GeotexInspector

An Image Processing Based Geotextile Analyzer

Rogério Galante Negri

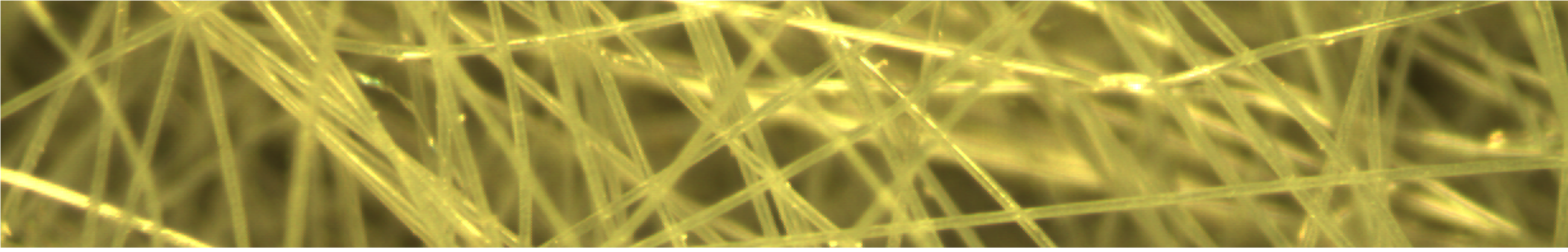
Manual translated by: Rodrigo Alves e Silva

República Federativa do Brasil - Ministério da Economia - Instituto Nacional da Propriedade Industrial - Diretoria de Patentes, Programas de Computador e Topografias de Circuitos Integrados - Certificado de Registro de Programa de Computador - Processo No: Br512019000211-3.

June 2018

GeotexInspector

1 GeotexInspector 5
1.1 The Software
1.2 Interface Items
1.3 Data format
1.4 Minimum requirements, Installation and Execution



1. GeotexInspector

1.1 The Software

The software `GeotexInspector` is the result of a research effort to investigate the pore size distribution (PSD) of nonwoven geotextiles using image-processing and pattern recognition techniques. The software was implemented between 2015 and 2017 with the use of functions available from the SLIC library (Small Library for Image Classification) [1]. PSD curves estimated using `GeotexInspector` were published by E Silva et al. [2].

The development of the presently described software was galvanized by the following main goals: (i) to offer a robust (and free-to-use) tool for teaching and research purposes; (ii) to build a relatively simple interface with respect to the adjustment of parameters for the experimental analyses; (iii) to provide researchers and industry professionals with a quick yet accurate method to examine pore openings in nonwoven geotextiles, particularly when they are applied to filtration functions.

1.2 Interface Items

The interface of `GeotexInspector` comprises a general menu (Figure 1.1) and a main window (Figure 1.2) where the user can specify the outputs that will be generated. The description of each interface item is provided below.

1.2.1 General Menu

Figure 1.1 depicts the items in the General Menu of `GeotexInspector`.

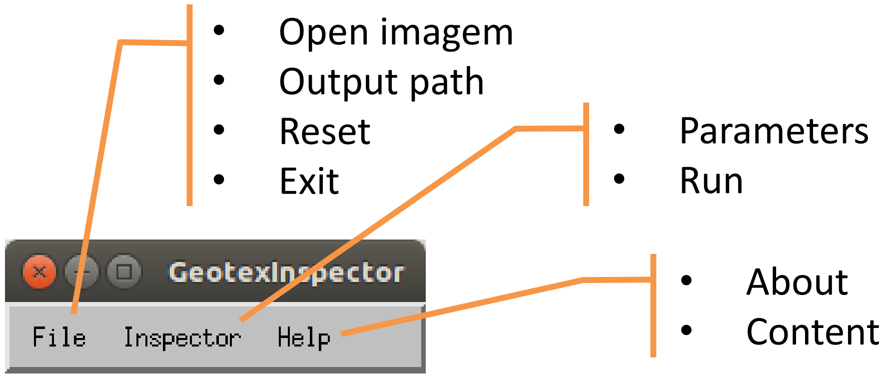


Figure 1.1: Options in the general menu

File: The user can open an image file, specify an output path, disassociate current files, or stop the current run and exit the program. The following features are associated with this option:

Open image: Selects an image from the user's library to be processed;

Output path: Defines a target path inside which the outputs from the image processing will be saved;

Reset: Disassociates current specifications (image file and output path) and clears the memory;

Exit: Stops the run and exits the program.

