3 EO Data Processing and Interpretation for Human Settlement Characterization

A Really Global Challenge

Paolo Gamba, Gianni Lisini, Gianni Cristian Iannelli, Inmaculada Dopido, and Antonio Plaza

CONTENTS

3.1 Introduction 35
3.2 EO Data and Urban Areas 36
3.2.1 From EO Data to Urban Features 36
3.2.2 From EO Data to Scalable Models 37
3.3 Open Global Challenges 41
3.3.1 Finding the Right Bit 42
3.3.2 Fusing Spaceborne, Airborne, and Ground Data 43
3.4 Conclusions 46
Acknowledgments 46
References 47

3.1 INTRODUCTION

The need for increasingly accurate models of the complex interactions between mankind and the environment calls for more precise monitoring of many different areas, from forests to oceans, from inland waters to urban areas. Specifically, human settlements appear to be the focus of a number of issues such as desertification and pollution as well as water, energy, and waste management. Since most of the population nowadays lives in urban areas, threats to human lives, such as diseases and man-made and natural disasters, are increasingly perceived as the causes of social and economic losses in urban areas. Using urban areas and some of their specific features as essential input information, scientists and researchers have developed models for climate change [1], earthquake risk [2], disease spread [3], and many others. To achieve this aim, global analyses, including more information than just the knowledge of urban

area locations, are mandatory. And yet the latest global data sets on urban areas are still incomplete or limited to the aforementioned sources of information. These involve spatial scales that are not completely useful to address intra-urban activities.

The task of extracting and managing urban datasets at different scales by using Earth Observation (EO) data is one of the great challenges of remote sensing data interpretation. However, many of the most promising techniques that have been proposed have been applied to small or limited data sets so far; this was originally due to the limited amount of available data, but currently it is related more to the complexity of designing data analysis procedures for multiple sensors at multiple spatial and spectral resolutions. Moreover, the finest spatial resolution available from EO sensors does not fit the requirements of all urban studies, and VHR data, with all their details, may be less suited to tasks like urban land use mapping. Finally, from a global perspective, issues come from the huge amount of data available as well as the need for an efficient methodology to extract useful information with a consistent approach in different geographical areas.

This chapter initially provides an overview of existing methodologies to address some of these issues as well as the challenges related to the implementation/realization of these methodologies. To explain the sort of "high-level" message included in these pages, a few examples from existing research by the authors are included.

The remainder of the chapter is organized as follows. Section 3.2 describes a general procedure to derive scalable models from EO data and urban features. Section 3.3 discusses several challenges still open for research in the aforementioned procedure. Section 3.4 concludes with some remarks.

3.2 EO DATA AND URBAN AREAS

This section provides an overview of available mechanisms to derive relevant, scalable models from EO data using features collected in urban environments. This includes a description of success cases and processing chains under different application scenarios. Remaining challenges are outlined in Section 3.3.

3.2.1 From EO Data to Urban Features

The standard approach to EO data exploitation in urban area characterization is to extract man-made features and use them as basic elements. By exploiting and combining these elements, more detailed analyses can be obtained. This is the case, for instance, in urban extent delineation—starting from elementary spatial patterns [4], road network extraction, grouping road candidates [5], or building two-dimensional and three-dimensional "builtscape" characterization to clustering 3D primitives [6]. The challenge in this type of approach is to design a sufficiently flexible algorithm, able to adapt to the multiple and different ways that these features may appear in EO data (particularly, in different geographical areas), and to recognize significant clusters as hints or proxies to urban elements. However, once the information is available, it may be exploited for a number of different applications, thus providing inputs to multiple models.

The main limitation of this processing framework is that the basic urban/artificial features are extracted at a given scale so that their use for multiple scale models is

not always immediate. For example, after 2D building information is extracted, a thermal model for each building may be considered, but a thermal model for a whole city would require some sort of reprocessing and clustering of the extracted data. Similarly, and as another example, road network extraction may be used for traffic modeling, mobility management, or noise pollution monitoring and prevention. To be useful, however, a road network extraction approach should be fast, efficient, and precise enough in addition to providing an output as scalable as possible.

To improve over existing approaches and propose a unitary framework, the methodology discussed in [7] may be considered. Specifically, the idea is to include in the information extraction from an urban scene many different features corresponding to multiple scales. These features can be used either to logically and spatially cluster them into elements at a lower (coarser) scale or to infer other features at a higher (finer) scale. If each feature extraction algorithm is designed within such a framework, it can incorporate enough flexibility to combine different features at different scales and thus simultaneously obtain more information. Similarly, this framework may be used to design techniques that can be tailored to work in multiple geographical areas.

