

# Measuring Statistical Geometric Properties of Tomographic Images of Soils

Siovani C. Felipussi, Jacob Scharcanski, and João L. D. Comba

**Abstract**—Porous media modeling is relevant in several applications, such as agricultural engineering, where soil compaction analysis requires the estimation of soil transport properties. For example, the prediction of root growing patterns and their environmental impact is usually measured by analyzing soil fluid infiltration capacity and water retention. Recently, tomographic images have been used in nondestructive tests of soil. However, using such images is challenging for two reasons: 1) Tomographic images are usually noisy, which complicates their segmentation, and 2) modeling the soil structure requires establishing adjacency relations among neighboring tomographic slices, which has a significant computational cost due to the combinatorial nature of this problem. In this paper, we propose a solution for both problems. The experimental results show that soil samples can be analyzed and classified with significant accuracy using our proposed approach.

**Index Terms**—Adaptive systems, feature extraction, geometric modeling, image processing, materials testing, porous media, soil measurements, statistical geometry, statistics.

## I. INTRODUCTION

**I**MAGE analysis is an important tool for the structural characterization of porous materials, and its applications can be found in several areas, such as oil reservoir modeling and estimation of soil permeability or bone density [1]–[4]. The structural properties of porous media can be represented by statistical and morphological aspects, such as distributions of pore sizes and pore shapes, or by topological properties, such as pore network interconnectivity.

Several approaches have been presented in the literature for porous media structural analysis. In [5], image analyses of porous media sections are used to measure pore size distributions. Using a similar approach, native forest and crop samples were analyzed using images of soil sample sections to identify the porosity and aggregation of soil particles [6]. However, destructive methods usually affect the measurements of porous media structure, and nondestructive methods are preferable. Particularly in the case of porous media sections obtained by computerized tomography, volumetric measures can be relatively accurate since the porous media connectivity is not disrupted. In [7], a more sophisticated statistical method for

analyzing the 3-D porous structure from a unique 2-D image of Berea sandstone is proposed. The main idea is to generate a Gaussian field directly from the measured autocorrelation function (on thresholded 2-D images). Although it preserves the pore correlation information, this approach does not capture the porous geometrical or topological structure. In [8], the swelling and shrinkage properties of loamy soils are evaluated by localizing heterogeneous regions in tomographic images of soil samples. In addition, different methods for studying volume variation in soils based on their hydration levels were proposed in [9]. According to the aforementioned works, the morphological and topological aspects are relevant for the study of soils.

A number of approaches have been proposed to characterize the geometrical and statistical aspects of porous media. In [10], the medial axis was used to detect throats in pore connections and thus determine throat and pore size distributions. However, the medial axis and skeletonization procedures are noise sensitive, and irregularities in the void-grain phase can have the same effect as dead-end pores [11], producing spurious branches in the topological analysis. In [12], Dodson and Sampson proposed the modeling of pore size distributions in a range of stochastic porous media using a family of gamma distributions, but their approach requires porous media classification based on prior image analysis and modeling.

In this paper, a new method for statistical characterization of soils as porous media based on tomographic images is proposed. Three-dimensional gamma-ray tomographic images obtained from soil samples are analyzed, and their properties are represented based on statistical and topological features (i.e., pore and throat size distributions, as well as pore connectivity and flow tortuosity). The interconnectivity of the pore network is analyzed using a network flow algorithm, avoiding the significant computational cost involved in enumerating all possible topological connections. Based on selected geometrical and statistical features, soil sample classification is performed. Our experimental results emphasize the analysis of the soil compaction resulting from different soil-preparation techniques commonly used nowadays.

## II. ADAPTIVE TOMOGRAPHIC IMAGE SEGMENTATION

Initially, some morphological concepts will be presented, and afterward, our adaptive image segmentation approach, as well as our proposed feature extraction scheme, will be described.

### A. Pores, Grains, and Throats

In general, porous media are constituted by two types of elements: 1) grains and 2) pore bodies. This idea is illustrated

Manuscript received May 8, 2007; revised December 13, 2007. First published June 10, 2008; current version published October 10, 2008.

S. C. Felipussi is with the Department of Computer Science, Federal University of São Carlos, 18052-780 Sorocaba, Brazil (e-mail: siovani@power.ufscar.br).

J. Scharcanski and J. L. D. Comba are with the Institute of Informatics, Federal University of Rio Grande do Sul, 90040-060 Porto Alegre, Brazil (e-mail: jacobs@inf.ufrgs.br; comba@inf.ufrgs.br).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIM.2008.925010

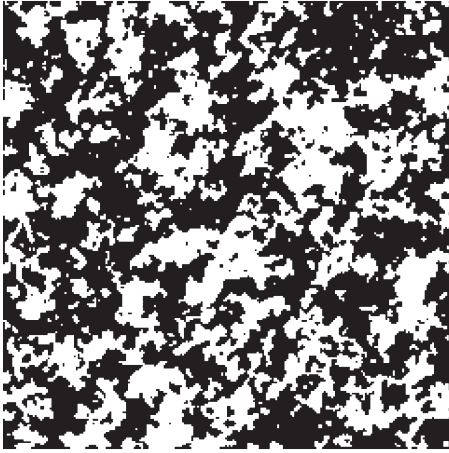


Fig. 1. Synthetic 2-D image section illustrating pores as clear regions and grains as dark regions.

for a generic porous medium in Fig. 1, using a synthetic image. Let  $S$  be a cross section of a porous medium given by a 2-D binary representation, such as the one shown in Fig. 1, where the pores (i.e., the void fraction) are represented by white regions, and the grains (i.e., the solid fraction) are represented by black regions. Phase function  $\Gamma$  is defined as [1]

$$\Gamma(\mathbf{x}) = \begin{cases} 1, & \text{when } \mathbf{x} \text{ belongs to the pore phase} \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where  $\mathbf{x}$  denotes the vector specifying a position in  $S$ . In fact,  $\mathbf{x}$  denotes the set of 2-tuples  $(x, y)$ , where  $x$  and  $y$  are integers and multiples of the measuring unit. In our case,  $S$  is represented by a 2-D image  $I$  of size  $N_h \times N_v$ , and  $(x, y)$  represents the pixel coordinates within this image. In gamma-ray tomography, each pixel represents the local attenuation of the porous medium matter to the gamma rays.

Within an image, we may identify sets of connected pixels (i.e., pixels belonging to the same region). A path from pixel  $U$  to pixel  $V$  is a set of pixels forming the sequence  $A_1, A_2, \dots, A_n$ , where  $A_1 = U$  and  $A_n = V$ , and  $A_{w+1}$  is a neighbor of  $A_w$ ,  $w = 2, \dots, n - 1$ . A region is defined as a set of pixels in which there is at least one path between each pixel pair of the region [13].

