Sensitivity of urban material classification to spatial and spectral configurations from visible to short-wave infrared

Arnaud Le Bris
Univ. Paris-Est, LASTIG STRUDEL, IGN, ENSG
Saint-Mande, France
arnaud.le-bris@ign.fr

Nesrine Chehata
EA G&E Bordeaux INP, Université Bordeaux Montaigne
Pessac, France
nesrine.chehata@ipb.fr

Abstract—Urban material maps are useful for several city modeling or monitoring applications and can be retrieved from remote sensing data. This study investigates the impact of spectral and spatial sensor configuration on urban material classification results, comparing several configurations corresponding to existing or envisaged airborne or space sensors. Images corresponding to such sensors were simulated out of an airborne hyperspectral acquisition. At the end, the relevance of an enhanced spectral configuration and especially providing bands from the SWIR domain was proven, as well as the need for a fine spatial resolution to retrieve urban objects. However, the (late) fusion of multispectral imagery at 2 m resolution with hyperspectral data at 8 m resolution was also proven to lead to good results.

Index Terms—Hyperspectral, Multispectral, Superspectral, Sensor assessment, Spatial resolution, Spectral configuration, Classification, Urban materials

I. Introduction - context

In urban areas, very fine land cover maps at material level, i.e. knowledge concerning the roofing materials or the different kinds of ground areas, are required for several city modeling or monitoring applications [1]. Remote sensing seems convenient for providing such information at a large scale but enhanced spectral configurations may be required. Within this context, this study aimed at comparing the impact of spectral and spatial sensor configurations on classification results. Several spectral and spatial configurations corresponding to existing or envisaged airborne or space sensors were compared. They were simulated out of a common hyperspectral acquisition

II. Data set and classification assessment strategy

A. Data set

Images corresponding to several sensor configurations were generated out of a common airborne hyperspectral image acquired at 2 m spatial resolution over Toulouse city (South Western France) in June 2015. This image was captured by an Hyspex sensor, providing 408 spectral bands ranging from 400 to 2500 nm. The study area covers a 2 km^2 area.

B. Classification and evaluation process

In further experiments, classifiers were trained out of samples extracted from images. Ground truth polygons were manually plotted by photo-interpretation using all available information (hyperspectral image, aerial red-green-blue-near infrared ortho-images at 12,5 cm spatial resolution, ortho-images from Géoportail, Google Earth and Bing Map as well as Street View images). However, the classes are present in very different proportions. For instance, roofs are mainly covered by clay tiles on the study area, while other roofing materials are less frequent, and sometimes difficult to identify. Concerning ground materials, many streets are in shadow or masked by trees, thus limiting the amount of available samples.

In further tests, a gaussian kernel support vector machine (svm) classifier was used [5]. Indeed, earlier tests have shown its ability to provide good results on hyperspectral data even with a very limited training set. Classification was performed per pixel and no spatial smoothing post-processing was applied to the results. Earlier experiments also showed that plotted ground truth polygons were not necessarily compatible with the simulated coarsest spatial configurations, for which they generated a too important amount of mixed pixels. Thus, such ground truth was first eroded in order to keep only its purest part. This eroded ground truth map was then used for all considered sensor configurations in order to have a same protocol for all of them: the same training set of 50 samples per class was used for each of them.
Conversely, obtained classification results were evaluated by comparison with the original ground truth map (before erosion). Indeed, quantitative evaluation consisted in comparing at pixel level the classification to this reference map. To have a fair comparison, all images were resampled (bilinear interpolation) at the original 2 m spatial resolution before being classified and evaluated. Several classic quality metrics were computed (User and producer accuracies and F1-Score) for each class. Indeed, per-class indices are here more relevant than global metrics because of the very imbalanced classes.

Furthermore it can be said that for some classes, it was difficult to find a sufficient number of samples for training and test that belong to different objects. Thus, the evaluation can sometimes be optimistic.