3.2.2 From EO Data to Scalable Models

A different and recently considered way to exploit EO data is related to the direct extraction of model-related features, less generic and more connected to the local/regional (and, eventually, global) models required by current studies. With respect to climate change, for instance, the thermal behavior of urban areas has been actively investigated [8] but without a strong link to global climate models. Similarly, risk analysis is a very important topic related to the impact of natural disasters that may occur in and around human settlements [9]. As shown in the aforementioned examples, on a city level (or more detailed scale), atmospheric circulation analyses, involving urban meteorological models as well as risk computations including physical and social vulnerability, have already been considered for one or more cities. They still need to be tuned and validated on a global scale as opposed to a case by case approach. As a preliminary step in this direction, there is a need to characterize every urban area at a global scale, according to land use/land cover typologies, which are peculiar to different environmental or risk models. This is the case of "urban climate zones" (areas with the same microclimatic behavior) for urban meteorology [10] or "uniformly built dwellings" (areas with buildings that have the same structural typology) for earthquake vulnerability [23].

Specifically, urban climate zones represent a comprehensive classification system for characterizing the urban environment with respect to urban meteorology, as reported in [10,11]. The same classification scheme was applied to a different environment in [12]. These works introduced the concept of "thermal climate zones" or "local climate zones," defined as regions with relatively uniform surface–air temperature distribution across different horizontal scales [10]. These climate zones can be differentiated by means of multiple characteristics from the urban 2D and 3D landscapes such as the built surface fraction, the building height-to-width ratio, the sky view factor (percentage of sky visible from the ground), the height of roughness elements, the anthropogenic heat flux, and the surface thermal admittance.

Most of these characteristics, ultimately connected to physical characteristics of the urban objects, can be extracted from remote sensing data.

The methodology reported here is based on two different processing chains. The first one is devoted to the extraction of spatially homogeneous urban areas within the scene, which may be labeled as "block." The idea is that these blocks may be then assigned, using the second part of the procedure, to one of the urban climate zone classes by considering a suitable combination of spatial and spectral indexes.

The first processing chain can be subdivided into two subsequent steps. The first one is the identification of the human settlement (as opposed to all the other land use classes in the area). To achieve this, we use the PanTex index proposed in [13], which proved to be effective to extract human settlement extents starting from panchromatic images at 2.5 m spatial resolution. The area identified as human settlement is further segmented into a homogenous zone using a spanning tree reduction scheme [14]. Alternatively, a more complex approach based on a combination of geometrical features into closed boundaries can also be used [15].

The employed processing chain, aiming at classifying each homogenous area into an urban climate zone class, is based on the joint analysis of a few indexes that, we feel, may capture most of the features listed in the introduction. We assume that a multiscale version of the same index used for urban area detection may be useful as it helps in enhancing spatial patterns at multiple geographical scales. To obtain a multiscale PanTex, the same textural feature (contrast) used to build the original index is now computed with different lag distances (which is equivalent to assuming a different spatial resolution of the data). Additionally, the original image and the results of an edge extraction technique (implemented using a Sobel filter) are included to insert spectral and edge density information, respectively.

Using these indexes, a decision tree classifier is designed using training data and is eventually applied to the whole data set. The decision tree structure used to label the segmented blocks and assign them to the different climate zones is obtained by a detailed analysis of a small sample of the blocks in the first test case described in the next section. Although this approach is apparently biased by a specific city structure and location, the same rules apparently work in different locations, as also discussed in the following paragraphs. The main rationale is that these rules refer to spatial indexes, which in turn describe quantitatively the spatial structure of the different parts of a town.

The decision tree, tuned with empirical tests, accepts as inputs three images:

1. The original image (OR)
2. The PanTex filter output with a kernel of 5×5 pixels applied at the full scale data (P1)
3. The PanTex filter output applied to a subsampled data set at 5 m/pixel (P5)

Experimental results were obtained on August 12, 2008 from a scene by the ALOS PRISM sensor with a spatial resolution of 2.5 m and depicting a portion of the town of Xuzhou in the Jiangsu province, People's Republic of China. The challenge, as highlighted in the previous section, was to use 2D data without spectral information to obtain spatial indexes allowing an analysis of different zones and their classification into thermal climate zones. Results depended on both segmentation and accuracy.