Inspector: The user can specify the adjustment parameters and start the analysis. The following features are associated with this option:

Parameters: Displays the window for the configuration of parameters related to the processing algorithm. (Section 1.2.3 – Figure 1.3);

Run: Starts the image processing.

Content: Items of informational purposes, which provide:

Help: Access to this manual;

About: Descriptions regarding authorship, funding and support.

1.2.2 Main window

Figure 1.2 depicts the main window of `GeotexInspector`, which includes:

1. **General menu**, discussed in Section 1.2.1;
2. **Outputs**, corresponding to a list of items which can be generated according to the user's preference, namely:
   (a) *Histogram:* graphical version of the frequency distribution of geotextile pores identified in the image-processing analysis;

Figure 1.2: Main window of `GeotexInspector`.

(b) *Image and Pores:* overlay of the original geotextile image and the representation of pore diameters as red circles;

(c) *Segmentation:* intermediate result from the segmentation step, i.e. distinction between pores and fibres (see Section 2.2);

(d) *Report:* text file containing information of the processed image, e.g. descriptive statistics and distribution of pore size diameters identified through the process.

1.2.3 Adjustment of parameters

The adjustment of parameters related to the identification and estimation of geotextile PSD curves can be done via `Inspector` → `Parameters`. Figure 1.3 depicts the interface with the available parameters. They are primarily associated with the classification method known as Support Vector Machine, or SVM (see Section 2.2), as well as with other processing steps adopted. The parameters within this interface comprise basically two groups:

3. **SVM parameters**, consisting of parameters inherently related to the SVM classification method:
   (a) *RBF Gamma:* parameter related to the radial basis function kernel adopted (RBF – *Radial Basis Function*), acting as a “flexing factor” of the hyperplane that distinguishes the pixels of a geotextile image between pores and fibres. Initial suggestions for this parameter are: 0.25, 0.5, 0.75, 1.0 or 1.25;
   (b) *Penalty:* parameter that acts as a tolerance during the search for a hyperplane and its decision boundaries to separate the classes. Initial suggestions for this parameter are: 100 or 1000.

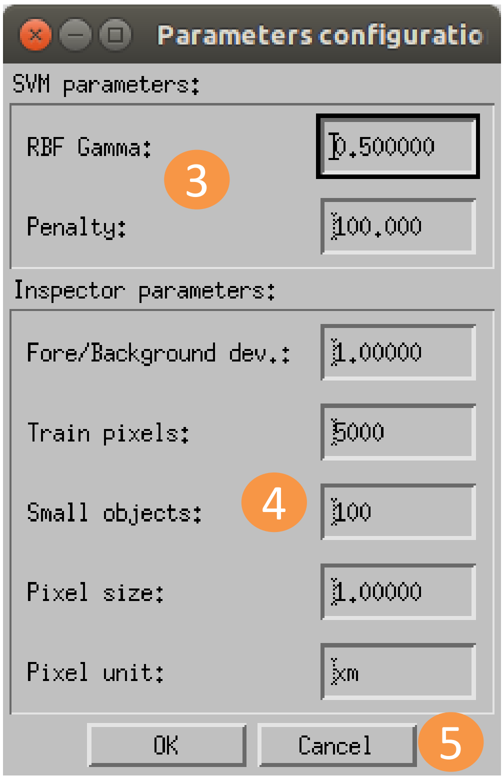


Figure 1.3: Interface for the adjustment of parameters

4. **Inspector parameters**, which allow the user to configure parameters such as:
   (a) *Fore/Background dev.:* multiplying factor that acts as a threshold in distinguishing "bright elements" from "dark elements", according to the rationale described in Chapter 2. Initial suggestions for this parameter are: 0.25, 0.5, 0.75 or 1.0;
   (b) *Train pixels:* corresponds to the number of pixels used as training examples of the SVM method. These equally distributed pixels are selected randomly, and further divided into two important sets: one containing examples of high intensity levels (e.g. fibres) and another containing examples of low intensity levels (e.g. pores). The selection of training samples is affected by the fore/background deviation previously described;
   (c) *Small objects:* determines the minimum pixel area for a given region to be classified as a geotextile pore;
   (d) *Pixel size:* length (horizontal measurement) of each pixel of the input image;
   (e) *Pixel unit:* unit associated to the pixel length.
5. **OK/Cancel button** either accepts or cancels the configuration of parameters established in this interface.

The parameter adjustment, either involved with the selection of training examples or with the training process of the SVM method (i.e. penalty and hyperplane flexing factor), must be conducted carefully, since it may interfere with the classification of pores and fibres and, therefore, with PSD estimation. The user is advised to check different combinations of parameter values, since geotextiles and image capturing devices will obviously be different from those employed by E Silva et al. [2]. Currently, there is no automatic method for obtaining the optimal set of parameters.

