New Local Thresholding Method for Soil Images by Minimizing Grayscale Intra-Class Variance

Recent advances in imaging techniques offer the possibility of visualizing the three-dimensional structure of soils at very fine scales. To make use of such information, a thresholding process is commonly implemented to separate the image into solid particles and pores. Despite the multitude of thresholding algorithms available, their performance is being challenged by the complexity of the soil structure. Experience shows that, to improve thresholding performance, existing methods require significant input from a skilled operator, making the thresholding subjective. In this context, this article proposes a new, operator-independent thresholding technique based on the analysis of the intraclass grayscale variance. The method extends the well-established Otsu technique, by applying first a preclassification of the voxels corresponding to the solid phase. Then, a threshold value is determined through minimization of the intraclass variance of the unclassified voxels. The method was implemented globally, then locally for a range of window sizes, with the optimal window size selected as that for which the standardized grayscale variances of the two voxel populations are equal. Results on the three-dimensional soil images investigated suggest that the proposed method performs noticeably better than Otsu’s method, and in particular is robust enough to variations in image contrast and soil structure. Tested on a synthetic image, the new method produces a misclassification of only 2% of voxels, compared to 4.9% with Otsu’s method. These results suggest that the proposed method can be very useful in the analysis of images of a variety of heterogeneous media, including soils.

Over the last decade, major technological advances in imaging techniques, among which are X-ray computed tomography (CT) and microfluorescence spectroscopy, have afforded the possibility to investigate and reconstruct the three-dimensional architecture of natural porous media at very fine scales. These techniques are becoming particularly valuable to soil scientists, who can make use of the information on the internal three-dimensional structure to develop predictive models for a range of physical, chemical, and biological processes in soil to improve understanding of soil functioning at microscopic scales (Kravchenko et al., 2011; Pajor et al., 2010; Falconer et al., 2011; Hapca et al., 2011).

In general, available imaging techniques provide the users with three-dimensional grayscale images. As part of the quantitative analysis process, segmentation of the resulting three-dimensional grayscale data into a binary image that separates specific features in the image from the background information is a crucial step for subsequent analysis results. Of particular interest for soil scientists is the use of grayscale images to extract the pore space within the three-dimensional volume, whose topology underpins many bio-chemical and physical processes in soil. However, experience so far has demonstrated that the binary segmentation of soil images represents a substantial challenge (Iassonov et al., 2009; Baveye et al., 2010; Iassonov and Tuller, 2010). The irregular distribution of soil minerals may lead to different levels of noise across the reconstructed images, whereas the complex spatial structure of soils, exceeding the resolution of the imaging system, produces a partial volume effect in the reconstructed image in which a large proportion of image elements consist of a mixture of pore and solid material phases.

Most of the existing segmentation methods are not designed for porous materials of this complexity and therefore cannot cope adequately, or require significant supervision by a skilled operator to achieve acceptable segmentations (Sezgin and Sankur, 2004; Iassonov et al., 2009; Wang et al., 2011). Iassonov et al. (2009) present an extensive review of.
a number of segmentation methods used on porous media. They compare 14 segmentation methods ranging from global methods, where a single thresholding value is determined and used to segment the image, to locally adaptive methods, where segmentation is performed for each voxel based on the information contained in the neighborhood of the voxel. Only a limited number of methods perform consistently well in all the tested soil images. Among the global techniques, only Otsu’s method, based on minimizing the intra-class variance (Otsu, 1979) and Ridler’s iterative method, also known as iso-data or k-means clustering (Ridler and Calvard, 1978) lead to adequate results. The local methods, however, prove to be more stable, suggesting that the use of local spatial image information is crucial to obtain good segmentation quality. The best overall segmentation results are provided by Oh and Lindquist’s (1999) indicator kriging method and the Bayesian segmentation algorithm proposed by Berthod et al. (1996). However, in addition to being very computationally intensive, the main limitation of both of these two methods as reported by Lassonov et al. (2009) is the requirement for significant supervision by a highly knowledgeable operator.

To overcome the computational demands imposed by these methods, Houston et al. (2013a) developed an improved version of the Indicator Kriging algorithm of Oh and Lindquist (1999). The key improvement consists of varying locally the parameters of the algorithm, depending on local image conditions. Houston et al. (2013a) demonstrated that this adaptation leads to significantly better segmentation performance for a range of soil images. However, the method relies heavily on an initial preclassification of the grayscale voxels into pores and solid, which in itself is a very challenging step due to a lack of clear bi-modality in the global intensity distribution of many soil images. To address this particular issue, Schlüter et al. (2010) introduced a method for the determination of a two-level intensity threshold based on edge detection of regions of interest through gradient masks. This leads to a partial classification of the image, after which segmentation is completed by “region-growing,” a simple form of local correlation-based classification (Vogel and Kretzschmar, 1996). Nevertheless, the method requires the selection and application of a filtering step before thresholding, which, according to Kaestner et al. (2008), may have significant effects on segmentation results. It also involves a number of tuning parameters that need to be set in advance, leading to some subjectivity in the use of the method, as with the methods by Berthod et al. (1996) and Oh and Lindquist (1999).