An incorrect segmentation may result in less precise classification of the urban blocks as the spatial indexes used by the decision tree are averaged for each block. However, the rules defined for the decision tree are more important because they allow assigning each block to a climate zone, once its spatial boundaries have been individuated by the segmentation step. For this reason and because the segmented urban image can be obtained by various means—for instance, by using available geographic information system (GIS) layers for a town—the evaluation described in the following paragraph will focus mostly on the second step of the procedure, without paying much attention to the approach used to achieve a correct segmentation.

For the test area, only five urban climate zones were considered, that is, those that were present in the scene settlements: "open set mid rise," "compact low rise," "open set low rise," "dispersed low rise," and "extensive low rise." By applying the aforementioned procedure to a first set of urban blocks, shown in Figure 3.1 together with the corresponding color legend, a relatively high overall accuracy at the object level is achieved (81%).

(a)

(b)

Typology	Color
Open set mid rise	Red
Compact low rise	Blue
Open set low rise	Green
Dispersed low rise	Yellow
Extensive low rise	Purple

(c)

FIGURE 3.1 **(See color insert.)** Experimental results for the urban climate zone extraction in a small subsample of the Xuzhou (People's Republic of China) scene: (a) the urban climate zone map to be compared with (b) a ground truth obtained by visual classification and superimposed on the original data set. Classes are identified by colors, according to the legends displayed in (c).

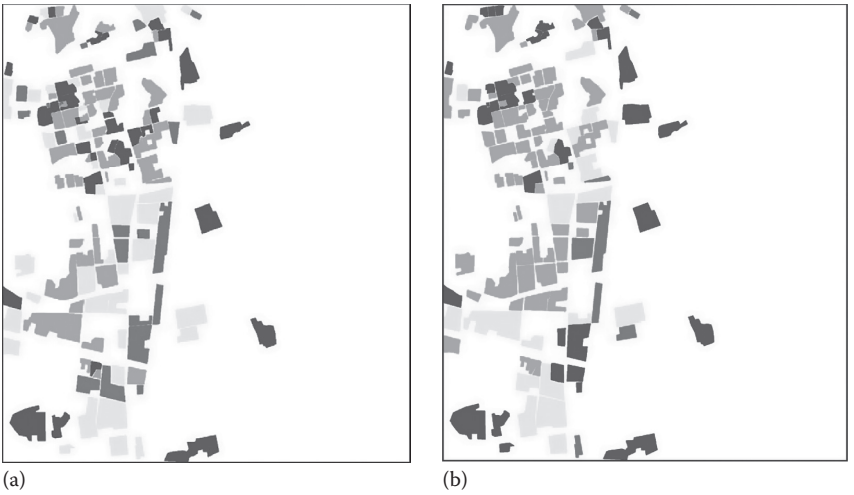

FIGURE 3.2 **(See color insert.)** Climate zone extraction for the town of Xuzhou, People's Republic of China: (a) final results of the proposed segmentation and classification procedure and (b) detailed ground truth obtained by manually delineating and labeling individual blocks.

In the previous example, the blocks were obtained manually as our focus was on the definition of the decision rules to be applied for block labeling. The same approach was however applied to the whole urban area, after performing an automatic segmentation, and the results are depicted in Figure 3.2a. These results should be compared with the detailed ground truth in Figure 3.2b. Although the color patterns appear visually similar, the overall accuracy at the block level is about 51% if computed regardless of the block size. Overall accuracy at the pixel level instead reaches 63%. The worst discrimination is achieved between the "open set low rise" and "dispersed low rise" classes. This may be due to the fact that the two typologies are very similar considering only two texture scales. Another option considering multiple spatial scales would consist of using differential attribute profiles [16].

One important consideration on these numbers is that the ground truth maps were not obtained by a meteorologist but by a remote sensing specialist using the panchromatic band only. Accordingly, we do not expect that the ground truth would be 100% accurate. In other words, a better validation procedure (including feedback from local experts) may be required.

To further illustrate the aforementioned observations, Figure 3.3 shows a panchromatic and the corresponding pansharpened images of a small portion of the area. The two additional images in this figure correspond to two different visual assessments of the urban climate zones made by two different experts and using the same color legend as in Figure 3.1. It is clear that the panchromatic image does not allow an easy discrimination between the classes, "open set mid rise" and "open set low rise." The color image may provide hints for discrimination, but these are connected to an a priori knowledge of building typologies and, thus not easily generalizable.