III. IMPACT OF SPECTRAL CONFIGURATION

A. Considered spectral configurations

To assess the impact of the spectral configuration, the next spectral configurations were considered :

- Spectral configuration of Pleiades sensor ;
- Original spectral configuration of Hyspex sensor limited to VNIR domain (400 nm - 1000 nm) after removal of noisy bands (affected by the atmosphere) ;
- Original spectral configuration of Hyspex sensor limited to VNIR domain and to the first part of SWIR domain (400 nm - 1800 nm) after removal of noisy bands ;
- Original spectral configuration of Hyspex sensor on the whole VNIR+SWIR domain (400 nm - 2500 nm) after removal of noisy bands ;
- 10 bands subset selected out of Hyspex bands. This band subset was selected automatically using a urban material spectral data set constituted out of several available reference spectral libraries, as in [2]. These spectral references are reflectance spectra acquired on the field or in laboratory, and thus, independent on the image scene processed. Besides, they correspond to a slightly different classification legend. Selected bands are : 440 nm, 560 nm, 600 nm, 665 nm, 781 nm, 1674 nm, 2107 nm, 2203 nm, 2311 nm and 2323 nm.

These configurations were explored, simulating them out of the original hyperspectral data set. All images were simulated for a 2 m spatial resolution, as the one of the original acquisition. Sensor SNR was taken into account. Except for the simulation of Pléiades configuration, kept in radiance, all images have undergone an atmospheric correction by Cochise method [6]. However, previous results have shown that when training a classifier directly from samples extracted from the image, such correction has a little effect on obtained results.

B. Results

Obtained results are presented on figure 1 and in table I. Differences between results obtained for the different spectral configurations are not so strong. However, it can be noticed that the best results are generally reached for enhanced spectral configurations. For instance, hyperspectral configuration limited to VNIR domain (400-1000 nm) leads to better results than for Pleiades configuration. Results are improved when adding bands from SWIR domain, and the best results are reached for the configuration covering the whole 400-2500 nm domain. The importance of SWIR domain can also be noted since improvements are observed on several classes for results obtained for a 10 band subset selected out of the 400-2500 nm domain, compared to Pléiades configuration, and even to hyperspectral VNIR 400-1000 nm configuration. From a qualitative point of view, such improvements can also be assessed visually : the finer the spectral configuration, the less noisy the results. It must here be reminded that these different results were obtained by per pixel classification and have undergone no spatial smoothing post-processing. The main remaining confusions concern next classes :

- Slates are hardly distinguished, as expected. Indeed, this class is represented by very few samples while it exhibits an important specular reflective behavior, leading to a strong intra-class variability and making it similar to other classes leading to confusions with shadows, water or metal.
- High and low vegetation tend to be confused. Indeed, these classes exhibit an intra-class variability (condition, species) as important as inter-class variability.
- Some confusions between asphalt, stone pavement and roofing gravels also occur.

However, this experiment remains limited as the study area is not wide, thus leading to a model overfitting risk, and as some classes are quite rare (training and evaluation will be performed on different samples but from the same areas).

IV. IMPACT OF SPATIAL RESOLUTION

A. Considered spatial configurations

The impact of spatial resolution was then investigated, especially within the specific context of an envisaged future French satellite Hypxim [4], for which several options are considered. Thus, images were simulated for two envisaged spatial resolutions 4 m and 8 m, and for the spectral configuration of this sensor (192 bands ranging from 418 to 2502 nm for a 10.9 nm band width), taking into account its SNR.

