A New Approach for the Morphological Segmentation of High-Resolution Satellite Imagery

Martino Pesaresi and Jon Atli Benediktsson, *Member, IEEE*

***Abstract*—A new segmentation method based on the morphological characteristic of connected components in images is proposed. Theoretical definitions of morphological leveling and morphological spectrum are used in the formal definition of a morphological characteristic. In multiscale segmentation, this characteristic is formalized through the derivative of the morphological profile. Multiscale segmentation is particularly well suited for complex image scenes such as aerial or fine resolution satellite images, where very thin, enveloped and/or nested regions must be retained. The proposed method performs well in the presence of both low radiometric contrast and relatively low spatial resolution. Those factors may produce a textural effect, a border effect, and ambiguity in the object/background distinction. Segmentation examples for satellite images are given.**

***Index Terms*—High-resolution satellite imagery, leveling, mathematical morphology, morphological segmentation.**

I. Introduction

In THIS paper, a new segmentation method is proposed. The proposed method uses the residuals of morphological opening and closing transforms based on a geodesic metric. The proposed approach may be considered analogous to region growing techniques. However, in contrast to the use of statistical local properties as in region growing approaches, the proposed method uses a pixel similarity rule based on the morphological characteristic of connected components in images. In the proposed approach the morphological residuals between the original grey-level function and the composition of a granulometry and an anti-granulometry by reconstruction are used to build a morphological profile function. Recent theoretical advances in mathematical morphology, such as the definitions of *leveling* and the *morphological spectrum*, form a theoretical framework for the formal definition of the morphological profile function. This function is interpreted as a fuzzy membership function related to a set of morphological characteristics of the connected components in the image. A labeling phase is then formalized using a decision rule based on the greatest value of the derivative of the morphological profile function.

The proposed approach is different from standard morphological segmentation approaches which are based on an edge-detection phase (watershed line extraction on a gradient image). The proposed method can be applied with both single-scale and multiscale approaches. Therefore, the original contribution of the paper is the definition of a morphological segmentation method, which avoids gradient computation and can be applied either to single-scale or multiscale image processing problems.

The proposed method is particularly well suited for the segmentation of complex image scenes such as aerial or fine-resolution satellite images. In segmentation of such scenes, very thin, enveloped and/or nested regions may have to be retained. Therefore, edge-detection based on gradient computation does not perform well for such scenes. The proposed method also performs well in the presence of both low radiometric contrast and relatively low spatial resolution, which may produce a textural effect, a border effect, and ambiguity in object/background distinction. All these factors are critical and lead to an instability effect if segmentation methods based on an edge-detection approach are applied.

The paper is organized as follows. In Section II, some basic concepts of the mathematical morphology approach to image analysis are reviewed. The general context that justifies the proposed method is discussed in Section III. Section IV contains the extension of the region growing approach to multiscale image segmentation. Application examples of segmentation processes applied to satellite images are given in Section V. Finally, conclusions are drawn in Section VI.

II. Image Processing by a Morphological Approach

Mathematical morphology is the name of a collection of operators based on set theory and defined on an abstract structure known as an infinite lattice. These operators were first systematically examined by Matheron and Serra in the 1960s and are an extension of Minkowski's set theory [1], [2]. Morphological operators include erosion, dilation, opening, closing, rank filters (including median filters), top hat transforms, and other derived transforms. These operations can be defined on binary or greyscale images in any number of dimensions. They can also be defined with Euclidean (isotropic) or non-Euclidean (geodesic) metrics.

The main application areas for the tools of mathematical morphology have been medical imaging, material sciences, and machine vision, where morphological transformations are particularly useful for image analysis and image enhancement. In the processing of satellite remote sensing data, mathematical morphology has until recently only had sporadic applications. There, it is mainly known for its binary and Euclidean operators that have been used for post-classification procedures like "salt-and-pepper" removal or other visual enhancement procedures.

Manuscript received October 28, 1999; revised May 2000.

M. Pesaresi is with the European Commission, Space Applications Institute, Environmental and Geo-Information Unit, Ispra, Varese, Italy (e-mail: martino.pesaresi@jrc.it).

J. A. Benediktsson is with the Department of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland (e-mail: benedikt@hi.is).

Publisher Item Identifier S 0196-2892(01)01159-7.

A. Morphological Transforms

1) Definitions and Notation: A complete lattice is a mathematical structure that can formalize an ordering relation, and the two basic operations infimum ($\wedge$) and supremum ($\vee$). For a particular set, the infimum is the greatest lower bound and the supremum is the smallest upper bound. Now let us consider transformations defined on a complete lattice.

A transformation ψ is called

1) *increasing* if and only if it preserves order,
2) *idempotent* if and only if $\psi\psi = \psi$,
3) *a morphological filter*, if it is increasing and idempotent.

Following a usual notation, G is an underlying digital grid of any type in the subset of $Z^n \times Z^n$. Also, $N_G(p)$ is the set of neighbors of a pixel p with respect to the grid G, and $\psi_N f(p)$ is the morphological transformation of $f(p)$ using N as a structuring element (SE). Since the subgraph of an n-dimensional image corresponds to a $n+$ one-dimensional (1-D) set, $n+$ 1-D SEs can be used to investigate the morphology of n-dimensional image structures. However, it is often recommended to use n-dimensional SEs in order to be independent from the image grey scaling, and in order to work with faster algorithms. These latter SEs are referred to as *flat* SEs because they are only of two dimensions in the case of two-dimensional (2-D) images. On the contrary, $n+$ 1-D SEs are called *volumetric*, *nonflat*, or *grey scale* SEs [3].

2) Euclidean Transforms: Let us assume that we have a flat structuring element that corresponds to the neighborhood $SE \equiv N_G(p)$. Then, the erosion ε_N of the grey level function using the structuring element N is defined by the infimum of the values of the grey level function in the neighborhood

$$\varepsilon_N f(p) = \{\wedge f(p') | p' \in N_G(p) \cup f(p)\}. \tag{1}$$

Dilation δ_N is similarly defined by the supremum of the neighboring values and the value of $f(p)$ as

$$\delta_N f(p) = \{\vee f(p') | p' \in N_G(p) \cup f(p)\}. \tag{2}$$

Classically, *opening* γ is defined as the result of erosion followed by dilation using the same SE

$$\gamma_N f(p) = \delta_N \varepsilon_N f(p). \tag{3}$$

Similarly, *closing* φ is defined as the result of dilation followed by erosion with the same SE

$$\varphi_N f(p) = \varepsilon_N \delta_N f(p). \tag{4}$$

One of the characteristics of opening and closing operators is that they erase structures that are smaller than the SE. If a greyscale image is interpreted as a topographical relief, then opening cuts peaks. In contrast, closing fills valleys that are smaller than the SE (have a thinner support). This effect can be observed by the computation of the residual between the filtered image and the original image in the *top hat* transform $\Gamma f(p)$

$$\Gamma f(p) = f(p) - \gamma_N f(p) = f(p) - \delta_N \varepsilon_N f(p) \tag{5}$$

and the *inverse top hat* (or bot-hat) transform $\Gamma' f(p)$

$$\Gamma' f(p) = \varphi_N f(p) - f(p) = \varepsilon_N \delta_N f(p) - f(p). \tag{6}$$

B. Different Metrics

Morphological transformations can either use a classical Euclidean metric or a non-Euclidean geodesic metric (with *geodesic* transforms and *reconstruction* morphological filtering). Reconstruction filters form an important class of morphological and connected filters [4], [5]. The reconstruction filters have been proven to be very useful for image processing [6] since they do not introduce discontinuities, and therefore, preserve the shapes observed in input images. In other words, the nonisotropic metric used in morphological transformations by reconstruction makes these transformations not sensitive to noise. Another effect of the nonisotropic metric is that their SE is adaptive and does not introduce much shape noise in both filtering and detection of structures.

1) Geodesic Transforms: Equations (3) and (4) use an isotropic metric where the shape and size of the SE is a constant for all points p in the domain of $f(p)$, independent of the structures present in the image. Basic dilation and erosion transforms can also be formalized with the notion of geodesic distance. Given a set X (or a *mask*), the geodesic distance between two pixels p and q is the length of the shortest path in X, which joins p and q. Let $X \subset Z^n \times Z^n$ be a discrete set and let $Y \subseteq X$. Then it is possible to define an elementary geodesic dilation (and similarly erosion) of Y inside X with an SE of minimal size; SE $= B$ (defined as only one step in the grid G). A standard dilation of size one followed by an intersection is defined by

$$\delta^X Y = (Y \oplus B) \cap X \tag{7}$$

where $(Y \oplus B)$ is the binary Euclidean dilation of the set Y using SE $= B$.

