy images – and to capture temporal distortions or irregularity between two series representing the same class of evolution. Actually, when several images per month are available – and therefore required for new applications – classical multi-temporal methods suffer from low regularity of the sampling at the scale of days. Therefore, only methods which deal with irregular temporal sampling are able to fully exploit the available acquisitions.

This article proposes to further study this measure and aims at addressing the computation issues raised by the amount of data that will be produced by upcoming satellites. This article relies on two key ideas. First, the use of a spatio-temporal segmentation method in order to reduce the amount of memory used for the storage of the satellite image time series. The spatio-temporal segmentation of the SITS reduces the vocabulary of data description. The segmentation method used is spatio-temporal and only slightly reduces the informativity of the data, but allows us to index the average radiometric levels of each region. The memory consumption for the analysis is thus sharply improved. However, handling spatio-temporal regions is quite difficult, since the representation of atomic geographic areas (x, y) is quite heterogeneous through the dataset: the radiometric series of evolution are packed. We propose to address this issue by using the ability of DTW to consistently compare time series of different lengths and sampling. The underlying idea is that DTW makes it possible to "unpack" the compacted values in a consistent way (*i.e.*, only when it makes sense).

Taking the most of the recent temporal extension of quasi-flat zones for region-based analysis [5] and of the DTW's ability to distort time series, we demonstrate that both runtime and memory issues can be addressed.

II. SPATIO-TEMPORAL QUASI-FLAT ZONES

The size reduction of the satellite image time series is achieved by spatio-temporal quasi-flat zones which have been successfully applied to video segmentation [5]. Quasi-flat zones are less constrained extension of connected components which are sets of adjacent pixels with *same* intensity values. Quasi-flat zones are sets of adjacent pixels with *close* intensity values. Multiple definitions have been proposed (see [6] for a complete review). Here, we use the quasi-flat zones definition using α and ω, where α is the local difference threshold and ω the global difference threshold. This means that, contrary to connected components, we accept between adjacent pixels of the same quasi-flat zones, an intensity value difference less or equal to α, and between all the pixels of the same quasi-flat zones, an intensity value difference less or equal to ω. The spatio-temporal extension of this method consists of two steps. First, we produce purely spatial quasi-flat zones. Then, we produce temporal quasi-flat zones on previous spatial quasi-flat zones to obtain spatio-temporal quasi-flat zones. This method takes into account both spatial and temporal information but avoid the side-effects of computing the SITS as a volume – which induces a under-segmentation. This leads to the partition of a satellite image time series in spatio-temporal intensity homogeneous regions, which induces a reduction of the initial series.

III. DYNAMIC TIME WARPING

When studying radiometric evolutions of sensed areas over time, the core of the process generally consists of comparing data in order to estimate (dis)similarity, whatever the method is. The distance tool provides an estimation of this similarity. It is a critical tool, on which results of analysis methods heavily rely. When the data is temporal, the choice of the distance is crucial since it completely defines the way of tackling the temporality of the data. This work focuses on the *Dynamic Time Warping* similarity measure introduced in [7]. This similarity measure can exploit the temporal distortions and compare shifted or distorted evolution profiles and whose time sampling is irregular, thanks to the optimal alignment of radiometric profiles.

DTW is able to find optimal global alignment between sequences and is probably the most commonly used measure to quantify the dissimilarity between sequences. It also provides an overall real number that quantifies the similarity between the two sequences. An example of two sequences aligned by DTW is presented in Figure 1.

DTW makes it possible to find the best global alignment between two numerical sequences. The cost of the optimal alignment between two sequences A and B – of respective lengths S and T – can be recursively computed by:

$$D(A_i, B_j) = \delta(a_i, b_j) + \min \begin{cases} D(\ A_{i-1}\ ,\ B_{j-1}\ ), \\ D(\ \ A_i\ \ ,\ B_{j-1}\ ), \\ D(\ A_{i-1}\ ,\ \ B_j\ \ ) \end{cases} \quad (1)$$

where A_i is the sub-sequence $\langle a_1, \ldots, a_i \rangle$. The overall similarity is given by $D(A_S, B_T)$. Providing the cost of this alignment, DTW is generally used as a dissimilarity measure