The images obtained by gamma-ray computerized tomography are slices (cross sections) of 3-D objects, and the intensity of a pixel represents the gamma-ray absorption in the spatial region represented by the pixel, which reflects the local material properties. This 3-D imaging technique has been recently applied to the analysis of porous media [14], [15]. In a slice  $z$ , the pore bodies (i.e., the pores) are disjoint regions  $P_i$ , where  $i = 1, \dots, N$ , and  $N$  is the number of pores in  $z$ , i.e.,

$$P_i(z) = \{I(x, y) | (x, y) \text{ connected}\}. \quad (2)$$

Let  $P_i(z)$  be a pore at slice  $z$  and  $P_j(z+1)$  be a pore in an adjacent slice  $z+1$ , with  $j = 1, \dots, M$  (where  $M$  is the number of pore bodies at slice  $z+1$ ). If there is at least one pixel in common between  $P_i(z)$  and  $P_j(z+1)$ , then there

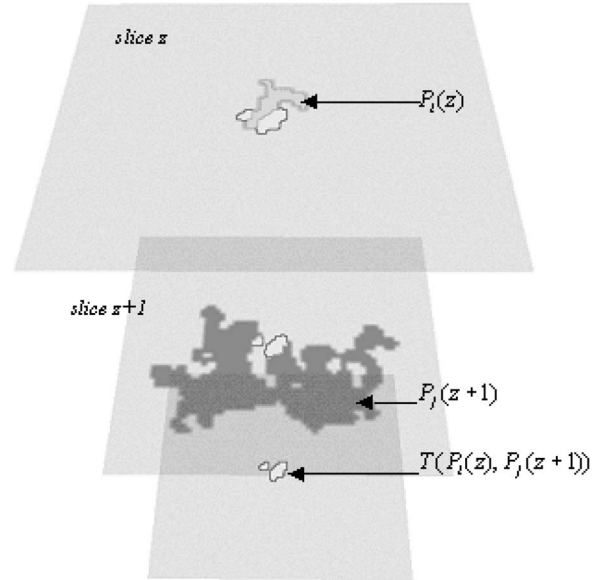


Fig. 2. Intersection between two regions of two slices and the resultant throat.

exists a throat connecting these pores, which is defined as follows:

$$T(P_i(z), P_j(z+1)) = \cap(P_i(z), P_j(z+1)). \quad (3)$$

Fig. 2 shows an illustration of two horizontal slices of a generic (synthetic) porous medium, where a pore  $P_i(z)$  in slice  $z$  and a pore  $P_j(z+1)$  in slice  $z+1$  are depicted as darker regions. The common regions are depicted as clearer areas and represent the throats for better visualization.

### B. Tortuosity

In hydrological surveys, tortuosity is one of the most informative parameters. However, it has different definitions in the literature. In [16], Prasher *et al.* introduced the concept of tortuosity  $\tau$  as the square of the ratio of the effective average path in the porous medium  $L_c$  to the shortest distance measured along the direction of the pore  $L$ . In this paper, as the analysis is based on the sequence of images of thin slices, the coefficient of tortuosity is evaluated based on three consecutive slices. Considering three interconnecting pores in slices  $n$ ,  $n+1$ , and  $n+2$ , the local tortuosity coefficient is estimated as the Euclidean distance between the centroids of their throats  $T(P_i(n), P_j(n+1))$  and  $T(P_j(n+1), P_k(n+2))$ , i.e.,

$$Tort = \sqrt{(\bar{x} - x)^2 + (\bar{y} - y)^2} \quad (4)$$

where the centroids  $(x, y)$  and  $(\bar{x}, \bar{y})$  correspond to the throats between pores at slices  $(n) - (n+1)$  and  $(n+1)(n+2)$ , respectively.

### C. Adaptive Tomographic Image Segmentation

Image segmentation is a very important step in porous media modeling, but it is one of the most difficult tasks in image analysis. Our goal is to obtain an image where the pore and solid phases of the original image are discriminated.

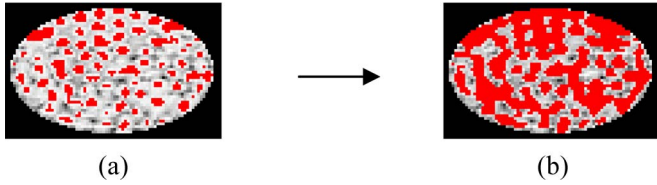


Fig. 3. (a) *Extended minima* with  $h = 1$  were applied to segment the brighter regions (in red). (b) Segmented image with  $h = 25$  (notice that some regions disappeared and that others grew by increasing  $h$ ).

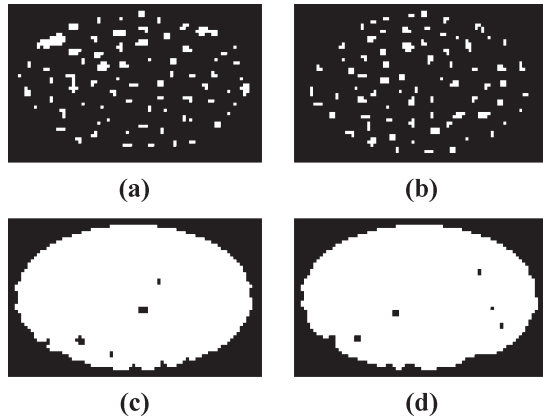


Fig. 4. (a) Segmentation of the first slice with  $k = 1$ . (b) Second slice of the stack also with  $k = 1$ . (c) and (d) First and second slices with  $k = 100$ . Note that correlation is high for large  $k$  and low otherwise.

In the literature, several authors have proposed the segmentation based on a global threshold as an early step in the porous media characterization [17]–[19]. Basically, the central question in these methods is the selection of a threshold to discriminate between pores and the solid phase based on the image histogram. This is typically interactively done by visual inspection of the results. Global thresholding generally results in serious segmentation flaws if the object of interest and the background are not uniform (as in our case). Next, we propose a locally adaptive image segmentation procedure to discriminate between pores and the solid phase, using *extended minima* or *maxima*. The concept of *extended minima/maxima* is presented next. The steps of our proposed segmentation method are illustrated in Figs. 3–5. This illustration is based on noisy and low-contrast nuclear magnetic resonance (NMR) slice images of glass spheres immersed in water, because, in this case, it is easier to visually verify the results since the regions of interest (i.e., spheres) are identifiable in the images. These NMR slice images have  $80 \times 50$  pixels, where each pixel corresponds to  $0.1 \times 0.1$  mm, the slice thickness is 0.25 mm, and the interslice spacing is 0 mm. This segmentation method will be used in tomographic images of soils based on gamma rays, which is the main goal of this paper, as will be discussed later.

1) *h-Maxima/Minima and Extended Maxima/Minima Transformations*: The  $h$ -maxima transform allows the reconstruction of  $I$  from  $I - h$ , where  $h$  represents an arbitrary gray-level constant. This way, the  $h$ -maxima transform provides a tool to filter the image using a contrast criterion. By definition, the  $h$ -maxima transform suppresses all maxima whose height is smaller than a given threshold level  $h$ . Therefore, the  $h$ -maxima detect peaks using a contrast parameter,

without involving any size or shape criterion. Let  $\rho_I$  denote the reconstruction function (see [20] for details), i.e.,

$$HMAX_h(I) = \rho_I(I - h). \quad (5)$$

Analogously, the  $h$ -minima transform reconstructs  $I$  from  $I + h$ , where  $h$  represents an arbitrary gray-level constant. By definition, the  $h$ -minima transform suppresses all minima whose depth is less than a given threshold level  $h$  [20], i.e.,

$$HMIN_h(I) = \rho_I(I + h). \quad (6)$$

The regional maxima  $RMAX(I)$  [regional minima  $RMIN(I)$ ] are peaks (or valleys) of  $HMAX_h(I)$  [ $HMIN_h(I)$ ] regions, producing an image with the relevant regions marked (i.e., a marker image). *Extended maxima (extended minima)* are obtained using  $I - \rho_I(I - h)$  in the computation of  $HMAX_h(I)$  [ $HMIN_h(I)$ ]. The *extended maxima EMAX (extended minima EMIN)* are the regional maxima (minima) of the corresponding  $h$ -extrema transformation [21], i.e.,

$$EMAX_h(I) = RMAX(HMAX_h(I)) \quad (7)$$

$$EMIN_h(I) = RMIN(HMIN_h(I)). \quad (8)$$

In our adaptive image segmentation approach, the choice of parameter  $h$  is a central issue, because the  $h$ -minima/maxima transform suppresses all minima/maxima regions whose depth/height is less than  $h$  (i.e., considering a gray-level image relief). Comparing the *extended minima* in Fig. 3(a) and (b), we can verify that, when  $h$  is increased, some objects were increased in size, and some disappeared.

