

MULTI-TEMPORAL OPTICAL VHR IMAGE FUSION FOR LAND-COVER MAPPING

Mathias Paget, Adrien Gressin, Clément Mallet

Université Paris-Est,IGN, SRIG, MATIS, 73 avenue de Paris, 94160 Saint Mandé, France
mathias.paget@ifsttar.fr, {adrien.gressin,clement.mallet}@ign.fr

ABSTRACT

Land-Cover databases (LC-DB) are very useful for environmental purposes, but need to be semantically detailed to provide robust and instructive spatial indicators. Moreover, remote sensed data allow to cover large areas with high temporal resolution. Such multi-temporal data are very useful input to discriminate LC classes. Nevertheless, automatic fusion method need to be developed to provide high quality LC-DB. In this paper, several fusion methods are proposed and introduced in an existing Land-Cover mapping framework. Those fusion methods allow to take advantage of multi-temporal data. Those methods are compared, and assessed thanks to a very high resolution LC-DB.

Index Terms— Remote sensing, change detection, land cover, satellite imagery.

1. INTRODUCTION

An increasing number of sensors and remote sensing (RS) data are available. In 2014, more than 80 civilian Earth observation mission are active and 12 satellites are launched every year [1]. Such recent sensors acquire images with increasing spectral, spatial and, in particular, temporal resolution. Multi-temporal data analysis offer new opportunities and challenging issues. Moreover, Land-Cover geodatabases (LC-DB) are inevitable inputs in various domains, such as for environmental monitoring or public policy planning. The users' needs increase, both in term of accuracy (semantically and geometrically), exhaustiveness and actuality. Thus, automatic methods are being designed to update such LC-DB. In this context, many studies have already shown that the fusion of multi-temporal images allow to improve Land-Cover classification both in terms of accuracy and semantics [2, 3]. For instance,

new classes can be discriminated by stacking multiple temporal cues or analysing the spectral profile through the time for the same location

In this paper, we propose and compare three different adaptations of an existing LC-DB updating framework [4], allowing to deal with such multi-temporal datasets. Namely, it consists in two feature selection strategies, and one class-level classification fusion method. Experimentations are performed on a large-scale LC-DB (composed of 5 classes and with a sub-metric planimetric resolution), covered by a multi-temporal Rapid-Eye acquisition.

2. METHODOLOGY

A hierarchical classification staking framework to update land-cover DB has been introduced in [4]. In this framework, three inspection levels exist: (1) a training pixel selection at the object level, (2) a classification fusion at the class level and (3) a decision is taken at the DB level. In this paper, three different adaptations are proposed. In all these strategies, the same features are extracted from each available image. The feature set is composed of 163 attributes describing spectral, textural and geometrical information for each pixel and its neighborhood.

First strategy consists in a overall feature selection (Figure 1-a). All available features are stacked into a unique feature set ($163 \times n$, where n corresponds to the number of available images), fed into the per-class selection step of the initial method. Only the 20 best features of each class are introduced in later classifications. Features are selected according to its associated importance from the Random Forest classifier. Then, the framework is computed as in the initial version. Unlike in the first one, in the second strategy (Figure 1-b), one

per-class feature selection is performed on each feature set from each image (corresponding to a specific epoch). Thus, for each image, the 20 most relevant features are selected, similarly to the first strategy. All selected features from all dates are stacked on one vector, and a second selection step is done on this features set. Finally, only the 20 best features are kept, and the method is also computed as in the initial framework. In the last strategy (Figure 1-c), the feature extraction step, the selection step and all the per object-classification step are carried out separately on each available image, resulting on one classification per object of the DB *and per image*. Thus, the fusion step takes into account all previously generated label maps in order to compute a single confidence map per class. The output of the process is again the same as in the initial framework. Finally, in order to compare the different proposed strategies to the initial framework, we apply the latter one, separately on each available image, and we select the most accurate classification results (Figure 1-d).

3. DATASETS

An area of interest near Tarbes, in Southern France, has been selected for the availability of both LC-DB and RS data described below, and for its geographic configuration: a peri-urban area covered by forest, grassland, and different urban landscapes (residential, industrial, ...). In this paper, a simplified LC-DB derived from the French National Reference LC-DB is generated. The 5 main classes have been extracted (namely *buildings, road and parking lots, water, forest and grassland*), and will be processed in our experiments.

The second part of the dataset is composed of 3 Rapid-Eye images. Each image is composed of five spectral bands (red, green, blue, red-edge and near infrared) at a resolution of 5 m, and was acquired between April and September 2011. In particular, we assume that the LC-DB and all RS datasets are correctly registered, even if the proposed framework is able to deal with small geometrical discrepancies.

4. VISUAL STUDY

Then, a visual study of a part of the image containing the 4 classes (namely *buildings, road and parking lots, forest and grassland*) have been conducted. The results are summarized in Figure 3, with the RapidEye image (acquired in September)

corresponding to this area (Figure 3a).

About the class *building*, the *overall selection method* seems to give more accurate results: there is less confusion with neighboring areas of buildings (mixed pixels) than other methods mergers (Figure 3b, 3d and 3e). However, the class *building* is always confused with the class *road*. Similar remarks can be made on the class *road*.