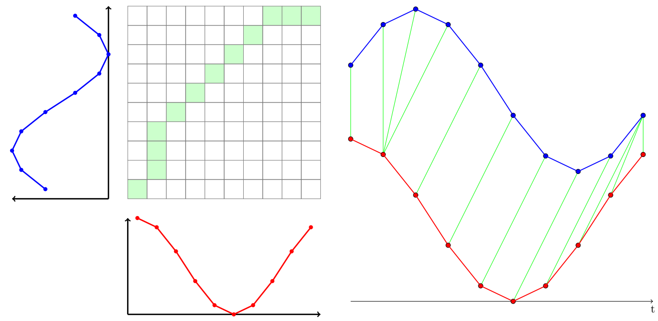


Fig. 1. Two 1D-sequences aligned with DTW. On the left: computation matrix of DTW – the warping path is depicted in green. On the right: resulting alignment of the sequences. Coordinates of the blue and red sequences have been, respectively, computed by $\cos(t)$ and $\cos(t + \alpha)$. For visualization purposes, the top sequence is drawn vertically shifted.

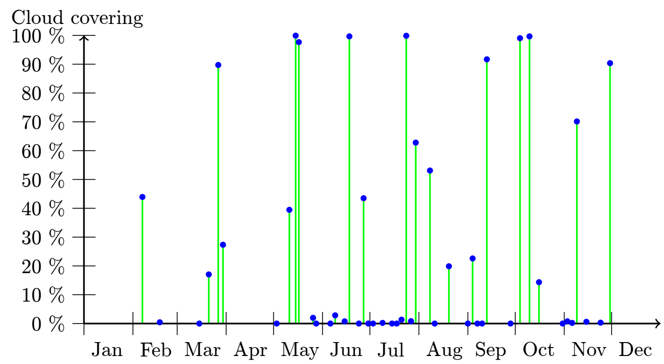


Fig. 2. Temporal distribution and cloud covering of the images from the studied SITS. Each spot represents an acquired image.

between two sequences. This measure enables to bring closer two sequences with time shifts and, more generally, time distortions. In this way, Dynamic Time Warping is a time-designed similarity measure, with several useful properties:

- gathering locally time-distorted sequences (*e.g.*, time shifts, local time distortions);
- being able to compare sequences with different lengths.

DTW has proven its accuracy for the analysis of SITS [2], [8].

IV. MATERIALS

The area of study for this work is located near the town of Toulouse in the South West of France. 46 FORMOSAT-2 images sensed over one cultural years is used; the temporal distribution and the cloud covering of these images is given in Figure 2. From these images, we use the multi-spectral product at a spatial resolution of $8\,\mathrm{m}$ and only the three bands Near-Infrared, Red and Green are kept, since the blue channel gives little information about vegetation and is very sensitive to atmospheric artifacts. Before being used in this work, the FORMOSAT-2 products have been orthorectified (guaranteeing that a pixel (x, y) covers the same geographic area throughout the image series) and the digital counts provided by the sensors are converted into physical magnitude. Moreover, we have a land cover reference map produced by the method described in [9] and using a comprehensive ground reference data set. Also, cloud masks are produced using the cloud screening procedure described in [10].

Then, each sequence – identified by its coordinates (x, y) – corresponds to a series of tuples (NIR,R,G) on which the temporal segmentation method as a pre-processing for the mining step.

V. EXPERIMENTS

These experiments aim at studying influence of the segmentation parameters (α and ω) in terms of (1) the memory consumption, (2) the results' quality and (3) the run-time. α and ω are set to the same value, which is a standard procedure [6]. The obtained results are presented in Table I. It can be noticed that the use of low values for α and ω (*e.g.*, 50 or 60), which produces an over-segmentation of the SITS, makes it possible to sharply reduce the memory consumption of the analysis. Furthermore, the segmentation step tends to smooth the radiometric heterogeneity and to capture spatial similarities, resulting in the improvement of the classification quality, compared to the classic DTW analysis. In addition, the overall computation time is slightly reduced, whatever the segmentation parameters are. Results obtained with α and ω set to 140 are also quite interesting, since they provide a classification with the same quality as the reference one, while significantly reducing the memory consumption compared to the configuration with α and ω set to 50. Finally, the quality remains quite stable with the increase of α and ω, which further assesses the relevance of the proposed methodology.