1.3 Data format

In order to guarantee compatibility with other digital image-processing techniques, both input and output data of `GeotexInspector` correspond to images represented by the RBG color scale and in TIFF format (Tagged Image File Format) [3]). This file format also assures flexibility, since it imposes no limit with respect to the number of bands and considers different types of data (e.g., *byte*, *integer*, *float*, *double*, *complex*, *double complex*, etc.) and methods of organizing image data (e.g. BIL, BIP and BSQ), among other advantages.

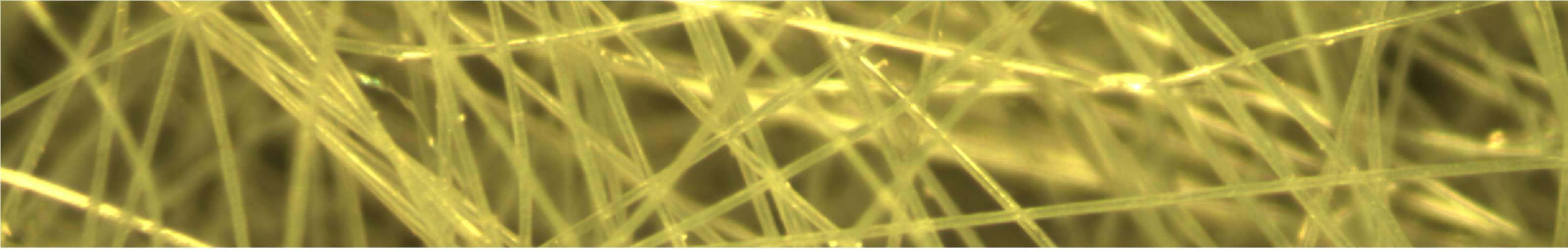
1.4 Minimum requirements, Installation and Execution

In order to use the `GeotexInspector` software, it is suggested the use of a platform with minimum of 2 gigabytes (GB) of internal (RAM) memory and Intel Core i3 processor (or superior). `GeotexInspector` can be executed in both Windows or Linux operating systems.

To install the software, one should simply copy the content from the directory `instGeotexInsp` to a local directory with access permission, access such directory and run the "GeotexInspector" executable file.

II Theoretical ground

2 Overview 13
2.1 Digital image processing for PSD curve estimation
2.2 Selection of pixels, training and classification
2.3 Identification of pore diameters and estimation of the pore size distribution curve



2. Overview

2.1 Digital image processing for PSD curve estimation

From digital images of geotextile specimens obtained through microscopy, four elementary concepts are performed: (1) identification and automatic selection of pixels related to both geotextile fibres and pores; (2) generation of a simplified representation based upon an image classification method known as Support Vector Machine (SVM), which demarcates the image considering the regions of fibres and pores; (3) removal of errors that eventually occurred during the classification and identification of polygons defining the pore regions; and (4) verification of the largest inscribing circle for the polygons recognized in the previous step. The frequency distribution of the circle diameters that are delimited through this process enabled the estimation of pore size distribution curves for a given geotextile planar section. Figure 2.1 depicts the preceding elementary concepts, which are further described in the next subsections.

2.2 Selection of pixels, training and classification

Since the focus at this stage was aimed at identifying geotextile pores through digital image analysis, the use of classification techniques posed as a very convenient starting point. In general, whenever a classification method is performed over the spectral behavior of pixels from a digital image, a second image is then generated with existing objects categorized in accordance with classes initially defined.

