Hierarchically exploring the width of spectral bands for urban material classification

Arnaud Le Bris*, Nesrine Chelata†, Xavier Briottet‡ and Nicolas Paparoditis*
*Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mande, France
†EA 4592, G&E, ENSEGID-IB, University of Bordeaux, France
‡ONERA, The French Aerospace Lab, Toulouse, France

***Abstract*—In urban areas, material maps, i.e. i.e. knowledge concerning the roofing materials or the different kinds of ground areas, are necessary for several city modelling or monitoring applications. Airborne remote sensing techniques appear to be convenient for providing them at a large scale but require an enhanced imagery spectral resolution. A superspectral sensor with a limited number of bands dedicated to urban materials classification could be a solution. Within this context, this study focused on the optimization of this band subset, considering both the position of the bands and their width. The used approach first builds a hierarchy of groups of adjacent bands, according to a relevance criterion to decide which adjacent bands must be merged. Then, band selection is performed at the different levels of this hierarchy. Several band configuration are thus explored within this hierarchy. This method was applied to a data set consisting of spectra generated from reflectance spectral signatures of 9 common urban materials collected from 7 spectral libraires. At the end, the potential of a superspectral sensor with wider bands was confirmed.**

I. Introduction

A. Some needs for urban materials maps

Needs for high resolution land cover data have been growing, to answer several societal, regulatory and scientific needs, to produce environmental indicators to manage ecosystems and territories, to monitor environmental or human phenomena, or to be able to have a picture of an initial situation and to evaluate the impacts of public policies. To answer these needs, national mapping or environment agencies have undertaken the production of such large scale land cover databases. Nevertheless, these databases provide a general classification and may not suit some specific applications requiring a finer semantic or geometric level of details. Indeed, in urban areas, both semantic and spatial finer knowledge about land cover, i.e. maps of urban materials, are required by several city modelling applications. Urban environment is indeed strongly influenced, in terms of ecology, energy and climate by the present materials. Such material maps would be useful to derive indicators to monitor public policies impacts, or to feed urban simulation models (such as micro-meteorology, hydrology, pollutants flow monitoring and ground perviousness monitoring). Several applications requiring materials maps are listed in [1] and [2] and reminded below.

1) Quantification of pollutant flows: By corrosion, some roofing materials can generate metallic or organic pollutant elements [3], [4]. Knowledge about the different roofing materials coverage areas is thus needed to quantify these emissions.

2) Monitoring of dangerous materials: asbestos-cement roofs: [1] Asbestos-cement can be dangerous for human health, especially when they are deteriorated.

3) Weather models: Urban land cover (in terms of materials, perviousness and vegetation) are required by micro weather (wind, temperature, ...) model simulators [1], [5].

4) Monitoring of ground perviousness: Monitoring the extension of impervious areas and checking their appliance to new legislation is important, since the development of impervious areas causes an aggravation of flooding events. Otherwise, perviousness maps are required as input data by (micro) hydrological models [1].

5) Determination of road type and monitoring of road condition: Maps of road types are useful for some of the mentioned applications (meteo, hydrology). A more complex one is the monitoring of road condition to plan of road network renovation projects to avoid expensive field investigation [6].

6) Monitoring of photo-voltaic (PV) development: Roofing material maps help to estimate the potential of a city to develop PV energy [7] and detecting installed panels enables to monitor the development of this technology.

B. Toward a superspectral camera ?

Thus, very high resolution urban land cover is required to provide knowledge about the roofing materials and the different kinds of ground areas. Airborne remote sensing appears to be convenient for obtaining such material map at a large scale. but requires an enhanced spectral resolution using superspectral or hyperspectral sensors. Hyperspectral imagery consists of hundreds of highly correlated contiguous spectral bands. Only a subset of well selected bands would be sufficient for urban materials classification [8]. A superspectral aerial camera system dedicated to urban material classification could then be designed from this band subset. Such system could offer some advantages compared to most hyperspectral sensors, combining the use of suitable spectral bands for a specific application with a higher spatial resolution and a larger swath. It would also be a photogrammetric system, enabling to capture multistereoscopic images and to derive BRDF models. In previous work [9], an automatic band selection framework was used to select optimal band subsets for urban materials classification. Experiments were performed on data sets generated from material reference reflectance spectra from available spectral libraries. This work was performed for a fixed spectral

band width. However, wider bands enable to collect more photons and thus to limit noise and increase spatial resolution, but there is a risk to loose discriminative information.
This new study investigates the optimization of spectral band width. The exploratory hierarchical approach from [10] is used, building a hierarchy of groups of adjacent bands and performing band selection at the different levels of this hierarchy.