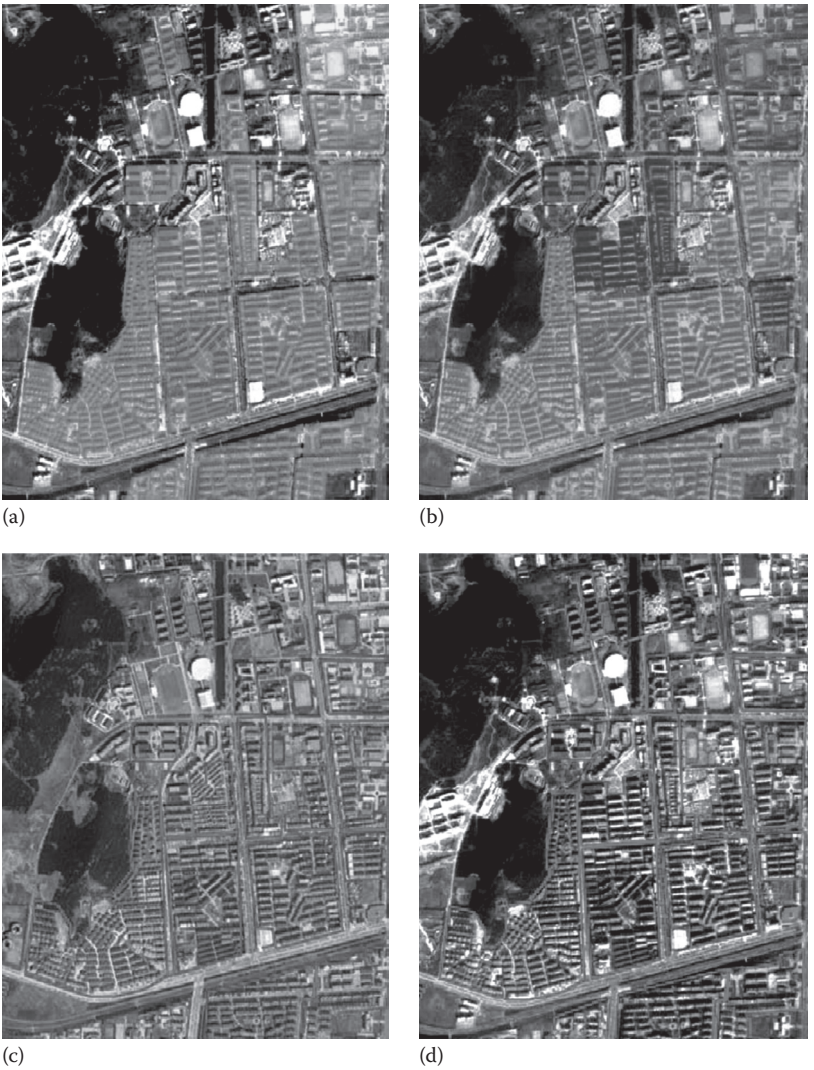

FIGURE 3.3 **(See color insert.)** Example of problematic assignment of urban blocks to urban climate zones: (a, b) two different visual interpretations by remote sensing experts, to be compared with the block borders superimposed on (c) the panchromatic and (d) the color image of the same area.

3.3 OPEN GLOBAL CHALLENGES

The few examples discussed in the previous section show that there are plenty of challenges still open for research. With the huge amount of EO data sets available and the need to consider existing information, many of these challenges can be grouped under two separate needs (1) to select and (2) to fuse the more relevant bits of information. Accordingly, and following the usual sequence of steps, feature extraction–feature

selection–feature fusion (but considering the need for multiple scales and a model-based output), urban remote sensing currently faces very interesting challenges.

3.3.1 Finding the Right Bit

Feature extraction and selection is a very important part of EO data exploitation to globally characterize urban areas. As mentioned in [7], a very promising methodology for the selection of urban scene features is the use of active learning approaches, which allow exploiting both the spectral and the spatial information in urban areas, thus enabling determination of the context that better characterizes a specific location.

Following the methodology introduced in [17], for instance, it is possible to develop a novel approach to perform semisupervised classification of urban hyperspectral images by exploiting the information retrieved with spectral unmixing. This is because many pixels in remotely sensed images are "mixed," that is, given by a combination of different substances that reside at the subpixel level. Within this framework, active learning techniques can be used for automatically selecting unlabeled samples in a semisupervised fashion. Specifically, the active learning approach in [17] selects highly informative unlabeled training samples in order to enlarge the initial (possibly very limited) set of labeled samples and perform semisupervised classification based on the information provided by well-established discriminative classifiers.

The proposed approach consists, therefore, of three main ingredients: semisupervised learning, spectral unmixing, and active learning.