B. Results

Visual analysis of the classification results (Figure 2) shows a quality decrease when the spatial resolution becomes coarser. Indeed, at 4 m, the scene can still be understood, while at 8 m, only the largest objects are still well classified. Quantitative results (Table II) are less strict. Some classes are even better retrieved when pixel size decreases from 4 to 8 m. However, it must be kept in mind that reference map is not a partition of the image, and that reference polygons have generally not been plotted on small urban objects. Thus, results at 8 m spatial resolution are coarser and smoother (less noise), and small objects can tend to disappear, being labelled as their largest neighbours to other classes, without a striking consequence in the quantitative evaluation. For both spatial resolutions, mixed pixels are not necessarily well classified. However classification is not necessarily the most relevant method at such resolutions and alternative approaches such as spectral unmixing should also be tested.

V. FUSION

Previous experiments showed that enhanced spectral configuration, and especially involving bands from the SWIR domain, improves urban material classification, but also that the spatial resolution can not be too coarse as for instance the one targeted for Hypxim satellite. However, a possible solution can consist in merging such hyperspectral data

with multispectral images exhibiting an enhanced spatial resolution. Although early fusion (i.e. pan/multi-sharpening) could be applied within this context, a late fusion process was retained, relying on posterior class membership probabilities generated by classifiers applied independently to each source. Indeed, it is more generic and could be applied even to images not acquired simultaneously or processed by specific land cover classification approaches (e.g. unmixing for hyperspectral data and classification for multispectral VHR one). The method described in [3] was applied. It consists in a per-pixel fusion based on a Dempster-Shafer evidential rules with masses calculated as in [3]).

It was here applied to the classification results of Hypxim 8 m and Pléiades 2 m simulations. It can be seen from Figure 2 and table 2 that results are improved. Qualitatively, the result is visually easier to interpret: the different objects are more clearly delineated than for original Hypxim 8 m classification. On the other hand, this result is also less noisy than Pléiades' original one. Besides, some errors are also corrected especially compared to Pléiades result. Quantitatively (see Table II, results are also generally better than for the independent classifications of original images.

TABLE I
F1-Score (in %) of the different classes for different spectral configurations.

Classes	Pleiades	400 - 1000 nm	400 - 1800 nm	400 - 2500 nm	10 bands
Slate roof	49.6	50.7	51.8	**54.6**	48.2
Asphalt	68.4	72.6	71.8	**75.1**	69.1
Cement	59.8	**74.4**	71.3	73.4	72.9
Water	99.3	99.3	99.1	99.2	99.2
Stone pavement	47.4	53.9	57.9	**65.6**	50.3
Bare ground	51.0	54.8	59.5	62.1	**61.8**
Gravel roof	70.2	71.6	75.0	**82.3**	76.6
Metal 1 roof	62.8	63.1	68.8	**70.9**	69.6
Metal 2 roof	43.4	44.9	56.3	**61.7**	60.0
Tile roof	90.9	90.5	95.5	**96.0**	95.6
Low vegetation	43.3	57.7	70.3	**71.4**	71.6
High vegetation	85.6	88.6	90.8	**91.0**	89.4
Railway	77.4	82.0	**84.3**	83.4	80.4
Plastic roof	76.9	76.7	83.2	**84.4**	82.1
Shadow	95.3	96.0	96.4	**96.5**	95.7

TABLE II
F1-Score (in %) of the different classes for different spectral and spatial configurations.

Classes	Hypxim 4m	Hypxim 8m	Pléiades 2m	Fusion
Slate roof	**56.9**	30.9	49.6	52.5
Asphalt	68.7	71.1	68.4	**76.6**
Cement	71.8	**79.0**	59.8	78.3
Water	97.9	96.9	**99.3**	98.7
Stone pavement	53.1	68.8	47.4	**70.5**
Bare Ground	60.9	**79.3**	51.0	77.8
Gravel roof	74.1	72.8	70.2	**82.4**
Metal 1 roof	68.3	80.6	62.8	**80.7**
Metal 2 roof	47.7	61.5	43.4	**65.5**
Tile roof	93.5	79.8	90.9	**94.6**
Low vegetation	**78.0**	64.7	43.3	74.4
High vegetation	89.6	84.7	85.6	**90.2**
Railway	82.3	66.1	77.4	**87.0**
Plastic roof	86.5	92.1	**76.9**	92.5
Shadow	**95.2**	77.5	**95.3**	94.6

data at 8 m resolution leads to results, better than for each source considered independently, and even better than for hyperspectral data at 4 m resolution.

