Journal of Mathematical Imaging and Vision 17: 109–129, 2002


# Shape Connectivity: Multiscale Analysis and Application to Generalized Granulometries*

COSTAS S. TZAFESTAS† AND PETROS MARAGOS

*National Technical University of Athens, School of Electrical and Computer Engineering, Division of Signals, Control and Robotics, Zografou Campus, Athens 15773, Greece*

ktzaf@softlab.ntua.gr

petros.maragos@cs.ntua.gr

**Abstract.** This paper develops a multiscale connectivity theory for shapes based on the axiomatic definition of new generalized connectivity measures, which are obtained using morphology-based nonlinear scale-space operators. The concept of connectivity-tree for hierarchical image representation is introduced and used to define generalized connected morphological operators. This theoretical framework is then applied to establish a class of generalized granulometries, implemented at a particular problem concerning soilsection image analysis and evaluation of morphological properties such as size distributions. Comparative results demonstrate the power and versatility of the proposed methodology with respect to the application of typical connected operators (such as reconstruction openings). This multiscale connectivity analysis framework aims at a more reliable evaluation of shape/size information within complex images, with particular applications to generalized granulometries, connected operators, and segmentation.

**Keywords:** shape analysis, mathematical morphology, multiscale connectivity measures, connected operators, reconstruction, generalized granulometries, soilsection image analysis, connectivity tree, hierarchical image representations, partitions

## 1. Introduction

In many image analysis problems, such as segmentation, a very important task is to extract particular regions of an image while preserving as much of the contour information as possible. Classical morphological operators perform local transformations, using one or more structuring elements, and may thus significantly modify boundaries within an image. Connected morphological operators are essentially different, since they act on the flat zone level [4], thus having the capacity to precisely identify and extract whole connected components in an image, which are treated as a whole without alteration of their boundaries. This very important property makes connected operators very attractive for many image processing and filtering tasks, especially when precise shape analysis is concerned. Typical connected operators are the reconstruction and area openings for binary or grayscale images [18]. In this paper, we develop a multiscale connectivity theory for shapes, obtained using morphology-based nonlinear scale-space operators. This theoretical framework and the resulting generalized connected operators aim at a more reliable evaluation of shape/size information within complex images, with particular applications to generalized granulometries and segmentation.

Multiscale shape analysis has been an active research area in computer vision. Some well-known approaches

*This work was supported by the Greek General Secretariat for Research and Technology and the European Union, under grant ΠΕΝΕΔ-99-ΕΔ164.

†Part of the work was performed while C.S. Tzafestas was with the Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", Athens, Greece.

include: the curvature Gaussian scale-space; the dynamic shape [6] obtained by thresholded Gaussian convolutions of the shape's binary indication function; and the reaction-diffusion scale-space [5] obtained via differential curve evolution and governed by Hamilton-Jacobi PDEs. Our approach to multiscale shape analysis in this paper is algebraic and based on lattice-theoretic formulations of connectivity and morphology.

The classical notion of connectivity is defined in the framework of topological spaces, as well as in graphs. Serra [14] has given a formal definition of connectivity class (or connection) in a complete lattice framework. Based on this definition and the equivalent concept of connected openings, several second-order (or second-generation) connections have been defined [15], usually based on some extensive morphological operator, like closing or extensive dilation. These are often called clustering connectivities, which in fact identify as connected, components that are "close enough" to each other. Numerous applications exist in the literature for this connectivity framework (including segmentation, motion compensation etc. [12]), as well as theoretical extensions like the set-oriented approach introduced in [11].

In this paper we modify this concept to cover a different class of second-order connections which, based on multiscale antiextensive morphological operators such as openings or antiextensive erosions, aims at differentiating between "strong" or "loose" connections in a set. That is, starting from a connectivity class $\mathcal{C}$, a set is treated as connected at a given scale if the application of an antiextensive operator at this scale yields a new set that also belongs to $\mathcal{C}$. A multiscale connectivity analysis framework is proposed based on the axiomatic definition of generalized connectivity measures that quantify the notion of a varying "degree" of connectivity of a set, like for instance a multiscale connectivity function defined using morphological adjunctions (erosion, dilation operators). The concept of connectivity-tree (C-tree) is introduced and an algorithm for its creation is described, which constitutes the core of the multiscale connectivity analysis. This hierarchical image representation corresponds, in fact, to a recursive partitioning of the image into progressively "stronger" connected components at each connectivity level, and can be used to define new generalized connected operators based on a decision criterion, which may for instance employ a thresholding connectivity profile chosen appropriately for a particular image analysis application.

The motivation for this generalized hierarchical connectivity framework resides on a well-known drawback related to the application of typical connected operators (such as reconstruction or area openings), which is often called "leakage" problem resulting in the creation of undesirable connections in an image due to the presence of thin connecting paths between large image components. One of the goals of the proposed multiscale connectivity framework is to control the effect of this problem by taking into account additional geometrical information related to the presence of "compound" shapes/structures and their interconnections within an image.

The generalized connectivity operators, defined based on the C-tree image representation, are used to establish new generalized granulometries to perform multiscale image analysis and evaluate morphological properties such as size distributions within an image. Granulometries constitute one of the most useful and versatile tools of morphological image analysis [8], with a wide range of applications described in the literature, including texture characterization [16], image segmentation etc., both for binary and grayscale images [19]. A particular application is considered in this paper that concerns granulometric analysis of soilsection images. Evaluation of soil structure is primarily concerned with detecting compound soil formations, differentiating them from void space and estimating pertinent morphological properties such as size/shape distributions. Extraction of such morphological features from complex sample soilsection images is a very demanding task. The boundaries preservation property of connected operators can be very useful in such situations where all homogeneous regions in an image have to be reliably and precisely identified. The granulometric analysis results obtained using typical connected operators are compared to the ones resulting from the application of the generalized connectivity operators introduced in this paper, demonstrating the power and the versatility of the proposed multiscale connectivity analysis framework.

Summarizing, this paper focuses on:

(a) the definition of new generalized connectivity measures, based on morphological lattice operators (Section 2),

(b) the introduction of a hierarchical, multiscale, connectivity analysis framework, based on the concept of connectivity-tree (C-tree) (Section 3). An algorithmic implementation for C-tree creation is also described in Section 3.2.
(c) the application of this hierarchical image representation for the definition of new generalized connectivity operators, which could prove more appropriate for a number of applications, like the development of generalized granulometries for reliable and precise evaluation of morphological properties such as size/shape distributions (Section 4),
(d) the implementation of this theoretical framework in a particular image analysis problem concerning multiscale granulometric analysis and evaluation of soilsection images (Section 4.4), demonstrating the ability of the proposed generalized operators to extract more reliable and accurate information about the shape/size structure within an image.

## 2. Generalized Morphological Connectivity Measures

### *2.1. Introduction to Lattice Operators*

In this section we recall some basic theoretical elements of lattice-based mathematical morphology that are used throughout this paper. For a more comprehensive discussion the reader may refer to [3].

We focus on the set of shapes (or binary images) that can be modelled by the power set $\mathcal{P}(E)$ (i.e. the collection of all subsets of $E$), where $E = \mathbb{R}^n$ or $\mathbb{Z}^n$ (in this paper we use mainly $n = 2$, but the concepts generally hold for $n \geq 2$). We view $\mathcal{P}(E)$, equipped with the partial order of $\subseteq$, as a complete lattice with supremum the union $\bigcup$ of sets and infimum the intersection $\bigcap$. Shape transformations can be then viewed as lattice operators on $\mathcal{P}(E)$. More important are the increasing operators $\psi$ that satisfy: $X \subseteq Y \Rightarrow \psi(X) \subseteq \psi(Y)$.

Four very useful increasing lattice operators are: the *dilation*, which distributes over $\bigcup$, the *erosion*, which distributes over $\bigcap$, the *opening*, which is increasing, antiextensive ($\psi(X) \subseteq X$) and idempotent ($\psi^2(X) = \psi(X)$), and the *closing*, which is increasing, extensive and idempotent. Classical examples of such operators on $\mathcal{P}(E)$ are the Minkowski dilation $\delta_B$ and erosion $\varepsilon_B$, defined as follows:

$$\delta_B(X) = X \oplus B \quad \text{and} \quad \varepsilon_B(X) = X \ominus B$$

where $\oplus$ and $\ominus$ are the Minkowski addition and subtraction respectively, and $B$ is a compact convex structuring element, such as the closed unit ball. The Minkowski opening and closing filters on $\mathcal{P}(E)$ are also defined as:

$$\gamma_B(X) = X \circ B = \delta_B(\varepsilon_B(X)) = (X \ominus B) \oplus B$$
$$\beta_B(X) = X \bullet B = \varepsilon_B(\delta_B(X)) = (X \oplus B) \ominus B$$

Multiscale operators can then be defined by replacing $B$ with a multiscale version $rB = \{rb : b \in B\}$ $(r \geq 0)$. Examples include the multiscale dilation $\delta_B^r$ and erosion $\varepsilon_B^r$: for $X, B \subseteq \mathbb{R}^n$,

$$\delta_B^r(X) = X \oplus rB, \quad \varepsilon_B^r(X) = X \ominus rB \tag{1}$$

For discrete shapes $X, B \subseteq \mathbb{Z}^n$, we define the multiscale dilation and erosion recursively:

$$\delta_B^r(X) = \delta_B\big(\delta_B^{r-1}(X)\big), \quad \varepsilon_B^r(X) = \varepsilon_B\big(\varepsilon_B^{r-1}(X)\big) \tag{2}$$

where $r = 1, 2, \ldots,$ and $\delta_B^0(X) = \varepsilon_B^0(X) = X$. Note that the two above definitions of multiscale dilation/erosion coincide in $\mathbb{R}^n$ if $B$ is convex and $r$ is an integer.

### *2.2. The Concept of Connectivity: Basic Definitions*

In this section we describe the general concept of connectivity and present the basic definitions, as well as the notation and formalism used throughout this paper. According to the classical definition of connectivity, a subset $X$ of a topological space is said to be connected when it cannot be partitioned into two nonempty closed (or open) sets. In an Euclidean topological space the concept of *arcwise* (*path-*) *connectivity* can also be defined, which proves to be more convenient. According to this definition, a set $X$ is said to be connected if, for every $a, b \in X$, there exists a continuous mapping $\psi$ from [0, 1] into $X$ such that $\psi(0) = a$ and $\psi(1) = b$ (i.e. a path from $a$ to $b$, belonging completely into $X$). In the sequel, $\mathcal{C}$ denotes the usual topological path-connectivity in $\mathcal{P}(E)$, which for digital image analysis problems resides on the 4- or 8-adjacency principle introducing elementary connections between neighboring pixels.

A basic result that follows from the classical definition of connectivity is that the union of two intersecting connected sets is also connected. This result has been used by Serra as a starting point to propose a different definition for connectivity [15].

