



HAL
open science

An interactive approach for extraction of urban patterns from multisource images

Camille Kurtz, Anne Puissant, Nicolas Passat, Pierre Gançarski

► To cite this version:

Camille Kurtz, Anne Puissant, Nicolas Passat, Pierre Gançarski. An interactive approach for extraction of urban patterns from multisource images. Joint Urban Remote Sensing Event (JURSE), 2011, Munich, Germany. pp.321-324, 10.1109/JURSE.2011.5764784 . hal-01719124

HAL Id: hal-01719124

<https://hal.univ-reims.fr/hal-01719124v1>

Submitted on 3 Mar 2018

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.

An interactive approach for extraction of urban patterns from multisource images

C. Kurtz^{*†}, A. Puissant^{*‡}, N. Passat^{*†} and P. Gancarski^{*†}

^{*}University of Strasbourg, 4 rue Blaise Pascal, 67000 Strasbourg, France

[†]LSIIT – Image Sciences, Computer Sciences and Remote Sensing Laboratory, UMR CNRS 7005
{ckurtz, passat, gancarski}@unistra.fr

[‡]LIVE – Image, City and Environment Laboratory, ERL CNRS 7230
anne.puissant@live-cnrs.unistra.fr

Abstract—The extraction of urban patterns from Very High Spatial Resolution (VHSR) images presents several challenges related to the size, the accuracy and the complexity of the considered data. In order to assist the end-user to efficiently carry out this task, a new approach is proposed for hierarchically extracting segments of interest from lower resolution data and finally determining urban patterns in these segments from VHSR ones. Based on an intuitive *modus operandi*, it allows the end-user to progressively segment images based on low level –spatial and radiometric– criteria. Experiments performed on urban images datasets provide encouraging results which may be further used for objects detection and classification purpose.

I. INTRODUCTION

The availability of a new generation of sensors of submetric resolution has led to the production of Very High Spatial Resolution (VHSR) images, and to an improved ability to accurately analyse urban scenes [1]. In particular, in such images, basic urban patterns (*e.g.*, individual houses, gardens, roads) are formed by different materials (*e.g.*, red or grey roofs, different asphalts or vegetations) while complex ones (*e.g.*, urban districts, urban blocks) generally contain different kinds of basic patterns. Thus, all of them are not necessarily composed of homogeneous pixels. These specific properties induced by VHSR images lead to new challenges, for human experts (since the size and complexity of the images make visual analysis a time consuming and error prone task), and for image analysis tools (since methods developed for lower resolutions, *e.g.*, region-based ones [2], [3], are generally designed to extract segments based on radiometric homogeneous hypotheses).

In this context, and due to the actual importance to analyse VHSR images in addition to lower spatial resolution ones (especially when such data are provided by different acquisition devices, providing complementary information at distinct radiometric bands), it is then useful to develop tools adapted to the extraction of potentially hierarchical patterns from such data, and in particular (low-level) segmentation ones. Such segmentation tools should allow the end-user to obtain satisfactory results, at possibly different levels (*i.e.*, scales) of pattern extraction, with minimal time (by automating the tasks which do not require human expertise), minimal efforts

(by reducing the parameters induced by *a priori* knowledge), and ergonomic interaction.

In order to do so, it is possible to take advantage of the data available at several resolutions (from Medium Spatial Resolution (MSR) to VHSR ones) [4], and to involve them in a hierarchical strategy which enables, at a given resolution, the exploration of the whole structure of an urban scene [5], [6]. By analysing first the image content at a coarse resolution and then gradually increasing this resolution [7], [8], it is possible to detect complex patterns (which structure the scene) while avoiding the semantic noise induced by the details [9].

Based on these considerations, a new hierarchical segmentation approach is proposed. It operates first on the low-resolution data, extracting from the image the global structure of the scene into a collection of regions, and subsequently enriches this description by adjoining finer patterns extracted by the high-resolution data (by opposition to the ascendant approaches generally proposed in the literature [10]). This approach, inspired from the principle of photo-interpretation and based on the skills of the end-user, aims at understanding the scene in the same way as the human vision system. This choice is also justified by complexity concerns, since the segmentation of huge VHSR images is facilitated if complex objects are localised from the low resolution data.

This article is organised as follows. Section II provides a general description of the proposed segmentation approach. Section III gathers experiments enabling to visually assess the relevance of this approach in the context of urban analysis. Conclusions and perspectives will be found in Section IV.

