

# CORRELATION-BASED MULTI-SHAPE GRANULOMETRY WITH APPLICATION IN POROUS SILICON NANOMATERIAL CHARACTERIZATION

Hae Yong Kim, Ricardo Hitoshi Maruta,  
Danilo Roque Huanca and Walter Jaimes Salcedo

Escola Politécnica, Universidade de São Paulo,  
Av. Prof. Luciano Gualberto, tr. 3, 158, 05508-010, São Paulo, Brazil.  
E-mails: hae@lps.usp.br, wsalcedo@lme.usp.br.

Image-based granulometry measures the size distribution of objects in an image of granular material. Usually, algorithms based on mathematical morphology or edge detection are used for this task. We propose an entirely new approach, using cross correlations with kernels of different shapes and sizes. We use pyramidal structure to accelerate the multi-scale searching. The local maxima of cross correlations are the primary candidates for the centers of the objects. These candidate objects are filtered using criteria based on their correlations and intersection areas with other objects. Our technique spatially localizes each object with its shape, size and rotation angle. This allows us to measure many different statistics (besides the traditional objects size distribution) e.g. the shape and spatial distribution of the objects. Experiments show that the new algorithm is greatly robust to noise and can detect even very faint and noisy objects. We use the new algorithm to extract quantitative structural characteristics of Scanning Electron Microscopy (SEM) images of porous silicon layer. The new algorithm computes the size, shape and spatial distribution of the pores. We relate these quantitative results to the fabrication process and discuss the rectangle porous silicon formation mechanism. The new algorithm is a reliable tool for the SEM image processing.

## INTRODUCTION

Granulometry is the process of measuring the size distribution of different objects/grains in a granular material. The *size distribution* is the histogram of objects as function of radius, also known as *granulometric curve* or *pattern spectrum*. Granulometry can be based on the analysis of an image or based on some other physical property, such as magnetism [1] or laser scattering [2]. In this paper, we will deal only with the digital image-based granulometries. There are two main groups of image-based granulometries in the literature:

- (1) Mathematical morphology-based algorithms;
- (2) Edge detection-based algorithms.

Morphology granulometry for binary images obtains the pattern spectrum of an image without explicitly segmenting it. It can be illustrated by sieving. In sieving, the granular sample passes through a series of sieves with decreasing hole sizes. As a consequence, grains in the sample are separated according to their sizes. The operation of passing the sample through a sieve of a certain size is described in morphology as an image

operator that returns the subsets of the image smaller than a specified size. Matheron has first proposed the morphology granulometry [3]. Dougherty et al. present a popular morphology granulometry for binary images [4]. Raimundo et al. used this algorithm to characterize porous material [5]. This algorithm can compute the size distribution, but it cannot compute the distribution of grains/objects by shape or orientation. Moreover, this algorithm is highly sensitive to noise: one single pixel inside grain/object with wrong color leads to a completely wrong size estimation of that grain/object.

Morphology-based granulometry for binary images can be extended to grayscale images using grayscale morphological operators. Vincent [6] presents an algorithm to accelerate grayscale morphology granulometry. A demonstration program of grayscale morphology granulometry with source code is available at [7]. Unfortunately, the output of grayscale morphology granulometry is highly non-intuitive, making its use difficult in practice. It represents the size distribution as the “sum of pixel values in opened image as a function of radius.” Like its binary version, the grayscale morphology granulometry is unable to detect the shape, rotation angle and spatial localization of each grain/object. It is also highly sensitive to noise (especially to impulsive noise).