II. Band extraction

Hyperspectral imagery, generates huge data volumes, consisting of hundreds of contiguous spectral bands being highly correlated to each other. Dimensionality reduction strategies aim at reducing data volume minimizing the loss of useful information. They belong either to **feature extraction** or **feature selection** categories. Feature extraction methods (e.g. Principal Component Analysis) consist in reformulating and summing up original information, reprojecting it in another feature space. On the opposite, feature selection (FS) methods applied to **band selection** select the most relevant band subset (among the original bands of the hyperspectral data set) for a specific problem. As hyperspectral adjacent bands are very correlated to each other, **band extraction**, that is to say the definition of an optimal set of spectral bands optimizing both their width and position along the spectra, can be considered as intermediate between feature extraction and individual band selection. Band selection/extraction approaches enable not to loose the physical meaning of the selected bands. They are adapted to the design of multispectral or superspectral sensors dedicated to specific land cover classifications for which only a limited band subset is relevant.

The approach used in this paper (see fig. 1) was described in [10]. It is an exploratory one consisting in first building a hierarchy of groups of adjacent bands. Then, band selection is performed at the different levels of this hierarchy. The hierarchy of groups of adjacent bands is used as a constraint for band extraction and a way to limit the number of possible combinations, contrary to some existing approaches such as [11] that extract optimal bands according to JM information using an adapted optimization method or [12] that directly use a genetic algorithm to optimize a wrapper score.

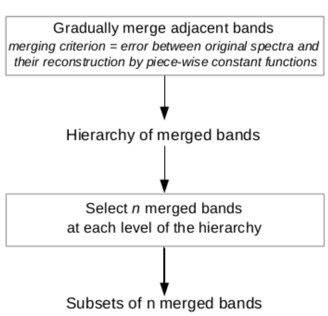

Fig. 1. Proposed band extraction approach

A. Hierarchical band merging

The first step of the proposed approach consists in building a hierarchy of groups of adjacent bands, that are then merged. The hierarchical band merging approach is a bottom-up one, starting from the original individual bands, and gradually merging adjacent bands according to a band merging criterion. More details about the algorithm can be found in [10].

A **spectra approximation error** [10] was used as merging score. It aims at preserving the shape of the spectra and relies on [13]'s method to decompose spectra into piece-wise constant functions (fig. 2). Adjacent bands are merged trying to minimize the reconstruction error between the original and the piece-wise constant reconstructed spectra.

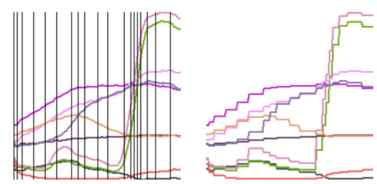

Fig. 2. Left: merged bands (black lines) superimposed on the original spectra. Right: piece-wise constant reconstructed spectra for these merged bands

B. Band selection

To optimize spectral configuration for a limited number of merged bands, band selection was performed at the different levels of the hierarchy of merged bands: a subset of a fixed number of merged bands was selected at each level of the hierarchy.
The used FS score is the one used in [9] : it is a wrapper score relying on Random Forests classifier and taking into account classification confidence.
It was here optimized at each level of the hierarchy using a modified version of Sequential Forward Floating Search (SFFS) [14] presented in [10]. The band merging hierarchy is considered within the FS process in a bottom-up approach : the band subset selected at the previous lower level is used as an initial solution when performing band selection at a new level of the hierarchy of merged bands. The less relevant band according to the FS score is then removed and classic SFFS is performed.