*Definition 1* (Connectivity class). A subset $\mathcal{C} \subseteq \mathcal{P}(E)$ is called a connectivity class if the following properties hold:

(i) $\emptyset \in \mathcal{C}$ and $\{x\} \in \mathcal{C}$ for every $x \in E$
(ii) if $\mathcal{X} \subseteq \mathcal{C}$ and $\bigcap \mathcal{X} \neq \emptyset$, then $\bigcup \mathcal{X} \in \mathcal{C}$

As shown in [14], the definition of a connectivity class is equivalent to the definition of a family of openings $\{\gamma_x, x \in E\}$, called *connectivity openings*, satisfying the following conditions:

(i) $\forall x \in E, \gamma_x(\{x\}) = \{x\}$
(ii) $\forall x, y \in E$ and $X \subseteq E, \gamma_x(X)$ and $\gamma_y(X)$ are either equal or disjoint
(iii) $\forall x \in E$ and $X \subseteq E,\ x \notin X \Rightarrow \gamma_x(X) = \emptyset$

Intuitively, $\gamma_x(A)$ extracts the connected component of $A$ containing element $x$, that is:

$$\gamma_x(A) = \bigcup \{B \in \mathcal{C} : x \in B \text{ and } B \subseteq A\} \tag{3}$$

The concept of connectivity can be extended using a variety of lattice operators. Let $\psi$ be an increasing and extensive operator on the lattice $\mathcal{P}(E)$. Then, it can be shown that a new connectivity class is obtained based on the following definition of connectivity openings:

$$\gamma_x^{\psi}(A) = \begin{cases} \gamma_x(\psi(A)) \cap A, & \text{if } x \in A \\ \emptyset, & \text{if } x \notin A \end{cases} \tag{4}$$

This is often called second-order (or clustering) connection since, starting from $\mathcal{C}$, a new connectivity class is created, where two components are considered as connected when being "close enough" to each other, in the sense that the application of the extensive operator $\psi$ yields a single connected component belonging to $\mathcal{C}$. Typical example of second-order connectivities use dilation and closing operators [15]. Based on the above concept of clustering connectivity, Braga-Neto and Goutsias in [1, 2] proposed the definition of a multiresolution connectivity measure as a non-negative function on the lattice of interest that quantifies the idea of a varying degree of connectivity.

*Definition 2* (Connectivity measure [1]). A function $\mu : \mathcal{P}(E) \to \mathbb{R}_+$ is defined as a connectivity measure[1] on $\mathcal{P}(E)$ if:

(i) $\mu(0) = \mu(x) = \sup\{\mu(A) : A \in \mathcal{P}(E)\}$, for $x \in E$
(ii) $\mu(\bigcup A_i) \geq \inf\{\mu(A_i)\}, \forall A_i \in \mathcal{P}(E) : \bigcap A_i \neq 0$

This definition implies that the union of some arbitrary intersecting sets is considered at least "as much connected" as any of the individual subsets (i.e. it has a connectivity measure at least equal or greater than any of the individual subsets). Families of multiresolution connectivity classes (more particularly, connectivity pyramids) can then be defined using such a measure of the connectivity of a set $A \in \mathcal{P}(E)$.

A special case of connectivity measure has been introduced in [1] using morphological dilation operators. This *dilation-based connectivity measure* has been defined as:

$$\mu_\delta(A) = m - \inf\left\{ r \in [0, m] : \delta_B^r(A) = A \oplus rB \in \mathcal{C} \right\} \tag{5}$$

where $m$ is an arbitrary positive real defining the maximum acceptable measure of connectivity (i.e. the maximum acceptable scale for dilations). This measure $\mu_\delta(A)$ quantifies in fact the notion of "how close" are the disconnected components of a set $A$, as interpreted by the number of dilations needed before $A$ becomes connected according to the usual definition of connectivity in an Euclidean topology.

However, what is needed in many image analysis problems (such as segmentation) is the inverse of the above, that is, to extract "strongly connected" (as opposed to "loosely connected") regions from an initially topologically connected set. The application of typical connected operators, such as the reconstruction openings/closings, leads to finding all connected regions of an image irrespective of the geometry of the path "tying together" these regions (that is, even if this path is "thin" and/or "long"). This is known as "leakage" problem, resulting in the creation of undesirable connections between large objects in an image due to the existence of thin connected paths between them (for instance, see [12]). To cover such situations, some new quantitative connectivity measures are introduced in the sequel, which will then support a multiscale hierarchical connectivity analysis.

*2.3. Generalization of Connectivity Measures*

In various image analysis problems we may be confronted with situations where the connected components of an image, containing a marker, have to be identified in order to evaluate some related morphological features. This is, for instance, the case when computing size distributions using reconstruction-based granulometries (as will be discussed in Section 4). In practice, however, such typical connected operators are known to present the drawback of reconstructing "too much", creating inappropriate connections in an image, between objects that should be intuitively considered as disjoint. This may be particularly undesirable in many situations where such loosely connected image components may need to be treated separately and differentiated from strongly connected ones. In such cases, the "degree of connectivity" has to be quantified, taking into account this form of additional geometrical information. In this paragraph, we propose the definition of generalized connectivity measures as a means to differentiate between strong or loose connections within an image and control the effect of the so-called "leakage" problem, resulting from the application of typical connected operators.

We illustrate this concept by an example, which will be used in the sequel as a means to validate the correctness of our approach in terms of defining appropriate connectivity measures. Let's consider, for instance, the three different sets $A_1, A_2, A_3 \subseteq \mathbb{R}^2$, shown in Fig. 1. Each one of these sets is initially topologically connected according to the pathwise definition of connectivity. What we need to define is a measure of the connectivity $\mu : \mathcal{P}(E) \to [0, 1]$ such that $\mu(A_1) > \mu(A_2) > \mu(A_3)$. In fact, $\mu(.)$ could be a non-negative function taking values $\mu(A) \to 0$ when $A$ is considered "nearly disconnected", and $\mu(A) \to 1$ when $A$ is considered "completely connected". We could thus define a *generalized connectivity measure* on $\mathcal{P}(E)$ as follows:

*Definition 3* (Generalized Connectivity Measure). A function $\mu : \mathcal{P}(E) \to [0, 1]$ constitutes a *generalized connectivity measure* on $\mathcal{P}(E)$ if:

(i) $\mu(\emptyset) = 0$ and
(ii) $\mu(\bigcup A_i) \geq \inf\{\mu(A_i), \mu(\bigcap A_i), \mu(A_i \setminus \bigcap A_i)\}$, $\forall A_i \in \mathcal{P}(E)$

The modification of condition (ii) above, with respect to Definition 2, states that the union of two sets cannot be "less connected" than the "least connected" element of the sets themselves and their intersection. The concept of *generalized connectivity class* can then be equivalently defined by strengthening condition (ii) of Definition 1, and introducing a more strict criterion to ensure that the union of some connected intersecting sets yields a new set that is itself "connected". Our goal is to extend Definition 1 to cover classes of "strongly connected" sets, i.e. sets that cannot be "easily" partitioned into disjoint components, for instance under the recursive application of an antiextensive morphological operator.

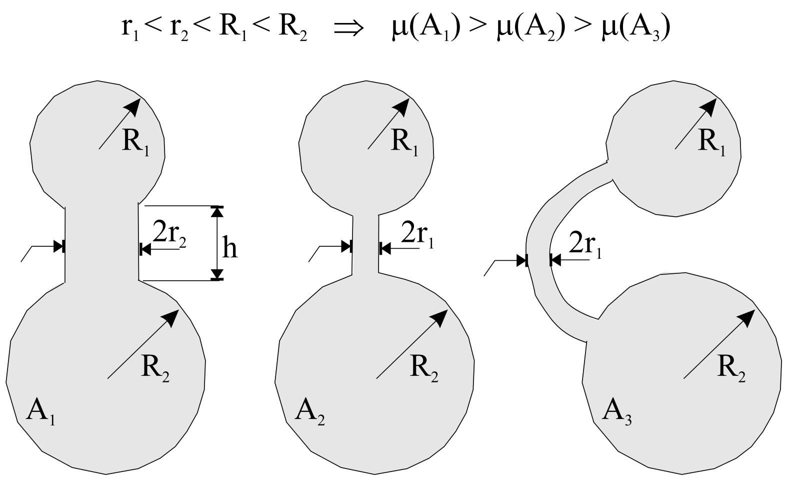



*Figure 1.* Generalized connectivity measure for three different sets: $\mu(A_1) > \mu(A_2) > \mu(A_3)$.

*Definition 4* (Generalized Connectivity Class). A subset $\mathcal{K} \subseteq \mathcal{P}(E)$ is called a *generalized connectivity class* if the following property holds:

$$X_i \in \mathcal{K}, \forall i \in I \quad \text{and} \quad \left\{ \bigcap_{i \in I} X_i, X_i \backslash \bigcap_{i \in I} X_i \right\} \in \mathcal{K} \setminus \emptyset$$
$$\Rightarrow \bigcup_{i \in I} X_i \in \mathcal{K}$$

where $I$ is an arbitrary index set.

The above condition states in fact that the union of two intersecting connected sets $(X_1, X_2)$ remains itself connected if the intersection $X_1 \cap X_2$ is "strong" enough to "adequately connect" the two sets. We can then associate a pyramid of generalized connectivity classes $\mathcal{K}_\mu$ with every connectivity measure $\mu$. For every $c \in [0, 1]$ we can define a *generalized connectivity class*:

$$\mathcal{K}_\mu^c = \{X \in \mathcal{P}(E) : \mu(X) \geq c\} \tag{6}$$

In other words, $\mathcal{K}_\mu^c$ contains all sets $X$ that have a generalized connectivity measure $\mu(X)$ greater than or equal to $c$. This means that:

$$\forall c_1, c_2 \in [0, 1],\ c_1 \leq c_2 \Rightarrow \mathcal{K}_\mu^{c_2} \subseteq \mathcal{K}_\mu^{c_1}.$$

The above definitions provide a unified theoretical framework enabling us to incorporate not only clustering connectivities (such as the dilation-based connectivity mentioned above) but also the inverse, that is, "partitioning connectivities" (like the erosion-based connectivity that will be discussed in the following paragraph), which can then form the basis for a multiscale connectivity analysis and support the definition of generalized connected operators.

### 2.4. Connectivity Measures Based on Morphological Operators

Let's start considering now some particular cases of generalized connectivity measures having the capacity to distinguish between the sets illustrated in Fig. 1. To differentiate between sets $A_1$ and $A_2$ we can define a connectivity measure based on some antiextensive morphological operator, like an antiextensive erosion $\varepsilon_B$ or opening $\gamma_B$. Such a connectivity measure $\mu_\varepsilon$ (or $\mu_\gamma$) would indicate in fact "how fast" a set $A \in \mathcal{C}$ becomes disconnected after the recursive application of an antiextensive operator. In the rest of the paper, we use some form of exponential function to define morphological connectivity measures. We can thus define an erosion-based connectivity measure as follows.

*Definition 5* (Erosion-based Connectivity Measure). Let $\varepsilon_B^r(X) = X \ominus rB$ denote a multiscale erosion on $\mathcal{P}(E)$, with $B$ being a compact convex structuring element. A function $\mu_\varepsilon : \mathcal{P}(E) \to [0, 1]$ defined as:

$$\mu_\varepsilon(X) = 1 - e^{-\lambda r_\varepsilon(X)} \quad \text{with}$$
$$r_\varepsilon(X) = \inf\{r \geq 0 : \varepsilon_B^r(X) \notin \mathcal{C} \setminus \emptyset\} \tag{7}$$

is called erosion-based connectivity measure, where $\lambda > 0$ is a parameter that determines the rate of the exponential function.

According to this definition, when $r_\varepsilon(X)$ (which we call *erosion-based connectivity degree*) equals zero, meaning that the set $X$ is already disconnected, then $\mu_\varepsilon(X) = 0$, while for compact sets $X$ with $r_\varepsilon(X)$ taking large values we get $\mu_\varepsilon(X) \to 1$.