The different methods give close results on the *forest* class. One can notice that the *best epoch* (Figure 3j) and the *per-date selection* method (Figure 3l) have, firstly, a greater confidence in forest areas (white areas to the right of the image) and, secondly, allow to discriminate hedges (on the left side of the image).

Finally, Figure 3n to 3q show that all the proposed methods fail to discriminate the class *grassland* from the classes *building* and *road*. This is due to the presence of urban parkland in the class *grassland*.

5. RESULTS

In our results, the overall accuracy for the different scenarios ranges from 0.54 to 0.66 (*cf.* Tab. 1). Counter-intuitively, single images provide better results than multi-temporal datasets. Choosing the best classification (*best epoch*) allows to discard lower quality single date classification stemming from other epochs. Finally, the fusion methods fails to catch the confusion introduced by these two additional dates. However, the current dataset is only composed of 3 images, which does not allow to really take advantage of such multi-temporal datasets.

Fusion	Overall Accuracy
<i>Best epoch</i>	0.66
<i>Overall selection</i>	0.54
<i>Per-date selection</i>	0.55
<i>Class-level fusion</i>	0.57

Table 1: Overall-accuracy comparison obtained by the different methods of multi-temporal data fusion.

Details of by class recall and accuracy, for each fusion method, are given in Table 2. Firstly, using one (best) epoch give the better overall accuracy, but to not give always the best by class recall and accuracy. For example, only the *grassland* class obtain a better accuracy and classes *road* and *grassland*

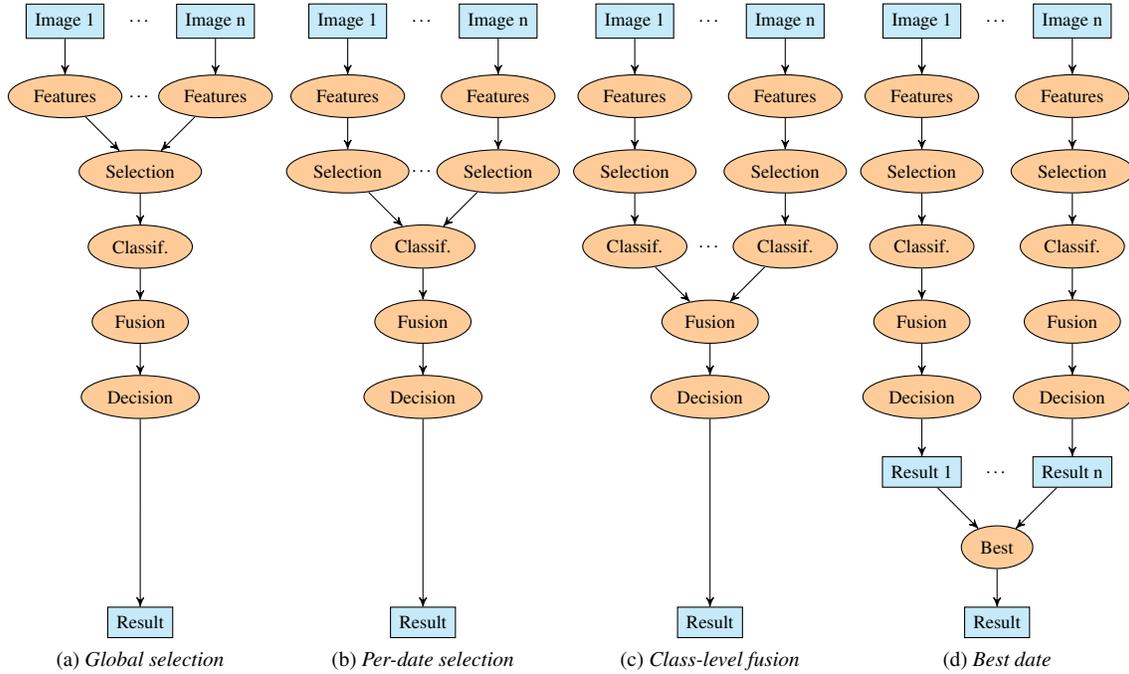


Fig. 1: Schematic comparison of different fusion methods: (a) Overall feature selection using all epochs, (b) Per-date feature selection, (c) Class-level classification fusion and (d) Best Epoch choice (no fusion).

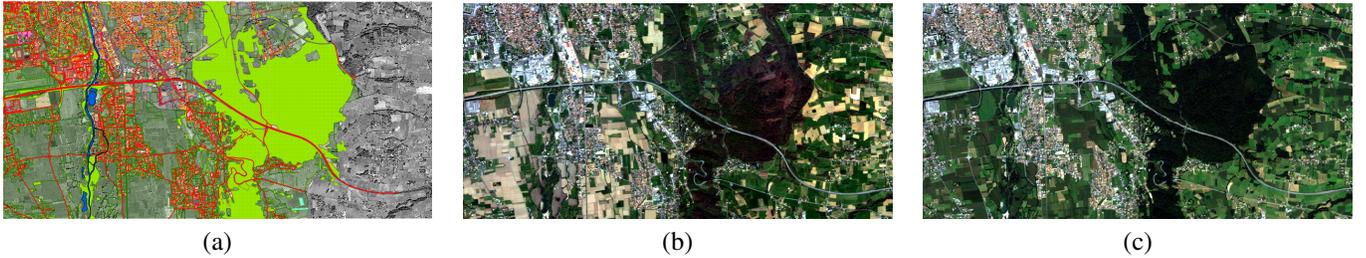


Fig. 2: (a) The French National Land-Cover Database, and two optical images (Rapid-Eye) acquired in April 2011 (b), and in September 2011 (c).

obtain a better recall. Otherwise, overall selection and class-level fusion allows to improve the accuracy, when per-date selection improve the recall.