III. Data set

Spectral optimization was performed from a library of reference spectra of urban materials. These spectra were collected from several available existing spectral libraries listed below.

- **ASTER Spectral Library** [1] [15]
- **SLUM** [2] [5] collected in London.
- **MEMOIRES** [3] **[16] and ONERA data**: mostly collected in Toulouse (France).

[1] http://speclib.jpl.nasa.gov/
[2] http://LondonClimate.info/LUMA/SLUM.html
[3] http://www.onera.fr/dota/memoires

- **Santa Barbara libraries** [4] [8], [6] collected (only field measures) in Santa Barbara.
- **Ben Dor spectral libray** [17] collected in Tel Aviv
- **DESIREX** [18] collected in Madrid.

All collected spectra were integrated into a common data base (DB), associating several attributes to each of them (e.g. material class, variety , colour, condition...), even though all this information was not always available.

Only **reflectance** spectra concerning both the Visible Near Infra-Red (VNIR) (400-1000 nm) and the Short Wave Infra-Red (SWIR) (1000-2400 nm) spectral domains were kept. The original spectral resolution of the spectra was comprised between 1 and 10 nm. Bands concerned by atmospheric absorption and other artefacts were removed.

Some classes were let aside from the DB, even if they can be important in urban land cover. Though important, vegetation was let aside since its discrimination from non vegetation is easy and its characterization was intended to be studied later. Water can vary a lot depending on depth and turbidity and was let aside since few spectra were available..

The number of available spectra per class can vary a lot and is generally not sufficient to correctly evaluate intra class variability. To cope with this insufficient number of available spectra, a random multiplicative factor was simply applied to reference spectra in order to generate more synthetic spectra from the DB, partly simulating intra-class variability.The proposed process to generate an experimental data set is described in fig. 3.

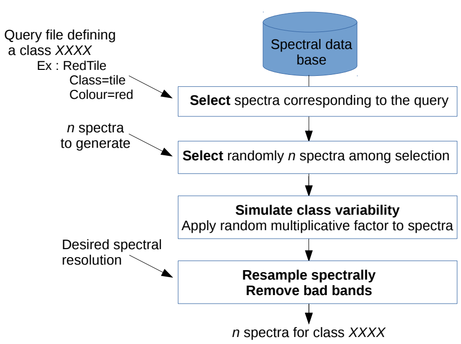

Fig. 3. Synthetic spectra collection generation scheme

Experiments were performed for next 9 items legend: *Slate - Asphalt - Cement/concrete - Gravel - Metal - Stone pavement bricks - Shingle - Bare ground - Tile*.
It consisted in the most common materials in the database and to other important classes (e.g. slate) frequently present in urban areas. In order to perform spectral optimization, a data set was generated from the data base according to this legend. It contained 100 training spectra and 500 test spectra, resampled at a 10 nm spectral resolution ranging from 420 to 2400 nm.

[4]http://www.ncgia.ucsb.edu/ncrst/research/pavementhealth/urban/

IV. Results

A. Hierarchy of merged bands

The hierarchy of merged bands is presented in fig. 4. Bands from 1150-1250 nm, 1500-1700 nm and 2100-2150 nm domains tend to be merged early, since the referance spectra of most classes are quite flat there. Conversely, bands from 420-600 nm and 2250-2350 nm domains are merged later.

Even though it is intended to be used to select an optimal band subset, this hierarchy of merged bands can also be a way to explore several band configuration with varying spectral resolution, that is to say with contiguous bands with different bandwidth. These configurations were evaluated considering the classification performance (measured by Kappa, mean and minimum FScore among all classes) reached using a RBF SVM classifier. Results are presented on fig. 5 : performances reached for several configurations are quite equivalent to the original spectral resolution, with quality rates being slightly increased or decreased but remaining in the same domain. The performance is decreased for configurations of less than 90 merged bands, corresponding to the fusion of original bands from 2250-2350 nm. But it is then improved for configurations of less than 40 bands corresponding to the fusion of the last original visible bands. Lower performances between these two events correspond to a over-representation of visible bands among these spectral configurations. Last, classification performance decreases strongly for spectral configurations of less than 22 bands, i.e. for configurations consisting of too wide bands leading to a loss of useful information.