II. INTERACTIVE SEGMENTATION OF URBAN PATTERNS

This section proposes an informal description of the segmentation approach presented in this article.

The proposed approach takes as input several images of a same (urban) scene, at successive spatial resolutions. In our case, three images are considered, namely a Medium Spatial Resolution (MSR, 30–5m) image, a High Spatial Resolution (HSR, 3–1m) image and a Very High Spatial Resolution (VHSR, less than 1m) one. (However, the method can be applied, without loss of generality, with $n \geq 2$ images.) It

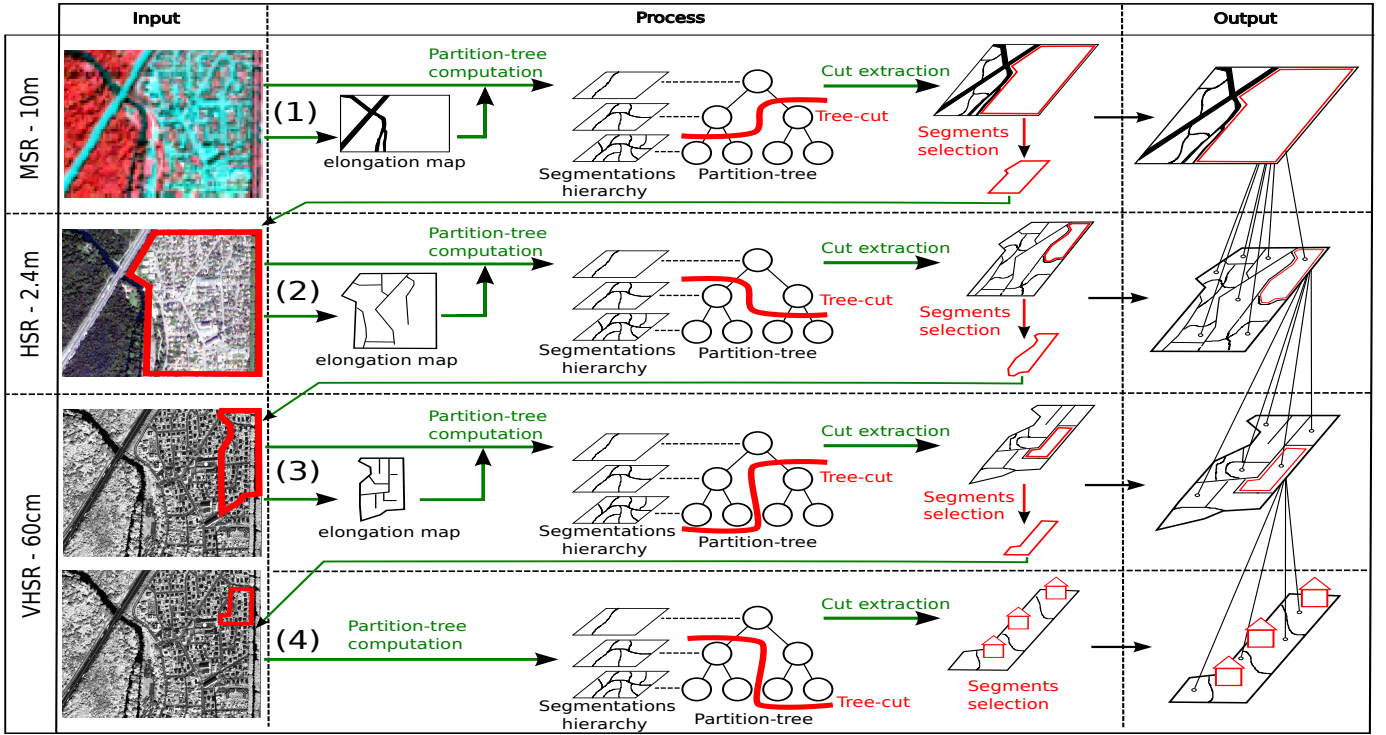


Fig. 1. Workflow of the proposed approach. In red: user interactions. In green: automatic processing.

provides as output a segmentation of some selected areas of the scene at different scales of interpretation.

In order to do so, the method performs in four steps, by successively processing (1) the MSR image, (2) the HSR image, (3,4) the VHSR image. The workflow proposed in Fig. 1 illustrates these successive steps. By lack of room, technical details (especially those related to elongation maps computation, the possible variants of partition trees, geometric and topological corrections of extracted regions, etc.) have been omitted. They will be described and discussed in a further article devoted to image processing issues related to this work.