In the greyscale case, the geodesic greyscale dilation of $f(p)$ inside $f(x) \in X$ based on the elementary SE is the infimum of the elementary dilation of $f(p)$ and the value of $f(x) \in X$

$$\delta^X_{(1)} f(p) = \delta_{(1)} f(p) \wedge f(x). \tag{8}$$

For computational purposes, it is interesting to note that geodesic dilation of a given size n can also be obtained by iterating n elementary geodesic dilations

$$\delta^X_{(n)} Y = \underbrace{\delta^X_{(1)} \cdot \delta^X_{(1)} \cdot \delta^X_{(1)} \cdot \delta^X_{(1)} \cdot \delta^X_{(1)} \cdot \delta^X_{(1)} \cdot \cdots \cdot \delta^X_{(1)}}_{n.times}(Y). \tag{9}$$

By duality, similar observations can be made for the erosion transform.

2) Reconstruction: The *reconstruction* $\rho^X(Y)$ of X from $Y \subseteq X$ is obtained by the iterative use of an elementary geodesic dilation of Y inside X until idempotence is achieved [7]

$$\rho^X(Y) = \bigcup_{n \geq 1} \delta^X_{(n)}(Y) \Big| \delta^X_{(n)} = \delta^X_{(n+1)}. \tag{10}$$

For a greyscale image with a flat SE, the union operator in (10) is equivalent to a supremum operation. Then the greyscale *reconstruction* $\rho^I(J)$ of the image I (also called the *mask*) from J (also called the *marker*) is obtained by the iterative use of elementary greyscale geodesic dilations of J under I until idempotence is reached, i.e.,

$$\rho^I(J) = \bigvee_{n1} \delta^I_{(n)}(J) \left| \delta^I_{(n)} = \delta^I_{(n+1)}. \right. \tag{11}$$

The definition of greyscale *dual reconstruction* by erosion is similar to (11). Let I and J be two greyscale images defined on the same domain and such that $I \leq J$. The dual greyscale reconstruction $\rho^{*I}(J)$ of the image I from the image J is obtained by iterating the greyscale geodesic erosion of J above I until idempotence is reached, i.e.,

$$\rho^{*I}(J) = \bigwedge_{n\geq 1} \varepsilon^I_{(n)}(J) \left| \varepsilon^I_{(n)} = \varepsilon^I_{(n+1)}. \right. \tag{12}$$

C. Watershed Segmentation

Watershed line detection [8] is the main tool of mathematical morphology used for image segmentation. Watershed segmentation was introduced in image analysis by Beucher and Lantuéjoul [9] and defined mathematically by both Meyer [10] and Najman and Schmitt [11]. However, except for a few simple cases where the target object is brighter than the background or vice versa, watershed segmentation cannot be applied directly. Generally, the method is applied to images that have been transformed by a gradient-like operator based on a measure of the local slope of the grey level function. Watershed extraction generally means the thinning of a gradient image with a homotopic transformation. It also involves the detection of basins as regions and crest lines as boundaries for these regions. For these reasons, a watershed approach generally leads to finding the structures in an image based on an edge-detection strategy. The standard approach in watershed segmentation causes severe over-segmentation, which may be difficult to overcome. This over-segmentation is due on the one hand to the presence of irrelevant local minima and local maxima in the image. On the other hand, it is also due to the presence of texture effects derived by the spatial interaction between the size of the object in the scene and the spatial resolution of the sensor. Since watershed segmentation depends on gradient calculus, a common (linear) approach to overcome the over-segmentation effect is to filter the input image with a low-pass filter prior to gradient extraction. By that, the final number of regions is decreased, but all relevant high frequency (spatial) information will be lost.

The standard nonlinear solution to the over-segmentation problem was introduced by Meyer and Beucher [8]. Their solution is a marker selection followed with flooding of the relief formed by the gradient obtained from these markers. The marker detection is the main problem with this approach. If there is no external information available, the marker detection problem is generally solved by morphologically filtering (usually by geodesic closing) the gradient image. Then the filtered gradient is thresholded. In (4), closing is defined by morphological dilation followed by erosion in (4). Therefore, the watershed-plus-marker approach assumes that the local minima of the gradient, which are smaller (thinner) than the SE, are not relevant. The same applies to grey level edges with values less than a given threshold.

All the abovementioned approaches assume that the region of interest for detection is large and homogenous relative to the spatial and spectral resolution of the sensor. Consequently, these approaches are very hard to apply in segmentation of textured or very complex scenes. They also often lead to results that are not stable.

III. Comparison of Opening and Closing by Reconstruction

The segmentation of urban scenes detected by satellite sensors is an excellent example of the inadequacy of the large-and-homogeneous region approach. In satellite remote sensing, the spatial resolution of commercial sensors currently reaches 5×5 to 2×2 m^2 per pixel. For these resolutions, the heterogeneity (different adjacent materials and spectral response) and geometrical complexity (small objects, 3-D/shadow effects) of the urban scenes may produce texture effects in images for structures that can be one to two pixels wide. Urban applications may also require the detection of very thin, or complex, elongated, and nested structures. At a given sensor resolution where there is not a clear distinction between the object and the background, the standard approach to morphological segmentation may be inadequate. In those situations, any attempt to find some kind of an edge of a structure has as an effect on the production of "surfaces of edges," where most of the pixels are connoted as "border pixels."

It is well known that there are two fundamentally different strategies for image segmentation: edge detection and region growing. Even though the standard approach to morphological segmentation is dependent on edge-detection, it is possible to consider a different morphological approach to the segmentation problem. The idea here is to try to characterize image structures by their morphological intrinsic characteristics instead of using their boundaries. In a hypothetical approach using a morphological region-growing technique, the border of a detected structure can be of size zero, thus avoiding the aforementioned surface-of-edges problems in segmentation of complex imagery. Therefore, a structure or an "object" in the image could be a connected component (region of pixels) with the same characteristics, measured by some kind of a morphological operator.

It is a common practice to use the opening and closing transforms in order to isolate bright (opening) and dark (closing) structures in images, where bright/dark means brighter/darker than the surrounding structures in the image. In order to isolate structures with a thinner support than a given SE, a widely applied technique is to take the residuals of the opening, closing, and original images by the *top hat* (5) and *inverse top hat* (6) transforms. Consequently, the idea is to use a composition of opening and closing transforms in order to build a definition for the morphological characteristics.

A. Definition of Leveling

The composition of opening and closing operations by reconstruction has recently been formalized as *leveling*. Leveling was introduced in [12]–[14] starting from the combination of two spatially connected operators: *monotone planings* and *flattenings* [15]. Applications of leveling techniques include object-oriented image compression, but an important feature of leveling is that it can be applied to an original image producing outputs at different levels of simplification. Object-oriented image compression also takes advantage of the fact that after leveling an image only has flat connected components. Leveling techniques are well documented in the literature. Therefore, they are only introduced here briefly for clarification purposes.

From (3), (4), (11), and (12), it is possible to define opening and closing transforms by reconstruction. *Opening by reconstruction* is obtained by using (11) for the reconstruction of the erosion under the original image. Similarly, *closing by reconstruction* is obtained by using (12) for dual reconstruction of the dilation above the original image. Based on (3) and (11), opening by reconstruction can be defined by

$$\gamma_N^* f(p) = \rho^{f(p)}(\varepsilon_N f(p)) = \mathrm{Rec}(\varepsilon_N f,\ f). \tag{13}$$

In a similar fashion, closing by reconstruction can be defined based on (4) and (12) by

$$\varphi_N^* f(p) = \rho^{*f(p)}(\delta_N f(p)) = \overline{\mathrm{Rec}}(\delta_N f,\ f). \tag{14}$$

Opening and closing by reconstruction can be considered as lower-leveling liv^- (opening) and upper-leveling liv^+(closing) operations [13]. A function g is an upper-leveling of a function f if and only if for any couple of connected neighboring pixels $(p,\ q)$, the following applies:

$$liv^+ \colon g_p > g_q \Rightarrow g_q \geq f_q. \tag{15}$$

Similarly, a function g is a lower-leveling of a function f if and only if for any couple of connected neighboring pixels $(p,\ q)$, the following applies:

$$liv^- \colon g_p > g_q \Rightarrow g_p \leq f_p. \tag{16}$$

It is possible to show [14] that all connected components in the upper-leveling function are flat where $g > f$, while in the lower leveling, the same occurs where $g < f$. A function g is called a leveling of a function f if it is both an upper and lower leveling of f.