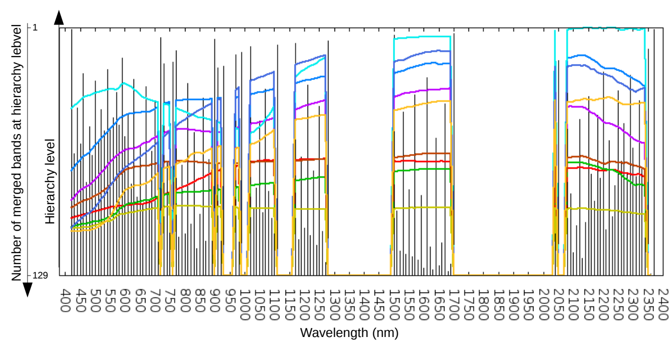

Fig. 4. Hierarchy of merged bands. Vertical black lines correspond to the border between merged bands (the higher in the hierarchy, the lower number of bands).

B. Selection of merged bands at the different level of the hierarchy

As the 10 bands were shown to be sufficient [9], subsets of 10 merged bands were selected at each level of the hierarchy. They are shown on fig. 6. Selected bands are quite stable in some spectral domaine (in visible and 1500-2400nm SWIR domain), but vary more in near infrared (750-1050 nm). These different band subsets were also evaluated considering the classification performance reached using a RBF SVM classifier. Results are shown on fig. 7. No strong improvement was observed using merged bands : quality rates slightly

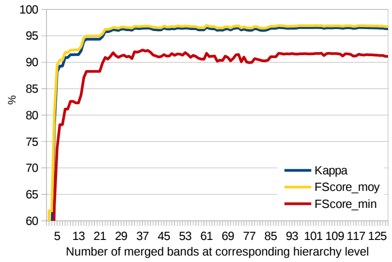

Fig. 5. Classification performance of the spectral configuration corresponding to the different hierarchy levels, before band selection.

increased for some configurations, or decreased for other ones. However, classification performance remain within the same magnitudes (<1%). Thus, it is possible to use bands wider than the original 10 nm width.

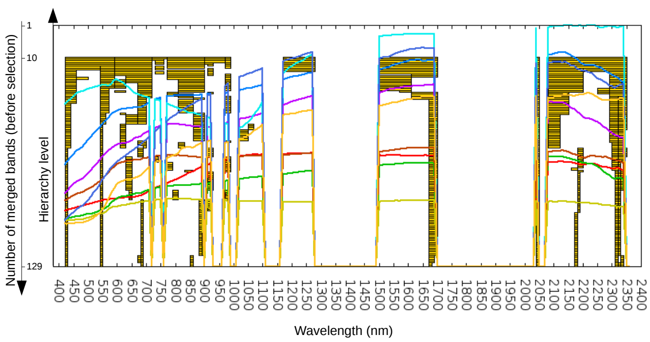

Fig. 6. Selected band subsets at the different levels of the hierarchy. Each line corresponds to a level in the hierarchy of merged bands.

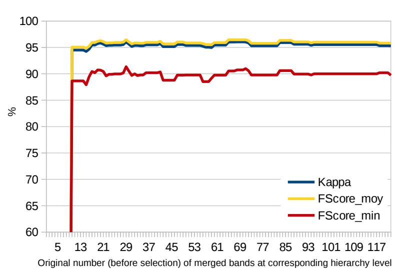

Fig. 7. Classification performance of the spectral configuration corresponding to the different hierarchy levels, before band selection.

V. Conclusion

Spectral optimization was performed in the context of designing a superspectral sensor dedicated to urban material map classification. An exploratory approach was used, building a hierarchy of merged bands and performing band selection at the different levels of this hierarchy. It enabled to consider several spectral configurations. 9 common urban materials were considered. Experiments have shown it is possible to some extent to use wider bands since classification performance reached using band subsets involving merged bands remain within the same magnitudes than using the original (10 nm width) bands. This is important since wider bands enable to collect more photons and thus to limit noise and increase spatial resolution.