This brief overview of the current state of the art of the application of thresholding techniques to images of complex systems, like soils, suggests that there is a dire need for the development of reliable, consistent, and fully automated (i.e., operator independent) thresholding algorithms for three-dimensional soil images. Another challenge faced by researchers in the field is to assess the quality of the resulting segmented images. A number of evaluation measures have been proposed in the past to achieve this goal. Commonly used measures are based on misclassification error, edge mismatch, relative foreground area error, region nonuniformity measure, or a combination of these (Zhang, 1996; Sezgin and Sankur, 2004; Zhang et al., 2008; Wang et al., 2011). However, with the exception of the region nonuniformity measure, existing approaches rely on the availability of a ground truth image, which unfortunately does not exist in the case of images of naturally occurring porous media. Attempting to circumvent this problem, scientists have tried instead to use morphological properties of scanned materials such as porosity, pore connectivity, or pore size distribution and compared the values provided by the segmented images with physical estimates of the different materials (Vogel et al., 2010, Wang et al., 2011, Pajor et al., 2010). While this is achievable for some porous materials, like synthetic soils or repacked soil samples, such metrics for real soils are generally not available.

To overcome the lack of “ground truth” information, and assist with the evaluation of different segmentation methods for soils, various techniques to simulate soil images have been proposed. For example, Wang et al. (2011) provided a comparison of image segmentation methods in simulated two-dimensional and three-dimensional images of soil aggregates. Simulated images were used to evaluate the performance of four segmentation methods including an entropy based method proposed by Sahoo et al. (1997), Ridler and Calvard’s (1978) iterative method, Otsu’s (1979) method, and Oh and Lindquist’s (1999) indicator kriging method. The study by Wang et al. (2011) reported that, when tested on simulated soil images, the Indicator Kriging method is optimal for images with a clear bimodal pattern, whereas for unimodal distributions, surprisingly Otsu’s method (Otsu, 1979) produced the lowest misclassification rate. The study by Wang et al. (2011) also concluded that, an evaluation criterion based on a region uniformity measure is not always adequate to select the best segmentation method (Wang et al., 2011).

In this general context, the present article describes a new, fully automated (operator-independent) method for the segmentation of three-dimensional images, where the threshold is determined by minimizing the intra-class variance associated with two classes of voxel populations. The proposed method extends the well-established Otsu technique by introducing a preclassification of voxels corresponding to the solid phase. The method is implemented first globally and then locally for different window sizes and is tested on a range of soil samples as well as on a three-dimensional synthetic image developed by Wang et al. (2011). A new evaluation measure is also proposed to assess the performance of the different local implementations and to help select the window size for which the method produces optimum segmentation results.
Theoretical Development of the Segmentation Method

The method developed in this study extends Otsu’s method (Otsu, 1979), which assumes that the grayscale values in the image correspond to two distinct classes (background and object). When producing the binary image, Otsu’s technique uses as threshold the value that minimizes the weighted sum of variances of the two classes, also known as the intra-class variance (Otsu, 1979).

From this vantage point, we propose to extend Otsu’s technique by first considering a preclassification step of voxels corresponding to the solid phase to accommodate the challenges imposed by the heterogeneity of the type of three-dimensional image we want to be able to threshold.

Basic Principles

For an 8-bit three-dimensional image, the grayscale value of each voxel ranges from 0 to 255. For a certain thresholding value, $t = 0 \ldots 255$, one obtains two classes of pore and solid voxels, for which the intra-class variance is defined as a weighted sum of variances of the two classes:

$$s^2_w(t) = \omega_0(t)s^2_0(t) + \omega_1(t)s^2_1(t)$$  \[1\]

where $\omega_i(t)$ are the weights of the two classes, and $s^2_i(t)$ are the variances of these classes for a given threshold value $t$. A common profile of the intraclass variance $s^2_w$, as a function of $t = 0 \ldots 255$ is presented in Fig. 1a. In Otsu’s technique, the threshold is given by the grayscale value, which minimizes the $s^2_w$ function (denoted by $T_{\text{Otsu}}$ in Fig. 1a). However, various tests on soil images show that this threshold produces a clear overestimation of the pore space (Baveye et al., 2010; Wang et al., 2011). An explanation for this is the fact that the two-class model underlying Otsu’s method is inappropriate for soils where more than two phases may occur.