The algorithm proposed in Meyer [13] for the computation of leveling requires a preliminary definition of two disjoint sets X and Y. Those two sets delimit two partitions of the image, where either reconstruction by opening or reconstruction by closing of a marker is applied. This is achieved by the use of the minimum value in the destination lattice for the opening and the maximum value for the closing. Consequently, the leveling is defined by

$$g = \begin{cases} \gamma^* f \colon X \\ \varphi^* f \colon Y \\ f \colon \overline{X \cup Y}. \end{cases} \tag{17}$$

The requirement of preliminarily partitioning the image into two sets X and Y is the main problem with the leveling approach. That problem is solved in [13] with a marker-gradient flooding method. However, the solution in [13] has the aforementioned problems related to the edge-detection approach. Therefore, some alternative solutions are presented in the next sections.

B. Definition of Morphological Characteristics

Following the approach proposed here, a structure or an "object" in an image is defined as a connected component of pixels sharing the same morphological characteristics. An idea proposed here is to use the residuals obtained from the original image and its leveling for a measure of those characteristics. The simpler and intuitive taxonomy of morphological characteristics could, for a given spatial domain, be the set $T = \{\text{“}flat,\text{”}\ \text{“}concave,\text{”}\ \text{“}convex\text{”}\}$. It is referred to here as the local curvature of the grey level function surface, where a given SE determines the spatial domain.

In order to define the morphological characteristics, we propose an alternative to Meyer's leveling algorithm. We will base this alternative on fuzzy logic, thus avoiding the problem of the ex-ante partition of the two complementary sets X and Y in (17). Let the residuals between any opening by reconstruction (or closing by reconstruction) and their original function be interpreted as a measure of the relative brightness of the structure (or relative darkness). Then, one membership function $\widehat{\mu}$ can be written relative to the class "convex" and another membership function $\breve{\mu}$ relative to the class "concave"

$$\widehat{\mu} = f - \gamma^*(f) \tag{18}$$

$$\breve{\mu} = \varphi^*(f) - f. \tag{19}$$

The leveling algorithm can now be rewritten as a decision rule based on the greatest value of the membership function

$$g = \Psi(f) = \begin{cases} \gamma^* f \colon \widehat{\mu} > \breve{\mu} \Rightarrow f - \gamma^*(f) > \varphi^*(f) - f \\ \varphi^* f \colon \breve{\mu} > \widehat{\mu} \Rightarrow \varphi^*(f) - f > f - \gamma^*(f) \\ f \colon \widehat{\mu} = \breve{\mu} \end{cases}. \tag{20}$$

In this perspective, given an image with a grey level function f and an SE $\equiv N$, the segmented image s of the characteristic can simply be a tessellation of three different labels like "*convex*" $\widehat{k}$, "*concave*" $\breve{k}$, and "*flat*" $\overline{k}$ as

$$s = \Phi_N(f) = \begin{cases} \widehat{k} \colon \Psi(f) < f \\ \breve{k} \colon \Psi(f) > f \\ \overline{k} \colon \Psi(f) = f. \end{cases} \tag{21}$$

For this segmented image, pixels where the lower leveling is strictly lower than the original image f are labeled $\widehat{k}$, and pixels where the upper leveling is strictly greater than f are labeled as $\breve{k}$. Pixels labeled $\overline{k}$ have maintained the same value of f in both the upper and lower leveling. Therefore, those pixels have been indifferent to the erosion/dilation-reconstruction process with a

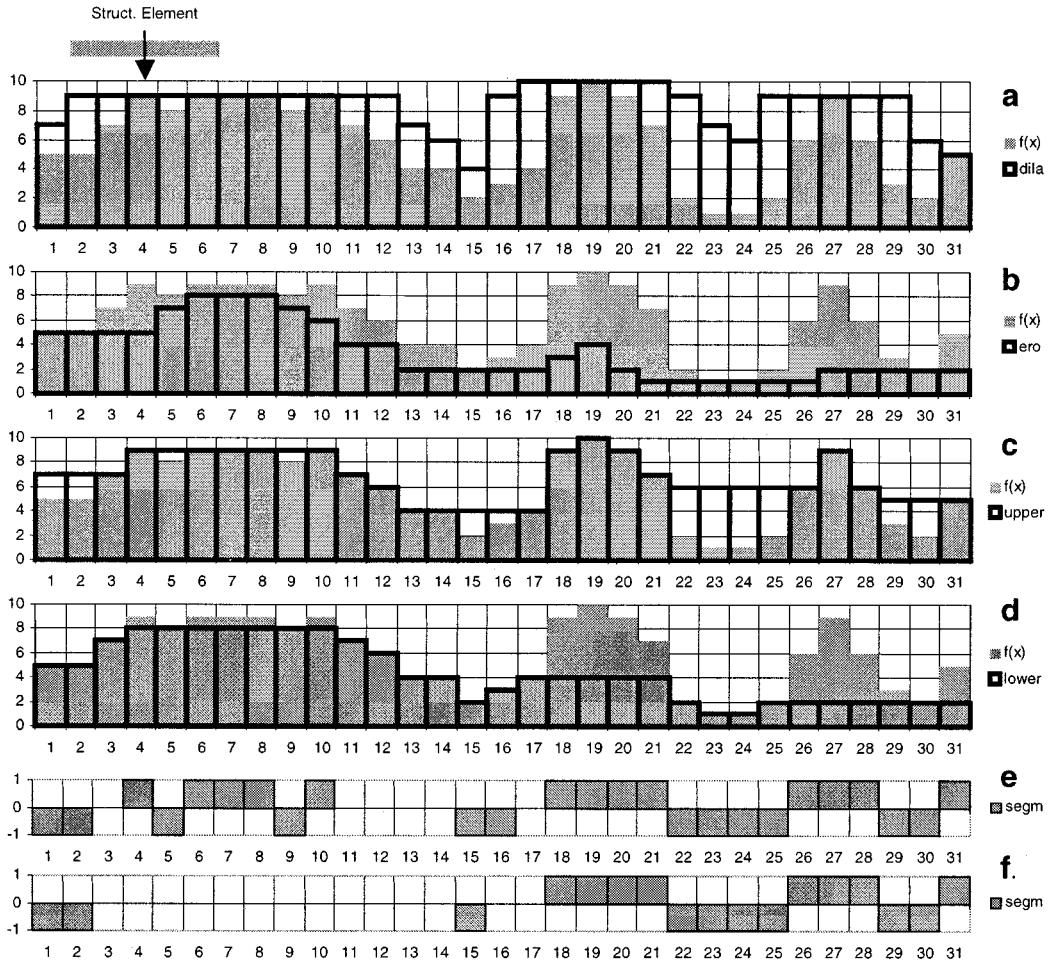

Fig. 1. Simulation of the proposed segmentation method over a strip of 31 pixels. (a) Dilation of $f(x)$ by a structuring element of five pixels. (b) Erosion of $f(x)$ by the same SE as in (a). (c) Upper-leveling of $f(x)$ as dual reconstruction of the dilation above $f(x)$. (d) Lower-leveling of $f(x)$ by reconstruction of the erosion under $f(x)$. (e) Segmentation based on (21) with $\widehat{k} = 1$, $\breve{k} = -1$, $\overline{k} = 0$. (f) Segmentation based on (22) with $\sigma = 2$.

given SE of size N, where N can be considered as the spatial domain of the characteristic function $\Phi_N(f)$.