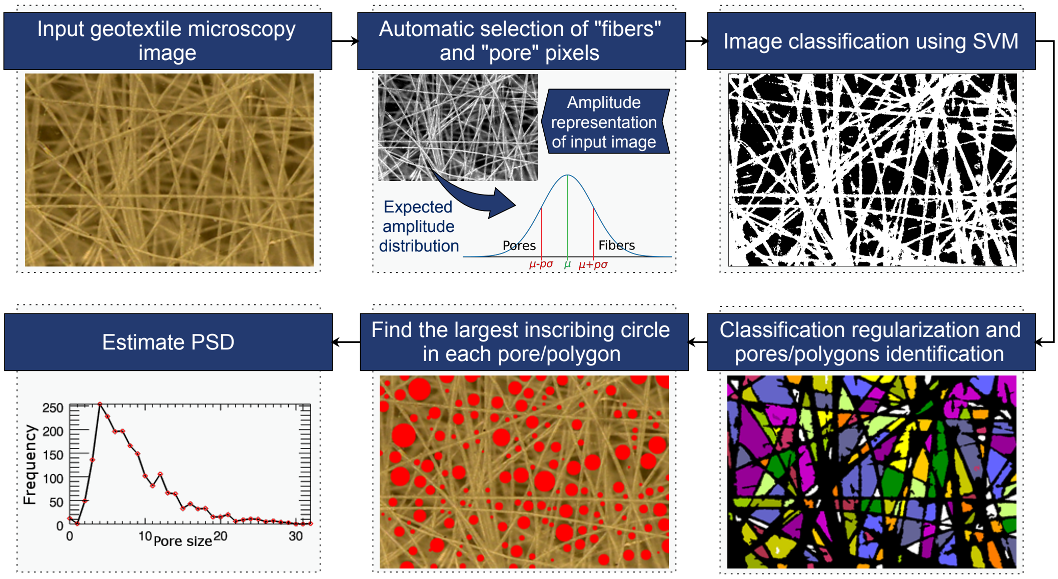


Figure 2.1: General overview of the pore size distribution curve estimation through GeotexInspector

Precisely, suppose $\mathscr{I}$ is an image whose spectral behavior of pixels, expressed in terms of different intervals from the electromagnetic spectrum or even by the channels of a color representation, is characterized by a vector $\mathbf{x} \in \mathscr{X}$, where $\mathscr{X}$ is a vector space termed attribute space. Furthermore, suppose $\Omega = \{\omega_1, \omega_2, \ldots, \omega_c\}$ as a set of classes related to the objects represented in the image $\mathscr{I}$. The classification of $\mathscr{I}$ according to the classes specified by Ω is obtained by applying a function $F : \mathscr{X} \to \Omega$ to the vector $\mathbf{x}$ associated to each pixel of $\mathscr{I}$.

The distinct classification techniques that have been proposed in literature correspond to distinct manners of modelling and expressing F. The parameters that determine the behavior of F, which in turn allow the adequate association between the elements of $\mathscr{X}$ to a single class of Ω, are obtained based upon training examples whose classes are known in advance.

Among an extensive variety of classification techniques previously proposed, SVM has proved robust in several application areas, exceeding many other methods when it comes to accuracy. Burges [4] describes a large portion of real applications employing the SVM method, among which facial recognition, text categorization, bioinformatics and other uses inside Civil and Electrical Engineering stand out. In general terms, the SVM method establishes a hyperplane capable of executing the separation between the elements of $\mathscr{X}$, and the parametrization of this separating hyperplane is optimally achieved based upon training examples. More details about the SVM method, Radial Basis Function and its parameters can be found in Webb and Copsey [5].

SVM is employed herein towards the classification of input images according to the classes "fibers" and "pores". Once the classes that contemplate the classification procedure are defined, it is then necessary to obtain the examples to train the SVM method. To this end, stemming from the premises that the geotextile fibres and the areas corresponding to pores have noticeable luminosity/amplitude differences and that the amplitude values (intensity) of pixels from a digital image can be represented by the norm of its attribute vector, the selection of training examples is done in consonance with the following rules:

$$
\begin{aligned}
\mathscr{D}_1 &= \{\mathbf{x}_i : \|\mathbf{x}_i\| > \mu + p\sigma;\ i = 1,\ldots,m\},\\
\mathscr{D}_2 &= \{\mathbf{x}_i : \|\mathbf{x}_i\| < \mu - p\sigma;\ i = 1,\ldots,n\}
\end{aligned}
\tag{2.1}
$$

where: $\|\cdot\|$ corresponds to the vector norm, μ and σ represent, respectively, the average and standard deviation of the amplitude values of pixels from $\mathscr{I}$, and p is a scale factor that controls the amplitude contrast between the elements $\mathscr{D}_1$ and $\mathscr{D}_2$. It is important to mention that $\mathscr{D}_1$ contains examples with high intensity levels of the class ω_1, which defines the "fibres", while $\mathscr{D}_2$ encompasses examples with low intensity levels of the class ω_2, defining the "pores". Additionally, the referred sets are composed by m and n examples, randomly selected from the pixels of $\mathscr{I}$ that satisfy the criteria established by Equation 2.1.

Sets $\mathscr{D}_1$ and $\mathscr{D}_2$ are put together in a single set $\mathscr{D}$, which is in turn employed by the SVM method in the classification of the input image. The step designated "Image classification using SVM" depicted in Figure 2.1 illustrates the classification of the input geotextile microscopy image into the predefined classes "fibres" (white) and "pores" (black). It is, at last, worth highlighting that pixels whose intensity levels fall in between $(\mu - p\sigma)$ and $(\mu + p\sigma)$ are not suitable examples for the SVM, since they may not provide an accurate perception as to what is actually *pore* or *fibre* elements.