In particular, the solid phase represents a mixture of minerals and organic matter, which may result in a large range of grayscale intensities in the image that cannot be represented by a single class. To help remove this bias, a preclassification of the solid population is required before estimating the threshold. A systematic inspection of a range of soil images suggested to use as upper bound for the solid population, the Otsu original thresholding value $T_{\text{solid}} = T_{\text{Otsu}}$, that is, the grayscale value where the intra-class variance $s^2_w$ function reaches the minimum (Fig. 1a). During the preclassification step, the grayscale values above $T_{\text{solid}}$ are assigned to the solid phase. The voxels with grayscale values below $T_{\text{solid}}$ remain unclassified, and they can represent pore space, organic matter, or partial volume voxels. After applying the preclassification step a new intraclass variance function $s^2_w$ is calculated on the unclassified voxels and the optimum threshold $T_{\text{optim}}$ is given by the value, which minimizes the new $s^2_w$ function (Fig. 1b).

Using $T_{\text{Otsu}}$ as the upper bound for the preclassification of the solid phase gave satisfactory results for most of the soil samples tested, in particular when the associated grayscale images displayed a unimodal or bimodal distribution. However, for the soil samples where the distribution of the grayscale intensities was characterized by more than two modes, the intra-class variance function $s^2_w$ displayed a more sinuous profile, possessing sometimes more than one local minimum (Fig. 2b). This situation was linked to the presence of more than two phases in the soil sample, either in the form of large amounts of organic matter or in the form of very dense soil minerals as part of the solid phase. In these cases, $T_{\text{Otsu}}$ was found inappropriate as it was underestimating considerably the solid phase, and the grayscale value where the first local minimum of the intra-class variance function was reached proved to be a better alternative value for $T_{\text{solid}}$ in the preclassification step. Further investigation of the profile of the intra-class variance function $s^2_w$ together with the profiles of the pore and solid

![Fig. 1.](image-url) (a) Typical profile of the intra-class variance function, which is used to determine the lower and upper bounds for the preclassification step. (b) The profile of the intra-class variance function on the unclassified voxels and the choice of the optimum threshold in the general approach.)
variances $s_0^2$ and $s_1^2$, respectively, revealed that for soil images where $s_w^2$ showed a clear unique minimum (coinciding with the global minimum), the profiles of $s_0^2$ and $s_1^2$ were crossing each other very close to that minimum (Fig. 2a). However, when $s_w^2$ displayed a more sinuous profile, then the profiles of $s_0^2$ and $s_1^2$ crossed each other at an earlier point value, close to the value of the first local minimum. In this situation, the point where the pore and solid variances becomes equal (denoted further by $T_{\text{equal var}}$), was found to be an appropriate value for $T_{\text{solid}}$. Consequently, to address the issue arising from soil sample images with multimodal grayscale distribution and to broaden the range of applicability of the method, the upper bound $T_{\text{solid}}$ in the preclassification step was chosen as $T_{\text{solid}} = \min\{T_{\text{equal var}}, T_{\text{Otsu}}\}$.

**Global and Local Implementations**

The method described above was implemented on images of size $256^3$, both globally and locally for a range of window sizes. First the three-dimensional image was decomposed into a cubic lattice, each cube of a fixed size ranging from $4^3$, $8^3$ up to $128^3$ voxels, and finally to the full image ($256^3$ voxels) corresponding to the global implementation of the method. The segmentation method described in the previous paragraph was then applied to calculate a threshold value for each cube in the lattice (Fig. 3b). As shown in Fig. 3b, due to the heterogeneous structure of soil samples, some of the cubes in the lattice produce threshold values very different from their neighbors and, therefore, using these discrete values to segment the grayscale image would produce discontinuities in the binary results. To address this, an interpolation method is applied to the resulting discrete thresholds in the lattice (Fig. 3b), to help obtain a smooth three-dimensional thresholding surface associated to each voxel in the image (Fig. 3c). Different interpolation methods were tested such as moving average, cubic spline, and the locally estimated scatterplot smoothing (LOESS, Cleveland and Devlin, 1988). However, in this case only LOESS gave satisfactory results, and therefore, it was used in the local implementation. So, for any point $x$ in the three-dimensional image the threshold is given by

$$T(x) = \sum w_i T(x_i)$$

**Fig. 2.** Typical profile of the intra-class variance function and corresponding pore and solid variance functions. (a) The situation when the pore and solid variance profiles cross each other very close to the point where the intra-class variance function reaches its global minimum. (b) A special profile of the intra-class variance function displaying two local minimum. For this case the pore and solid variance function crosses each other very close to the first local minimum of the intra-class variance function.