In our case, the images are noisy and present low contrast, making the choice of parameter  $h$  critical. Therefore, let us consider the union of the pixel sets of the regions obtained by the *extended minima/maxima* when  $h$  varies between 1 and  $k$ ,  $k \in [1, \dots, 255]$ . Formally, we have

$$T_{\text{MIN}}^k(I) = \bigcup_{h=1}^k EMIN_h(I) \quad (9)$$

$$T_{\text{MAX}}^k = \bigcup_{h=1}^k EMAX_h(I). \quad (10)$$

If parameter  $k$  is small, we obtain small regions centered on the regional minima/maxima of the image. If  $k$  increases, the regions grow and can be merged. Important issues in the soil tomography analysis result from the fact that such images tend to be noisy and present low contrast, affecting the spatial continuity in three dimensions (since these structures are 3-D). We evaluate the spatial continuity between each pair of slices in the image stack, using correlation analysis, given an appropriate  $k$  value. If  $k$  is small, the correlation is low, because the pixels of the adjacent images do not present spatial continuity [Fig. 4(a) and (b)], and if  $k$  is large, the correlation is higher

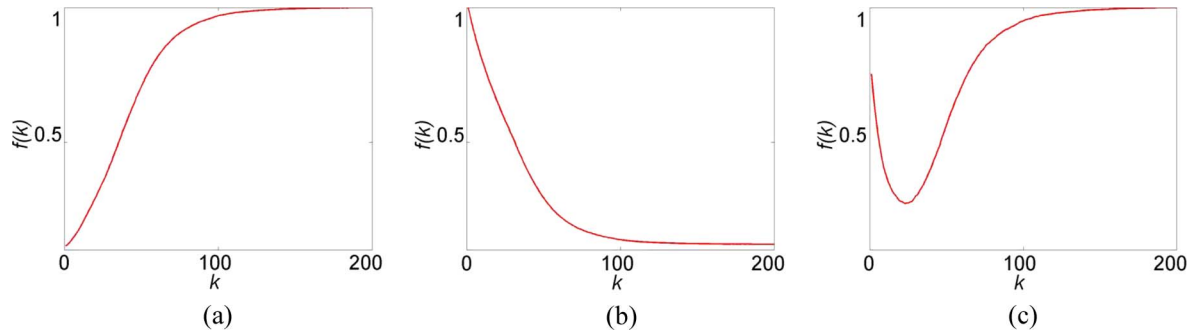


Fig. 5. (a) Function  $f(k)$  with  $a = 0.9$  and  $b = 0.1$ . (b) The complement of areas is prioritized with  $a = 0.1$  and  $b = 0.9$ ; in both cases,  $r_n(\theta)$  is small, but operator  $D$  is not adequate for segmentation. (c) Plot of  $f(k)$  with  $a = b = 1$ .

[Fig. 4(c) and (d)]. To select the *best* value of  $k$  for void segmentation, we use the following function:

$$f(k) = a(1 - \text{area}) + b(\text{correlation}) \quad \forall a, b \in [0, 1]. \quad (11)$$

This function provides a relationship  $(1 - \text{area}) \times (\text{correlation})$  and is shown in Fig. 5(c). The “correlation” parameter is estimated as the ratio between the number of pixels representing voids that intersect in adjacent images (for all images) and the number of image pixels. The “area” parameter is the ratio between the number of pixels representing voids in the segmented regions and the number of image pixels (corresponds to the void fraction). It is necessary for an optimization algorithm to find the proper values for parameters  $k$ ,  $a$ , and  $b$ , given the tomographic image data [13]. We propose to estimate parameters  $k$ ,  $a$ , and  $b$  using the minimax approach [19]. The minimax approach requires that we estimate the values of  $k$ ,  $a$ , and  $b$  as parameter values leading to the smallest risk/error in void segmentation, considering all the operators  $D(k)$  that can be obtained (notice that  $k$  defines our operator).

Let  $\theta$  be the set of all slice image segmentations. We estimate the *true void segmentation*  $f \in \theta$  based on the original noisy images using an operator  $D$ . The segmentation error of this estimation  $\bar{F} = DX$  is given by

$$r(D, f) = E \{ \|DX - f\|^2 \}. \quad (12)$$

Unfortunately, the expected error cannot precisely be computed, because the probability distribution of all slice image segmentations in  $\theta$  is not known *a priori*. Therefore, the minimax approach proposes to minimize the maximum error

$$r(D, \theta) = \sup_{f \in \theta} E \{ \|DX - f\|^2 \} \quad (13)$$

and the minimax error is the lower bound computed over all operators  $D$ , i.e.,

$$r_n(\theta) = \inf_{D \in O_n} r(D, \theta) \quad (14)$$

where  $O_n$  is the set of all operators.

In practice, we must find a decision operator  $D$  that is simple to implement (small  $k$ ), such that  $r(D, \theta)$  is close to minimax

error  $r_n(\theta)$ , and we have

$$r_n(\theta) = \inf_{D \in O_n} \left\{ \sup_{f \in \theta} E \{ \|DX - f\|^2 \} \right\}. \quad (15)$$

In our case, we must approximate the error function based on problem constraints, i.e., we wish to have slice continuity along the slice stack, and at the same time, we want to segment the largest (maximum) number of voids. These are obviously conflicting constraints because of the very nature of such tomographic images (i.e., they have low contrast and are noisy, and at the same time, the porous media structure may be quite random). Therefore, we approximate the error function by the measurable data described by (11), with the additional constraint that  $|a - b| \leq \delta$ , with  $\delta \in [0, 1]$ . When this constraint is not considered, function  $f$  prioritizes just one of the constraints and returns an inadequate result (Fig. 5).

An overview of the segmentation algorithm is outlined here.

- 1) Compute the image  $\rho_I(\bigcup_{h=1}^k EMIN_h(I))$  by performing the grayscale reconstruction with the union of the sets of the *extended* minima to all  $k \in [1, \dots, 255]$ .
- 2) Compute each  $k$ , i.e., the correlation between each pair of slices of the stack of images.
- 3) Compute  $T_{\text{MIN}}^k(I)$  for all slices of the image stack.
- 4) Compute the function of (11) for all  $k$  values in the interval.
- 5) Through the minimax principle, find a decision operator  $D(k)$  leading to the smallest maximum error (i.e., deepest valley in the plot).

The image  $\rho_I(\bigcup_{h=1}^k EMIN_h(I))$  is our void segmentation estimate. This segmentation algorithm was applied to 465 images of soil samples (some examples are shown in Fig. 6) and to 58 images of glass spheres immersed in water. As mentioned before, the images of glass spheres were obtained by NMR tomography and were only used for illustration purposes. However, this imaging technique was not used for the soil samples since soils usually contain magnetic materials such as iron oxide, which interfere with the homogeneity of the NMR magnetic field. The soil sample images were acquired with a portable gamma-ray tomography equipment, using a source of Americium-241, 60 keV, and activity of  $1.11 \times 10^{10}$  Bq. Each soil sample is represented by 31 slice images, with an interslice spacing of 5 mm, and the thickness of each slice is 0.25 mm.