The map corresponding to one of the results is depicted in Figure 3(a). It can be noticed that the clusters extracted are spatially regular and connected which gives a first quality trend of this result. Moreover, the orange (resp. yellow) class corresponding to corn (resp. wheat) crop fields, as well as the dark green class corresponding to hardwoods, are well separated. However, the grassland (light green) and the waste land (gray) classes are gathered in the same cluster, since they are radiometrically close over time. Finally, the soybean class (red) is also quite mixed up with the sunflower one (purple) ; a similar comment can be made between the rape class (pink) and the wheat one.

$\alpha = \omega$	Kappa %	F-Measure %	Memory in MB	Run-time in hh:mm
#reference	*23.1*	*28.8*	*758*	*1:41*
50	**23.5**	**29.2**	**142**	**1:26**
60	**23.3**	**29.2**	**128**	**1:24**
70	21.4	27.0	117	1:23
80	21.5	27.2	109	1:22
90	21.3	27.0	103	1:21
100	21.6	27.3	98	1:21
110	22.6	28.4	95	1:21
120	22.6	28.4	92	1:21
130	21.1	27.2	89	1:21
140	**23.1**	**29.2**	**87**	**1:21**
150	21.2	27.2	86	1:20
160	22.1	28.0	85	1:20
170	22.7	28.7	84	1:20
180	20.9	27.0	84	1:20
190	21.5	27.4	84	1:20
200	21.1	26.9	84	1:21
210	19.8	25.7	81	1:20
220	21.8	27.7	83	1:20
230	20.2	26.4	80	1:20
240	18.9	25.1	80	1:20
250	16.0	22.0	80	1:20
260	20.7	26.6	84	1:20
270	22.5	28.7	80	1:20
280	20.8	27.1	79	1:20
290	12.8	18.9	79	1:20
300	14.3	20.5	77	1:21

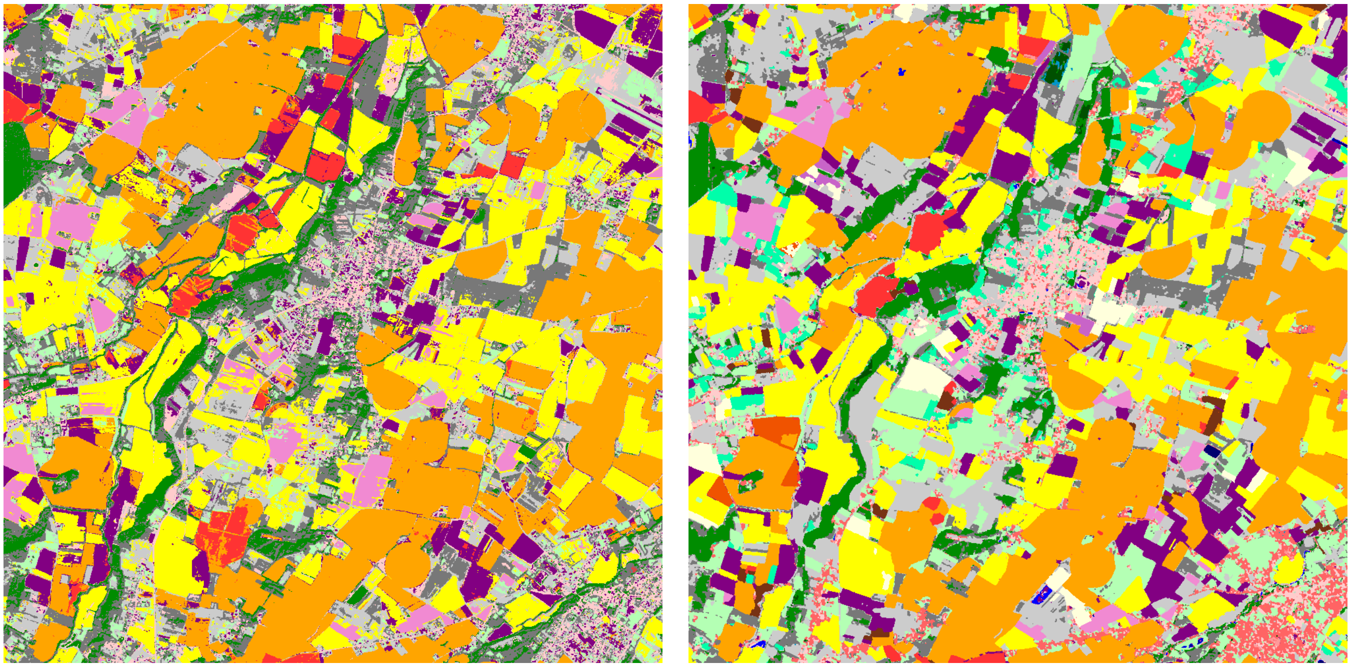
Table I. Influence of the segmentation parameters. The results can be compared to the reference ones obtained with the DTW similarity measure without a segmentation step.