1) Softer Object/Background Distinction: In case of uncertainty or ambiguity in distinguishing between scene foreground and background, it is also possible to soften the conditions of the morphological characteristics by rewriting (21) as

$$s = \Phi_N^{\sigma}(f) = \begin{cases} \widehat{k}\colon f - \psi(f) > \sigma \\ \breve{k}\colon \psi(f) - f > \sigma \\ \overline{k}\colon f - \psi(f) \le \sigma; \qquad \psi(f) - f \le \sigma \end{cases} \tag{22}$$

for a given level of contrast σ ($\sigma \ge 0$). With $\sigma = 0$, (21) and (22) are equivalent. On the other hand, by increasing the value of σ, the level of necessary contrast is increased in order to avoid the labeling of the pixels with the "flat" label. Thus, the level σ could be interpreted as a threshold used in distinguishing between the foreground and background of the image.

C. Example

In Fig. 1, the application of the proposed method is shown for a strip of 31 pixels having associated signal values in the discrete range [1, $\cdots$, 10]. Fig. 1(a) and (b) shows the selection phase for the dilation [Fig. 1(a)] and erosion [Fig. 1(b)] markers of the original grey level function. Fig. 1(c) and (d) show the dual reconstruction of the dilation and the reconstruction of the erosion as upper-leveling [Fig. 1(c)] and lower-leveling [Fig. 1(d)] of the original function. Fig. 1(e) and (f) represent the segmentation of the original function with $\sigma = 0$ [Fig. 1(e)] and $\sigma = 2$ [Fig. 1(f)], where $\widehat{k} = 1$, $\breve{k} = -1$, $\overline{k} = 0$. Note that pixels with indices from 11 to 14 are strictly flat at the given size of SE (five pixels). Also, note the increase of the size of the flat areas with the increasing value of σ. Pixels with indices from four to ten show a low contrast for this SE size. They are assigned a "flat" label if the value of σ is augmented. Because of the nonlinear nature of this approach, structures that have a greater contrast than σ maintain exactly the same behavior in the segmentation (pixels with indices from 18 to 31). This mechanism can be useful if we want, for a given spatial domain, to distinguish the response due to the presence of structures in an image from the signal produced by noise.

IV. Multiscale Extension

As described previously, segmentation using residuals obtained from the original grey level function and a composition of operations of opening and closing by reconstruction requires the definition of the spatial domain where the method is applied

in terms of a SE size. Some structures may have a high response for a given SE size and a lower response for other SE sizes. That depends on the interaction between the SE size and the size of the structure. Sometimes we know ex-ante the size of the structures that is to be detected. However, that is often not possible, and then a single-SE-size approach appears to be too simplistic. Therefore, it may be a good idea to use a range of different SE sizes in order to explore a range of different hypothetical spatial domains. Consequently, the best response of the structures in the image will be used for segmentation. Given our proposed segmentation method, it is straightforward to extend the same concept to multiscale processing.

Theoretically, the intuitive idea of multiscale morphological characteristics can be interpreted as a variation of the notion of a morphological spectrum. It can be defined as an extension of the *opening spectrum* studied by Haralick *et al.* [16] or the *pattern spectrum* defined in Maragos [17]. Both are based on the definition of some kind of granulometry [18] for the opening spectrum. This means that an image sequence is created by the computation of the differences between successive images in a granulometry generated by a flat SE family with an integral index set. Applications of variations of the morphological spectrum have been proposed for image noise reduction [19] and pseudo band-pass image decomposition [20].

The definition of leveling is stricter than that of the morphological spectrum. The reason is that leveling requires opening and closing operations made by reconstruction. In contrast, the morphological spectrum can also be extracted by composition of Euclidean opening and closing operations. Given the aforementioned theoretical similarities, the idea of the multiscale segmentation based on the derivative of the morphological profile (DMP) is developed in the following section.

A. Definition

Let the vector $\Pi\gamma(x)$ be the *opening profile* at the point x of the image I defined by

$$\Pi\gamma(x) = \{\Pi\gamma_\lambda\colon \Pi\gamma_\lambda = \gamma_\lambda^*(x),\ \forall\,\lambda \in [0, \cdots, n]\} \tag{23}$$

and let the vector $\Pi\varphi(x)$ be the *closing profile* at the point x of the image I defined by

$$\Pi\varphi(x) = \{\Pi\varphi_\lambda\colon \Pi\varphi_\lambda = \varphi_\lambda^*(x),\ \forall\,\lambda \in [0, \cdots, n]\}. \tag{24}$$

Here $\gamma_0^*(x) = \varphi_0^*(x) = I(x) \Rightarrow \Pi\gamma_0(x) = \Pi\varphi_0(x) = I(x)$ for $\lambda = 0$ by the definition of opening and closing by reconstruction. Given (23) and (24), the opening profile can also be defined as a granulometry made with opening by reconstruction, while the closing profile can be defined as an antigranulometry made with closing by dual reconstruction. The derivative of the morphological profile is defined as a vector where the measure of the slope of the opening-closing profile is stored for every step of an increasing SE series.

The derivative of the opening profile $\Delta\gamma(x)$ is defined as the vector

$$\Delta\gamma(x) = \{\Delta\gamma_\lambda\colon \Delta\gamma_\lambda = |\Pi\gamma_\lambda - \Pi\gamma_{\lambda-1}|,\ \forall\,\lambda \in [1, \cdots, n]\}. \tag{25}$$

By duality, the derivative of the closing profile $\Delta\varphi(x)$ is the vector

$$\Delta\varphi(x) = \{\Delta\varphi_\lambda\colon \Delta\varphi_\lambda = |\Pi\varphi_\lambda - \Pi\varphi_{\lambda-1}|,\ \forall\,\lambda \in [1, \cdots, n]\}. \tag{26}$$

In general, the derivative of the morphological profile $\Delta(x)$, or the DMP can be written as the vector

$$\Delta(x) = \left\{ \begin{array}{c} \Delta_{c+\lambda}\colon \Delta\gamma_\lambda,\ \forall\,\lambda \in [1, \cdots, n] \\ \Delta_{c-\lambda+1}\colon \Delta\varphi_\lambda,\ \forall\,\lambda \in [1, \cdots, n] \end{array} \right\} \tag{27}$$

for an arbitrary integer c with n equal to the total number of iterations.

Given all of the above, the morphological multiscale characteristic Φ of the image I at the point x can be defined as the SE size with the greatest associated value in the following Δ function:

$$\Phi(x) = \{\varepsilon\colon \Delta_\varepsilon(x) = \vee\Delta(x),\ \forall\,\varepsilon \in [c-n+1, \cdots, c+n]\}. \tag{28}$$

Equation (28) can be rewritten in order to maintain information about the type of structure that is to be detected. In order to do that, the *multiscale-opening characteristic* $\Phi\gamma(x)$ of the image I at the point x can be defined by

$$\Phi\gamma(x) = \{\lambda\colon \Delta\gamma_\lambda(x) = \vee\Delta\gamma(x)\}. \tag{29}$$

Similarly, the *multiscale-closing characteristic* can be defined by

$$\Phi\varphi(x) = \{\lambda\colon \Delta\varphi_\lambda(x) = \vee\Delta\varphi(x)\}. \tag{30}$$

With these definitions, an algorithm for multiscale segmentation of the image I, based on its characteristic, can be written by a generalization of (21) and (22) as

$$\Phi(x) = \begin{cases} \widehat{k}_\lambda = \Phi\gamma(x)\colon\ \vee\,\Delta\gamma(x) > \vee\Delta\varphi(x) \\ \widecheck{k}_\lambda = \Phi\varphi(x)\colon\ \vee\,\Delta\gamma(x) < \vee\Delta\varphi(x) \\ \overline{k} = 0\colon \vee\,\Delta\gamma(x) = \vee\Delta\varphi(x). \end{cases} \tag{31}$$

In (31), the label of the morphological characteristic is the iteration code of the opening or closing series that correspond to the greatest value of the derivative. If this greatest derivative value is strictly equal for both the opening and the closing series, the "flat" label $\overline{k}$ is applied. In this sense, an image structure is a set of connected pixels or a connected component with the same value of Φ. The function Φ takes the following values.

1) In the range $\Phi(x) \in [\widehat{k}_1, \cdots, \widehat{k}_n]$ in case of prevalently "convex" regions.
2) In the range $\Phi(x) \in [\widecheck{k}_1, \cdots, \widecheck{k}_n]$ in case of prevalently "concave" regions.
3) $\Phi(x) = \overline{k} = 0$ in case of prevalently "flat" or morphologically "indifferent" regions for all the used sizes of SE $\in [1, \cdots, n]$.