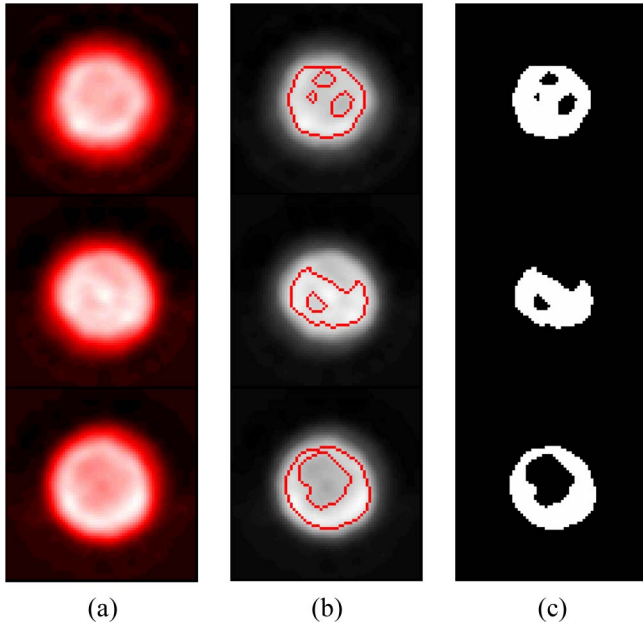


Fig. 6. (a) Original soil sample. (b) Boundaries of the segmented regions (red). (c) Segmented image.

Each slice image has  $71 \times 71$  pixels, and each pixel corresponds to  $1 \times 1$  mm.

#### D. Geometric and Topological Features

The first step is to label and store the geometric location of pores in each segmented image. In the next step, individual features of pores and throats are extracted. Each throat also stores the identification of its two associated pores (i.e., the IDs of the source and target pores). The extracted features for pores and throats are area, perimeter, diameter, centroid, major and minor axes, orientation, Euler number, and eccentricity.

These features were selected because of the important role of porous media characterization. The areas of pores and throats are calculated by simply counting their pixels(16)

$$A_k = \sum_{i=1}^n \sum_{j=1}^m P_k(i, j), \quad \text{where } k \text{ is a region with } i, j \in N. \quad (16)$$

The sum of pore areas  $A_k$  divided by the total area covered by the image  $A_{\text{tot}}$  gives the total porosity of the medium. The angle between the  $x$ -axis and the major axis of the pore gives the orientation, which is an important feature in the percolation process. The eccentricity gives an estimate of the pore circularity and is calculated as the ratio between the foci of the ellipse and the major axis length, with the resulting value in the interval  $[0,1]$  (i.e., a pore whose eccentricity is 0 is circular and a pore whose eccentricity is 1 is a line segment). Elongated pores, such as cracks, affect the physical properties of the medium in a different way than round pores [5]. The Euler number is a scalar that represents the number of objects in the region minus the number of holes in those objects and is relevant to the prediction of the (pore) network connectivity [22]. A network is defined as a set of pores that are interconnected such that there is a passage from any part

of the set to every other part. The importance of the network connectivity is related to the “fingering” phenomenon and the trajectory of the fluid or gas during the percolation process. The “fingering” phenomenon is described by means of the displacement of a gas in a porous medium by a high-viscosity fluid [23]. The “fingering” phenomenon occurs in fluid flows in soils, and in this case, this phenomenon is dominated by the force of gravity [24].

We represent the percolation and the “fingering” phenomenon by a multistage graph. In [25], a multistage graph is defined as a directed graph  $G = (V, E)$  in which vertices are partitioned into  $k \geq 2$  disjoint sets  $V_i$ ,  $1 \leq i \leq k$ . In addition, if  $\langle u, v \rangle$  is an edge in  $E$ , then  $u \in V_i$  and  $v \in V_{i+1}$  for some  $i$ ,  $1 \leq i \leq k$ . Let  $s$  and  $t$  be the vertices in  $V_1$  and  $V_k$ ;  $s$  is the source, and  $t$  is the sink. Each set  $V_i$  defines a stage in the graph. Because of the constraints on  $E$ , every path from  $s$  to  $t$  starts in stage 1 and goes to stage 2 and, in succession, goes until stage  $k$ .

The multistage graph is defined as follows: Each slice is represented as one stage, and each vertex is a pore at that graph level. The edges connect vertices in adjacent levels, i.e., connect vertices (pores) that share a throat.

We build this graph by incrementally processing the vertices (pores) of each stage and by defining a connection (edges) to a pore in the next stage if a throat is defined by the two pores. Since we are only concerned with paths that go from the first to the last level (also called valid paths), it was important to devise an efficient solution that discards invalid paths. Fig. 7 illustrates a path that starts in the first (top) slice and develops into other paths as we move through the next eight slices toward the last slice. In each level, the circles have different colors; the circle area is proportional to the area of the respective pore.

It is easy to establish that the number of paths that can occur can be exponential. Let  $P$  be the number of pores of the slice with the largest amount of pores,  $S$  be the number of pores of the slice with the smallest amount of pores, and  $n$  be the number of slices. Then, the maximum number of paths that could occur is  $O(P^n)$ , and the smallest number of paths that could occur is  $O(S^n)$ . Obviously, in moderate-to-large graphs, the computational cost makes it impracticable to enumerate all possible paths that emanate for every pore.

#### E. Statistical Analysis

The porous medium statistics are represented using contingency tables [26]. Each dimension of the contingency table captures one aspect of the porous medium, i.e., its geometrical and topological attributes, such as pore or throat areas, tortuosity, and depth of the media. In this paper, we focus on the following relations between the attributes: pore areas *versus* throat areas; pore areas *versus* depth; throat areas *versus* tortuosity; throat areas *versus* depth; and tortuosity *versus* depth.

Among the preceding attributes, tortuosity is the only one that involves computation in more than two slices. The connectivity information is obtained by traversing the multistage graph previously described.

In this paper, tortuosity is calculated based on three adjacent graph levels, forming a subpath with three vertices. However,

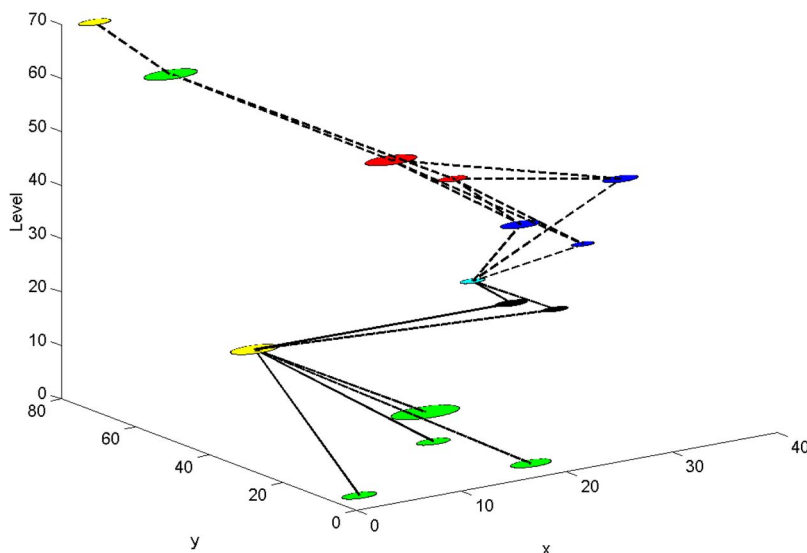


Fig. 7. Path developing in eight slices (synthetic data).

as mentioned before, searching and enumerating all possible paths are exponential, so we use a network flow algorithm with a dynamic programming approach to count the number of possible paths. The different paths in the graph can be separated in the following groups.