1. For the semisupervised part of our approach, the multinomial logistic regression (MLR) classifier [18] provides probabilistic outputs, which play an essential role in our active learning process. Furthermore, a sparsity-inducing prior is added to the regressors to obtain sparse estimates. As a result, most of the components of the regressors are zero. This allows controlling the complexity of the proposed techniques and their generalization capacity. Finally, we use LORSAL algorithm [19] to learn the MLR classifier as it is able to learn the posterior class distributions directly and deal with the high dimensionality of hyperspectral data in a very effective way. This is very important for semisupervised learning since, ultimately, we would like to include as many unlabeled samples as possible, a task which is difficult for normal algorithms from the viewpoint of computational complexity.
2. The unmixing strategies considered in the second step include those attempting to consider spatial information within the extraction procedure. The first one is the fully constrained linear spectral unmixing (FCLSU), which first assumes that labeled samples are made up of spectrally pure constituents (endmembers) and then calculates their abundances and provides a set of fractional abundance maps (one per labeled class). An alternative approach is mixture tuned matched filtering (MTMF), which also assumes that the labeled samples are made up of spectrally pure constituents (endmembers) but then calculates their abundances by means of the MTMF method, which is a hybrid between target detection and unmixing,

thus providing a set of fractional abundance maps (one per labeled class) without the need to know the full set of endmembers in the data.

3. The third ingredient of our proposed method consists of using active learning to improve the selection of unlabeled samples for semisupervised learning. In our proposed strategy, the candidate set for the active learning process (based on the available labeled and unlabeled samples) is inferred using spatial information (specifically, by applying a first-order spatial neighborhood on available samples) so that high confidence can be expected in the class labels of the obtained candidate set. This is similar to human interaction in supervised active learning, where the class labels are known and given by an expert. In a second step, we run active learning to select the most informative samples from the candidate set. This is similar to the machine interaction level in supervised active learning, where in both cases the goal is to find the samples with higher uncertainty. Due to the fact that we use a discriminative classifier (MLR) and spectral unmixing techniques, active learning algorithms, which focus on the boundaries between the classes (which are often dominated by mixed pixels), are preferred. This way, we can combine the properties of the probabilistic MLR classifier and spectral unmixing concepts to find the most suitable (complex) unlabeled samples for improving the classification results through the selected active learning strategy. It should be noted that many active learning techniques are available in the literature [20]. In this work, we use the well-known breaking ties (BT) [21] to evaluate the proposed approach. This algorithm finds the samples minimizing the distance between the first two most probable classes.

Results for the hyperspectral ROSIS Pavia dataset (13 m spatial resolution, 610 × 340 pixels, 103 spectral bands, 9 ground-truth classes [22]) are shown in Figure 3.4. The use of BT alone leads to a mapping result (see Figure 3.4b) with an overall accuracy of 75.5%, definitely larger than the one achievable considering a standard supervised approach (63.6%). The joint use of unmixing information further improves this result, reaching an accuracy value of 79.3% in case FCLSU is used (see Figure 3.4c) while MTMF has a slightly worse performance (79.1%).

3.3.2 Fusing Spaceborne, Airborne, and Ground Data

The process of using EO data to characterize urban areas cannot avoid the fact that, in urban areas, much information will increasingly be collected and stored. Accordingly, the challenge is to include and combine the relevant existing information with the EO extracted features to obtain the multiscale model input mandatory for global models. In doing so, spaceborne remotely sensed data should somehow be "fused" with available spaceborne data, GIS layers, as well as ancillary information collected on the ground (e.g., by means of sensors or sensor networks). As an example of this procedure, we provide here a quick introduction to the approach developed to map exposure within the Global Exposure Database for the Global Earthquake Model (GED4GEM) project [23].