Finally, in a similar fashion to what was done when (22) was derived, (31) can be extended for the case of a contrast threshold σ for a structure

$$\Phi^\sigma(x) = \begin{cases} \widehat{k}_\lambda = \Phi\gamma(x)\colon\ \vee\,\Delta\gamma(x) > \vee\Delta\varphi(x) > \sigma \\ \widecheck{k}_\lambda = \Phi\varphi(x)\colon\ \vee\,\Delta\gamma(x) < \vee\Delta\varphi(x) > \sigma \\ \overline{k} = 0\colon\ \vee\,\Delta\gamma(x) = \vee\Delta\varphi(x) \le \sigma. \end{cases} \tag{32}$$

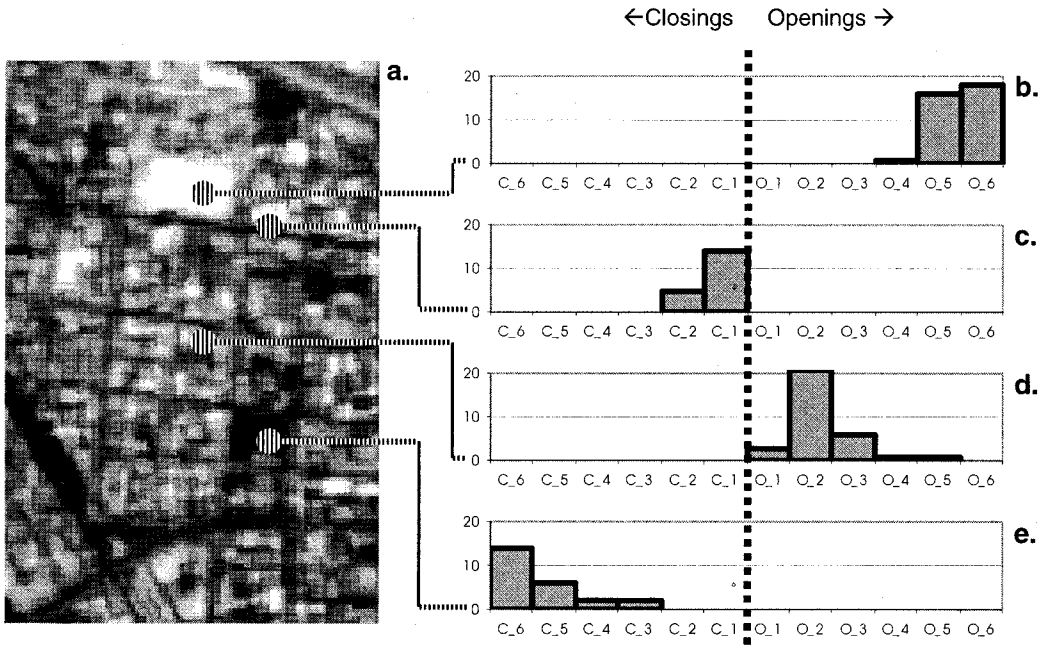

Fig. 2. Derivative of the morphological profile relative to different points in a densely built-up area. (a) Original piece of IRS-1C satellite scene with 5×5 m of spatial resolution. (b) Commercial building. (c) Small street (d) Residential building. (e) Small green area.

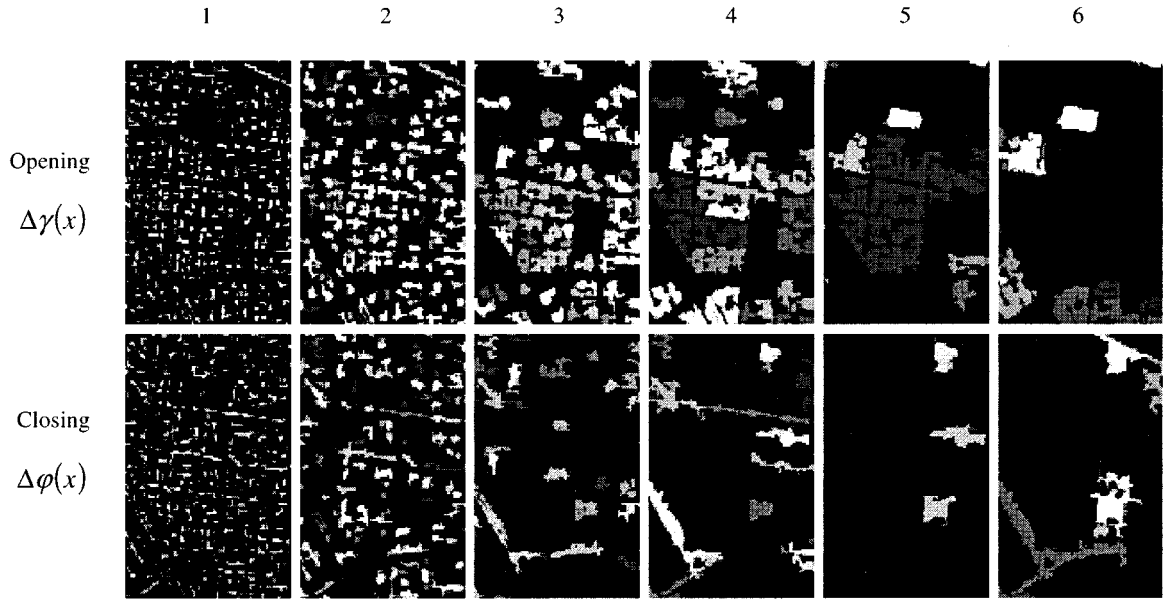

Fig. 3. Morphological decomposition of the image in Fig. 2 by using the derivative of the opening and closing profiles. The images have been visually enhanced. The derivative has been calculated relative to a series generated by six iterations of the elementary SE (size $= 3 \times 3$ pixels).

B. Example

The general idea underlying the proposed segmentation method is that the derivative curve $\Delta(x)$ is some type of a structural or morphological signature that can be used for the pixel discrimination by morphological criteria. The structural or morphological signature is analogous to the spectral signature approach of the multispectral satellite images. Figs. 2 and 3 show the multiscale derivative of the morphological profile for a small sample of satellite image from a densely built-up area (sensor: IRS-1C panchromatic, 5×5 m^2 of spatial resolution). In this example, a range of sizes of SEs was used. This range was derived based on six iterations of the elementary eight-connected SE [using (9)], which produces a SE size range from 3×3 up to 13×13 pixels.

Fig. 2 shows the derivative curve for four representative objects: a commercial building, a street inside urban texture, a residential building, and a green area. The histograms on the right show the level of the derivative relative to the opening and closing series for every step of the iteration. Note how the behavior of the derivative curve $\Delta(x)$ stores information about both the type and the size of the connected components inside the image. Connected components that are brighter than their adjacent components have a $\Delta(x)$ function that is unbalanced to the right (opening series), while darker connected components