- 1) FL paths correspond to the paths starting at a node in the first level and ending in a node located in the last level (or, symmetrically, from the last to the first level).
- 2) F paths correspond to paths that start at a node in the first level but do not reach a node in the last level.
- 3) L paths correspond to paths starting at a node in the last level but do not reach a node in the first level.
- 4) Air paths are intermediate paths that do not reach either the first or last levels.

Since we are only interested in counting FL paths, we need to find a way to discard the other path types. We proceed as follows: Starting at the last level of the multistage graph, we proceed upward one stage at a time, counting for each vertex of the graph the number of its children nodes. Similarly, we start at the first level and proceed downward to compute the number of its parent nodes. Vertices at internal stages without parent or children are not part of the FL path (invalid vertices) and are thus discarded. Once this information is computed and invalid vertices are deleted, counting the number of valid paths that pass through a vertex requires just multiplying its number of parent and children vertices.

To compute tortuosity in subpaths, we sweep the graph from the first to the last level. At this point, the number of children and parent nodes is computed and stored for each node in the graph, as previously described. A node of the first level is chosen, and a depth-first transversal of only three levels is performed. The current subpath is accepted if the current node has children at least two levels below. Therefore, each accepted subpath has three nodes in adjacent levels of the graph (i.e., slices). The subpath tortuosity is calculated, as shown in (4), based on the distances between the centroids of nodes located in adjacent levels. However, we should keep in mind that we are actually interested in the tortuosity of a path that starts in

the first level of the graph, and the current subpath is part of a larger path starting in a node in the first level of the graph. The subpath calculated contributes to the overall path tortuosity (seen from the node in the first level). This subpath contributes once for each branch developing from the subpath and is part of a path that has the starting node in the first level in common (i.e., each subpath associated with a child node shares the same path and the subpath for which the tortuosity was calculated). Therefore, at this point, the subpath tortuosity contribution to a path starting in the first level is multiplied by the number of child nodes. The algorithm is recursively applied to all vertices in the graph until the last graph level. The complexity of the depth-first traversal with a depth of 3 applied to several stages is  $O((n - 2)P^3)$ , where  $P$  is the number of pores of the slice with the largest number of pores, and  $n$  is the number of slices. As an illustration of the computational performance we obtain, we tested our algorithm in a computer equipped with a 3.4-GHz Dual-Core Pentium with 2 GB of random access memory. The experiment we used was the computation of tortuosity for the simulated data, which produced  $6.694 \times 10^{12}$  different paths, and was tested using a Matlab implementation. While the brute-force algorithm took over a day to evaluate the tortuosity, the proposed dynamic programming method required 47 s to complete. In a scenario where a field specialist is performing soil analysis, we believe that such performance is acceptable. The complete algorithm is outlined here.

- 1) Starting at the last level, compute the number of child nodes for each node until the first level of the graph is reached.
- 2) Starting at the first level of the graph, compute the number of parent nodes of each node until the last level of the graph is reached.
- 3) Tortuosity calculation: Perform a depth-first transversal for each node of the current level up to two levels below, and find subpaths: one for each leaf node.
- 4) Compute the tortuosity using the three nodes of the subpath.

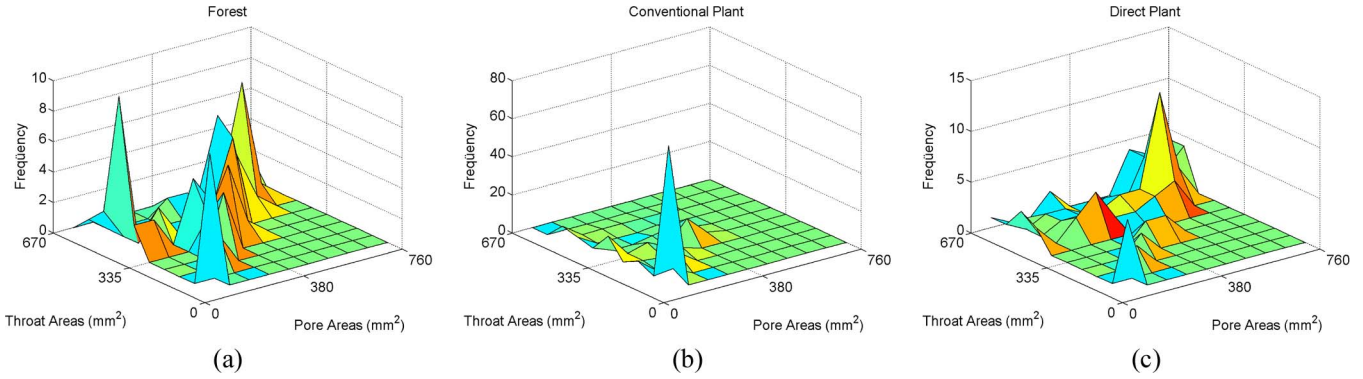


Fig. 8. Pore areas versus throat areas contingency table in ten intervals. (a) Forest. (b) Conventional planting. (c) Direct planting.

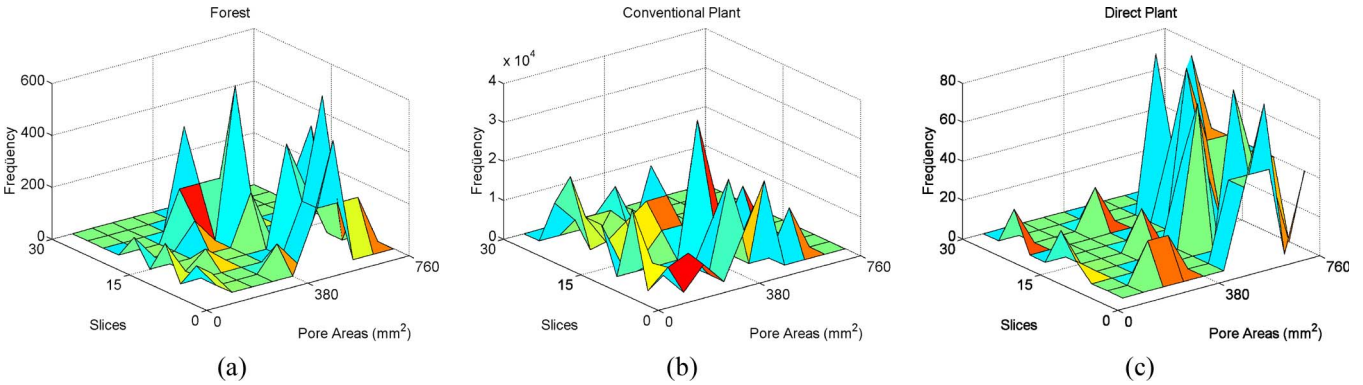


Fig. 9. Pore areas versus depth contingency table in ten intervals. (a) Forest. (b) Conventional planting. (c) Direct planting.

- 5) In the first and the third nodes of the subpath, multiply the tortuosity by the number of parent nodes and child nodes, respectively, to calculate the number of occurrences of the calculated tortuosity value in porous media, and then, update the contingency tables where tortuosity is used.
- 6) If there are pores in the current level that has yet to be visited, go to step 1.
- 7) If the current level has more than two levels below before the last level of the graph and all pores in the current level have been visited, make the current level be the level below, and go to step 1.

F. Porous Medium Sample Classification

Given the contingency tables (i.e., statistics) of the porous medium analysis, we are interested in assigning samples to particular porous media classes [27]. In this paper, the maximum likelihood approach is used, as described by Gonzalez and Woods [28]. This method is supervised and tends to improve its classification rate when the number of training samples increases. In our case, 465 samples were used for training, and 90 samples were used for the classification test.