In other words, the erosion-based connectivity measure indicates "how fast" a set becomes disconnected (or vanishes) under the recursive application of an erosion operator, thus corresponding to the width of the "narrowest" path between the major connected components of an image. In fact, this definition leads to a class of "second-order" (partitioning) connections. This means that, starting from a given connectivity class $\mathcal{C}$, a new erosion-based generalized connectivity class is created, where a set is considered as connected if it cannot be "easily partitioned" into two (or more) non-empty components through the application of an antiextensive morphological—erosion—operator. This concept is similar to the one underlying the definition of the dilation-based second-order clustering connections, which identify as connected the components that are "close enough" to each other, in the sense that they can be "easily clustered" together through the application of an extensive morphological—dilation—operator. Applying now Definition 5 for the two sets $A_1$ and $A_2$ of Fig. 1, we get:

$$\mu_\varepsilon(A_1) = 1 - e^{-\lambda r_2} \quad \text{and} \quad \mu_\varepsilon(A_2) = 1 - e^{-\lambda r_1},$$
$$\text{and since } r_1 < r_2 \Rightarrow \mu_\varepsilon(A_1) > \mu_\varepsilon(A_2).$$

The erosion-based connectivity, as defined above, does not distinguish though between the two sets $A_2$ and $A_3$ of Fig. 1. For these sets we have: $r_\varepsilon(A_2) = r_\varepsilon(A_3) = r_1$ and hence $\mu_\varepsilon(A_2) = \mu_\varepsilon(A_3)$. Therefore, a different kind of connectivity measure must be defined to cover such situations. Such a measure must take into account, not only the "width" but also the "length" of the narrowest path within a topologically connected set. One could think, for instance, of employing some form of dilation-based connectivity measure, such as the one mentioned in Section 2.2. However, this measure should be here extended using conditional dilation operators, in order to take into account connectivity information (path geometry etc.) contained in the original set. One way of doing this is to recursively apply some form of anti-extensive morphological operator, until the set becomes disconnected (as it has already been performed in Definition 5), and subsequently use some form of extensive operator (e.g. conditional dilations, with the original set as a mask) until the set becomes once again connected.

This consecutive application of one anti-extensive morphological operator followed by an extensive one, in order to quantify some sort of connectivity properties, could lead to the definition of a connectivity measure based on adjunctions. Let for instance $\alpha = (\varepsilon_B, \delta_B)$ denote an adjunction on $\mathcal{P}(E)$. Then we could define an *adjunctional connectivity measure* as a function $\mu_\alpha : \mathcal{P}(E) \to [0, 1]$ such that:

$$\mu_\alpha(X) = e^{-\lambda r_\alpha(X)} \tag{8}$$

with

$$r_\alpha(X) = \inf\{r \in \mathbb{N} : \delta_B^r(\varepsilon_B^{r_\varepsilon(X)}(X) \mid X) \in \mathcal{C}\}$$

where $r_\varepsilon$ refers to the erosion-based connectivity degree of $X$ (Definition 5), and $\delta_B(X \mid Y)$ denotes the conditional dilation of set $X$ using $Y$ as a mask:

$$\begin{aligned} &\delta_B^r(X \mid Y) = \delta_B(\delta_B^{r-1}(X \mid Y) \mid Y) \quad \text{and} \\ &\delta_B(X \mid Y) = (X \oplus B) \cap Y \end{aligned} \tag{9}$$

We then get for the example-sets of Fig. 1:

$$r_\alpha(A_2) < r_\alpha(A_3) \Rightarrow \mu_\alpha(A_2) > \mu_\alpha(A_3)$$

However, $r_\varepsilon(A_1) = r_2 > r_1 = r_\varepsilon(A_2)$, and thus: $r_\alpha(A_1) = r_2 + (h/2) > r_1 + (h/2) = r_\alpha(A_2)$, which leads to an undesirable result: $\mu_\alpha(A_1) < \mu_\alpha(A_2)$.

In other words, applying the connectivity measures defined above in the case of the example illustrated in Fig. 1, we conclude that the erosion-based connectivity measure succeeds in differentiating only between sets $A_1$ and $A_2$, while the adjunctional connectivity measure is successful only for sets of the form $A_2$ and $A_3$ failing to correctly discern sets $A_1$ and $A_2$. In order to cover all these situations successfully based on a single morphological connectivity measure, a multiscale connectivity function is introduced in the following paragraph, extending the definition of adjunctional connectivity measure. An average adjunctional connectivity measure can then be defined that leads to the desired results, as will be illustrated using the example sets of Fig. 1.

### 2.5. *Multiscale Connectivity Function*

We now extend the morphological connectivity measures introduced in the previous paragraph and define a multiscale connectivity function as follows:

*Definition 6* (Adjunctional Multi-Scale Connectivity Function). Let $\alpha = (\varepsilon_B, \delta_B)$ denote an adjunction on $\mathcal{P}(E)$. A function $\mu_\alpha : \mathcal{P}(E) \times \mathbb{R}_+ \to [0, 1]$ defined as:

$$\begin{aligned} &\mu_\alpha(X, s) = e^{-\lambda r_\alpha(X,s)} \quad \text{with} \\ &r_\alpha(X, s) = \inf\{r \in \mathbb{N} : \delta_B^r(\varepsilon_B^s(X) \mid X) \in \mathcal{C}\backslash\emptyset\} \end{aligned} \tag{10}$$

is called adjunctional connectivity function and gives a measure of the connectivity of a set $X$ at scale $s$.

Applying this definition for the sets $A_1, A_2, A_3$ of Fig. 1 we obtain the three connectivity profiles shown in Fig. 2, where we have taken $\lambda = 0.05$ and $r_1 = 4, r_2 = 10, R_1 = 20, R_2 = 25, h = 25$ (in arbitrary length units), while a unit disk has been used as structuring element for the morphological operators. We may note here that:

$$\begin{gathered} \mu_\alpha(A_i, s) = 1, \text{ for all scales } s: \\ 0 < s < r_1 \text{ and } R_1 < s < R_2 \\ r_\varepsilon(A_1) = r_2 > r_1 = r_\varepsilon(A_2) = r_\varepsilon(A_3) \end{gathered}$$

(the connectivity function equals 1, for large scales $s: R_1 < s < R_2$ where there is only one connected component left in the image, while the rest of it has vanished under the application of $\varepsilon^s$).

For the sets $A_1$ and $A_2$ we have:

$$\begin{aligned} &\mu_\alpha(A_1, s) = 1 > \mu_\alpha(A_2, s), \ \forall s : r_1 < s < r_2 \quad \text{and} \\ &\mu_\alpha(A_1, s) = \mu_\alpha(A_2, s), \ \forall s : r_2 < s < R_1 \end{aligned}$$

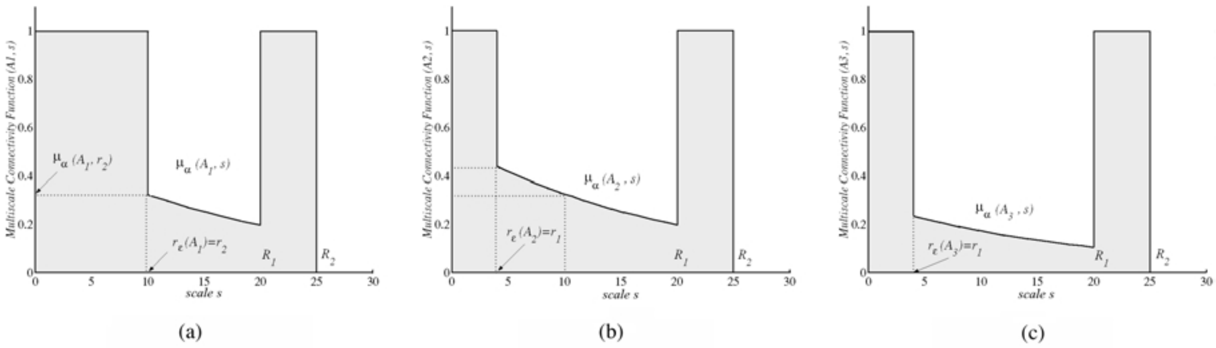



*Figure 2.* Adjunctional connectivity functions for the three sets $A_1$, $A_2$, $A_3$ of Fig. 1. (a) $\mu_\alpha(A_1, s)$; (b) $\mu_\alpha(A_2, s)$; and (c) $\mu_\alpha(A_3, s)$.

For the sets $A_2$ and $A_3$ we also have:

$$r_\alpha(A_2, s) < r_\alpha(A_3, s) \Rightarrow \mu_\alpha(A_2, s) > \mu_\alpha(A_3, s), \quad \forall s : r_1 < s < R_1$$

Thus, for all scales $s > 0$ we have:

$$\mu(A_1, s) \geq \mu(A_2, s) \geq \mu(A_3, s)$$

which means that for the integrals $\int \mu_\alpha(A_i, s)\, ds$, represented by the shaded regions in Fig. 2, we have:

$$\int_s \mu_\alpha(A_1, s)\, ds > \int_s \mu_\alpha(A_2, s)\, ds > \int_s \mu_\alpha(A_3, s)\, ds$$

This observation leads to the definition of a new average connectivity measure, as follows.

*Definition 7* (Average Adjunctional Connectivity Measure). Let $\alpha = (\varepsilon_B, \delta_B)$ denote an adjunction on $\mathcal{P}(E)$ and $\mu_\alpha : \mathcal{P}(E) \times \mathbb{R}_+ \to [0, 1]$ an adjunctional multi-scale connectivity function. We call average adjunctional connectivity measure the function $\bar{\mu}_\alpha(.) : \mathcal{P}(E) \to [0, 1]$ defined as:

$$\bar{\mu}_\alpha(X) = \frac{\int_{s=0}^{s_{\max}} \mu_\alpha(X, s)\, ds}{s_{\max}} \tag{11}$$

where $s_{\max}$ is a normalizing factor indicating the maximum applicable scale (for erosions), and can be defined as:

$$s_{\max} \in \mathbb{R}_+ : \; \varepsilon_B^s(X) = \emptyset, \; \forall s > s_{\max}, \; X \in \mathcal{P}(E).$$

The connectivity function, as defined above, contains useful geometrical information related to the connectivity structure (shape/size) of a set at multiple scales. The connectivity profile of a set incorporates important morphological cues, interpreting how "easily" the set becomes partitioned into disjoint components and providing a measure of the "distance" between these principal connected components. In the sequel, the adjunctional connectivity function and its algorithmic computation form the basis of a multiscale connectivity analysis framework. The concept of *connectivity tree* is introduced and an algorithm for its creation is presented, leading to a hierarchical partitioning of a set into connected components with progressively increasing average connectivity measure.