VI. CONCLUSION

This paper showed how the Dynamic Time Warping similarity measure can be used with a segmentation step in order to lower both the memory and the time consumption for SITS analysis. We showed that the combination of quasi-flat zones segmentation and DTW makes it possible not only to address this issue, but to improve the quality of the temporal classification. We believe this work opens up a number of research directions. One of these directions can be the use of the extension of quasi-flat zones to interactive segmentation [11], in order to further improve the quality of the corresponding step.

ACKNOWLEDGMENTS

The authors would like to thank the Foster project for partially supporting this work as well as the French Space Agency (CNES) and Thales Alenia Space for supporting F. Petitjean's work under research contract n°1520011594. The authors would also like to thank the colleagues from CES-BIO (Danielle Ducrot, Claire Marais-Sicre, Olivier Hagolle



(a) (b)

Fig. 3. (a) Clustering result obtained with $\alpha = \omega = 50$ on the SITS. This map has been recolored according to the land cover reference map (b).

and Mireille Huc) for providing the land-cover maps and the geometrically and radiometrically corrected FORMOSAT-2 images.

VII. REFERENCES

[1] G. Bell, T. Hey, and A. Szalay, "Beyond the data deluge," *Science*, vol. 323, no. 5919, pp. 1297–1298, 2009.

[2] F. Petitjean, J. Inglada, and P. Gançarski, "Satellite Image Time Series Analysis under Time Warping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, Aug. 2012.

[3] ——, "Temporal Domain Adaptation under Time Warping," in *IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 3578–3581.

[4] F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for Dynamic Time Warping, with applications to clustering," *Pattern Recognition*, vol. 44, no. 3, pp. 678–693, Mar. 2011.

[5] J. Weber, S. Lefevre, and P. Gançarski, "Spatio-temporal quasi-flat zones for morphological video segmentation," in *10th International Symposium on Mathematical Morphology*, ser. Lecture Notes in Computer Science, vol. 6671. Springer, Jul. 2011, pp. 178–189.

[6] P. Soille, "Constrained connectivity for hierarchical image partitioning and simplification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1132–1145, Jul. 2008.

[7] H. Sakoe and S. Chiba, "A dynamic programming approach to continuous speech recognition," in *Proceedings of the Seventh International Congress on Acoustics*, vol. 3, 1971, pp. 65–69.

[8] F. Petitjean, J. Inglada, and P. Gançarski, "Clustering of Satellite Image Time Series under Time Warping," in *IEEE International Workshop on the Analysis of Multi-Temporal Remote Sensing Images*, 2011, pp. 69–72.

[9] S. Idbraim, D. Ducrot, D. Mammass, and D. Aboutajdine, "An unsupervised classification using a novel ICM method with constraints for land cover mapping from remote sensing imagery," *International Review on Computers and Software*, vol. 4, no. 2, pp. 165–176, March 2009.

[10] O. Hagolle, M. Huc, D. V. Pascual, and G. Dedieu, "A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images," *Remote Sensing of Environment*, vol. 114, no. 8, pp. 1747–1755, 2010.

[11] J. Weber, S. Lefèvre, and P. Gançarski, "Interactive video segmentation based on quasi-flat zones," in *International Symposium on Image and Signal Processing and Analysis*, Sep. 2011, pp. 265–270.