III. EXPERIMENTAL RESULT

In our experiments, tomographic images of soils were employed. Samples are composed of purple dystrophic latosol, which is classified as loamy soil composed of 45% of clay, 15% silt, and 42% sand, submitted to three types of treatment: soil of natural forest (Phantom1), conventional treatment before

planting (Phantom2), and direct planting (Phantom3). All were obtained at horizon A, from the first layer, near the surface at a depth in the range of 0–20 cm. The images were obtained with a computerized tomograph of second generation in the laboratories of Cnpdia-Embrapa, São Carlos, Brazil, and each sample generated 31 slices with an interslice distance of 5 mm. The experimental results indicate that the type of soil treatment has an influence on the size distributions of pores and throats. In addition, considering that pores at adjacent tomographic slices tend to be interconnected by throats, the areas of pores and their associated throats are correlated (i.e., larger pores tend to be associated with larger throats, and *vice versa*). Fig. 8 further clarifies the existing correlation between the sizes of pores and throats, and a relative similarity between forest and direct planting can be observed.

Fig. 9 indicates that conventional soil treatment produces a more heterogeneous porous medium, which may possibly affect the soil permeability because of the higher concentration of small pores in deeper layers. In addition, the distribution of pores in forest and direct planting samples is more uniform, which implies better connectivity and less throats with small width (i.e., narrow). We also calculated the joint distribution of pore areas versus the depth of analysis and obtained experimental evidence that smaller pores tend to occur near the surface in the case of conventional soil treatment, which is an indication of soil compaction. Again, Fig. 10 shows that conventional soil treatment produces less homogeneous throat distributions; however, this does not confirm a decrease in soil permeability. Therefore, we also calculated the joint distribution of throat areas versus the depth of analysis and obtained experimental



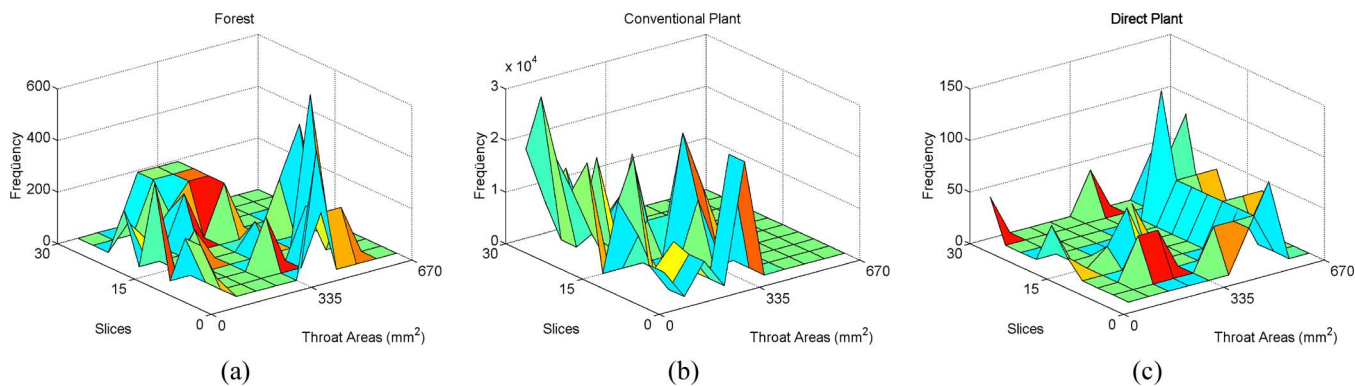


Fig. 10. Throat areas versus depth of analysis contingency table in ten intervals. (a) Forest. (b) Conventional planting. (c) Direct planting.

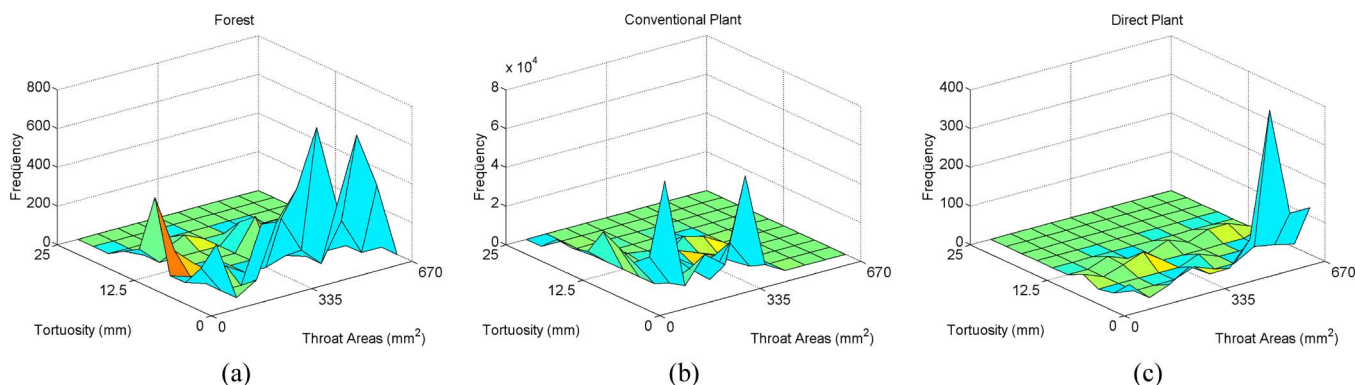


Fig. 11. Tortuosity versus throat areas contingency table in ten intervals. (a) Forest. (b) Conventional planting. (c) Direct planting.

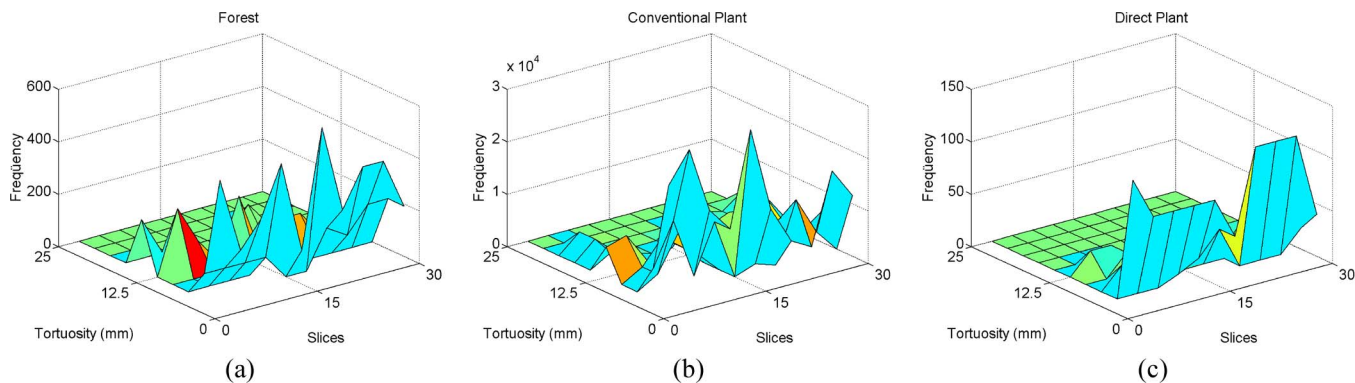


Fig. 12. Tortuosity versus depth of analysis contingency table in ten intervals. (a) Forest. (b) Conventional planting. (c) Direct planting.

evidence that smaller throats tend to occur near the surface, indicating that soil compaction exists in this case, which has a negative impact on soil permeability.

Fig. 11 shows that a considerable number of throats with large tortuosity coefficient values occur in the case of conventional planting, making it more difficult for solutes to flow in these soils, which can be caused by soil compaction. Now considering the depth of analysis, Fig. 12 shows that tortuosity tends to follow a similar trend for any depth of analysis in the case of forest and direct planting; however, in the case of conventional soil treatment, tortuosity tends to be more homogeneous near the surface.