## 3. Multiscale Hierarchical Connectivity Analysis

### 3.1. *The Connectivity Tree*

In this section, the concept of multiscale connectivity function, together with the generalized morphological connectivity measures introduced above, are used to establish the theoretical framework for a hierarchical connectivity image analysis. The basic idea lies on the following observation: it is well known that a binary image (i.e. a set $X \subseteq \mathbb{R}^2$, which can in fact correspond to the thresholding of a grayscale image at a particular gray-level) can be decomposed into a collection of disjoint connected components $\{X_i \subseteq X : X_i \in \mathcal{C}, \; \bigcap X_i = \emptyset$ and $\bigcup X_i = X\}$, which constitutes a *partition* of $X$ [13]. For each one of these components, a multiscale connectivity function can be computed providing useful geometrical cues related to its "connectivity structure", as explained in the previous section. Based on this information, each component $X_i$ can be further partitioned into a set of new connected components $\{Y_j^i : j \in J_i\}$ such that: $\bigcup_j Y_j^i = X_i$ and

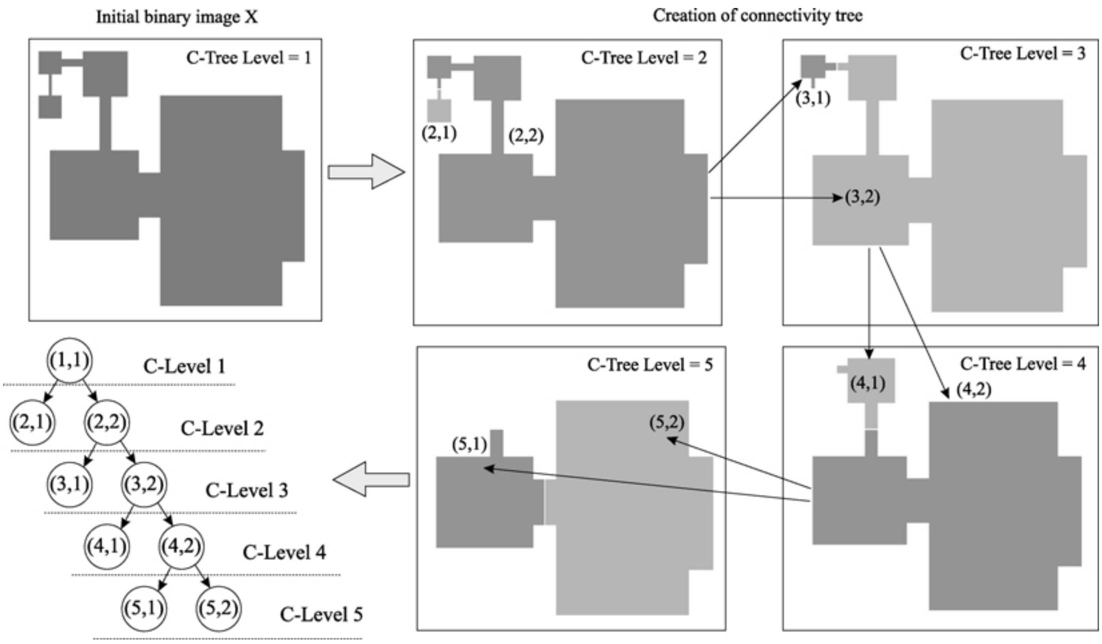



*Figure 3*. Connectivity Tree (C-Tree) decomposition for an example binary image-set.

$\mu(Y_j^i) \geq \mu(X_i)$, $\forall j \in J_i$ (where $J_i$ is an appropriate index set). In other words, each $Y_j^i$ deriving from $X_i$, is more strongly connected than its predecessor $X_i$. This hierarchical decomposition into progressively stronger connected components leads in fact to the creation of a tree representation of the initial binary image, where at each level $k$:

$$\left\{X_i^k\right\}_{i=1,\dots,n_k} \mapsto \left\{X_j^{(k+1)}\right\}_{j=1,\dots,n_{k+1}}$$

with $X_i^k$ denoting the $i$th node at level $k$, and $n_k$ the total number of nodes at this level. This representation is called *Connectivity-Tree* (or C-Tree).

The concept of connectivity tree can be better explained by an example, illustrated in Fig. 3. At level 1 the C-tree contains the initial decomposition of the original image $X$ into connected components (that is, only one component for the example set of Fig. 3). At each level $k$, the C-tree contains nodes $\{(k, i)\}_{i=1,\dots,n_k}$ ($n_k = 2$, for all C-tree levels of Fig. 3), corresponding to the set of connected sub-components $\{X_i^k\}_{i=1,\dots,n_k}$ of $X$. Each node is recursively partitioned into a set of new connected components, with progressively increasing generalized connectivity measure. As will be explained in the following paragraph (describing an algorithm for the creation of C-trees), this is accomplished through the recursive application of an antiextensive morphological operator, such as erosion $\varepsilon_B^s(\cdot)$ with progressively increasing scale $s$, following a procedure similar to the one used in the previous section to compute adjunctional multiscale connectivity functions. The leaf-nodes of the C-tree representation then contain binary image components that cannot be further partitioned, that is, vanish completely after the recursive application of such an antiextensive morphological operator.

This hierarchical image representation enables the definition of new generalized connected operators, like for instance, modified connected openings, which treat as connected image components $X_i^k$ that either correspond to leaf-nodes or satisfy the criterion: $\mu(X_i^k) \geq \theta$, where $\theta$ is a given thresholding connectivity measure. We can thus define extensions of typical connected operators, such as reconstruction or area openings/closings, where any of the generalized connectivity measures $\mu$ introduced above can be used to define different criteria. When the adjunctional multiscale connectivity function of Definition 6 is used, this criterion becomes:

$$\mu\left(X_i^k, s\right) \geq \theta(s), \quad \forall s \geq 0 \tag{12}$$

where $\theta : \mathbb{R}_+ \to [0, 1]$ in this case defines a thresholding multiscale connectivity profile. In the following paragraph, we present an algorithm for the creation of C-tree structure, integrating the computation of multiscale connectivity function for each tree-node.

### *3.2. Algorithmic Implementation*

The objective of this section is to describe an algorithmic implementation for the creation of connectivity trees (C-tree creation). This implementation relies mainly on two basic procedures:

(a) A procedure that finds and labels the connected components of a binary image-set (using 4- or 8-connectivity).

  Function: *find-connected-components*$(\cdot)$
  Input: binary image $X$,
  Output: a partition of $X$ consisting of a set $\{Y_j\}_{j=1,\ldots,n_c}$ of connected and disjoint marked components, that is: $Y_j \in \mathcal{C}$, $\bigcap Y_j = \emptyset$ and $\bigcup Y_j = X$.

(b) A conditional wavefront expansion algorithm that, given a binary image-set $X$ (mask) and a set $\{Y_j\}_{j=1,\ldots,n_c}$ (acting as a marker) of connected (and disjoint) marked image components, reconstructs the original image $X$ by concurrently expanding the boundaries of all $Y_j$, thus resulting in a new partition $\{Z_j\}$ of $X$. For each pixel belonging to the current boundary of $Y_j$, its neighbors inside $X$ are found and marked as belonging to the corresponding reconstructed set $Z_j$ (if yet unmarked). This procedure repeats itself recursively, and terminates when all the pixels of the original image $X$ are visited.

  Function: *wavefront-expansion*$(\cdot)$
  Inputs: $\{Y_j\}_{j=1,\ldots,n_c} : Y_j \in \mathcal{C}, \bigcap Y_j = \emptyset$ and $\bigcup Y_j \subseteq X$ and binary image $X$ (mask),
  Output: a partition of $X$, that is: a set $\{Z_j\}_{j=1,\ldots,n_c}$ of connected and disjoint components, such that: $Z_j \in \mathcal{C}$, $Y_j \subseteq Z_j, \bigcap Z_j = \emptyset$ and $\bigcup Z_j = X$.

The basic element that needs to be specified for the algorithmic implementation of the C-tree creation is the data-structure of each node (C-node) and the information it should contain. The fields used to represent the *C-node structure* corresponding to the connected image component $X_j^k$, are described hereafter.

**C-node structure** (corresponding to connected component $X_j^k$):

| | |
|---|---|
| c-level: | C-node level $k$ within the C-tree structure |
| c-index: | integer $j$ uniquely identifying the C-node within level $k$ |
| Image: | contains the binary image component $X_j^k$ |
| size: | maximum scale $s_{\max} : \forall s > s_{\max}$, $\varepsilon_B^s(X_j^k) = \emptyset$ |
| area: | Area$(X_j^k)$ |
| c-function[.]: | multiscale adjunctional connectivity function $\mu_\alpha(X_j^k, s)$, $0 < s < s_{\max}$ |
| num-of-children: | the number of children C-nodes in the hierarchical C-tree structure (0, if C-node (k, j) is a leaf-node) |
| child[.]: | pointers to the children C-node structures |

The algorithm for C-tree creation then resides on a recursive procedure (called *create-C-tree*(...)) that takes as input a C-node structure, constructs the children C-nodes and recursively calls itself creating the lower part of the C-tree hierarchy. This procedure thus consists of four main steps:

*Step 1*: Perform erosion $X_\varepsilon = \varepsilon_B^s(X)$ on the input image ($X$ = C-node $\rightarrow$ Image), with progressively increasing scale $s$ until $X_e$ is partitioned into $n_c > 1$ disjoint connected components $Y_j$ ($j = 1, \ldots, n_c$) (if $X$ vanishes completely for a particular scale $s$ without being partitioned into separated connected components, then the current C-node is a leaf node, with: C-node $\rightarrow$ size $= s$ ).
*Step 2*: Perform conditional wavefront expansion on the partition $\{Y_j\}$ of $X_\varepsilon$, to reconstruct a partition $\{Z_j\}$ of $X$, that is, a new set of disjoint connected components $Z_j$ such that: $\bigcup Z_j = X$.
*Step 3*: Create children C-node structures (child[j], for $j = 1, \ldots, n_c$). Call recursively the *create-C-tree*(child[j]) procedure.
*Step 4*: Compute the adjunctional multiscale connectivity function $\mu_\alpha(X_j^k, s)$.

The basic structure of the C-tree creation algorithm is described hereafter.

**Algorithm: C-Tree Creation**

```
Procedure   Create-C-Tree (C-node)
    /* Initialization: */
    X=C-node→Image;   level=C-node→c-level;   Xε = X; continue=TRUE;   s = 0;
    while (continue)
/* Step 1: Partition X into connected sub-components */
            s = s + 1 ;   Xε = εB(Xε);
            find-connected-components (Xε) ↦ {Yj}j=1,...,nc
                (i.e. Yj ∈ C :  ⋂ Yj = ∅  and  ⋃ Yj = Xε)
            if (nc == 0)          (i.e. Xε == ∅)
                continue = FALSE;
                C-node →  {size = s;   num-of-children = 0; }
            else if (nc == 1)
                C-node →  c-function[s] = 1;      //(i.e. μα(s) = 1)
            else          /* nc > 1,  in which case:
                                s = rε (erosion-based connectivity degree) */
                continue = FALSE;
/* Step 2: Reconstruct a partition {Zj} of X, starting from the partition {Yj} of Xε */
                Wavefront-Expansion ( {Yj} | X ) ↦ {Zj}
            /* {Zj} is a partition of X, i.e.: ∀j,   Yj ⊆ Zj,  ⋂ Zj = ∅  and  ⋃ Zj = X */
/* Step 3: Recursive creation of C-tree children-nodes */
                C-node →  num-of-children = nc;
                For all j = 1, . . . , nc
                    new-C-node  →  {c-level = level + 1;  Image = Zj;
                                        area = Area(Zj)}
                    Create-C-Tree (new-C-node);
                    C-node →  child[j] = new-C-node;
/* Step 4: Compute adjunctional connectivity function  →  μα(X, s) */
                while (Xε ≠ ∅)
                    p=1;      Xδ = Xε;      not-connected = TRUE;
                    while (not-connected)
                        Xδ = δB(Xδ | X)      /* conditional dilation */
                        if (Xδ is connected) then
                            not-connected = FALSE;
                            C-node → c-function[s] = exp(-lambda*p);
                              //(i.e. μα(s) = e^(−λ·p) )
                        else    p = p + 1;
                    end-while
                    Xε = εB(Xε);    s = s + 1;
                end-while
            end-if
    end-while
```

The time devoted by the *Create-C-Tree*$(\cdot)$ procedure for the computation of the multiscale connectivity function (Step 4 of the algorithm), can be drastically reduced if we take into account information provided by the evaluation performed for the children C-nodes, and then apply a simple approximative heuristic based on the fact that:

$$\mu_\alpha(X, s) \leq \mu_\alpha(child(X), s), \ \forall s \in (0, size(child(X)))$$

The information stored in each C-node structure during the C-tree creation procedure, can be used for the implementation of new connected morphological operators, such as generalized-connectivity openings/closings. These operators can then form the theoretical basis for the definition of generalized (connected) granulometries, as described in the following section. Such generalized morphological operators can be also applied to image segmentation problems. For instance, at a post-processing stage following the marker extraction phase, they could constitute a useful tool to improve the connectivity structure of markers, prior for instance to the application of a watershed procedure [17]. The goal in this case would be to reduce undesirable effects of the so-called "leakage" related to the creation of "loose" connections that may result from the application of typical connected operators.