In Step 1, the processing at the MSR is done on the whole image. An elongation map (emphasising the linear zones of the image) is first computed. This map enables to characterise the linear structures (roads, rivers, railways, etc.) which generally divide an urban scene into large regions. From both MSR and elongation images, a partition tree [6] is computed. This tree proposes hierarchical subdivisions of the image, with respect to a cost function taking into account both radiometric homogeneity (from the input MSR image) and (low) elongation criteria (from the elongation image). Broadly speaking, each cut of the tree corresponds to a specific segmentation (in particular, the segmentation formed by the leaves of the tree corresponds to the initial partition, where each region is composed by one pixel, while the remaining nodes of the tree represent regions that are obtained by merging neighbouring regions with similar spatial and spectral properties). Moreover, the cuts/segmentations which are close to the leaves (resp. close to the root) present segmented regions with homo-

geneous (resp. heterogeneous) radiometric/elongation values. This tree can then be interactively traversed, in a “threshold-like” fashion, thus allowing the user to interactively and easily determine the most satisfactory segmentation (globally, and/or by refining one or several branches). Among the areas of the segmentation result, the user can then select one or several ones in order to refine them thanks to the others resolutions. These selected segments correspond to the parts of the image where the next steps of the segmentation will be carried out.

Steps 2 and 3 are similar to Step 1, with the only difference that Step 2 (resp. Step 3) is carried out on the HSR (resp. VHSR) image, and that the process only deals with the selected segments defined at the end of Step 1 (resp. Step 2). (More generally, this step is applied $n - 2$ times when n images are considered as input.)

Step 4 is performed, once again, on the VHSR image. However, by opposition to Step 3, the considered cost function only deals with radiometric homogeneity. Indeed, the segments extracted at Step 3 generally correspond to urban blocks, and the segmentation of their content will then no longer require to consider roads, and more generally, linear structures.

The tree-cut performed by the user at Step 4 will finally provide the last output of the method.

III. EXPERIMENTS

A. Data

Experiments have been performed on a multiresolution set of images, presenting a part ($6\ 500\text{m} \times 5\ 400\text{m}$) of the urban area of Strasbourg, France (see Fig. 2).

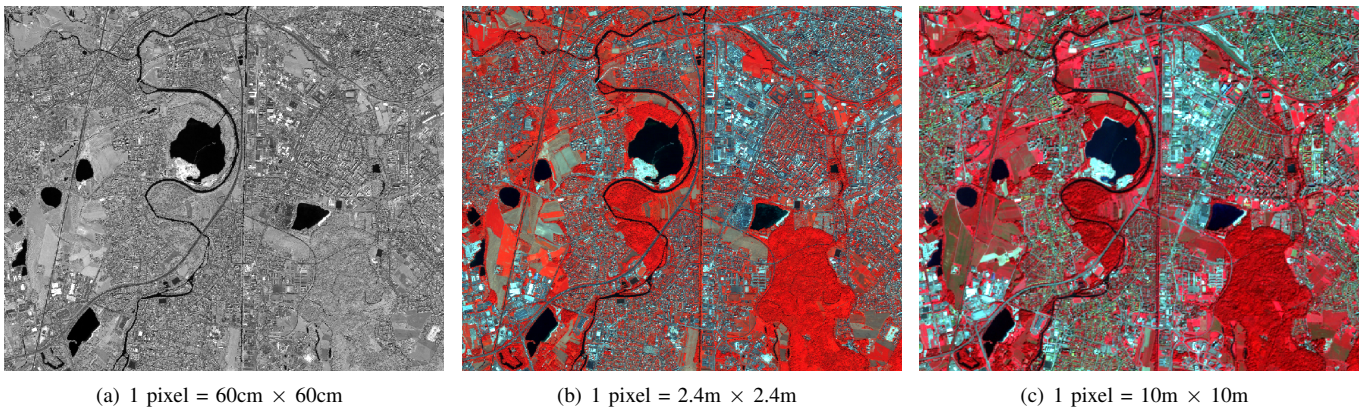


Fig. 2. Satellite images of the same geographical area (6 500m × 5 400m) with different spatial resolutions. (a) VHSR. (b) HSR. (c) MSR.