**Fig. 3.** Illustration of the steps involved in the implementation of the local method: (a) a two-dimensional slice of the original grayscale image; (b) decomposition of the original images into a cubic lattice with a thresholding value calculated for each cube; (c) LOESS interpolation method applied to the discrete thresholding values in Fig. 3b to obtain a smooth thresholding surface; (d) resulting binary image.
where \( T(x_i) \) is the discrete threshold value produced by the \( i \)th cube in the lattice when applying the thresholding method described in Basic Principles, and \( x_i \) is the center of the cube. The weights in the interpolation formula (2) are calculated using the tri-cubic weight function, that is,

\[
    w_i = \left(1 - \frac{\text{dist}(x, x_i)}{R}\right)^3
\]

where \( R \) is the maximum distance from \( x \) to the \( x_i \) in the lattice. The weights \( w_i \) are further renormalized to sum to 1 before being used in Eq. [2]. Finally, a 60% majority filtering on a \( 3^3 \) voxel centered neighborhood matrix, which is commonly used in image segmentation, to smoothen the solid–pore interface and to remove the solid voxels (Oh and Lindquist, 1999; Wang et al., 2011), was applied to the resulting binary images.

**Development of a Selection Criterion to Determine the Best Window Size for the Local Implementation**

Most of the local segmentation methods are implemented for a specific value of the window size (White and Rohrer, 1983; Niblack, 1986; Oh and Lindquist, 1999; Sauvola and Pietikäinen, 2000), which can be either encoded in the segmentation algorithm or is presented as an input parameter that needs to be specified by the image analyst. In most cases, an evaluation of the effect of the window size on the thresholding results is lacking, and therefore, the choice of the window size parameter remains subjective. For the new method described above, it was found that the window size affects strongly the results, with a decrease in window size leading in general to an increase in segmented pore space (Fig. 4). Therefore, classical evaluation measures like region nonuniformity, defined as the ratio between the grayscale variance in the pore space and the total grayscale variance (Zhang et al., 2008), are not appropriate as they tend to favor the global method (the one that achieves the lowest porosity). Neither are measures such as the intra-class variance appropriate. Due to the fact that Otsu’s method is based on minimizing the intra-class variance, none of the proposed local implementations could have outperformed Otsu’s method if the intra-class variance measure was used for evaluation. To overcome this problem and help select an appropriate window size for the local implementation, a new selection criterion is developed here based on minimizing both pore space and solid phase variances. The approach is similar to the one used in predictive statistics for binary variables to find a trade off between the sensitivity and specificity of a classification model (Hastie et al., 2009).

When the proposed method is applied with different window sizes, we found as a general trend that an increase in the window size produces a decrease in the segmented pore space and a concomitant decrease in the corresponding pore space grayscale variance, whereas the solid phase variance is gradually increasing. Although

![Fig. 4.](image-url)
this pattern is common to most of the tested soil samples, small deviations from this trend are observed in some cases, which may be accounted for by the scale of structural heterogeneity. As a result, a correction was applied to the series of binary images resulting from the different local implementations. For a given window size, the correction was achieved by successively multiplying (element by element) the corresponding three-dimensional binary matrix with the three-dimensional binary matrices resulting from the application of the segmentation method with larger window sizes. This way ensures a consistent increase in the pore space with the decrease of the window size, which is essential for the selection criterion of the optimum window size described in what follows.

If $s_{0,i}^2$ and $s_{1,i}^2$ denote the variances of the grayscale values in the pore space and solid phase, respectively, for the binary images corresponding to window sizes $i = 4, 8, ..., 256$, then (following the correction described above) the largest variance in the pore space is given by $s_{0,4}^2$ while the smallest variance is $s_{0,256}^2$. Thus, the variances $s_{0,i}^2, i = 4, 8, ..., 256$, can be further standardized to take values between 0 and 1, by applying the following formula:

$$c_{0,i} = \frac{s_{0,i}^2 - s_{0,256}^2}{s_{0,4}^2 - s_{0,256}^2} \quad i = 4, 8, ..., 256$$  \[4\]

As a result, the standardized variances $c_{0,i}$ of the pore space decrease from 1 to 0 with the increase of the window size $i = 4, 8, ..., 256$ (Fig. 4i).