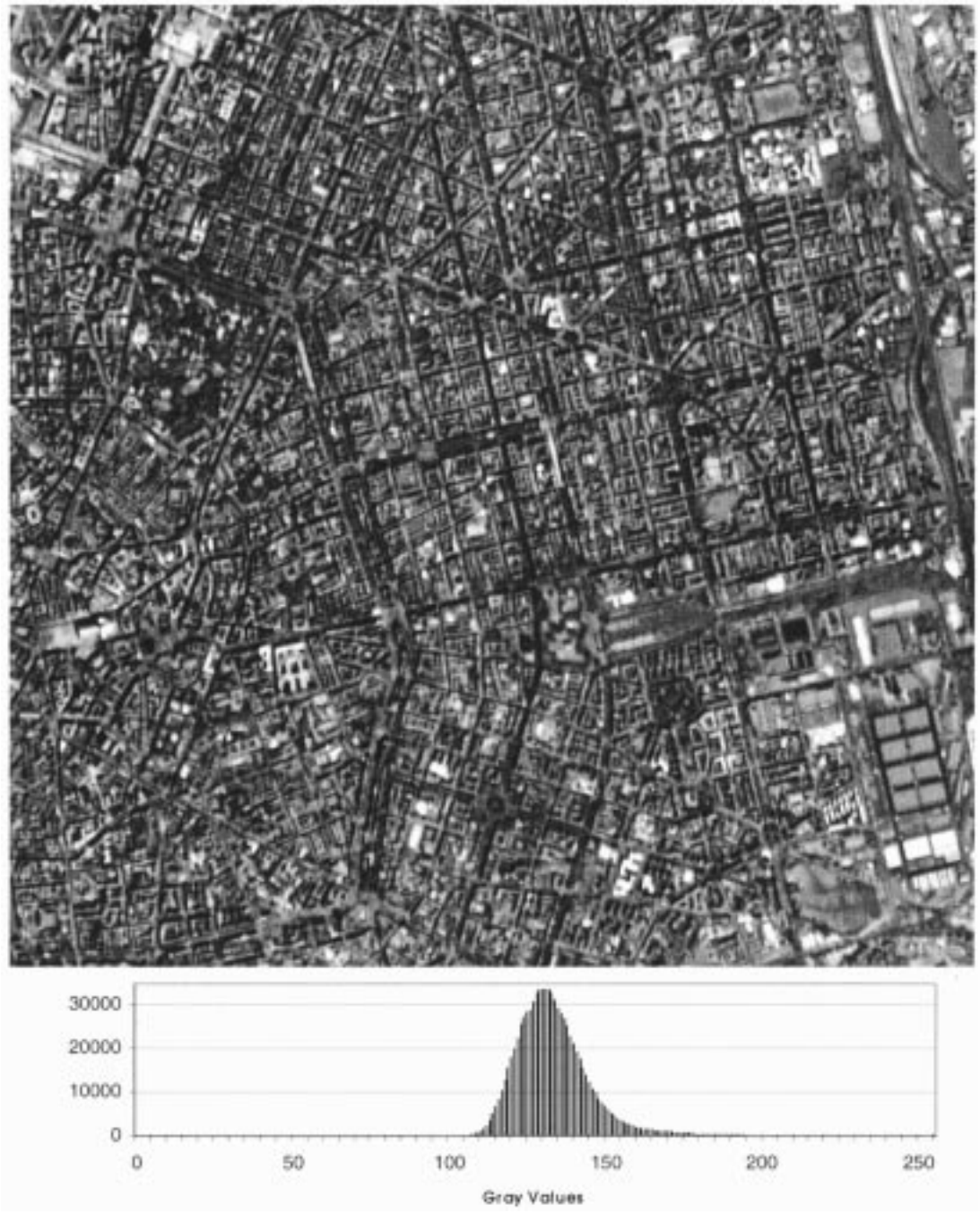

Fig. 4. (Top) Portion of an IRS-1C panchromatic scene over a densely built-up area from Milan, Italy (1998, ANTRIX, SIE, Euromap). (Bottom) The histogram of the IRS-1C image. The image covers an area of 4×4 km^2 with a resolution of 5×5 m (800×800 pixels). The image has been enhanced for visualization by min–max histogram stretching.

show a $\Delta(x)$ function that is unbalanced to the left (closing series). The point where $\Delta(x)$ takes the maximum value is used to record the size of SE, which gives maximum response. Consequently, this point gives a good indication of the size of the structure in the given spatial domain range, which can be used in the definition of the proposed $\Phi(x)$. That is in contrast to the decomposition of electromagnetic signals by spectral slices, which is common for multispectral images. It can be noted in Fig. 3 how the proposed multiscale derivative approach decomposes the image by the morphological criteria.

V. Applications

A. Previous Experiments

The idea to use a composition of opening transforms for a morphological segmentation of satellite data was proposed some time ago for the detection of different urban structures [21], [22]. In the experiments in [21], [22], segmentation labels were obtained after the arithmetic summing of an opening series with an increasing SE. The method is only applicable to Boolean maps (binary or 2 grey-level images), and it does

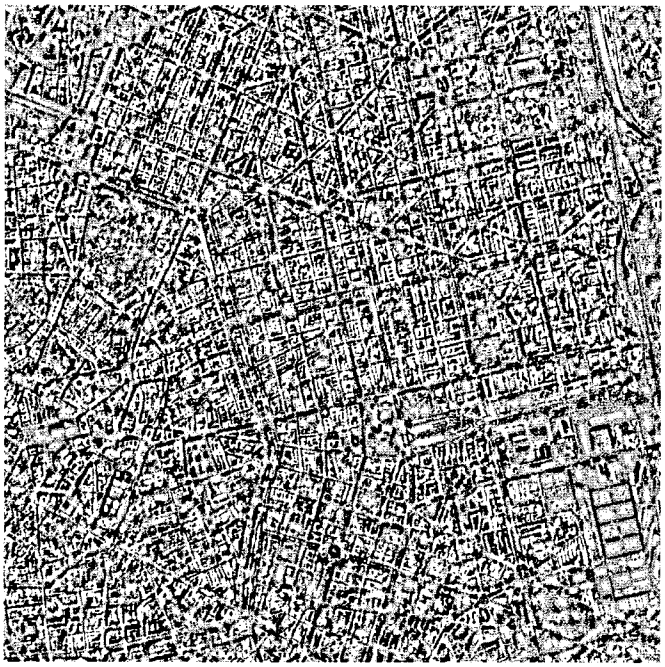

Fig. 5. Segmentation of the image in Fig. 4 by single-scale leveling (21). The utilized SE is an octagon of diameter $d = 35$ m (seven pixels) with $\widehat{k}$ = white, $\breve{k}$ = black, and $\overline{k}$ = medium grey.

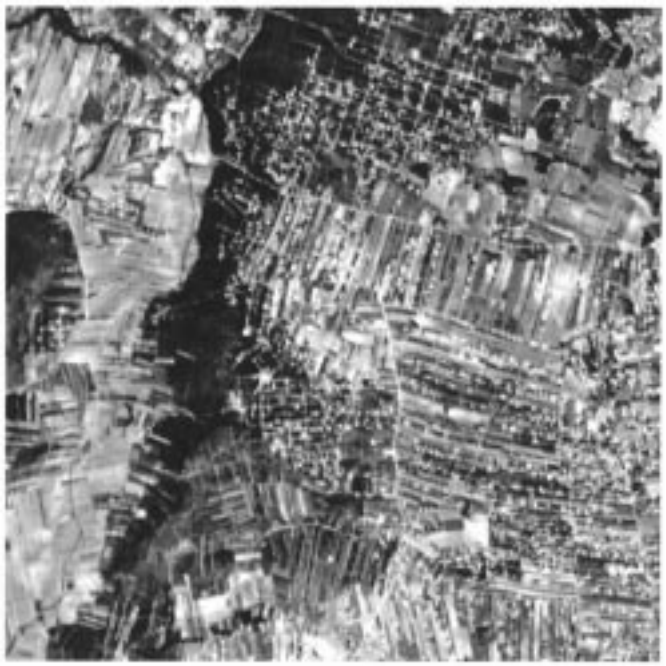

Fig. 6. Portion of IRS-1C panchromatic scene over an agricultural and scattered settlement area NE of Athens, Greece, 1998 (ANTRIX, SIE, Euromap). The image covers an area of 4×4 km^2 with a resolution of 5×5 m meters (800×800 pixels). The image has been enhanced for visualization by min–max histogram stretching.

not use geodesic metric. More recently, Pesaresi and Kanellopoulos [23] used a composition of geodesic opening and closing operations of different sizes in order to build a morphological profile. Then, they used a neural network approach for the classification of structures. The difference between their method and the method proposed here is that in [23], the absolute residual between the original image and the opened or closed one was used as a morphological characteristic function. Therefore, the method in [23] cannot be used for multiscale segmentation since it limits the explored spatial domain by a restraint range of SEs. The method proposed here is both more general and more robust than all the methods above. That will be demonstrated by the following application examples.

B. Application Examples

Examples of the application of the proposed method are now given for the segmentation of two satellite high-resolution data sets. The satellite imagery is taken from a densely built-up area (Milan, Italy) and from an agricultural area connoted by scattered settlement (NE of Athens, Greece). Both data sets were recorded by the IRS-1C panchromatic sensor, which has ground spatial resolution of 5×5 m^2. The images used in the experiments are a sub-sample of 800×800 pixels (4 km $\times$ 4 km) of surface from an original scene of about $15\,000 \times 15\,000$ pixels (75 km $\times$ 75 km).