This set is composed by a couple of HSR QUICKBIRD images (© DigitalGlobe Inc.) acquired in July 2008 and a single SPOT-5 MSR one (© CNES – ISIS program) acquired in September 2002. The couple of QUICKBIRD images is composed by a VHSR panchromatic image with a spatial resolution of 60cm (Fig. 2(a)) and by a HSR multispectral image (blue, green, red and near-infrared) with a spatial resolution of 2.4m (Fig. 2(b)). The MSR SPOT-5 multispectral image (green, red, near-infrared and middle-infrared) has a spatial resolution of 10m (Fig. 2(c)). All the data are georeferenced in the same local cartographic projection (Lambert I).

These images present a typical suburban area with water surfaces (in black, centre of the image), forest area (in red on the MSR/HSR images, south of the images), industrial areas (in grey on the MSR/HSR images, centre/east of the images), individual or collective housing blocks (in red, blue and white textured on the MSR/HSR images), agricultural zones with different spectral responses due to the seasons (bare soil at the beginning of summer on the HSR image/red at the end of summer on the MSR image). Land cover/use changes are not taken into account in this paper and do not disturb the results of the proposed approach.

B. A case study

To assess the efficiency of the proposed approach, several tests were performed to help end-users to extract (interactively) a hierarchy of different complex urban patterns (urban districts, urban blocks, urban objects, etc.) from the data described above. Fig. 3 shows the result of a case study by considering each step of the method described previously.

Step 1 was first applied to separate the largest urban structures of the scene in order to extract an urban district from the MSR image. During this step, a partition tree was generated for the whole MSR image enabling the user to select a relevant cut through this structure. Fig. 3(a) shows the partition corresponding to the selected cut. One can observe that this partition is composed of large regions matching with urban districts and numerous tiny regions forming linear structures and covering vegetation areas. This over-segmentation problem is due to the spatial criteria used by the algorithm

(the elongation one) which does not consider the vegetation areas. Once the cut selected, the user has selected a segment (region depicted in yellow on Fig. 3(a)) corresponding to a mixed urban district composed with individual and collective housing structures and industrial ones.

Step 2 was then applied (on the HSR image) to refine the district extracted at the previous step. Note that for this step (and for the next), the elongation map and the partition tree were only computed for an area of the image corresponding to the segment previously extracted. At this intermediate level, a relevant cut of the partition tree is a partition which enables to separate a district into different sub-ones. Fig. 3(b) shows the partition corresponding to the selected cut. The district previously extracted was split into four large regions. The largest one is still a mixed urban district. The two smaller districts correspond to commercial or industrial sub-districts and the latest one (region depicted in red) regroups several housing blocks with individual houses. Linear structures are gathered in numerous small regions.

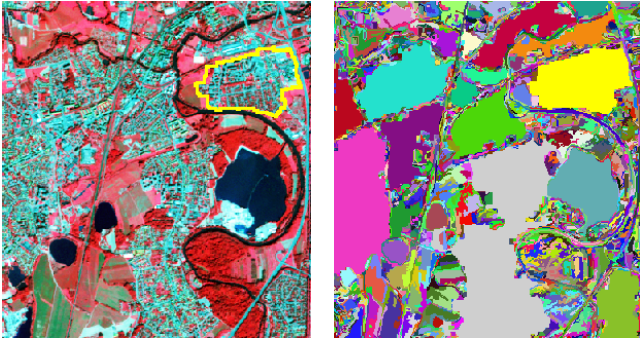
Step 3 was then applied (on the VHSR image) to split the urban sub-district in red (previously extracted) into different urban blocks. Fig. 3(c) shows the partition selected from the partition tree. The selected region seems to correspond to an individual housing block (depicted in light green).

Finally, Step 4 was performed on VHSR image and enables to extract “basic” urban objects (Fig. 3(d)) from the urban block previously extracted. Individual houses in the urban block are fairly well extracted even if the boundaries are not perfect. Note that all the urban structures extracted at the different steps are hierarchically linked.

C. Discussion

This subsection proposes a brief discussion about the main advantages/drawbacks related to the proposed method.

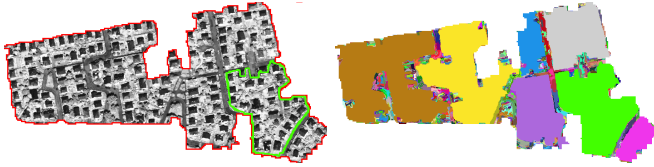
First, the method is globally parameter-free. Indeed, by using an interactive/ergonomic approach (which enables to fully exploit the end-users skills), end-users can choose different levels (scales) of partition to extract the most relevant segments linked to their needs without tuning (non-intuitive) parameters or threshold values.