A similar procedure is applied to the range of variances within the solid phase $s_{1,i}^2$, by using the formula:

$$c_{1,i} = \frac{s_{1,i}^2 - s_{1,4}^2}{s_{1,256}^2 - s_{1,4}^2} \quad i = 4, 8, ..., 256.$$  \[5\]

This time, the standardized variances of the solid phase $c_{1,i}$ increases from 0 to 1 with the increase of the window size $i = 4, 8, ..., 256$ (Fig. 4i). The optimum window size for the local implementation can be then chosen at the point where the two lines cross each other, or where the two coefficients $c_{0,i}$ and $c_{1,i}$ are the closest (Fig. 4i). However, in some cases the profiles of the two standardized variances may cross each other halfway between two window sizes, in which case both sizes can be considered as optimum. To address this problem, the segmentation method can be further refined by applying an interpolation procedure to calculate a suitable thresholding surface based on the thresholding surfaces generated by the two optimum window sizes. If $T_i$ and $T_{i-1}$ represent the thresholding surfaces generated by the two optimum windows of sizes $i$ and $i+1$, respectively, then an optimum thresholding surface $T_{\text{optim}}$ can be calculated as follows:

$$T_{\text{optim}} = \alpha_i T_i + \alpha_{i+1} T_{i+1}$$  \[6\]

where the interpolation weights are taken to be inversely proportional to the absolute difference between the standardized pore and solid variances corresponding to each window size, that is,

$$\alpha_i = \frac{c_{1,i+1} - c_{0,i+1}}{c_{0,i} - c_{1,i} + c_{1,i+1} - c_{0,i+1}}$$ \[7\]

$$\alpha_{i+1} = 1 - \alpha_i$$

**Materials, Method Implementation, and Testing**

**Image Data**

The performance of the segmentation method proposed in this study was assessed based on a series of real three-dimensional soil images as well as simulated soil images.

**Soil Images**

Three-dimensional soil images were obtained in house scanning facilities consisting of a Metris X-Tek X-ray micro-tomography system (NIKON Metrology, UK). Use of this equipment requires a series of operational decisions during the image acquisition and data preprocessing stages; decisions that can have a great impact on the segmentation results. In particular, as part of the data preprocessing step, the 32-bit floating-point representation of the original topographic reconstruction is mapped to unsigned 8-bit representation, so that the information can be easily displayed using simple computer graphics techniques. Due to the heterogeneity of the soil chemical structure, and in particular the presence of soil minerals, the intensity range in the original 32-bit floating-point representation can be very wide, with a small proportion of very high intensities. As a result, lower and upper limits need to be applied a priori to the grayscale values of the original image to determine a suitable interval for mapping. The choice of the lower and upper limits of this interval involves a subjective decision of the image analyst through a visual inspection of the original image histogram. When no limits are applied, a very poor contrast in the 8-bit image is obtained. The majority of grayscale values are clustered within a short interval, and there is little discrimination between grayscale values corresponding to pore and solid populations.

To assess the effect of interval mapping choice on segmentation results, undisturbed soil samples from a previous study by Sun et al. (2011) were used. The soil samples, obtained in cylindrical plastic containers (2-cm tall, 3-cm diameter), originated from an experimental site at the James Hutton Institute, Dundee, UK. Samples were scanned at 160 kV and 201 µA using a 0.1 mm Al filter to obtain 3003 angular projections (based on a 360° rotation), each of which was determined using the average of four exposures per frame. Radiographs were reconstructed into a three-dimensional volume using CT-Pro v.1.6 (NIKON Metrology, UK) and then imported into the image analysis software VGStudio-Max v.2.1 (Volume Graphics, Germany), where operator-selected regions of

---

www.VadoseZoneJournal.org  p. 6 of 13
interest were converted to unsigned 8-bit precision and exported as BMP format image stacks. All soil samples were scanned and reconstructed into three-dimensional volumes at a resolution of 30 μm. Three samples of different bulk densities (high, intermediate, and low) were selected from the larger set of samples and used for this study. When converting the original 32-bit reconstructed images into 8-bit images, two choices for the mapping interval were explored. First, the interval was chosen following a procedure commonly used in statistics to reject outliers (Zar, 1999), whereas the second mapping was done without applying any constraints to the 32-bit data. As a result, two 8-bit images were prepared for each of the three samples, giving a total of six images. Comparison of the quality of resulting paired images, in terms of image contrast, was done based on the root mean square (RMS) contrast measure, which is given by the ratio between the standard deviation and the range of the image grayscale values (Peli, 1990; Houston et al., 2013b).