Acknowledgment

References

[1] H. Shafri, E. Taherzadeh, S. Mansor and R. Ashurov, "Hyperspectral remote sensing of urban areas: an overview of techniques and applications", Research Journal of Applied Sciences, Engineering and Technology, vol. 4(11), pp. 1557-1565, 2012.

[2] A. Le Bris, N. Chehata, X. Briottet and N. Paparoditis, "Hierarchically exploring the width of spectral bands for urban material classification", Proc. of JURSE 2017, 2012.

[3] W. Ouerghemmi, A. Le Bris, N. Chehata and C. Mallet, "A two-step decision fusion strategy: application to hyperspectral and multispectral images for urban classification", International Archives of the Photogrammety, Remote Sensing and Spatial Information Sciences, vol. XLII-1-W1, pp. 167–174, 2017.

[4] S. Michel, M.-J. Lefèvre-Fonollosa and S. Hosford, "HYPXIM - a hyperspectral satellite defined for science, security and defence users", Proc. of the 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) 2011, Lisbon, Portugal, 2011.

[5] B. Schölkopf and A. J. Smola, "Learning with kernels: Support Vector Machines, Regularization, Optimization and Beyond", MIT Press, 2002.

[6] L. Poutier, C. Miesch, X. Lenot, V. Achard and Y. Boucher,"COMANCHE and COCHISE: two reciprocal atmospheric codes for hyperspectral remote sensing", Proc. of the 2002 AVIRIS Earth Science and Applications Workshop, 2002.