## 4. Generalized Connectivity Granulometries

The goal in this section is to apply the multiscale connectivity analysis framework proposed above, in order to define new connected operators, with a specific application in mind, that of granulometric image analysis [8, 10]. Performance of these new generalized granulometries, as compared to the use of typical connected operators (such as reconstruction openings), is demonstrated in a particular problem concerning morphological evaluation of sample soilsection images.

### 4.1. *Granulometry and Size Distribution: Introductory Elements*

Granulometry constitutes a very useful and versatile tool of mathematical morphology. It is a parameterized family $\{\gamma_s\}_{s=0,1,\ldots}$ of openings that satisfy:

$$s_1 \geq s_2 \Rightarrow \gamma_{s_1}(X) \subseteq \gamma_{s_2}(X),\ \forall s_1, s_2 \geq 0,\ X \in \mathcal{P}(E).$$

A useful granulometry is obtained by applying typical morphological openings and setting: $\gamma_s(X) = X \circ sB$ $(s = 1, 2, \ldots)$. For discrete scales $s$ the structuring element $sB$ can be defined recursively as: $sB = (s-1)B \oplus B$ ($B$ is a basic finite structuring element, e.g. unit ball). The application of such consecutive opening operators leads to a progressive smoothing (filtering) of the image, successively cutting off the sharp "light" (white) areas of the image not "large" enough to "contain" the structuring element $sB$. Therefore, evaluating the evolution of these signals $\gamma_s$ in multiple scales $s$ can give useful information on the "power" (area or volume) of the light (dark) areas of the image in each scale $s$, and can lead to the extraction of very important features concerning the distribution of sizes within the image.

The dual equivalent of the above is the parameterized family $\{\beta_s\}_{s=1,2,\ldots}$ of closings, with $\beta_s(X) = X \bullet sB$, for $s = 1, 2, \ldots,$ such that: $s_1 \geq s_2 \Rightarrow \beta_{s_1} \geq \beta_{s_2}$. This is known as *antigranulometry* associated with $\{\gamma_s\}_{s=0,1,\ldots}$. The two families $\{\gamma_s\}$ and $\{\beta_s\}$ $(s = 1, 2, \ldots)$ may be considered as a unified sequence of nonlinear morphological filters, leading to a multiresolution decomposition of an image $X$:

$$\Gamma(X) = \{\ldots, \beta_2(X), \beta_1(X), X, \gamma_1(X), \gamma_2(X), \ldots\} \tag{13}$$

This image analysis methodology using nonlinear morphological operators provides information not in the field of frequency, as is the case of the classical linear operators, but in relation with variable "sizes" (scales), in the sense that the variations of $\gamma_s(X)$ and $\beta_s(X)$ in multiple scales $s$ indicates the distribution of respective sizes in the image $X$, depending on the form of the structuring element $B$.

The concept of multiscale granulometric analysis thus leads to two important tools of mathematical morphology known as *size distribution* and *size density*. Let's consider for instance a binary image $X$. Size distribution can be defined as:

$$S_X(s) = \frac{1}{\text{Area}\,(W)} \begin{cases} \text{Area}\,(\gamma_s(X)) & \text{for } s \geq 0 \\ \text{Area}\,\big(\beta_{|s|}(X)\big) & \text{for } s \leq -1 \end{cases} \tag{14}$$

where Area$\,(X)$, indicates in fact the number of white pixels for a binary image $X$, and $W \supseteq X$ is the analysis window of $X$.

The *size density* of an image $X$ can be then defined as:

$$D_X(s) = \frac{1}{\text{Area}\,(W)} \times \begin{cases} \text{Area}\,(\gamma_s(X) - \gamma_{s+1}(X)) & \text{for } s \geq 0 \\ \text{Area}\,\big(\beta_{|s|}(X) - \beta_{|s|-1}(X)\big) & \text{for } s \leq -1 \end{cases} \tag{15}$$

It is clear from the above that the size distribution $S_X(s)$ reflects the "weight" of each individual component $\gamma_s$ and $\beta_s$ ($s = 1, 2, \ldots$) in the multiscale decomposition $\Gamma(X)$ of image $X$, while the size density $D_X(s)$ gives information on the "weight" of elements $\{\gamma_s - \gamma_{s+1}\}$ and $\{\beta_{s+1} - \beta_s\}$ ($s = 1, 2, \ldots$), which form a new *size-density multiscale decomposition* of image $X$ defined as:

$$\Delta(X) = \{\ldots, \beta_2(X) - \beta_1(X),\ \beta_1(X) - X,\ X - \gamma_1(X),\ \gamma_1(X) - \gamma_2(X), \ldots\} \quad (16)$$

Granulometries can be applied in a variety of image analysis problems obtaining useful results. In this paper, we study the application of such methods in the particular problem of analyzing geological images taken from sample soil-sections, in order to investigate the possibility of developing a novel computer-vision based tool for automatic analysis and quality assessment of soil regions. Evaluating the differences of consecutive openings/closings in multiple scales $s$ applied in a sample image, and analyzing the evolution of the signal $D_X(s)$, may lead to the identification of characteristic sizes (related to the scale) in the image. These data may facilitate significantly the extraction of interesting conclusions regarding the structure of the particular soilsection, and may thus constitute a useful tool for analysing and evaluating the quality of the respective soil sample.

### 4.2. Connected Operators and Granulometries

The granulometric analysis described in the previous paragraph is based on the application of typical morphological operators (openings/closings). This procedure leads to the computation of the size density signal $D_X(s)$, which, as discussed above, provides useful multiscale information concerning the presence of characteristic sizes within the image. A major issue, however, is to ensure that the variations of the size density signal are as representative as possible of such morphological features, indicating with precision and reliability the presence of characteristic elements in the image. For this reason, the use of connected operators could prove more appropriate.

A connected operator is an operator that coarsens the partition of an image into (foreground/background) connected components [4]. The concept of partition of a space $A$ is defined as a set of connected components $\{A_i\}$ that are disjoint, and for which: $\bigcup A_i = A$ [13]. A partition $\{A_i\}$ is said to be coarser (finer) than another one $\{B_j\}$ if any pair of points belonging to the same $B_j$ ($A_i$) also belong to a unique $A_i$ ($B_j$). The main property of connected operators is that, as opposed to classical morphological operators that perform "local" functions using a structuring element, they do not change values at individual pixel level but, instead, treat entire connected components as a whole, operating on the flat zone level. A direct consequence is that connected operators do not affect boundaries in an image, thus more accurately preserving contour (size/shape) information. This characteristic property of connected operators can prove of importance for many applications, including granulometric analysis and size density computation that constitutes the focus of this section.

A typical example of connected operator is the (conditional) *reconstruction opening*, which can be obtained by iterating conditional dilations:

$$\rho(Y \mid X) = \lim_{n\to\infty} \delta_B^n(Y \mid X) \quad (17)$$

where $Y$ is in fact a marker that is used to reconstruct (part of) the original image-set $X$ (mask). It can be proved that $\rho(Y|.)$ constitutes an opening that extracts all the connected components of image $X$ intersecting marker $Y$. A very common implementation of reconstruction is obtained using as marker the result of a typical opening $\gamma_s(X) = X \circ sB$ at a particular scale $s$. This operator is often called *multiscale opening-by-reconstruction*, defined as follows:

$$\rho_s(X) = \rho(\gamma_s(X) \mid X) = \lim_{n\to\infty} \delta_B^n(\gamma_s(X) \mid X) \quad (18)$$

The dual closing-by-reconstruction operator can be defined similarly.

These opening/closing-by-reconstruction filters are very useful since they simplify the input image eliminating completely the elements of size less than the scale $s$ of the filter, while preserving exactly all larger connected components. Another typical connected operator is the area opening/closing, which extracts all connected components that have an area (number of pixels for the binary case) not below a given threshold. Any type of such connected morphological operators can be used to define new granulometries. For instance, *reconstruction granulometries* can be defined using a family $\{\rho_s\}$ of reconstruction operators. Similarly, area granulometries can be obtained using area openings/closings.

All these typical connected operators operate on an image by finding and extracting with precision (boundary preservation property) all connected components that satisfy a criterion (i.e. they have a size large enough to contain a structuring element, in the opening-by-reconstruction case, or their area is not below a threshold, for the area opening case). Although this may be useful in many image analysis problems where exact contour information is important, the major drawback of these typical connected operator lies on the fact that they cannot distinguish between "loosely" and "strongly" connected image components. Such additional information related to the "degree" of connectivity of an image, could be quantified using the generalized connectivity measures defined in Section 2. In the sequel, the multiscale connectivity analysis framework introduced in Section 3, is applied to this problem of defining generalized granulometries for morphological size/shape evaluation.

### *4.3. Multiscale Connectivity and Generalized Granulometries*

As we have described in Section 3, a binary image-set $X$ can be represented by a hierarchical tree-structure called connectivity tree (or C-tree). At each level $k$, the nodes of the C-tree correspond to connected components $\{X_j^k\}_{j=1,\ldots,n_k}$. This set of connected components at each C-tree level, constitutes in fact a partition for the opening-by-reconstruction at a particular scale $s_k \geq k$, that is: $\bigcap_j X_j^k = \emptyset$ and $\bigcup_j X_j^k = \rho_{s_k}(X)$.