Morphology granulometry assumes implicitly that every pixel inside object/grain is darker (or brighter) than the surrounding background. In some applications, this assumption does not hold. In this case, granulometry based on object/grain boundary detection may be used. According to Maerz et al. [8], mining companies can use boundary-based granulometry to quantify the size of rock fragments and forestry companies can use it to measure the wood chip sizes. However, there are only few papers on the boundary-based granulometry in the literature and they present only overview of the technique. This seems to indicate that boundary-based granulometries depend on ad hoc implementations using many heuristics. Maerz et al. [8] present a boundary-based granulometry that detects the edges of the objects using conventional gradient operators and thresholding. Then, it delimitates the objects using the edges. Edge-detection is a noise-sensible operation and may not be reliable in noisy or blurred images. In many applications, objects/grains (such as rock fragment or wood chip) may present internal edges that can be erroneously taken as object boundaries. Also, some object/grain boundaries may not be clearly visible in the image. So, to use this technique, special lighting and careful imaging may be necessary to obtain good results.

This paper presents a new approach for the granulometry, based on cross correlations. In other words, our technique is based on template matching, a technique used in image processing for finding small parts of a search image that match a template image [9, 10]. One of the authors of this paper has recently proposed two fast rotation and scale-invariant template matching techniques [11, 12]. Unfortunately, these techniques are not adequate for finding very simple and small shapes that may appear faintly in the image. In granulometry, we are highly concerned with the accuracy, and not deeply concerned with the computer performance. So, we use a simple idea that is not adequate for real-time applications but yields high accuracy: compute the correlations with kernels that represent all the shapes, eccentricities, orientations and sizes. Our technique is appropriate for applications where an object/grain has darker (or brighter) *average* grayscale than the *average* background grayscale. This is a requirement much easier to fulfill in practice than the requirement of morphology granulometries (every pixel inside object/grain must be darker than the background). Unlike the morphology granulometry, our algorithm spatially localizes each individual grain/object with its shape, size and rotation angle, providing much more information than morphology granulometries. Experiments show that the

new algorithm is greatly robust to noise and can detect even very faint, noisy and partially intersecting objects. In order to obtain scale-invariance with low computational cost, we use pyramidal structure. We have implemented the proposed algorithm in C++ using OpenCV image processing library.

We use the proposed technique to compute the quantitative structural characteristics of the porous silicon layer applied to Scanning Electron Microscopy (SEM) images. These results allow us analyzing the square pores formation correlated to thermal annealing process of the silicon wafer that was metalized previously to the anodization process. The new algorithm is a reliable tool for the SEM image processing.

Recently, mathematical morphology-based granulometry has been used to analyze 3D images, most of them obtained using X-ray tomography or focused ion beam tomography [13-15]. We think that the fundamental ideas presented in our paper can be applied to 3D images. However, the implementation details must be adapted to each particular application.

## CORRELATION-BASED GRANULOMETRY

### *Cross Correlation and Correlation Coefficient*

The discrete cross correlation between two real-valued images  $T$  and  $A$  is defined by expression:

$$R_{\text{corr}}(x, y) = T(x, y) \circ A(x, y) = \sum_i \sum_j T(i, j) A(x+i, y+j). \quad (1)$$

In practice, images are defined only inside a rectangular domain. An image is considered to be filled with zeros outside of its domain. Cross correlation can be efficiently computed using FFT (Fast Fourier Transform). In our case, we will use cross correlation to find a (usually small) template image  $T$  inside a (usually large) image to analyze  $A$ , an operation known as *template matching*. Let  $t$  be the vector obtained by copying the pixel values of  $T$  and let  $a(x, y)$  be the vector obtained by copying the pixel values of  $A$  inside the domain of  $T$  translated to pixel  $(x, y)$ . With these definitions, the cross correlation can be rewritten as:

$$R_{\text{corr}}(x, y) = t \cdot a(x, y) \quad (2)$$

where “ $\cdot$ ” is the dot product (also known as scalar product). To use cross correlation for template matching, a small modification must be introduced in the cross correlation, yielding the mean-corrected cross correlation:

$$R_{\text{mcc}}(x, y) = \tilde{t} \cdot a(x, y) \quad (3)$$

where  $\tilde{t} = t - \bar{t}$  is the mean-corrected vector obtained by subtracting the mean grayscale  $\bar{t}$  from each element of vector  $t$ . The peaks of image  $R_{\text{mcc}}$  correspond to the occurrences of  $T$  in  $A$ .