1) Example 1: Fig. 4 shows the subsampled image from Milan. It is easy to note the relatively poor dynamic range of the data recorded by the sensor. As a consequence, the frequency histogram is concentrated around grey level 50. Relevant objects in urban remote sensing applications, such as buildings and streets, can have a thickness of 1 to 2 pixels in images with a comparable resolution to the resolution of this image. Most of the pixels in such images are placed between the borders of different ground objects. Therefore, the dynamic range can be critical if it is needed to apply standard morphological segmentation based on contour detection.

Fig. 5 shows the results of the application of the proposed method for segmentation. In this experiment, the single-scale approach in (21) was used. Consequently, the reconstruction of the erosion under the original image and the dual reconstruction of the dilation above the original image were composed using only one SE size. Both erosion and dilation were done with a SE equal to an octagon of diameter $d = 7$ pixels (35 m on the ground), which is close in dimension to the most relevant objects in the scene (buildings, roads). White, black, and medium grey pixels were labeled by the segmentation procedure as $\widehat{k}$, $\breve{k}$, and $\overline{k}$, respectively. By looking at Fig. 5, it can be noted that the segmentation method appears to correctly detect most of the relevant regions. In the figure, the urban structure is well highlighted even though it is very complex. It is also interesting that buildings generally appear labeled as radiometric "convex" curvature, while roads appear to be labeled as radiometric "concave" curvature. This can be useful for a successive automatic classification phase. Finally, note that most of the commercial buildings on the bottom right of the image have been labeled as radiometric "flat." That occurred because the dimension of the utilized SE is smaller than the size of the object.

2) Example 2: Fig. 6 shows the subsampled image from a scattered settlement on the N-E of Athens, Greece. It is easy

to note the co-presence in the scene of objects/regions of different sizes. As a consequence, the multi-scale approach defined in (31) was applied. Fig. 7 shows the segmentation results. The spatial domain explored in this experiment was given by a range of 10 increasing octagonally-shaped SE's with a diameter ranging from 7 pixels (45 meters) up to 61 pixels (305 meters). The step from iteration λ to iteration $\lambda + 1$ was equal to 6 pixels (30 m.). Consequently, the final number of labels was $n = 10 + 10 + 1 = 21$, including the "flat" label. In Fig. 7, the function $\Phi(f)$ took values as follows:

1) In the range $\Phi(x) \in [\widehat{k}_1, \cdots, \widehat{k}_{10}]$ in case of prevalently "convex" regions.
2) In the range $\Phi(x) \in [\breve{k}_1, \cdots, \breve{k}_{10}]$ in case of prevalently "concave" regions.
3) $\Phi(x) = \overline{k} = 0$ in case of prevalently "flat" or morphologically "indifferent" regions for all the used sizes of SE$\in$[1, $\cdots$ 10].

In the segmented image, both large regions and small ones were retained without an undesired loss of detail. Similar findings were observed for nested, thin, and complex regions. It is interesting to note that the proposed multiscale approach seemed to have a hierarchical effect. Large regions appear to have the same label. Also, no over-segmentation effect was detected due to the presence of nonrelevant local minima and local maxima, which is usual in classical segmentation by the watershed approach.

Fig. 8 shows a comparison between the proposed approach and classical watershed segmentation for a 100 $\times$ 100 pixel area of the Athens data set. The subsample is placed in the center of Athens and is taken over a compact urban area with an internal vegetated area (park). Subimage 1 shows the original radiometric data enhanced with min-max histogram stretching for visualization purposes. The classical morphological approach requires the detection of the border of the regions, and subimage 2 is the direct application of a morphological gradient transform (defined as the difference between dilation and erosion) to the original data. In this complex context, it is possible to observe that attempting to start from edges of regions leads to the production of "surfaces of edges," where most of the pixels are connoted as "border pixels." Another evident problem with the classical approach is the over-segmentation generated by nonrelevant local minima of the gradient function. Subimage 3 shows the gradient of the filtered data where a morphological filter was applied, defined as the opening of the closed image with a flat SE equal to 3 $\times$ 3 pixels. Consequently, the situation in subimage 3 appears to be simpler than in subimage 2. Subimage 4 shows the results of watershed segmentation using the gradient image in subimage 3. Subimage 5 shows the output of the multiscale morphological segmentation defined in (31). We can note that the proposed approach retains a better description of the original structural information, introducing less shape noise than the classical watershed segmentation approach. Another positive characteristic of the proposed method is the intrinsic hierarchy that reduces dramatically the over-segmentation effect. This can be detected in the case of the green area that is labeled as only one region by the proposed method but a set of nonhomogenous regions by the classical approach.

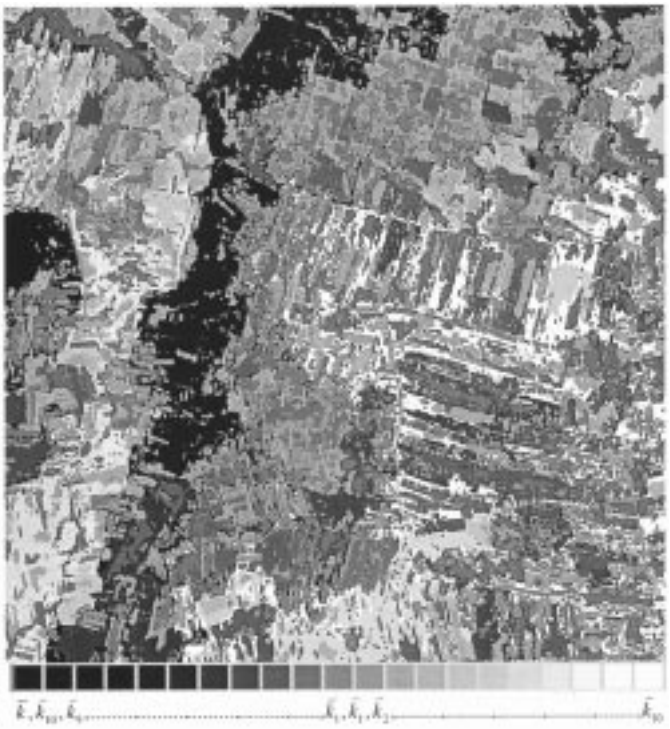

Fig. 7. Multiscale segmentation of the image in Fig. 6 obtained by (31). The explored spatial domain ranges from an octagon of seven pixels (45 m) to an octagon of 61 pixels (305 m), with ten steps of six pixels (30 m) each.

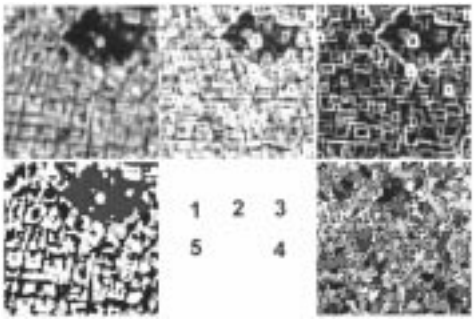

Fig. 8. Comparison of the proposed morphological segmentation approach and classical watershed segmentation. The subimages are ordered as shown in the middle of the bottom row. 1) Original radiometric information (IRS-1C panchromatic sensor) after linear histogram stretching, 2) gradient of the original data, 3) gradient of the filtered data, 4) image obtained by watershed segmentation, and 5) image obtained by multiscale segmentation as defined in (31).

VI. Conclusions

Morphological segmentation by the derivative of the morphological profile was proposed. The proposed method is based on the use of residuals from opening and closing by reconstruction. In experiments, the proposed method demonstrated excellent performance even where the classical morphological approach had problems. In particular, the proposed approach gives a better shape description than the classical approach. It also retains small but significant regions in images, and has an effect of

intrinsic hierarchy that reduces dramatically the over-segmentation problem of the classical approach.

The drawback of the proposed method concerns the necessity of looking at a range of increasing opening and closing by reconstruction operations, which may cause a heavy computational burden. As a consequence, for images with very large and homogeneous regions, it is possible that a gradient-plus-watershed approach may be more efficient, since it does not need to explore a very wide range of different SE sizes. For the above reasons, the method presented here is particularly suited for segmentation of complex image scenes such as aerial or satellite images where very thin, enveloped, and/or nested regions must be retained. It is also well suited for images with low radiometric contrast and relatively low spatial resolution, which produce textural effects, border effects, and ambiguity in the object/background distinction. All these factors are critical and can lead to instability effects if segmentation methods based on the edge-detection approach are used.