Using the information stored at each C-node, as described in paragraph 3.2 during the recursive C-tree creation procedure, new generalized multiscale connectivity operators can be defined that extract all C-node components satisfying a particular criterion. Such criteria can be based on the *size* (i.e. max scale $s_{\max} : \varepsilon^s(X_j^k) = \emptyset,\ \forall s > s_{\max}$) and on a *generalized connectivity measure* $\mu$ of each C-node (e.g. the *multiscale connectivity function* $\mu_\alpha(X_j^k, s)$). Generalized connectivity operators can then be defined as follows:

$$\phi_s(X \mid \theta) = \bigcup \left\{ X_j^k \in \text{C\_Tree}(X) : \text{size}\left(X_j^k\right) \geq s \text{ and } \mu\left(X_j^k\right) \geq \theta \right\} \tag{19}$$

where $\theta$ is a given thresholding connectivity measure, and *size*$(X)$ can be for instance defined as the maximum scale $s_{\max} : \forall s > s_{\max},\ \varepsilon_B^s(X) = \emptyset$. Applying these operators on a multiscale basis, we can now define a class of *generalized connectivity granulometries* $\{\phi_s\}$, leading to a new size-density multiscale decomposition $\Phi$ of image $X$:

$$\Phi(X \mid \theta) = \{X - \phi_1(X \mid \theta), \phi_1(X \mid \theta) - \phi_2(X \mid \theta), \phi_2(X \mid \theta) - \phi_3(X \mid \theta), \ldots\}$$

The new *generalized-connectivity size-density* of image $X$ is then defined as:

$$D_X(s \mid \theta) = \frac{\text{Area}\,(\phi_s(X \mid \theta) - \phi_{s+1}(X \mid \theta))}{\text{Area}\,(W)} \tag{20}$$

Let's apply now this generalized connectivity granulometry to the example-set of Fig. 3. The multiscale connectivity profiles for two nodes (2,2) and (4,2) are shown in Fig. 4. The size-densities resulting

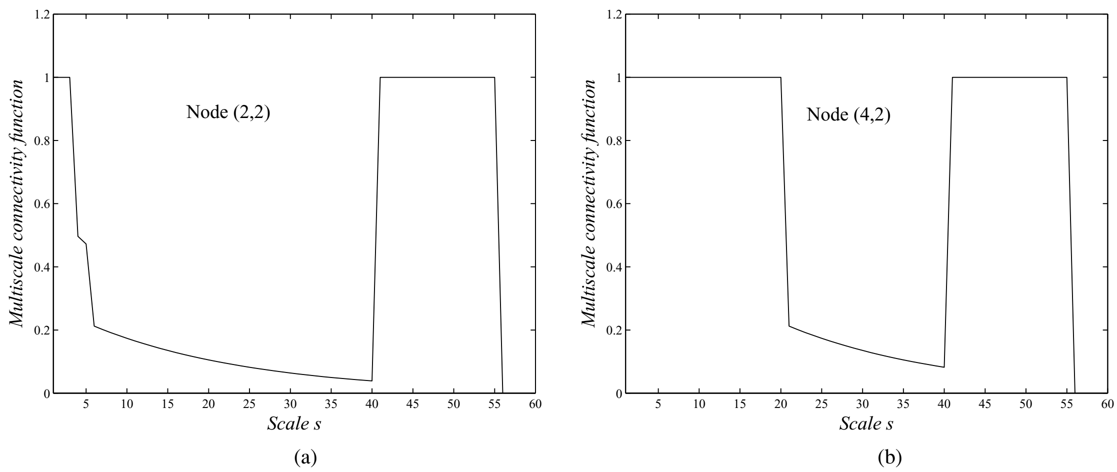



*Figure 4.* Multiscale connectivity functions for two C-nodes of the example image-set of Fig. 3. (a) C-node (2,2) and (b) C-node (4,2).

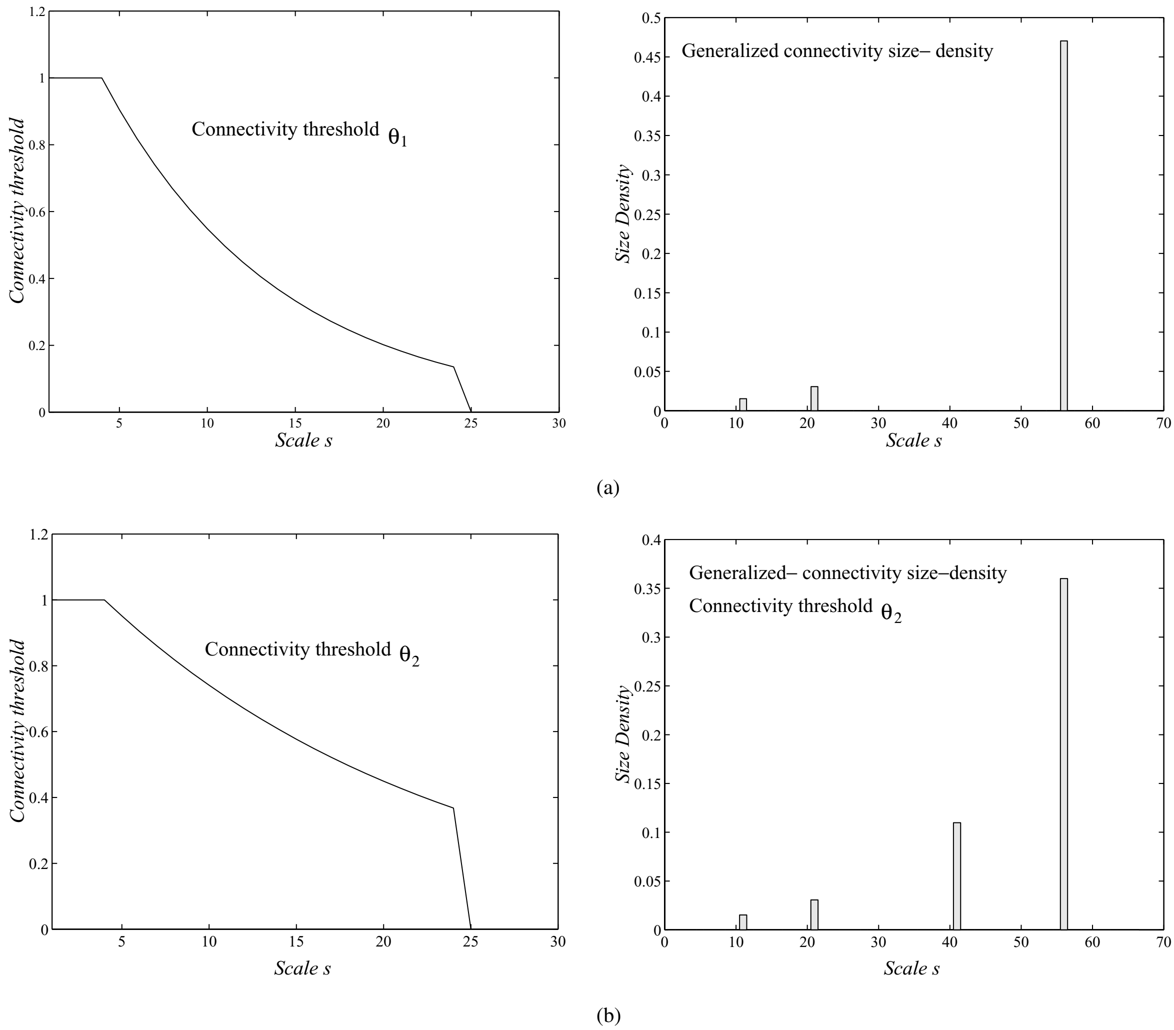



*Figure 5*. Size densities for the example image-set of Fig. 3 using generalized-connectivity granulometries with two different thresholding connectivity profiles. (a) Generalized size density with thresholding connectivity profile $\theta_1$ and (b) Generalized size density with thresholding connectivty profile $\theta_2$.

from the application of generalized-connectivity granulometries, for two different thresholding connectivity profiles $\theta_1(s)$ and $\theta_2(s)$, are shown in Fig. 5(a) and (b) respectively. We may note that, depending on the "strictness" of the thresholding connectivity profile, the functionality of the operator change, treating for instance C-node (4,2) either as connected (case a, connectivity threshold $\theta_1$), or as disconnected (case b, with connectivity threshold $\theta_2$), in which case the resulting size density provides indication on all four characteristic sizes included in the test image-set of Fig. 3. This example provides a first illustration of the power and versatility of the proposed generalized-connectivity multiscale operators as related to the reliable preservation of shape/size information within an image. To further demonstrate the functionality of these operators, the generalized granulometric image analysis framework is applied to a particular problem concerning morphological evaluation of soilsection images, as described in the following paragraph.

### 4.4. *Application to Soilsection Image Analysis*

In automated soilsection image analysis, a very important task is to detect compound soil formations (e.g.

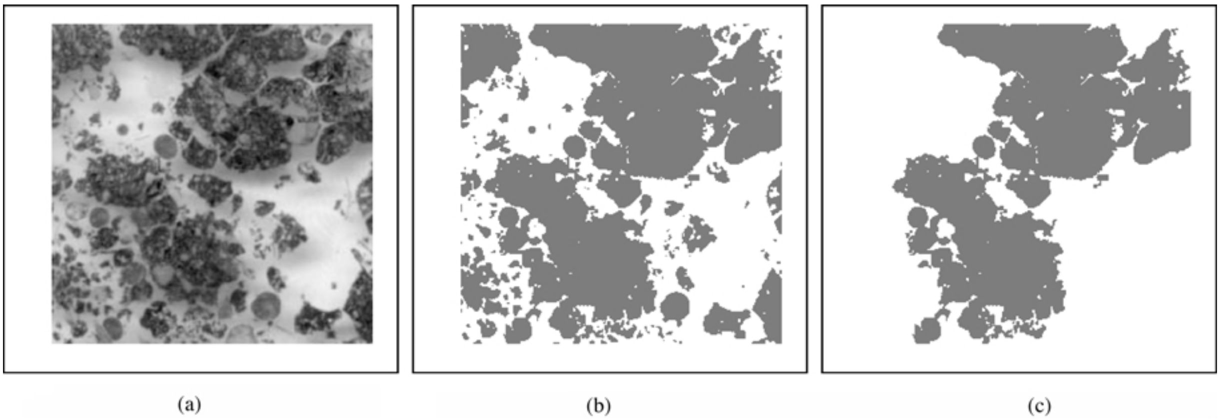



*Figure 6*. Sample soilsection-1: test image for granulometric analysis experiments. (a) Original grayscale image; (b) Thresholded binary image; and (c) A single connected component.

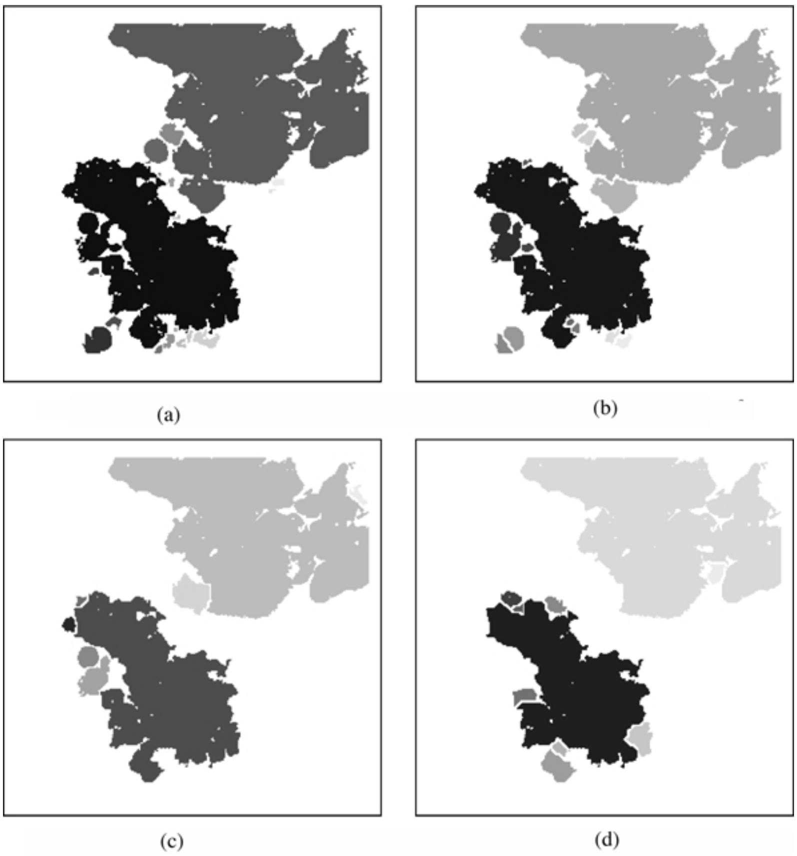



*Figure 7*. Multiscale connectivity partition: Four different C-tree levels for the sample soilsection-1. (a) Connectivity level = 1; (b) Connectivity level = 2; (c) Connectivity level = 3; and (d) Connectivity level = 4.

elementary objects, grains etc.), differentiating them from void space, in order to evaluate pertinent morphological properties such as size/shape distributions. Soil structure is concerned with the arrangement of primary particles and voids and the variations of size/shape characteristics. Soilsection images exhibit a great variety of geometric features which can be either 1D, such as edges or curves, or 2D such as light or dark blobs (small homogeneous regions of uncertain shape, which sometimes seem to be randomly distorted circles or ellipses). Extraction of such morphological features and estimation of their numerical properties, like size density, can thus provide useful information for the evaluation of soil structure quality. In order to obtain reliable and representative measures of such properties, based for instance on morphological granulometries, the operators applied need to detect with precision all important homogeneous regions constituting soil formations of interest within the image. The use of connected operators in a granulometric image analysis framework can thus prove a good choice, globally treating connected regions in an image while preserving important contour information.