To assess the effect of the physical structure on the segmentation results, images of repacked soil samples made of different aggregate sizes were used. This second set of soil samples, used in previous work by Schmidt et al. (2012) consisted of a Eutric Cambisol derived from undifferentiated sandstone, made of sand 71%, silt 19%, and clay 10%. The soil was air-dried and then sieved to 4 to 2, 2 to 1, and 1 to 0.5 mm. Cylindrical plastic containers (10-cm tall; 3-cm diameter) were loosely packed (i.e., filled) with soil sieved to the three aggregate fractions, resulting in samples with bulk densities varying between 0.9 Mg m⁻³ (for the samples made of larger aggregate) to 1 Mg m⁻³ (for the samples made of smaller aggregates). Samples were wetted to a matric potential of −0.03 MPa and stored at 20°C in darkness for 24 h before they were scanned (145 kV and 140 μA, resolution 34.9 μm, 2855 projections). A 0.1 mm aluminum filter was used to reduce beam hardening. Three replicates per aggregate fraction were prepared giving a total of nine samples.

All the reconstructed 8-bit images were cropped to obtain equally sized volumes for all samples of size 256³ voxels.

Synthetic Images
For further assessment of segmentation performance, the synthetic three-dimensional soil image generated by Wang et al. (2011) was also used in this study. The simulation protocol developed by these authors is based on a ground-truth binary image of soil aggregates, to which a combination of several image layers including simulated partial volume effect, differences in attenuation values of different solid materials and noise that accompanies the scanning, are added (Wang et al., 2011). The size of the simulated soil image provided by Wang et al. (2011) is 50³ voxels; however, to match the protocol developed in this study, a sub-image of 48³ voxels was cropped from the original image provided by Wang et al. (2011) and used to test the method.

Method Implementation and Testing
The segmentation method was implemented in Matlab v.7.11 (MathWorks, UK) for various window sizes and tested on the three-dimensional soil images as well as the synthetic image. For the soil image data, the local method was implemented with window of sizes 4³, 8³, 16³, 32³, 64³, 128³, and 256³ voxels, respectively, while for the synthetic image, due to the relative small size, the window in the local implementation was of sizes 4³, 8³, 16³, 24³, and 48³ voxels. Otsu’s thresholding method was also re-implemented in Matlab and applied on the same set of images to allow for method comparison. To provide a valid comparison between the two segmentation methods, images thresholded using Otsu’s technique were also subjected to a 60% majority post-filtering procedure. To assess segmentation performance, porosity measures were calculated for each segmented image and used to compare the performance of the two methods.

Method performance in response to variation in image contrast was assessed by comparing the profiles of the standardized variances (for the different window sizes of the local implementation) between paired images, as well as by calculating the percentage of voxels that were misclassified when comparing binary images resulting from respectively good-contrast images and poor-contrast ones.

The performance of the method in response to differences in the soil physical structure was evaluated by comparing segmentation results of the nine images of loosely repacked soil samples made of aggregates of different sizes (Image Data). Of particular interest was the impact of soil structure on the performance of the selection criterion for the optimum window size. The robustness of the local implementation was determined by examining whether the same window size was selected for the different replicates of the same aggregate size treatment.

Finally, the performance of the new selection criterion proposed in Development of a Selection Criteria was also evaluated on the three-dimensional synthetic image, by calculating the misclassification error between the ground-truth image and the binary image, which is given by the ratio between the number of misclassified voxels and the total number of voxels in the image (Sezgin and Sankur, 2004).

Results
Segmentation of Three-Dimensional Soil Data
Results of the local implementation of the segmentation method show a continuous shrinkage of the pore space when the window size is increased (Fig. 4). This behavior is consistent in most of the tested soil images. Small deviations from this pattern were observed on the images of repacked soils with larger aggregate size (2–1 and 4–2 mm), which then got corrected following the procedure described in Development of a Selection Criteria. In addition, regardless of the quality of the soil image or the type of soil structure involved, visual inspection of the resulting segmented images
shows that the global implementation produces an underestimation of the pore space, whereas the local implementation with the smallest window size, $4^3$, overestimates it. The variance within the pore space decreases as the window size gets larger, whereas the variance within the solid phase steadily increases. The point where the profiles of the two variances cross each other seems to be a reasonable objective criterion to select an optimum window size. In Fig. 4, this optimum window size corresponds to $64^3$ voxels. Results suggest that the choice of the optimum window size depends on the quality of the images and the structure of the soil sample.

