DYNAMIC SEGMENTATION FOR IMAGE INFORMATION MINING

Giuseppe Masi, Raffaele Gaetano, Giuseppe Scarpa, Giovanni Poggi

University "Federico II" of Naples, Italy
Department of Biomedical, Electronic and Telecommunication Engineering
Via Claudio 21, 80125, Naples, Italy
{firstname.lastname}@unina.it

1. PROBLEM STATEMENT

The last thirty years have witnessed the deployment of a large number of remote sensing systems of widely different nature and characteristics, based on optical and microwave technology, satellite or air-borne, etc., which gather more and more data by the day. Therefore, the problem of how to exploit effectively and efficiently such a wealth of data is becoming very urgent, all the more so considering the advances in sensor technology which make now available images with spatial resolution well below one meter. It is worth underlining that increasing spatial resolution does not just pose a problem of scale (not trivial by itself), but it also requires a full change of paradigm in data modeling, and the need to enrich the acquired information with new features from their very early stages of existence. In short, the main focus in remote sensing is definitely shifting from data acquisition to data management, including in this generic term all phases of data life, with special emphasis on image mining and retrieval which stand at the core of efficient user-oriented systems.

As a matter of fact, there is an intense research activity in this field, as testified by several IIM/EO (Image Information Mining for Earth Observation) systems recently proposed in the literature. A good example is the Geospatial Information Retrieval and Indexing System (GeoIRIS) by Shyu at al. [1] which, relies on the creation of a feature database with an indexing structures used to process the queries. Once the features (spectral, textural, or object-based) are computed, stored and indexed, the original data are no more used for answering any query. Other systems (a survey of CBIR systems can be found in [2]) resort to a similar data management approach. All of them, therefore, keep the image description at the pixel-level, although much enriched with many valuable features, making no effort to fill the semantic gap between the acquired data and the user needs. The user is hence required to carry out his/her own algorithms, including typically some forms of segmentation, in order to extract more high-level synthetic information on the observed image.

This choice, though unsatisfactory, is justified by the tough challenge represented, in a general purpose system, by the segmentation process, which is required to deal with wildly different scenarios depending on sensor technology, spatial and spectral resolution of images, information content of the scene and, of course, application goals. A reliable segmentation step would provide a significant added value to any such system, but no single algorithm can be expected to manage well all these situations.

2. PROPOSED SOLUTION AND PROOF OF CONCEPT

Following this line of reasoning, it is natural for us to focus on a *dynamic* segmentation approach, where the overall problem is split in elementary tasks, each of which can be addressed by a different specific algorithm which takes into account the peculiar characteristics of the current step. This general framework would then accommodate the wild variety of approaches and techniques proposed for specific problems, ranging from spectral-based clustering algorithms to the most complex knowledge-based methods, each one contributing to solve a different level of the problem.

The need to define a dynamic segmentation strategy comes from our experience with several types of segmentation algorithms and their sometimes disappointing results. Techniques based on the random Markov fields model [3], which used to work very well on low-resolution images, provided instead poor results on high-resolution data. Likewise, our recent texture-based technique [4, 5], which succeeds in telling apart large complex regions based on their spatial and contextual properties, turns out to be quite inaccurate when going into fine details.

Beyond the disappointment, such results underline the fact that no single segmentation algorithm can address tasks that differ profoundly under so many possible respects. If one aims at providing a general segmentation framework, many different tools must be available for use in subsequent passes of the process, different source of information (including different resolutions) must taken into account, and an accurate tuning of parameters must be carried out.

As an example, in a first step one can aim at isolating some large homogeneous areas of the image, for which a low-resolution version of the data might be used and feed to a simple spectral-based clustering algorithm. Such a step divides the image in several fragments/classes, then, each new fragment can be further segmented if necessary in a recursive fashion. Homogeneous regions may be considered as atomic, but complex regions, such as those corresponding to urban areas, or mixed urban-rural, should be further analyzed. A further step could then be, for example, the extraction of the road network, for which dedicated algorithms have been proposed in the literature [6]. Other segmentation engines might work instead on a high-resolution version of the image in order to single out elementary objects, like buildings (in which case a strong prior knowledge can be invoked [7, 8]), or trees, or other relevant structures. Needless to say, such a process would benefit from side information conveyed on the scene by Geographic Information Systems, which could be used, for example, to select the segmentation engines to be used and their sequence.