However, as it has already been mentioned above, typical connected operators such as reconstruction or area openings, present an important drawback, known

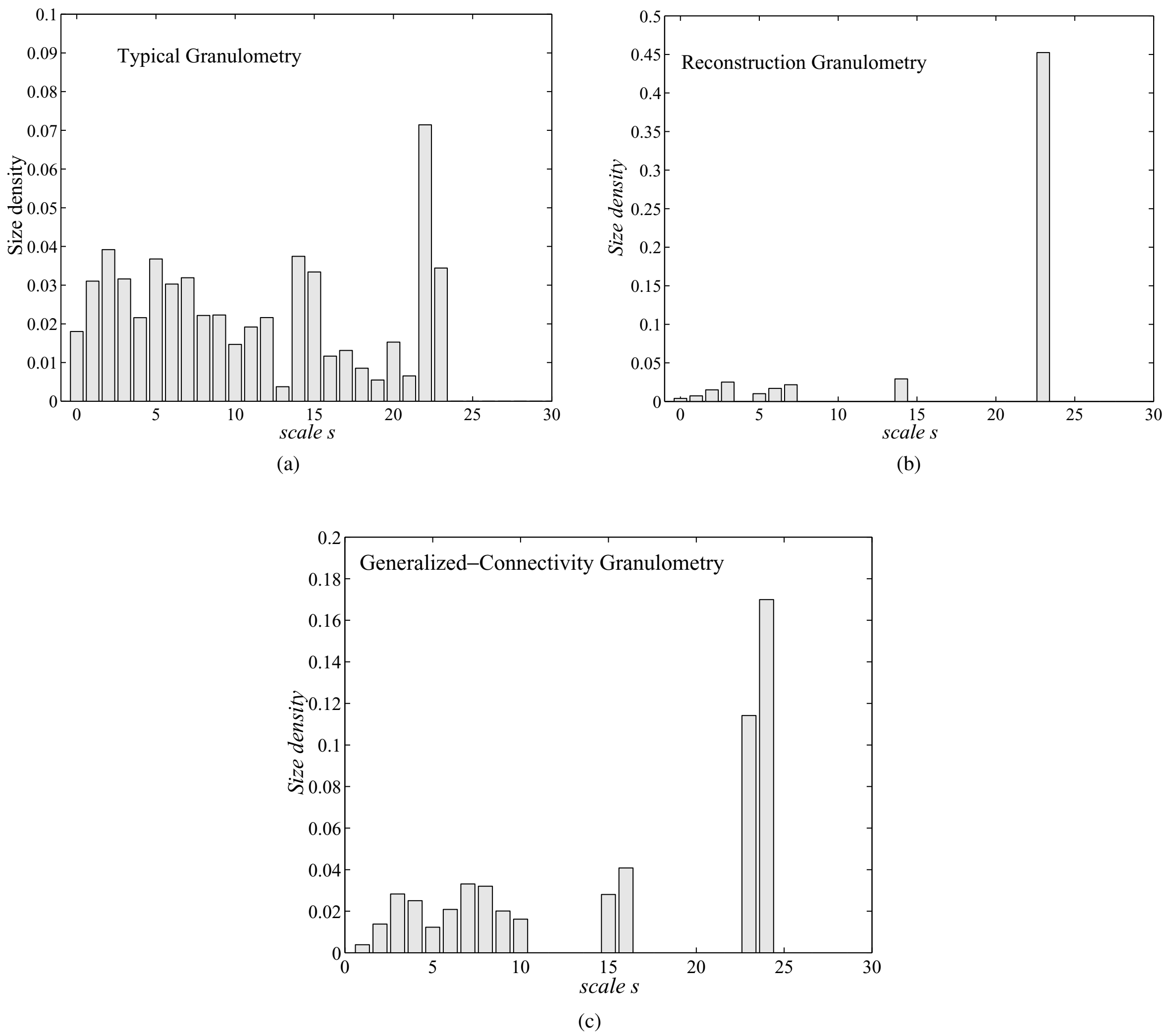



*Figure 8.* Size densities for soilsection-1. Comparative results using typical, reconstruction and generalized-connectivity granulometries. (a) Typical granulometry; (b) Reconstruction granulometry; and (c) Generalized granulometry using multiscale connectivity analysis.

as leakage problem, which leads to the creation of undesirable connections within an image, when thin connecting paths between large objects exist. This problem is particularly hindering for soilsection image analysis and evaluation tasks, where strongly connected soil formations need to be identified and differentiated from loosely connected regions (which in fact should be partitioned into a set of finer disjoint connected components). This is illustrated in Fig. 6, where we present a sample soilsection that will be used in the sequel as the first test image for the granulometric analysis experiments. Figure 6(b) shows the binary image resulting from the thresholding of the original grayscale soilsection-1 (at gray-level value = 32, with the gray-levels of the original image ranging from 0 to 255). However, what is particularly important to point out is illustrated in Fig. 6(c), which shows a single connected component resulting from the application of a simple

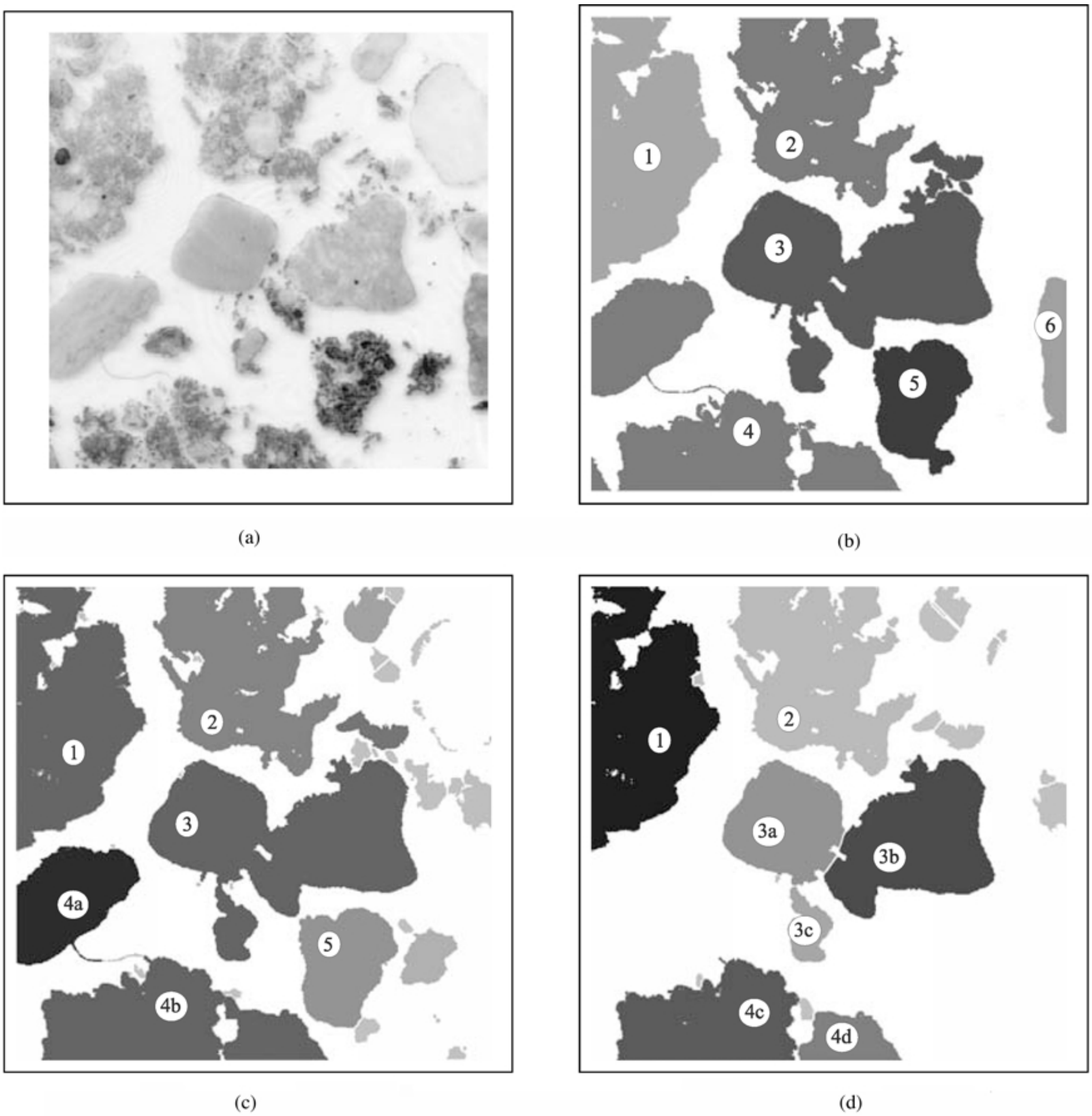


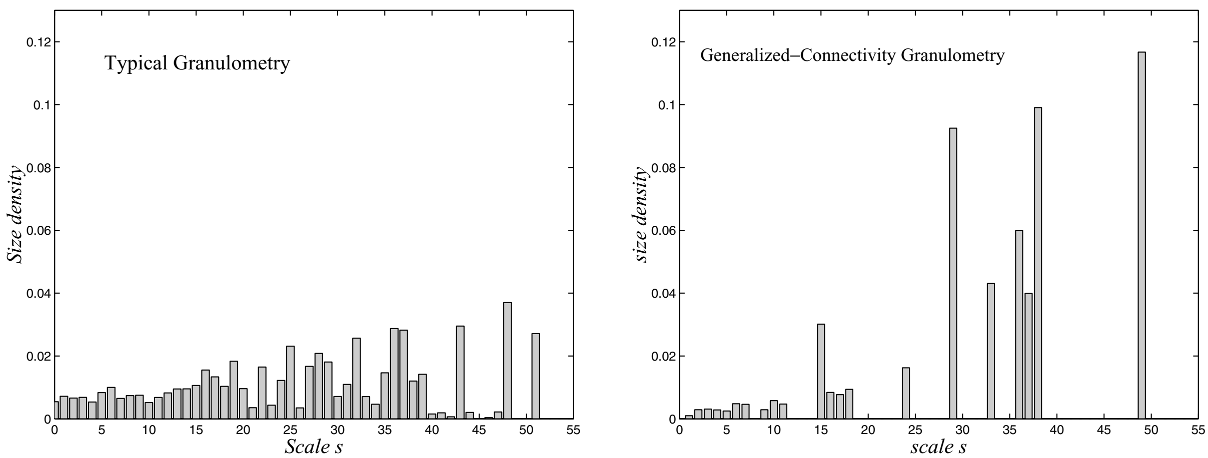

*Figure 9.* Sample soilsection-2: Multiscale connectivity partitioning. (a) Soilesection 2: Original grayscale image; (b) Thresholded image: Connected components; (c) Connectivity level = 2; and (d) Connectivity level = 4.

algorithm using 8-connectivity. This characteristic example demonstrates the problems that may result from the application of typical connected operators, treating as connected components object formations that should be intuitively considered as separate.