In this study, the variability of some classes could not be completely considered since they were represented by few spectra. Thus obtained quantitative evaluations are optimistic and must be considered carefully. New urban material spectra measurement should be integrated in the data base. Experiments will also be carried out using aerial hyperspectral scenes to bring more realistic evaluation results, since such data will be perturbed by sensor and atmospheric noise, while in this paper only clean reflectance data measured on the field or in laboratory were used.

References

[1] W. Heldens, U. Heiden, T. Esch, E. Stein, and A. Muller, "Can the future EnMAP mission contribute to urban applications? a literature survey," *Remote Sensing*, vol. 3, pp. 1817–1846, 2011.

[2] 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, no. 11, pp. 1557–1565, 2012.

[3] D. Lemp and U. Weidner, "Improvements of roof surface classification using hyperspectral and laser scanning data areas," *ISPRS Archives*, vol. 36, no. 8/W27, pp. CD–Rom, Mar. 2005.

[4] P. Robert-Sainte, "Contribution des matériaux de couverture à la contamination métallique des eaux de ruissellement," Ph.D. dissertation, Université Paris-Est, 2009.

[5] S. Kotthaus, T. Smith, M. J. Wooster, and C. Grimmond, "Derivation of an urban materials spectral library through emittance and reflectance spectroscopy," *IJPRS*, vol. 94, pp. 194–212, 2014.

[6] M. Herold, D. Roberts, O. Smadi, and V. Noronha, "Road condition mapping using hyperspectral remote sensing," in *Proceedings of the AVIRIS Workshop*, Mar. 2004.

[7] A. Jochem, B. Höfle, M. Rutzinger, and N. Pfeifer, "Automatic roof plane detection and analysis in airborne lidar point clouds for solar potential assessment," *Sensors*, vol. 9, pp. 5241–5262, Jul. 2009.

[8] M. Herold, D. Roberts, M. Gardner, and P. Dennison, "Spectrometry for urban area remote sensing–development and analysis of a spectral library from 350 to 2400 nm," *RSE*, vol. 91, no. 3-4, pp. 304 – 319, 2004.

[9] A. Le Bris, N. Chehata, X. Briottet, and N. Paparoditis, "Spectral band selection for urban material classification using hyperspectral libraries," in *ISPRS Annals*, vol. 3(7), jul 2016, pp. 33–40.

[10] ——, "Extraction of optimal spectral bands using hierarchical band merging out of hyperspectral data," in *ISPRS Archives*, vol. 15(3/W3), sep 2015, pp. 459–465.

[11] S. B. Serpico and G. Moser, "Extraction of spectral channels from hyperspectral images for classification purposes," *IEEE TGRS*, vol. 45, no. 2, pp. 484–495, Feb. 2007.

[12] J. Minet, J. Taboury, M. Pealat, N. Roux, J. Lonnoy, and Y. Ferrec, "Adaptive band selection snapshot multispectral imaging in the vis/nir domain," *Proc. of SPIE*, vol. 7835, p. 10, 2010.

[13] A.-C. Jensen and A.-S. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE GRSL*, vol. 4, no. 4, pp. 547–551, Oct. 2007.

[14] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, pp. 1119–1125, Nov. 1994.

[15] A. M. Baldridge, S. Hook, C. Grove, and G. Rivera, "The ASTER spectral library Version 2.0," *RSE*, vol. 113, pp. 711–715, 2009.

[16] C. Martin and B. Rosier, "Status of "MEMOIRES" - a thermal and optical data base," in *Fraunhofer IOSB*, Ettlingen, Allemagne, 2012.

[17] E. Ben-Dor, N. Levin, and H. Saaroni, "A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4-1.1 μm). A case study over Tel-Aviv, Israel," *IJRS*, vol. 22, no. 11, pp. 2193–2218, Nov. 2001.

[18] J. Sobrino, "DESIREX 2008 Final Report," ESA, Tech. Rep., 2008.