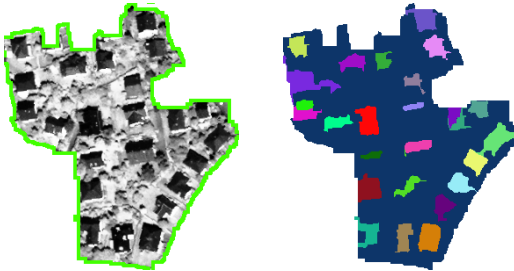
(a) Step 1: extraction of an urban district (depicted in yellow).



(b) Step 2: extraction of an aggregate of housing blocks (depicted in red).



(c) Step 3: extraction of a single urban housing block (depicted in light green).



(d) Step 4: extraction of "basic" urban objects.

Fig. 3. Results of the proposed approach: hierarchical interactive extraction of urban objects. The colours assigned to the regions have no semantics and have been randomly chosen.

Moreover, by pre-processing both elongation map and initial partition of the partition tree (by using the flat zones of the image), this segmentation approach can be used as a real-time application. It allows the end-users to extract hierarchical urban structures with a small time between two interactions (less than 30s CPU on an Intel® Core™2 Quad at 2.8 GHz).

However, it has to be pointed out that this interactive approach is still dedicated to the semi-automated extraction of a *limited set of hierarchical patterns*. The next step of this work will then consist of finding a way to use implicit information provided by these interactions to reproduce automatically such multiscale patterns extraction to the whole images.

IV. CONCLUSION AND PERSPECTIVES

The hierarchical segmentation approach proposed in this article enables to extract complex urban patterns from a set of multisource/multiresolution satellite images. This methodology combines the possibilities offered by a multiresolution analysis and the efficiency of a multilevel approach in order to help and assist human experts in the context of the understanding and of the mapping of urban areas. This multiresolution strategy enables to extract complex structures from an urban scene in a descendant fashion (from low to high resolution) by hierarchically splitting the scene using linear structuring elements.

This article has presented the first results obtained with this method on a multiresolution dataset. Experiments have shown that the quality of the extracted urban patterns seems sufficient to further accurately perform both classification or object detection. This seems to validate (1) the relevance of the proposed method and (2) the soundness of the semi-automation of the photo-interpretation approach.

In a next issue, these results will be assessed by quantitative comparisons with a certified groundtruth map and will be used for unsupervised classification purpose. The spatial/hierarchical relationships of these extracted patterns could be used as new properties to describe these objects in order to enhance the classification process.

REFERENCES

- [1] A. Puissant and C. Weber, "The utility of Very High Spatial Resolution images to identify urban objects," *Geocarto International*, vol. 17, no. 1, pp. 33–44, 2002.
- [2] M. Baatz, C. Hoffmann, and G. Willhauck, "Progressing from object-based to object-oriented image analysis," in *Object-Based Image Analysis*, ser. Lecture Notes in Geoinformation and Cartography, T. Blaschke, S. Lang, and G. J. Hay, Eds. Springer, 2008, ch. 1.2, pp. 29–42.
- [3] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 2–16, 2010.
- [4] R. Gaetano, G. Scarpa, and G. Poggi, "Hierarchical texture-based segmentation of multiresolution remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2129–2141, 2009.
- [5] H. G. Akcay and S. Aksoy, "Automatic detection of geospatial objects using multiple hierarchical segmentations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2097–2111, 2008.
- [6] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 561–576, 2000.
- [7] S. Mallat, "Wavelets for a vision," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 604–614, 1996.
- [8] F. Yuying, D. Peijun, and Z. Guangyun, "Hierarchical urban object extraction from high resolution remotely sensed imagery based on wavelet transform and support vector machine," in *Proceedings of the Joint Urban Remote Sensing Event*, 2009, pp. 1–6.
- [9] W. Sun, V. Heidt, P. Gong, and G. Xu, "Information fusion for rural land-use classification with high-resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 4, pp. 883–890, 2003.
- [10] M. J. Barnsley and S. L. Barr, "Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system," *Computers, Environment and Urban Systems*, vol. 21, no. 3, pp. 209–225, 1997.