Let's apply now the multiscale connectivity analysis framework, introduced above, for the binary image component of soilsection-1 shown in Fig. 6(c). Figure 7 illustrates the recursive C-tree creation procedure, showing four different connectivity levels (c-level = 1 to 4) and the respective partitions of the original binary image. This figure demonstrates the successive partitioning of a typical "loosely-connected" component into progressively stronger connected particles. The information stored in the C-nodes during this C-tree creation procedure can then be used by a generalized-connectivity granulometry to evaluate size density measures, as described in the previous section. Such generalized size-densities can be "tuned" to provide more accurate and reliable connected-shape information, by selecting appropriate thresholding connectivity profiles or by choosing a different decision criterion in Eq. (19), depending on the application.

Figure 8 shows comparative granulometric analysis results for the soilsection-1 binary test-image, using typical, reconstruction and generalized-connectivity operators. The size densities resulting from the application of typical openings and reconstruction openings (Fig. 8(a) and (b) respectively) constitute in fact two extreme cases, with the first one containing a wide spectrum of sizes, while the latter contains a limited number of spikes corresponding to typical connected components in the image. In other words, we are presented with two situations: (a) the use of classical morphological openings, which results in size densities containing a large amount of information that may, however, not be representative of actual sizes/shapes included in the image (similar to performing an image over-segmentation), and (b) the application of typical connected operators (e.g. reconstruction openings), which create granulometries that actually filter-out the results of (a), but cut-off important size information (similarly to performing an excessive over-filtering on the results of (a)). As opposed to the above two situations, the generalized connectivity granulometries result in size-densities (see Fig. 8(c)) that constitute an intermediate solution, which can be tuned to preserve the advantages and reduce the drawbacks of each one of the above two extreme cases. The results shown in Fig. 8(c) are more representative of characteristic sizes contained in the original soilsection binary image of Fig. 6(b), potentially facilitating the subsequent image modeling and soil-structure evaluation steps.

Figures 9 and 10 show the results obtained for a different sample test image, soilsection-2. Figure 9(a) shows the original grayscale image, while Fig. 9(b) shows the major connected components (labeled 1–6) for the binary image that results from thresholding the original soilsection-2 image at gray-level 53 (gray-levels of soilsection-2 containing values ranging again




(a)

(b)

*Figure 10.* Granulometric analysis of soilsection-2. (a) Typical granulometry and (b) Generalized-connectivity granulometry.

from 0 to 255). Figure 9(c) and (d) illustrate the C-tree creation procedure at two different connectivity levels, $c\text{-}level = 2$ and $c\text{-}level = 4$ respectively. It is particularly important to note how the typical "loosely connected" region 4 (at the particular gray-level of Fig. 9(b)) is recursively partitioned into regions 4a, 4b (Fig. 9(c)), and 4c, 4d (Fig. 9(d)). Similarly, region 3 is partitioned into the "strong" connected components 3a, 3b and 3c (at $c\text{-}level = 4$, Fig. 9(d)). If one refers to the original soilsection image, this hierarchical multiscale decomposition into connected regions seems more appropriate, at least with respect to a granulometric image analysis. Figure 10(a) and (b) show the size densities obtained for soilsection-2, using typical openings and generalized-connectivity granulometries, respectively. One may notice once again how the latter better captures the size information contained in the original grayscale image.

The proposed multiscale connectivity analysis thus leads to image representations that contain more accurate and reliable size information, incorporating complex multiscale geometrical cues related to the connectivity structures and the presence of compound object formations (shape information) within an image. Moreover, the proposed theoretical framework can accommodate a variety of connectivity-related criteria, similar to the definition of "adaptable fuzzy connectivity measures", which means increased versatility and adaptability to a variety of image processing and computer vision applications.

## 5. Conclusion and Future Work

This paper has introduced a multiscale connectivity analysis framework based on an axiomatic definition of generalized morphological connectivity measures, such as an adjunctional multiscale connectivity function. This function in fact incorporates geometrical cues quantifying the notion of "how strongly connected" is a set, that is, "how easily" it can be partitioned into disjoint non-empty sub-sets through the recursive application of an antiextensive morphological operator (like opening or antiextensive erosion).

The concept of connectivity tree (C-tree) has then been presented, leading to a hierarchical representation of binary images. The information incorporated in the C-tree structure has been used to establish new generalized connected operators. Based on the definition of appropriate decision criteria (including the use of thresholding connectivity profiles), these generalized operators can be tuned to differentiate between "strong" and "loose" connections within an image, thus controlling the undesirable effects of the so-called "leakage" problem, related to the application of typical connected operators (such as reconstruction or area openings).

The generalized connectivity operators introduced in this paper have been used to define new generalized granulometries, aiming at a more accurate and reliable evaluation of morphological properties, such as characteristic size/shape distribution within an image. Comparative results obtained for a particular problem of soilsection image analysis demonstrate the power and versatility of the proposed multiscale connectivity analysis framework.

Potential future applications of such generalized connected operators, based on the proposed multiscale connectivity analysis and the C-tree concept, besides an improved granulometric image analysis described in this paper, also include: (a) image segmentation to improve the "connectivity properties" at the marker extraction phase [17]; (b) reliable shape/size representation for statistical image analysis and modeling (employing for instance GRFs [16], or MRFs [7]); (c) other applications of connected operators, like in motion detection/analysis and gesture recognition. In our on-going work, we plan to extend these operators for analyzing grayscale images and use them for applications related to image segmentation and statistical image modeling and classification.

## Note

1. In fact $\mu$ is just defined as a non-negative function and does not necessarily comply with the formal definitions in the measure-theoretic sense.

## References

1. U.M. Braga-Neto and J. Goutsias, "Multiresolution connectivity: An axiomatic approach," in *Mathematical Morphology and its Applications to Image and Signal Processing*, J. Goutsias, L. Vincent, and D.S. Bloomberg (Eds.), Kluwer Academic Publishers: Dordrecht, 2000. (*Proceedings of the ISMM'2000*).
2. U.M. Braga-Neto and J. Goutsias, "A complete lattice approach to connectivity in image analysis," Tech. Report JHU/ECE 00-05, Dept. ECE, Johns Hopkins University, Baltimore, MD, Nov. 2000.
3. H.J.A.M. Heijmans, *Morphological Image Operators*, Academic Press: San Diego, 1994.

4. H.J.A.M. Heijmans, "Connected morphological operators for binary images," *Computer Vision and Image Understanding*, Vol. 73, No. 1, pp. 99–120, 1999.
5. B.B. Kimia, A.R. Tannenbaum, and S.W. Zucker, "Shapes, shocks, and deformations I: The components of two-dimensional shape and the reaction-diffusion space," *Int. J. Computer Vision*, Vol. 15, pp. 189–224, 1995.
6. J.J. Koenderink and A.J. van Doorn, "Dynamic shape," *Biological Cybernetics*, Vol. 53, pp. 383–396, 1986.
7. B.S. Manjuna and R. Chellappa, "Unsupervised texture segmentation using Markov random field models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 5, pp. 478–482, 1991.
8. P. Maragos, "Pattern spectrum and multiscale shape representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 701–716, 1989.
9. P. Maragos, "Morphological signal and image processing," in *The Digital Signal Processing Handbook*, V.K. Madisetti and D.B Williams (Eds.), CRC Press/IEEE Press: Bocaraton, New York, Ch. 74, 1999, pp. 1–30.
10. G. Matheron, *Random Sets and Integral Geometry*, Wiley: New York, 1975.
11. C. Ronse, "Set-theoretical algebraic approaches to connectivity in continuous or digital spaces," *Journal of Mathematical Imaging and Vision*, Vol. 8, pp. 41–58, 1998.
12. P. Salembier, A. Oliveras, and L. Garrido, "Antiextensive connected operators for image and sequence processing," *IEEE Trans. Image Processing*, Vol. 7, No. 4, pp. 555–570, 1998.
13. P. Salembier and J. Serra, "Flat zones filtering, connected operators, and filters by reconstruction," *IEEE Trans. Image Processing*, Vol. 4, No. 8, pp. 1153–1160, 1995.
14. J. Serra, *Image Analysis and Mathematical Morphology: Theoretical Advances*, Academic Press: New York, 1988.
15. J. Serra, "Connections for sets and functions," *Fundamenta Informaticae*, Vol. 41, Nos. 1/2, pp. 147–186, 2000.
16. K. Sivakumar and J. Goutsias, "Morphologically constrained GRFs: Applications to texture synthesis and analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, No. 2, pp. 99–113, 1999.
17. A. Sofou, C. Tzafestas, and P. Maragos, "Segmentation of soil-section images using connected operators," in *Proc. IEEE Intern. Conf. on Image Processing (ICIP'2001)*, Thessaloniki, Greece, Sep. 2001.
18. L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. Image Processing*, Vol. 2, No. 2, pp. 176–201, 1993.
19. L. Vincent, "Granulometries and opening trees," *Fundamenta Informaticae*, Vol. 41, No. 1/2, pp. 57–90, 2000.


**Costas S. Tzafestas** holds an Electrical and Computer Engineering Degree from the National Technical University of Athens, as well as a D.E.A. and Ph.D. Degrees on Robotics from the Université Pierre et Marie Curie (Paris 6), France. He is currently a Lecturer on Robotics at the School of Electrical and Computer Engineering of the National Technical University of Athens. He has previously worked as a Research Associate at the Institute of Informatics and Telecommunications of the National Center for Scientific Research "Demokritos", Athens, Greece. His main research interests include virtual reality and human-machine interaction with applications in the field of telerobotics. He has also worked on robust, adaptive and neural control with applications in co-operating manipulators and walking robots. He is a member of the IEEE and of the Greek Technical Chamber.

**Petros Maragos** received his Ph.D. from Georgia Institute of Technology, Atlanta, USA, in 1985. From 1985 until 1993 he worked as professor of electrical engineering at the Division of Applied Sciences of Harvard University, Cambridge, Massachusetts. In 1993, he joined the faculty of the School of ECE at Georgia Tech. During parts of 1996–1998 he also worked as a senior researcher at the Institute for Language and Speech Processing in Athens. Since 1998, he has been working as professor of electrical & computer engineering at the National Technical University of Athens. His current research and teaching interests include the general areas of signal processing, systems theory, control, pattern recognition, and their applications to image processing and computer vision, and computer speech processing and recognition. His research work has received several NSF and IEEE Awards. In 1995, he was elected Fellow of IEEE.