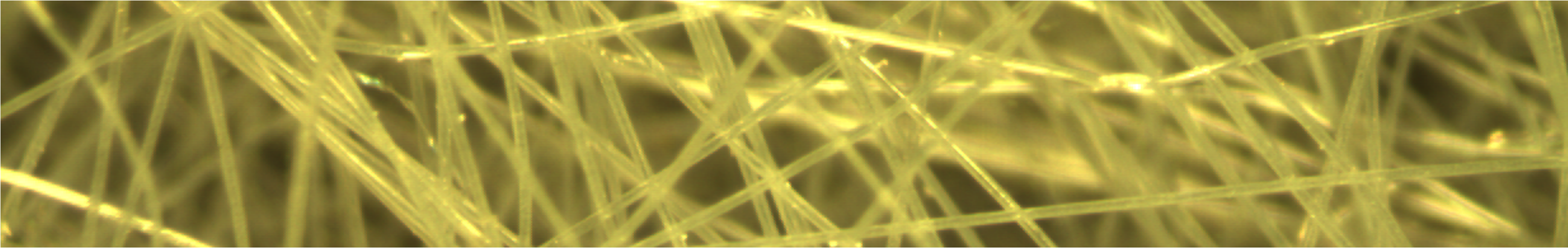
2.3 Identification of pore diameters and estimation of the pore size distribution curve

As discussed in the previous subsection, a representation distinguishing the regions of fibres and pores of the geotextile can be achieved through the classification of each pixel from an input image. Such representation plays a very important role in the estimation of the pore size distribution curve. However, it must be noted that effects derived from the individual classification of pixels may be detrimental. These effects are associated to pixels whose classes differ from their immediate neighbours, an aspect commonly known as "salt-and-pepper" noise and that may induce misidentification of sparse and small regions.

In order to eliminate the pixels whose classes diverge from their neighbours, a convolution filter with kernel of 3×3 pixels wide is applied to substitute the class of the central pixel for the higher frequency inside the kernel. This process is frequently referred as "mode filter" [6]. In view of this regularized classification and considering only the pixels associated to the class "pores", the identification of each polygon delimiting the geotextile pores is carried out with the use of a connected component labelling technique [7]. Finally, polygons eventually identified as having an area smaller than a predefined reference value τ are aggregated to neighbouring polygons (i.e. those having common borders) of higher area.

After performing the previous steps, which regularizes the classification result of an input image in view of the classes "fibres" and "pores" and identifies the different polygons associated to pore regions (depicted as "Classification regularization and pores/polygons identification" in Figure 2.1), the inscribing circles are verified for each polygon. This process, exemplified as "Find the largest inscribing circle in each pore/polygon" in Figure 2.1, is conducted in an exhaustive manner, starting from circles whose diameter progressively increases until only the internal area of the polygons were simultaneously occupied. At the end of this procedure, it is possible to quantify the cumulative frequency $f(d)$ of circles whose diameter are equal or less than a certain value d. Consequently, the pore size distribution curve for the concerned geotextile input image is estimated.

The processes described in Sections 2.2 and 2.3 delineate the implementation of `GeotexInspector`, available from:
`https://github.com/rogerionegri/GeotexInspector`.



Bibliography

[1] R. G. Negri. "Máquina de Vetores de Suporte Adaptativa ao Contexto: formalização e aplicações em Sensoriamento Remoto". PhD thesis. National Institute for Space Research, 2013, p. 166.

[2] R. A. E Silva, R. G. Negri, and D. M. Vidal. "A New Image-Based Technique for Measuring Pore Size Distribution of Nonwoven Geotextiles". In: *Geosynthetics International* 26.3 (2018), pp. 261–272.

[3] ADOBE. *TIFF: Revision 6.0*. 2013. URL: `https://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf/`.

[4] Christopher J.C. Burges. "A Tutorial on Support Vector Machines for Pattern Recognition". In: *Data Mining and Knowledge Discovery* 2 (1998), pp. 121–167.

[5] A.R. Webb and K.D. Copsey. *Statistical Pattern Recognition*. 3rd. Malvern, UK: Wiley, 2011.

[6] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. 4th. New York, NY, USA: Pearson, 2017, p. 1184.

[7] Linda G. Shapiro. "Connected Component Labeling and Adjacency Graph Construction". In: *Machine Intelligence and Pattern Recognition* 19 (1996), pp. 1–30.