In this abstract we will present just a proof of concept of the proposed approach resorting to quite a limited set of segmentation engines, an MRF-based contextual technique [3] which we will call here basic-TFR because of a few later refinements, and the Texture fragmentation and reconstruction (TFR) technique [4, 5], based on texture modeling and recognition, improved in its turn by including a geometry-based processing step [9], and hence called advanced-TFR. The reader is referred to the references for all details.

We use these two techniques for the segmentation of a very-high resolution image (Fug.1a) of the city of San Francisco (USA). Based on prior knowledge of the scene, possibly gathered by a GIS, it seems reasonable to carry out a first step in which ocean and land are divided. To this end we resort to basic-TFR, which does a pretty good job (see Fig.1b) despite the fact that even the sea region is not so homogeneous, and in fact is obtained as the merging of several elementary classes. In this step, a low-resolution version of the image was considered (4x4 subsampling), using the full-resolution image only for local refinements. Following the dynamic segmentation philosophy, the second step concentrates on just on of the two fragments found before, the land region, which is quite complex and does not lend itself to an easy subdivision in large areas. For this task we resort to the texture-based TFR, which solves in its turn several subtasks, performing first a color-based initial fragmentation (Fig.1c), then devoting special attention to potentially disrupting background fragments [9] which are associated to a region-adjacency graph (Fig.1d) and further fragmented based on geometric properties (Fig.1e), and finally merging all textural fragments until only two regions are obtained (Fig.1f). As a result, the land region has been segmented pretty accurately in two subregions with clearly (for a human being) different characteristics, say, residential and industrial.

Of course, the process is meant to go on to further segmentation, using ad hoc engines, more complete results will be available at the conference.

3. REFERENCES

[1] Chi-Ren Shyu, Matt Klaric, Grant J. Scott, Adrian S. Barb, Curt H. Davis, and Kannappan Palaniappan, "Geoiris: Geospatial information retrieval and indexing systemcontent mining, semantics modeling, and complex queries," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 4, pp. 839–852, April 2007.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Mechine Intelligence*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[3] C. D'Elia, G. Poggi, and G. Scarpa, "A tree-structured Markov random field model for Bayesian image segmentation," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1259–1273, October 2003.

[4] G. Scarpa, M. Haindl, and J. Zerubia, "A hierarchical finite-state model for texture segmentation," in *In Proc. of IEEE ICASSP'07*, Honolulu, HI (USA), April 2007, vol. 1, pp. 1209–1212.