6. CONCLUSION

In this paper, we proposed three different adaptations of an existing Land-Cover geodatabases (LC-DB) updating, allowing to deal with multi-temporal datasets. Firsts results tend to show that, on our dataset, introducing several images does not improve results obtained with a single image. However, in these preliminary results, only three image composed the input dataset, so that no general conclusion can be drawn so

	<i>Best</i>	<i>Overall</i>	<i>Per-date</i>	<i>Class-level</i>
	Accuracy			
<i>building</i>	0.36	0.4	0.33	0.37
<i>road</i>	0.25	0.25	0.19	0.26
<i>forest</i>	0.94	0.98	0.96	0.98
<i>grassland</i>	0.58	0.55	0.57	0.56
	Recall			
<i>building</i>	0.96	0.92	0.97	0.95
<i>road</i>	0.89	0.88	0.88	0.86
<i>forest</i>	0.82	0.80	0.85	0.80
<i>grassland</i>	0.83	0.63	0.72	0.69

Table 2: Details of by class recall and accuracy obtained for each fusion method.

far. Thus, we have plain to carry out several experimentations to confirm or disprove those firsts results. For that purpose, a larger dataset, composed of a larger number of image acquired the same year, will be introduced.

7. REFERENCES

- [1] A. Belward and J. Skoien, “Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. in Press, Apr. 2014.
- [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, pp. 3081–3095, Aug. 2012.
- [3] F. Bovolo, L. Bruzzone, and R.L. King, “Introduction to the special issue on analysis of multitemporal remote sensing data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1867 – 1869, 2013.
- [4] A. Gressin, C. Mallet, N. Vincent, and N. Paparoditis, “Updating the new French national land cover database,” in *IEEE International Geoscience and Remote Sensing Symposium*, Québec, Canada, 2014.

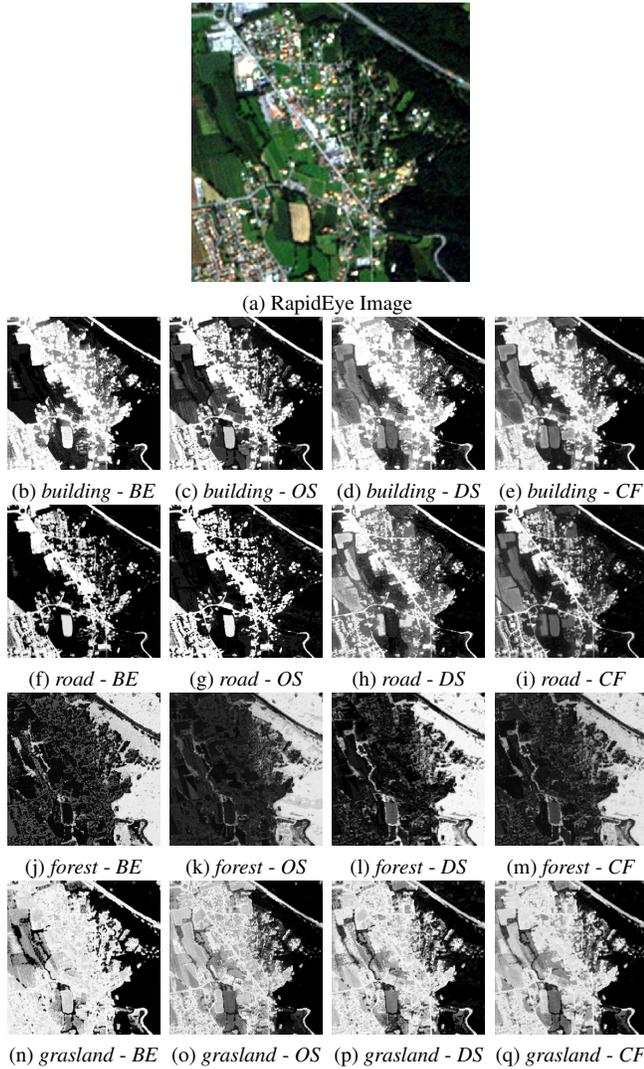


Fig. 3: Classifications details of the classes *building*, *road*, *forest* and *grassland*, obtain with the proposed multi-temporal fusion methods: *best epoch (BE)*, *overall selection (OS)*, *per-date selection (DS)* and *class-level fusion (CF)*. Each figure is colored from white (strong belonging to the class) to black (low belonging to the class).