Visual comparison between the performance of Otsu's method and the new local method for undisturbed soil images as well as images of packed soils (Fig. 5) shows that in all the tested images Otsu's method overestimated the pore space compared to the new approach. In addition, in both undisturbed and packed soil samples, the estimated porosity based on Otsu's segmented images does not have much physical meaning, with porosity of the intermediate density soil sample (porosity = 21.9%) being lower than porosity of the high density soil sample (porosity = 34.6%). At the same time, according to Otsu's method the porosity of packed soils decreases with the increase of aggregate size. Given that the aggregates were packed at similar bulk densities, we expect the porosity of the samples to be very close (samples made of smaller aggregates were packed at 1 Mg m$^{-3}$ bulk density and therefore should have a porosity slightly lower than those made of larger aggregates, which were packed at 0.9 Mg m$^{-3}$). However, small pores (smaller than the resolution of the scanning), which occur more often in the soil samples made of smaller aggregates, are not picked up by the scanning system, and therefore they are not represented correctly in the grayscale image, resulting in an underestimation of the total porosity, in particular for the samples made of small aggregates. This trend in porosity of soils packed with aggregates of different sizes was well reproduced by the newly proposed thresholding method, whereas Otsu's method yielded the opposite effect. These qualitative findings show that Otsu's technique is less reliable when applied to a range of images of soils with different physical structures.

Further comparison of segmentation performance between good-contrast and poor-contrast undisturbed soil images shows a change in optimum window size (Fig. 6) subject to changes in image contrast, calculated based on the RMS contrast measure (Table 1). With the decrease of image contrast the optimum window size shifts consistently toward larger values. Despite this difference in optimum window size between good- and poor-contrast images, comparison of the binary images obtained by using these respective window sizes shows a very good agreement for all three pairs of images. Only 1% of image voxels were classified differently for each of the three undisturbed soil samples (Fig. 6), and small difference in the resulting porosity between good- and poor-contrast images were obtained as indicated by very small standard errors in

Fig. 5. Comparison between performance of Otsu’s segmentation technique and the new proposed method. Porosity measures of the resulting binary images are also provided. For the undisturbed soil samples the mean porosity of good- and poor-contrast images is provided together with the corresponding standard error ($n = 2$); for the packed soils, mean porosity over the three replicates per each aggregate size is provided together with the corresponding standard error ($n = 3$).

Fig. 5. These results suggest that the proposed method is remarkably robust with respect to fluctuations in image quality (such as contrast), which might happen during the reconstruction process, due to the different reconstruction settings that are manipulated.
Application of the segmentation method on images of repacked soil samples made of aggregates of different sizes (Fig. 7) leads to standardized variance profiles that are almost identical for the three replicates of each of the soil aggregate sizes. At the same time, the choice of the best window size increases consistently from below $8^3$ voxels for the 1 to 0.5 mm aggregate size treatment to close to $16^3$ voxels for the 4 to 2 mm aggregate size treatment, indicating that the optimum window size is also dictated by the scale of the structural heterogeneity in the image.

![Fig. 6. Comparison between segmentation results of good contrast (sample a) and poor contrast (sample b) images of undisturbed soils with high density (sample 1), intermediate density (sample 2) and low density (sample 3) soil content.](image)

Table 1. Comparison between good contrast and poor contrast images of undisturbed soils as indicated by the RMS contrast values.

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Image quality</th>
<th>Good contrast RMS</th>
<th>Poor contrast RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>High density</td>
<td>0.136</td>
<td>0.037</td>
<td></td>
</tr>
<tr>
<td>Intermediate density</td>
<td>0.179</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>Low density</td>
<td>0.189</td>
<td>0.067</td>
<td></td>
</tr>
</tbody>
</table>
Method Validation by Synthetic Images

In the case of the synthetic image by Wang et al. (2011), the optimum segmentation results are obtained by interpolating between the thresholding surfaces corresponding to window sizes 43 and 83 (Fig. 8g). This segmentation choice is validated by a misclassification error of 2%, which is the smallest among the range of different local implementations (Table 2). In addition, the porosity of the resulting binary image is the closest to the porosity of the ground-truth image (16.86% as compared to 17.53% porosity, Table 2). In turn, Otsu’s technique performed less well than the proposed local method (Fig. 9), with a misclassification error of 4.92% and an overestimated porosity of 22.20% (Table 2). This is in accordance with the results on soil images obtained previously.

Discussion and Conclusions

This article presents a new method for bi-level segmentation of images, which aims to address a number of challenges encountered when dealing with images of heterogeneous systems, such as three-dimensional X-ray CT images of soils. The proposed approach is an extension of the well-established Otsu method (Otsu, 1979) in the sense that it builds on the analysis of the intra-class variance of the two classes, but unlike with Otsu’s method, in this work a preclassification step is performed before thresholding. Tested on a series of soil sample images, the proposed thresholding method shows a consistent improvement in segmentation results as compared to Otsu’s method. The underperformance of Otsu’s technique on the tested soil images can be explained by the fact that the method is implemented globally without taking into account local spatial correlations. In addition, it is based on the unrealistic assumption that the grayscale image values correspond to two distinctive groups (pore and solid). In reality, the solid phase is a mixture of organic matter and different minerals and, therefore, can cover a large range of intensity values in the grayscale image. Thus, the solid phase can be made up of several distinct subgroups,

![Fig. 7. Performance of the newly developed criterion for the selection of the optimum window size in the local implementation applied to images of soil samples packed at three aggregate sizes, with three replicate for each aggregate size.](image-url)
and this bias is removed in the present method by performing a preclassification of the solid population before thresholding.