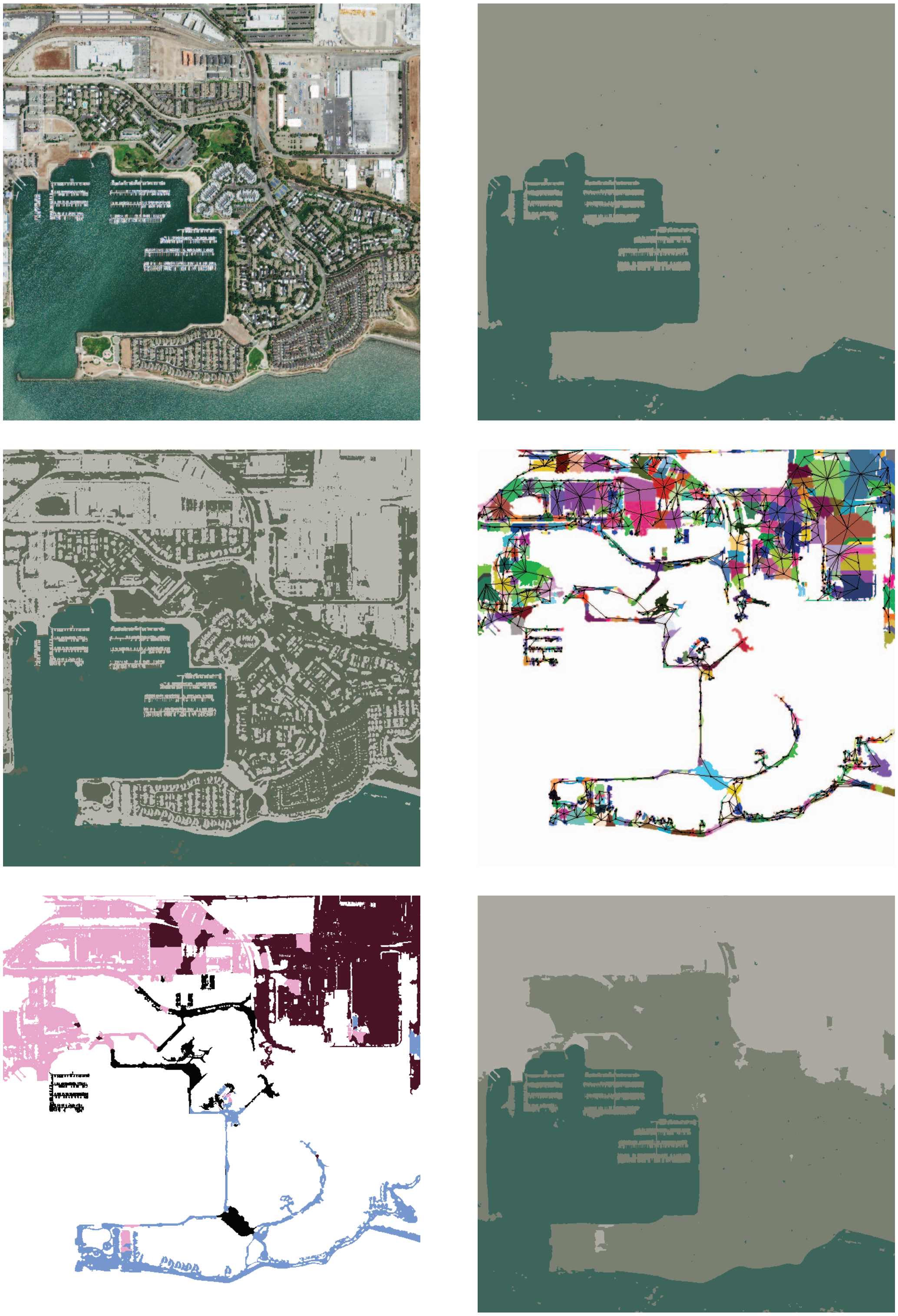
[5] G. Scarpa, R. Gaetano, M. Haindl, and J. Zerubia, "Hierarchical multiple markov chain model for unsupervised texture segmentation," *IEEE Transactions on Image Processing*, pp. 1–14, 2009, to appear.

[6] E. Binaghi, I. Gallo, and M. Pepe, "A cognitive pyramid for contextual classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2906–2922, December 2003.

[7] P. Gamba, F. Dell'Acqua, G. Lisini, and G. Trianni, "Improved VHR urban area mapping exploiting object boundaries," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2676–2682, August 2007.

[8] H.G. Akçay and S. Aksoy, "Morphological segmentation of urban structures," in *4th IEEE GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Paris (France), April 11-13 2007.

[9] Raffaele Gaetano, Giuseppe Scarpa, and Giovanni Poggi, "Advances in texture-based segmentation of high resolution remote sensing imagery," in *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, 2009.



(a) **(b)**

(c) **(d)**

(e) **(f)**

Fig. 1. Test experiment: (a) remotely sensed GeoEye-1 image portraiting an area of San Francisco, 3000×3000, 0.5m geometric resolution; (b) binary segmentation by basic TFR; (c) intermediate color-based binary segmentation of the "ground" class, operated at second round of TFR; (d) a background region decomposed in the Geometric Level with superposition of related RAG; (e) background reconstruction; (f) 3-class segmentation at the end of the second TFR (advanced one) iteration.