Currently, our work is focused on the extension of the proposed approach by improving the morphological characteristic detection. The proposed formula for the morphological characteristic assumes a "simple" behavior of the morphological profile where each structure is supposed to have only one significant derivative maximum. In more complex environments, it is possible that some structures may have more than one significant derivative maximum. These complex environments can have a greater explored spatial domain range, occurrence of nested regions at different grey levels, and/or presence of spatial periodicity. In these cases, it is possible to make the characteristic function Φ more sophisticated by using any distribution-free classification procedures such as neural network classifiers.

ACKNOWLEDGMENT

The authors would like to thank C. Lavalle of the EC JRC Space Applications Institute, SSSA Unit, for providing the IRS-1C data used in the Milano experiment.

REFERENCES

[1] J. Serra, *Image Analysis and Mathematical Morphology*. London, U.K.: Academic, 1982.
[2] ——, "Introduction to mathematical morphology," *Comput. Vis., Graph., Image Process.*, vol. 35, pp. 283–305, 1986.
[3] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin, Germany: Springer-Verlag, 1999.
[4] J. Serra and P. Salembier, "Connected operators and pyramids," *Proc. SPIE*, vol. 2030, pp. 65–76, 1993.
[5] J. Crespo, J. Serra, and R. Schafer, "Theoretical aspects of morphological filters by reconstruction," *Signal Process.*, vol. 47, pp. 201–225, 1995.
[6] P. Salembier and J. Serra, "Multiscale image segmentation," in *Proc. SPIE*, vol. 1818, 1992, pp. 620–631.
[7] L. Vincent, "Graphs and mathematical morphology," *Signal Process.*, no. 16, pp. 365–388, 1989.
[8] F. Meyer and S. Beucher, "Morphological segmentation," *J. Vis. Commun. Image Represent.*, vol. 11, pp. 21–46, 199.
[9] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," in *Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*. Rennes, France: CCETT/IRISA, Sept. 1979.
[10] F. Meyer, "Integrals, gradients and watershed lines," in *Proc. Mathematical Morphology and Its Applications to Signal Processing*, Barcelona, Spain, May 1993, pp. 70–75.
[11] L. Najman and M. Schmitt, "Definition and some properties of the watershed of a continuous function," in *Proc. Mathematical Morphology and Its Applications to Signal Processing*, Barcelona, Spain, May 1993, pp. 75–81.
[12] F. Meyer, "Morphological connected mapping with scale space properties," Tech. Rep., Centre Morphologie Math., Fontainebleau, France, 1997.
[13] ——, "From connected operators to levelings," in *Mathematical Morphology and Its Applications to Image and Signal Processing*, H. J. A. M. Heijmans and B. T. M. Roerdink, Eds. Dordrecht, The Netherlands: Kluwer, 1998, pp. 191–198.
[14] ——, "The levelings," in *Mathematical Morphology and its Applications to Image and Signal Processing*, H. J. A. M. Heijmans and B. T. M. Roerdink, Eds. Dordrecht, The Netherlands: Kluwer, 1998, pp. 199–206.
[15] G. Matheron, "Les nivellements," Tech. Rep., Centre Morphologie Math., Fontainebleau, France, 1997.
[16] R. M. Haralick, E. R. Dougherty, and P. L. Katz, "Model-based morphology: The opening spectrum," in *in Advances in Image Analysis*, Y. Mahavieh and R. C. Gonzales, Eds. Bellingham, WA: SPIE, 1992.
[17] P. Maragos, "Pattern spectrum and multiscale shape representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 701–716, 1989.
[18] E. J. Kraus, H. J. A. M. Heijmans, and E. R. Dougherty, "Greyscale granulometries compatible with spatial scaling," *Signal Process.*, vol. 34, pp. 1–17, 1993.
[19] R. A. Peters, "A new algorithm for image noise reduction using mathematical morphology," *IEEE Trans. Image Processing*, vol. 4, pp. 554–568, 1995.
[20] ——, "Morphological pseudo bandpass image decomposition," *J. Electron. Imag.*, vol. 5, pp. 198–213, 1996.
[21] M. Pesaresi, "Analisi numerica dello spazio edificato nella città diffusa," Tech. Rep., IUAV DAEST, Venice, Italy, 1993.
[22] A. Bianchin and M. Pesaresi, "Outils de morphologie mathématique appliquée aux images satellite pour l'analyze de l'urbanization diffuse," in *Proc. Conf. EGIS-MARI 94*, Paris, France, Mar. 29–Apr. 1, 1994, pp. 2085–2094.
[23] M. Pesaresi and I. Kannellopulos, "Detection of urban features using morphological based segmentation and very high resolution remotely sensed data," in *Machine Vision and Advanced Image Processing in Remote Sensing*, I. Kanellopoulos, G. G. Wilkinson, and T. Moons, Eds. Berlin, Germany: Springer-Verlag, 1999, pp. 271–284.

Martino Pesaresi received the M.A. degree in town and regional planning ("Tesi di Laurea in Urbanistica") from the Istituto Universitario di Architettura di Venezia (IUAV), Venice, Italy, in 1992.

From 1993 to 1997, he worked as Individual Consultant, specializing in the application of G.I.S. and remote sensing technologies in the fields of urban analysis and urban planning. In particular, he collaborated on a variety of research projects and teaching activities with IRES, the former Centro Interdipartimentale di Cartografia e Fotogrammetria (CICaF), the Dipartimento di Analisi Economica e Sociale del Territorio (DAEST), and the Dipartimento di Urbanistica (DU) of the IUAV, and the Dipartimento di Progettazione Urbanistica (DPU), Pescara University, Pescara, Italy. Since 1996, he has been a Lecturer of Spatial Statistics, Corso di Diploma in Sistemi Informativi Territoriali (CDSIT), IUAV. Among others, he has had advice activity with the Veneto Region Administration, Italy, the Master Plan of Durazzo, Albania, and the Master Plan of El Salvador, El Salvador. Other collaborations with private companies include INSIEL S.P.A., Udine, Italy, and T.S.A. S.R.L., Padova, Italy, for design and implementation of advanced image processing procedures, and GRETA Econometrics S.R.L., Venice, Italy, for design of G.I.S.-based geomarketing product. Since 1998, he has been working with a research contract at the European Commission, Space Applications Institute, Environment and Geo-information Unit, Ispra, Italy. His research interests are pattern analysis, mathematical morphology, image fusion, and visualization applied to high resolution remotely sensed data.

Mr. Pesaresi received the ERASMUS study grant from the Centre d'Analyze et de Mathématique Sociales (CAMS) of the École des Hautes Etudes en Sciences Sociales (EHESS), Paris, France, from 1991 to 1992, and the COMETT study grant from the Laboratoire d'Informatique Appliquée (LIA) of the ORSTOM, paris, France, from 1992 to 1993.

Jon Atli Benediktsson (S'84–M'90–SM'99) received the cand. scient. degree in electrical engineering from the University of Iceland, Reykjavik, in 1984, and the M.S.E.E. and Ph.D. degrees from the School of Electrical Engineering, Purdue University, West Lafayette, IN, in 1987 and 1990, respectively.

He is currently Professor of electrical and computer engineering, University of Iceland, and was with the University's Laboratory for Information Technology and Signal Processing from 1984 to 1985. From 1985 to 1991, he was affiliated with the School of Electrical Engineering and the Laboratory for Applications of Remote Sensing (LARS), Purdue University. He has held visiting positions with the Joint Research Centre, European Commision, Ispra, Italy, Denmark's Technical University (DTU), Lyngby, Denmark, and the School of Electrical and Computer Engineering, Purdue University. His research interests are in pattern recognition, neural networks, remote sensing, image processing, and signal processing, and he has published extensively in those fields.

Dr. Benediktsson received the 1991 Stevan J. Kristof Award from LARS as outstanding graduate student in remote sensing, and from 1996 to 1999, he served as the Chairman of the IEEE Geoscience and Remote Sensing Society's Technical Committee of Data Fusion. He was the recipient of the 1997 Icelandic Research Council's Outstanding Young Researcher Award. He is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is a member of the International Neural Network Society (INNS) and Tau Beta Pi.