The visual analysis using *Volview* [29] confirms these conclusions obtained from the statistical analysis of the data. Fig. 13 shows 3-D sample reconstructions for each soil treatment

considered in this paper. To better visualize the information contained in each slice, we used a feature of this software that allows the specification of the thickness of each slice, and increased resolution is obtained by interpolation between slices. In addition, to improve visualization, we used the negative, where the pores are the dark regions, with  $2.5 \text{ mm}^2/\text{pixel}$  (square pixels). The pore connectivity can be observed in Fig. 13(a) and (c), which allows fluids to flow in the columns; However, Fig. 13(b) shows that less connectivity is obtained in the case of conventional planting. The forest soil, which is untreated, illustrates a pore connectivity that allows the natural infiltration of water in the soil, whereas the soil after direct planting has a reduced connectivity compared to the forest soil because of the compactness produced by this method (i.e., planting is made by light machinery, animals, and people; heavy



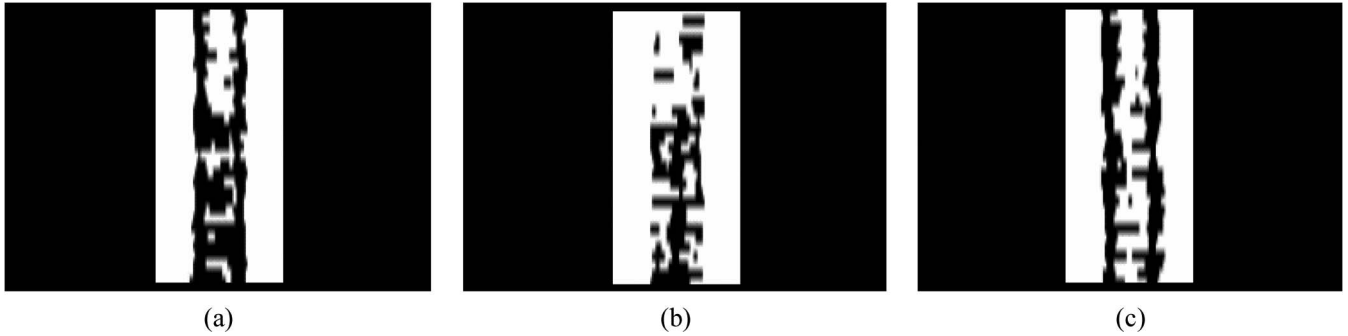


Fig. 13. Transversal view of a 3-D reconstruction of soil samples (with 31 slices per sample, an interslice spacing of 5 mm, a slice thickness of 0.25 mm, each slice having  $71 \times 71$  pixels, and each pixel corresponding to  $1 \times 1$  mm). (a) Forest. (b) Conventional planting. (c) Direct planting.

TABLE I  
CLASSIFICATION RATES USING INDIVIDUAL FEATURES

8 x 8	Pore Area X Throat Area			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	37%	87%	47%	57%
In Sections	77%	100%	77%	85%

8 x 8	Pore Area X Depth			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	47%	30%	27%	35%
In Sections	73%	87%	73%	78%

8 x 8	Throat Area X Depth			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	20%	93%	37%	50%
In Sections	73%	100%	97%	90%

8 x 8	Throat Area X Tortuosity			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	3%	100%	3%	35%
In Sections	73%	100%	73%	82%

8 x 8	Depth X Tortuosity			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	17%	83%	3%	34%
In Sections	77%	93%	73%	81%

TABLE II  
CLASSIFICATION RATES BASED ON THE FEATURE COMPOSITIONS

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tortuosity)			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	27%	60%	13%	33%
In Sections	83%	90%	80%	84%

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tortuosity) X (Pore Area x Depth)			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	37%	97%	17%	50%
In Sections	97%	100%	83%	93%

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tortuosity) X (Throat Area x Depth)			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	37%	97%	10%	48%
In Sections	93%	100%	80%	91%

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tortuosity) X (Pore Area x Depth) X (Throat Area x Depth)			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	37%	97%	10%	48%
In Sections	93%	100%	77%	90%

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tort.) X (Pore Area x Depth) X (Throat Area x Depth) X (Depth x Tort.)			
	Forest	Conv. Plant	Direct Plant.	Average
All Slices	10%	93%	3%	35%
In Sections	77%	100%	73%	83%

machinery is not used as in conventional planting). In the case of conventional planting, our experiments indicate that a higher compactness degree is expected (probably resulting from the use of agricultural machines and tractors in breeziness and cultivation), consequently increasing the mechanical resistance to root growing. In addition, soil pore connectivity tends to decrease, decreasing water infiltration into the soil.

Our image analysis experimental results confirmed the field analysis conclusions obtained by soil specialists, i.e., that the direct planting method has less of an impact on the natural soil porosity and leads to less soil compactness.

#### IV. CLASSIFICATION OF SOIL SAMPLES

In our sample classification experiments with a maximum likelihood classifier, we tried to use individual slices, but we

found that slices of conventional planting samples are too heterogeneous, and a small correct classification rate was obtained (see Table I). To improve the robustness of the classification process, we considered groups of three slices as one sample, reducing the effect of individual slice heterogeneity/homogeneity. The notation *all slices* and *in sections* in Table I denote classification using all the slices available and groups of three slices, respectively, and  $8 \times 8$  represents the contingency table dimensions (other dimensions were tested but produced sparse matrices and small classification success rates).

Different classification rates were obtained using different feature combinations, and the best classification results for all soil types were based on feature combinations, considering groups of slices as one sample (see Table II). Compositions of four features resulted in the best classification rates but needed more samples for training (as expected, since compositions of

TABLE III  
CLASS CONFUSION MATRIX FOR THE BEST CLASSIFICATION RESULT

8 x 8	(Pore Area x Throat Area) X (Throat Area x Tortuosity) X (Pore Area x Depth)		
<i>In Sections (93% Correct Classification Rate)</i>			
	Forest	Conv. Plant	Direct Plant
Forest	97%	0%	3%
Conv. Plant	0%	100%	0%
Direct Plant	0%	17%	83%

more features produce larger contingency tables). The correct classification rate decreased when five (or more) feature combinations were used, probably because the sparseness of the data in the contingency tables tends to increase when more feature combinations are used.

The best correct classification rate we obtained was 93%. Table III indicates that, even in this case, some classes were confused in the sample classification. For example, among all samples assigned to the forest class, 3% was direct planting samples that were incorrectly classified as forest samples. The direct and conventional planting classes were also confused, and 17% of the samples assigned to the direct planting class were, in fact, misclassified conventional planting samples. All samples assigned to the conventional planting class were correctly classified.

## V. CONCLUDING REMARK

This paper proposed a method for representing porous media using morphological and topological features, which were estimated using relatively low complexity algorithms. The proposed porous medium representation is based on a directional graph and an efficient algorithm for graph traversal. The proposed feature set can be used for porous media sample classification. In the case of compacted soils, our proposed image analysis method could potentially help in deciding which soil recuperation technique should be used. Our preliminary results based on soil samples were promising and were confirmed by the field analyzes of soil specialists.

As future work, we intend to investigate which feature combinations provide the best classification results for different types of soils. The applications of this paper in other porous media will be a theme of future investigation, as will be the adaptation of our approach to large-scale problems. We would also like to incorporate into the system our own set of 3-D visualization techniques. Although, to some extent, the software we used (*Volview* [29]) allows other direct imaging possibilities, having our own volume-rendering algorithm allows more customizable imaging from the input soil images and might allow better insights into the data. We are looking into other direct 3-D imaging possibilities and implementation using next-generation graphics hardware.