VI. Conclusion

Different spectral configurations were compared for a same 2 m spatial resolution, showing that the finer the spectral configuration, the better the classification quality : classes are better discriminated and results are less noisy even though some confusions between some classes persist. Different spatial resolution (4 and 8 m) were also compared for a same spectral configuration (Hypxim sensor), showing the need for a (very) high spatial resolution to retrieve individual urban objects. Indeed, the visual quality decreases when the resolution becomes coarser and at 8 m, only the biggest objects are well classified. However, it was also shown that a late fusion of multispectral imagery at 2 m resolution with hyperspectral
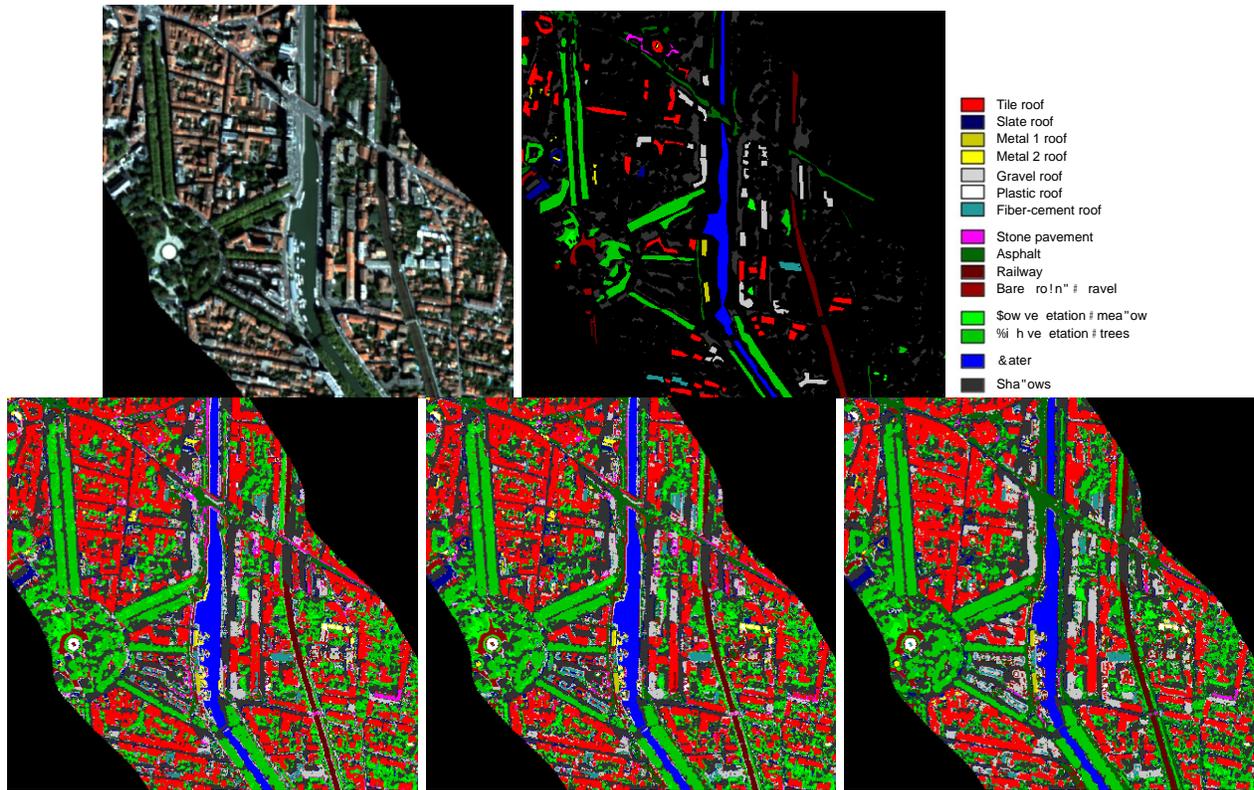
Fig. 1. From left to right and from top to bottom, ortho-image on a part of the test zone, associated ground truth, classification results for Pléiades, Hyspex VNIR and Hyspex SWIR spectral configurations at 2 m spatial resolution
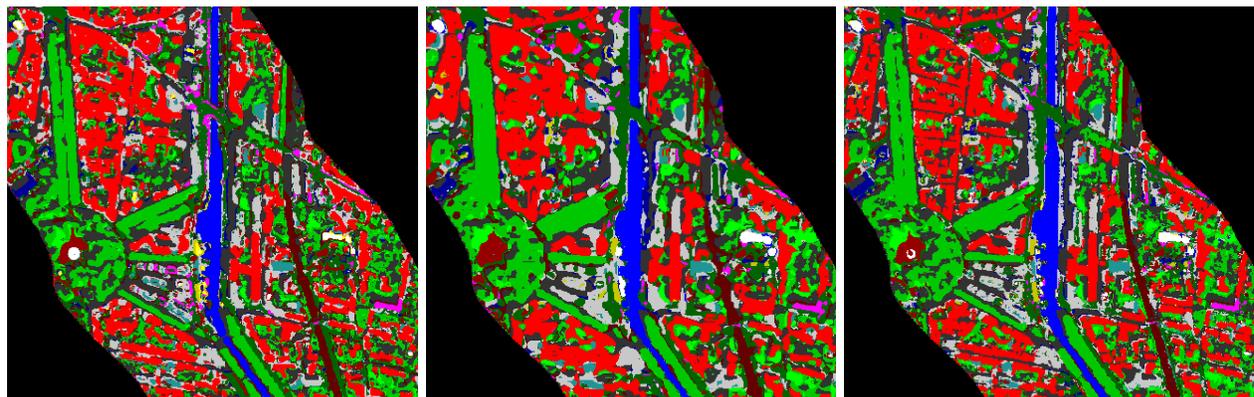

Fig. 2. From left to right and from top to bottom, classification results for Hypxim spectral configuration at 4 m and 8 m spatial resolution, and results of the fusion of Hypxim at 8 m with Pléiades at 2 m spatial resolution