One could argue that for a complete soil architecture characterization, in addition to the pore space identification, separation of the different components of the solid phase (such as soil organic matter or other soil minerals) should be aimed for. However, due to the heterogeneous nature of both the physical and chemical structure of the solid phase, segmentation of the different solid phases is very difficult. In particular, when trying to separate the soil organic matter, segmentation is challenged by the fact that the range of the associated image intensity values (grayscale values) overlaps that of the pore–solid interface, where image elements consist of a mixture of pore and solid material phases. Existing multiphase segmentation methods cannot cope with such complex structures (Kulkarni et al., 2012), and therefore under the current practice, quantification of the different phases is still achieved by manually labeling a substantial part of the image elements to help identify suitable thresholding values for the different soil phases (De Gryse et al., 2006; Sleutel et al., 2008; Elyeznasni et al., 2012).
Recent research efforts into the development of multiphase segmentation methods for three-dimensional soil images have resulted in a semi-automated algorithm based on Bayesian Markov Random Fields (Kulkarni et al., 2012). However, the proposed method is at an early stage of development, given that it has only been tested on materials of known structure, where the different solid phases are well separated. In addition, it relies on some form of manual input during the segmentation process, which makes the process subjective. Whereas it is acknowledged that the development of reliable multiphase segmentation methods for soil images is desirable to achieve a full characterization of soil architecture, this cannot be reached without a valid binary segmentation method that would be able to separate in a first instance the pore space from the solid phase in an objective and fully automated manner. The thresholding method proposed in this study represents progress in this direction.

Based on a preclassification step applied to the solid phase, to adequately model the scale of the structural heterogeneity of the soil sample, the method was implemented locally for a series of window sizes ranging from $4^3$, $8^3$ up to $256^3$ and a selection criterion was developed to help chose the optimum window size. As a result, in its current state the proposed method is very computational intensive. Large computational costs in the current implementation are associated with the use of a global interpolation method to produce a smooth thresholding surface (Global and Local Implementations). With the decrease of the window size, the discrete lattice becomes bigger and bigger and therefore a lot of resources are required to calculate the distances from a given point to all the nodes in the lattice. In turn, if the interpolation method is implemented locally on a neighborhood of fixed size, this would control the computational cost regardless of the size of the lattice. However, further testing is required to evaluate the effects of local interpolation on the segmentation results and to determine an appropriate neighborhood size for the local interpolation that would ensure the proper functionality of the method. Ways to improve the computational costs consist also of applying the local method on a restricted set of window sizes. This can be determined using an iterative approach by first generating results for the extreme window sizes $4^3$ and $256^3$ and an intermediate value such as $32^3$. Then, depending on the values of the standardized variances, the search can be narrowed down in the $4^3$ to $32^3$ interval or $32^3$ to $256^3$ interval, and so on, until the optimum window size is determined.

For the method developed here, it was found that the optimum window size depends greatly on the quality of the images as well as on the soil structure. In general, it is difficult to untangle what specifically in the sample image has contributed to the choice of the optimum window size, but it appears that this choice is related to the scale of important pore-solid features in the image. Despite this being a relatively intuitive result, the majority of local segmentation methods lack a proper evaluation of the effect of window size on thresholding results (White and Rohrer, 1983; Niblack, 1986; Oh and Lindquist, 1999; Sauvola and Pickettinen, 2000). In most of the local implementations, the window size is either presented as an input parameter, which needs to be specified by the image analyst, or it is encoded as a fixed value in the algorithm. Results suggest that the criterion proposed here to select the optimum window size is robust to changes in image contrast as well as changes in the physical structure of the soil. In addition, the method has been successfully validated on a synthetic three-dimensional soil image. Further testing, subject to different acquisition and reconstruction parameters, such as resolution of the scanning, noise level reduction, or beam hardening, should be considered, however, in a follow-up study to fully validate the robustness of the proposed binarization method.

While it is clear that the method described in this article is geared toward images of soil samples and shows great promise in that context, nothing prevents its use in a much broader context. Indeed the theoretical perspective on which it is based could be easily adapted to images of other materials, by identifying in the preclassification step thresholding values appropriate for those situations.

References