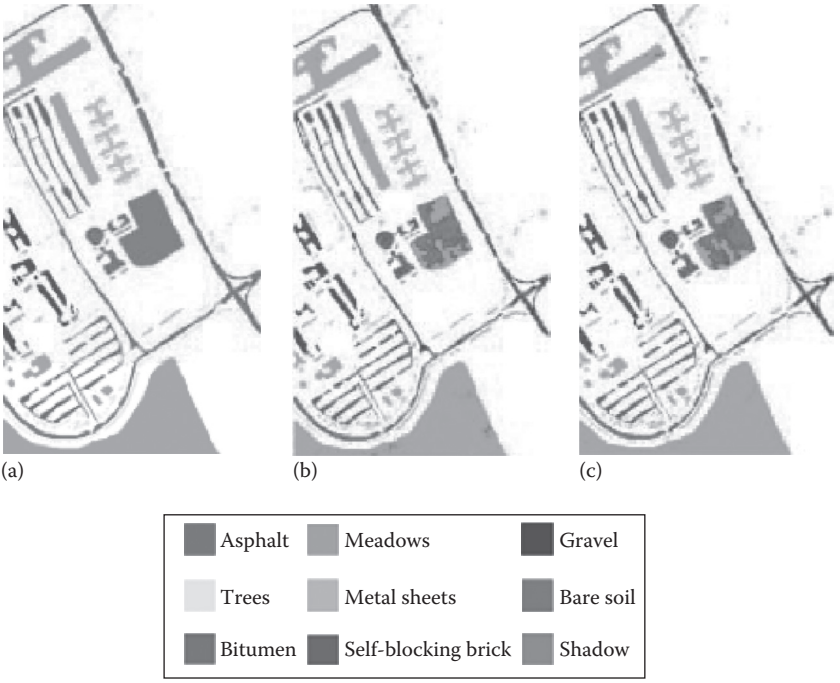

FIGURE 3.4 **(See color insert.)** Urban land use mapping results in Pavia, Italy: (a) ground truth for the hyperspectral ROSIS data sets, (b) mapping results using semisupervised classification but no unmixing information, and (c) mapping results jointly considering a semisupervised technique and unmixing information.

Obtaining building exposure is typically a problem related to the combination of information from multiple spatial scales. Indeed, a map of all buildings is currently out of reach using EO data only, either because we do not have a full geographical coverage or because the methods to extract building features (useful for identifying vulnerability) from EO data require inputs from multiple sensors, and this is not feasible for wide areas [9]. One possible solution for a globally suitable and manageable approach is to combine available EO information with ancillary data at different scales.

The idea is graphically represented in Figure 3.5, and includes the use of globally available EO data at coarse-resolution or existing maps to select the built-up areas to focus on, VHR EO data to extract building counts, GIS and census/survey data available from local/international databases to extract dwelling/building fractions according to building typologies, and a good deal of a priori knowledge to logically connect all these features.

The part of the methodology relying on EO data can be implemented according to a procedure for urban spatial pattern recognition. Specifically, the artificial composition of built-up structures and gaps among them (mostly, but not only, roads) is a general and globally valid assumption in all urban areas and usually results in a higher local contrast within these areas than in any natural environments. Accordingly, an option to

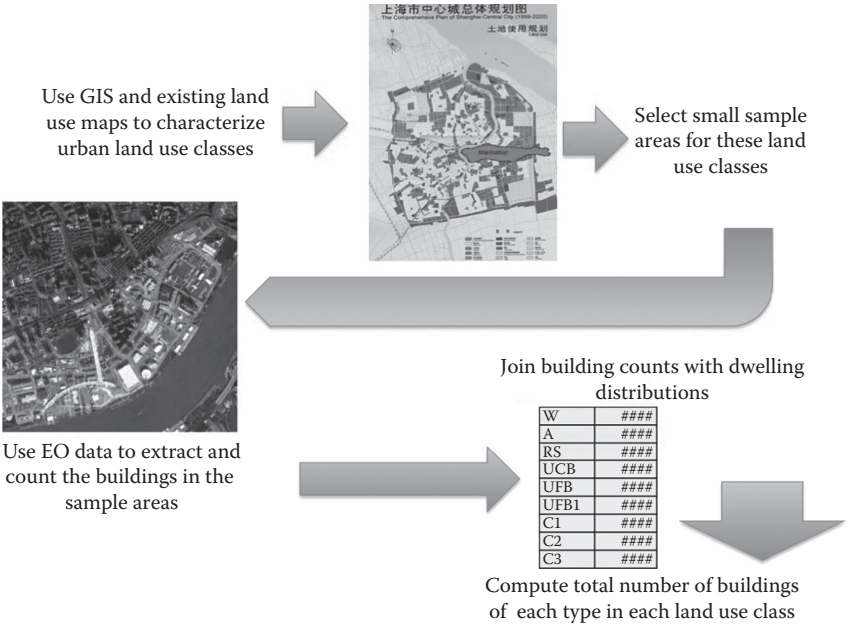

FIGURE 3.5 Graphical representation of the proposed procedure to obtain building exposure data and earthquake risk vulnerability data exploiting EO images, ancillary (GIS/map) information, and available ancillary data.

detect urban areas, widely used on EO analysis and already proved reliable at the global level for both optical (panchromatic) and SAR data, is the use of the textural features [13,24,25]. Here, we specifically refer to the *range* textural feature [26], computed starting from the occurrence matrix and defined as the difference between the maximum and the minimum value of the reflectance in a 5 × 5 pixel kernel moving over the image. After range extraction, a few postprocessing steps are performed, as shown in Figure 3.6.