## ACKNOWLEDGMENT

The authors would like to thank Dr. S. Crestana, Dr. J. de Mendonça Naime, and Dr. C. M. P. Vaz (Cnpdia-Embrapa, São Carlos, Brazil) for providing the data used in this paper and Dr. M. Biasussi for the helpful discussions.

## REFERENCES

- [1] C. P. Fernandes and F. S. Magnani, "Multiscale geometrical reconstruction of porous structures," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 54, no. 2, pp. 1734–1741, Aug. 1996.
- [2] H. J. Vogel and A. Kretzchmar, "Topological characterization of pore space in soil—Sample preparation and digital image-processing," *Geoderma*, vol. 73, no. 1/2, pp. 23–38, Sep. 1996.
- [3] H. Q. Zhao, I. F. Macdonald, and M. J. Kwiecen, "Multi-orientation scanning: A necessity in the identification of pore necks in porous media by 3-D computer reconstruction from serial section data," *J. Colloid Interface Sci.*, vol. 162, no. 2, pp. 390–401, 1994.
- [4] L. Zhi-Qiang, L. Hui, and S. Dance, "Image processing techniques for quantitative bone image analysis," in *Proc. Int. Symp. Signal Process. Appl.*, Gold Coast, Australia, Aug. 25–30, 1996, pp. 431–432.
- [5] F. S. Anselmetti and S. Luthi, "Quantitative characterization of carbonate pore systems by digital image analysis," *AAPG Bull.*, vol. 82, no. 10, pp. 1815–1836, Oct. 1998.
- [6] L. A. C. Jorge and M. F. Guimarães, "Estudo de lâminas delgadas de solo: Porosidade, agregados e fissuras, com o auxílio de processamento de imagens digitais," in *Proc. Nat. Symp. Agropecuary Instrum.*, São Carlos, Brazil, Nov. 1997, pp. 284–289.
- [7] Z. R. Liang and C. P. Fernandes, "A reconstruction technique for three-dimensional porous media using image analysis and fourier transforms," *J. Pet. Sci. Eng.*, vol. 21, no. 3/4, pp. 273–283, Nov. 1998.
- [8] M. Biassusi, "Estudo da deformação de um vertissolo através da tomografia computadorizada de dupla energia simultânea," Ph.D. dissertation, Univ. Federal do Rio Grande do Sul, Porto Alegre, Brazil, Feb. 1996.
- [9] E. A. Pualetto, "Metodologias para a determinação da variação de volume em solos deformáveis," in *Proc. Nat. Symp. Agropecuary Instrum.*, São Carlos, Brazil, Nov. 1997, pp. 128–135.
- [10] W. B. Lindquist and A. Venkatarangan, "Investigating 3D geometry of porous media from high resolution images," *Phys. Chem. Earth*, vol. 25, no. 7, pp. 593–599, 1999.
- [11] S. C. Felipussi and W. L. Roque, "A construção do esqueleto de formas bidimensionais utilizando o diagrama de voronoi," PPGC-Univ. Federal do Rio Grande do Sul, Porto Alegre, Brazil, Mar. 2001. Internal Rep.
- [12] C. T. J. Dodson and W. W. Sampson, "Flow simulation in stochastic porous media," *Simulation*, vol. 74, no. 6, pp. 351–358, 2000.
- [13] M. Sonka and H. Hlavac, *Image Processing, Analysis, and Machine Vision*, 2nd ed. Pacific Grove, CA: PWS, 1999.
- [14] S. Crestana, "A tomografia computadorizada como um novo método para estudos da física da água no solo," Ph.D. dissertation, Univ. São Paulo, São Paulo, Brazil, Aug. 1985.
- [15] H. Taud and T. Martinez-Angeles, "Porosity estimation method by x-ray computed tomography," *J. Pet. Sci. Eng.*, vol. 47, no. 3/4, pp. 209–217, Jun. 2005.
- [16] S. O. Prasher, J. Perret, and A. Kantzas, "Three-dimensional quantification of macropore networks in undisturbed soil cores," *Soil Sci. Soc. Am. J.*, vol. 63, no. 6, pp. 1530–1543, 1999.
- [17] R. Gosine and X. Zhao, "Automated image analysis for applications in reservoir characterization," in *Proc. 4th Int. Conf. Knowledge-Based Intell. Eng. Syst. Allied Technol.*, Brighton, U.K., Sep. 2000, pp. 620–623.
- [18] A. W. J. Heijs, J. Lange, J. F. Schoute, and J. Bouma, "Computed tomography as a tool for non-destructive analysis of flow patterns in macroporous clay soils," *Geoderma*, vol. 64, no. 3/4, pp. 183–196, Jan. 1995.
- [19] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.
- [20] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 176–201, Apr. 1993.
- [21] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin, Germany: Springer-Verlag, 1999.
- [22] N. Roberts, M. Reed, and G. Nesbitt, "Estimation of the connectivity of a synthetic porous medium," *J. Microsc.*, vol. 187, no. 2, pp. 110–118, Aug. 1997.
- [23] K. J. Maloi, J. Feder, and T. Jossang, "Viscous fingering fractals in porous media," *Phys. Rev. Lett.*, vol. 55, no. 24, pp. 2688–2691, Dec. 1985.
- [24] R. N. Onody, A. N. D. Posadas, and S. Crestana, "Experimental studies of the fingering phenomena in two dimensions and simulation using a modified invasion percolation model," *J. Appl. Phys.*, vol. 78, no. 5, pp. 2970–2976, Sep. 1, 1995.
- [25] E. Horowitz, *Computer Algorithms/C++*. Rockville, MD: Comput. Sci., 1996.

- [26] U. Grenander and M. I. Miller, "Representations of knowledge in complex systems," *J. R. Stat. Soc. B*, vol. 56, no. 4, pp. 549–603, 1994.
- [27] E. Gose, *Pattern Recognition & Image Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [28] R. C. Gonzalez and R. E. Woods, *Processamento De Imagens Digitais*. Reading, MA: Addison–Wesley, 1992.
- [29] *Volview Software*, (2002, Apr.). Kitware Inc. [Online]. Available: [www.kitware.com](http://www.kitware.com)



**Siovani C. Felipussi** received the B.S. degree in mathematics and the M.Sc. degree in applied mathematics from the State University of São Paulo, São José do Rio Preto, Brazil, in 1995 and 1998, respectively, and the Ph.D. degree in computer science from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 2003.

He is currently an Associate Professor of computer science with the Department of Computer Science, Federal University of São Carlos, Sorocaba, Brazil. His research interests are image processing and analysis, visual information retrieval, and porous media.



**Jacob Scharcanski** received the B.Eng. degree in electrical engineering and the M.Sc. degree in computer science from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 1981 and 1984, respectively, and the Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1993.

He is currently an Associate Professor with the Institute of Informatics, Federal University of Rio Grande do Sul. He also held research and development positions in Brazilian and North American Industry. He has authored and coauthored more than 80 papers in journals and conference proceedings. His research interests are image processing, pattern recognition, and computer vision.



**João L. D. Comba** received the B.Sc. degree in computer science from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, the M.S. degree in computer science from the Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, and the Ph.D. degree in computer science from Stanford University, Stanford, CA.

He is currently an Associate Professor of computer science with the Institute of Informatics, Federal University of Rio Grande do Sul. His current projects include the development of algorithms for large-scale scientific visualization, data structures for point-based modeling and rendering, and general-purpose computing using graphics hardware. He has authored and coauthored more than 50 papers in journals and conference proceedings. His research interests are graphics, visualization, spatial data structures, and applied computational geometry.