As shown in Figure 3.5, another step in the exposure mapping procedure proposed in GED4GEM is building count extraction from VHR data, either for the whole area or for some samples. This step can be then performed in many

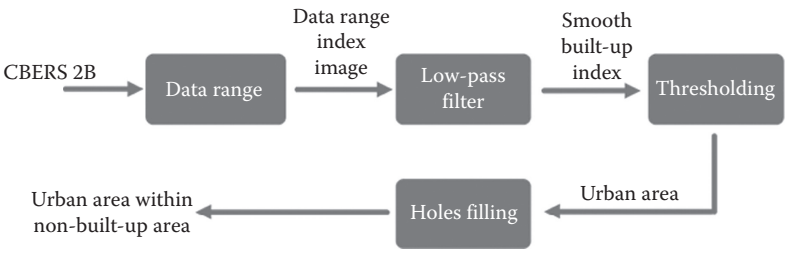

FIGURE 3.6 Data processing chain aimed at urban extent extraction from VHR panchromatic/SAR data.

different ways, according to the trade-off between accuracy and adaptivity to different geographical areas and building styles. A first option, as suggested in [24], is to classify the multispectral image available for the area of interest as delineated by the previous step and then segment the classification map into significant elements with the areas assigned into building classes. A second (less precise) option is to have a very rough estimate of the building counts by means of the same approach described in Figure 3.6 but this time, looking for the peaks of the *range* index [25]. The choice between these two options depends on the accuracy of the ancillary data (e.g., the dwelling distributions into building typologies) and their statistical significance. For instance, if the dwelling distribution is available at the country level, there is no point in precisely extracting building counts at a spatial resolution of a few meters.

3.4 CONCLUSIONS

This chapter briefly touches the huge range of challenges and opportunities brought by the use of EO data to monitor urban areas all around the world. These challenges are both in the data processing domain as well as in the domain of theoretical information extraction. Specifically, it has been made clear by a few examples that the use of data from satellites may improve our knowledge of urban areas and provide invaluable inputs to models that attempt to capture the interaction between artificial and natural environments. In many situations, these inputs must be at multiple scales, and EO-related information must be combined with data from other sources, according to the final application under study.

Moreover, the need to design algorithms and information extraction procedures that are valid at the global scale includes the necessity to address settlements and artificial elements of the landscape with very different spatial and spectral properties in different parts of the world. Distilling what is common to all these properties is one of the most serious issues to be considered.

Notwithstanding the many open issues highlighted by this work, this chapter includes a few examples of successful application of the multiple scale and multiple feature framework recently proposed in [7] for EO data information extraction in urban areas. Moreover, the efforts currently channeled through the Global Earth Observation working plan for 2011–2015 and, especially, the task related to the design and management of a Global Urban Observatory will help to provide some answers to the burning questions that are still open.

ACKNOWLEDGMENTS

The authors acknowledge support provided by the TOoLs for Open Multi-risk assessment using Earth Observation data (TOLOMEO) project, funded by the European Research Agency under the International Research Staff Exchange Scheme (Contract no. 269115) for some of the research efforts proposed in this chapter. Similarly, the support of the Global Earthquake Model initiative for risk-related analyses is gratefully acknowledged.

REFERENCES

1. E. Bartholome and A.S. Belward, GLC2000: A new approach to global land cover mapping from Earth observation data, *International Journal of Remote Sensing*, 26(9), 1959–1977, 2005.
2. A. Smolka, GEM—The Global Earthquake Model, *EGU General Assembly 2009*, Vienna, Austria, April 19–24, 2009, p. 4104.
3. C. Linard and A.J. Tatem, Large-scale spatial population databases in infectious disease research, *International Journal of Health Geographics*, 11, 7, 2012.
4. P. Gamba and M. Stasolla, Spatial indexes for the extraction of formal and informal human settlements from high resolution SAR images, *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, 1(2), 98–106, June 2008.
5. M. Negri, P. Gamba, G. Lisini, and F. Tupin, Junction-aware extraction and regularization of urban road networks in high resolution SAR images, *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), 2962–2971, October 2006.
6. Q.-Y. Zhou and U. Neumann, Complete residential urban area reconstruction from dense aerial LiDAR point clouds, *Graphical Models*, 75(3), 118–125, available online October 11, 2012. doi:10.1016/j.gmod.2012.09.001.
7. P. Gamba, Human settlements: A global challenge for EO data processing and interpretation, *Proceedings of IEEE*, 101(3), 570–581, March 2013.
8. Q. Weng, H. Liu, B. Liang, and D. Lu, The spatial variations of urban land surface temperatures: Pertinent factors, zoning effect, and seasonal variability, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(2), 154–166, 2008.
9. D. Polli, F. Dell'Acqua, and P. Gamba, First steps towards a framework for earth observation (EO)-based seismic vulnerability evaluation, *Environmental Semeiotics*, 2(1), 16–20, 2009.
10. I. Stewart and T. Oke, Classifying urban climate field sites by 'local climate zones': The case of Nagano, Japan, *Proceedings of the Seventh International Conference on Urban Climate*, Yokohama, Japan, 2009. Available online at http://www.ide.titech.ac.jp/~icuc7/extended_abstracts/pdf/385055-1-090515165722-002.pdf (last accessed in November 2013).
11. I. Stewart, Newly developed 'thermal climate zones' for defining and measuring urban heat island magnitude in the canopy layer, *Proceedings of the Eighth Symposium on Urban Environment*, Phoenix, AZ, 2009. Available online at http://ams.confex.com/ams/pdfpapers/150476.pdf (last accessed in November 2013).
12. I.C. Nduka and A.I. Abdulhamed, Classifying urban climate field sites by 'thermal climate zones' the case of Onitsha metropolis, *Research Journal of Environmental and Earth Sciences*, 3(2), 75–80, 2011.
13. M. Pesaresi, A. Gerhardinger, and F. Kayitakire, A robust built-up area presence index by anisotropic rotation-invariant textural measure, *IEEE Journal of Selected Topics in Earth Observations and Applied Remote Sensing*, 1(3), 180–192, 2008.
14. P.R. Marpu, I. Niemeyer, and R. Gloaguen, Unsupervised image segmentation by identifying natural clusters, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, *IGARSS'07*, Barcelona, Spain, July 23–27, 2007, pp. 1903–1904.
15. G. Lisini, F. Dell'Acqua, P. Gamba, and W. Thompkinson, Image interpretation through problem segmentation for very high resolution data, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, *IGARSS'05*, Seoul, South Korea, July 2005, pp. 5634–5637.
16. M. Dalla Mura, J.A. Benediktsson, B. Waske, and L. Bruzzone, Morphological attribute profiles for the analysis of very high resolution images, *IEEE Transactions on Geoscience and Remote Sensing*, 48(10), 3747–3762, 2010.

17. A. Plaza, I. Dopido, J. Li, and P. Gamba, Semi-supervised classification of hyperspectral data using spectral unmixing concepts, *Proceedings of the Tyrrhenian Workshop 2012 on Advances in Radar and Remote Sensing*, Naples, Italy, September 2012, unformatted CD-ROM.
18. D. Boehning, Multinomial logistic regression algorithm, *Annals of the Institute of Statistical Mathematics*, 44, 197–200, 1992.
19. J. Bioucas-Dias and M. Figueiredo, Logistic regression via variable splitting and augmented Lagrangian tools, Instituto Superior Tecnico, TULisbon, Portugal, Technical Report, 2009.
20. D. Tuia and G. Camps-Valls, Urban image classification with semisupervised multiscale cluster kernels, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1), 65–74, March 2011.
21. T. Luo, K. Kramer, D.B. Goldgof, S. Samson, A. Remsen, T. Hopkins, and D. Cohn, Active learning to recognize multiple types of plankton, *Journal of Machine Learning Research*, 6, 589–613, 2005.
22. P. Gamba, A collection of data for urban area characterization, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, *IGARSS'04*, Anchorage, AK, September 2004, Vol. I, pp. 69–72.
23. P. Gamba, K. Jaiswal, D. Cavalca, C. Huyck, and H. Crowley, The GED4GEM project: Development of a global exposure database for the Global Earthquake Model initiative, *Proceedings of the 15th World Conference on Earthquake Engineering*, Lisbon, Portugal, September 2012, unformatted CD-ROM.
24. T. Esch, H. Taubenböck, A. Roth, W. Heldens, A. Felbier, M. Thiel, M. Schmidt, A. Müller, and S. Dech, TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns, *Journal of Applied Remote Sensing*, 6(1), 1–21, 2012.
25. P. Gamba, M. Aldrighi, M. Stasolla, and E. Sirtori, A detailed comparison between two fast approaches to urban extent extraction in VHR SAR images, *Proceedings of JURSE 2009*, Shanghai, China, May 22–24, 2009, unformatted CD-ROM.
26. F. Dell'Acqua, P. Gamba, and G. Lisini, Rapid mapping of high resolution SAR scenes, *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5), 482–489, September 2009.