Template matching using mean-corrected cross correlation is brightness-independent but contrast-dependent. That is, if many instances of  $T$  with different contrasts occur in  $A$ , the instances with large contrasts will correspond to the high peaks in  $R_{\text{mcc}}$ . The mean value (brightness) of the instances does not change the values in  $R_{\text{mcc}}$ . For image pro-

cessing or computer vision applications where the brightness/contrast of the images can vary due to illumination and exposure conditions, contrast-dependency is an undesirable property. So, to obtain brightness/contrast-invariance, the normalized cross correlation (or normalized correlation coefficient) is usually applied instead [10]:

$$R_{\text{ncc}}(x, y) = \frac{\tilde{t} \cdot \tilde{a}(x, y)}{\|\tilde{t}\| \|\tilde{a}(x, y)\|} \quad (4)$$

where  $\|\cdot\|$  is the length of the argument vector, that is,  $\|\tilde{t}\| = \sqrt{\tilde{t} \cdot \tilde{t}}$ . However, in our application, the contrast-dependency of the mean-corrected cross correlation allows us to distinguish clear objects (a high contrast instance in  $A$  yields high peak in  $R_{\text{mcc}}$ ) from faint objects (a low contrast instance in  $A$  yields so low peak in  $R_{\text{mcc}}$ ). An appropriate value to threshold image  $R_{\text{mcc}}$  allows us detecting only the clear objects (high threshold) or even faint low-contrast objects (low threshold). Consequently, we will use the non-normalized cross correlation with mean-corrected templates.

### ***Correlation-Based Granulometry with Circular Kernels***

To obtain the pattern spectrum of an image  $A$  of granular material with only circle-shaped grains/objects, we propose to compute cross-correlations of  $A$  with circular kernels of different radii  $T_1, T_2, \dots, T_n$  (the leftmost column of figure 1). The sizes of the kernels increase in geometric progression. We have noticed experimentally that 5 kernels per octave are enough to obtain scale-invariance. The radii  $r_i$  and  $R_i$  of inner and outer circles are chosen to satisfy  $R_i \cong r_i \sqrt{2}$ . This makes the number of black pixels (that correspond to the negative values) nearly equal to the number of white ones (that correspond to the positive values). Gray pixels are zeros. In each kernel the sum of all negative pixels is  $-0.5$  and the sum of all positive pixels is  $+0.5$ . So, if the grayscale of the image  $A$  ranges (for example) from 0 to 1, the image resulting from the correlation will range from  $-1$  to  $+1$ . Let us denote the images resulting from the cross correlation between  $A$  and  $T_i$  as  $C_i$ , that is:

$$C_i(x, y) = T_i(x, y) \circ A(x, y), \quad 1 \leq i \leq n \quad (5)$$

Note that, by construction, the template  $T_i$  is equal to the mean-corrected template  $\tilde{T}_i$ , so the cross correlation is equal to the mean-corrected cross correlation. Let us define the maximal correlation image as:

$$C(x, y) = \underset{i=1}{\overset{n}{\text{MAX}}}[C_i(x, y)]. \quad (6)$$

A pixel  $(x, y)$  is a peak in  $C$  if it is greater than or equal to its eight spatial neighbor pixels. Each peak  $C(x, y)$  corresponds to an approximately circular object/grain in  $A$ . The scale of the detected object is given by the argument of the maximal correlation:

$$G(x, y) = \underset{i=1}{\overset{n}{\text{ARGMAX}}}[C_i(x, y)] \quad (7)$$

The radius of the object at  $(x, y)$  is the radius of the corresponding kernel  $T_{G(x, y)}$ . Usually, this process detects many non-existing grains together with the real objects. We filter out the false objects using two parameters chosen by the user:

The first parameter  $\tau$  ( $0 \leq \tau \leq 1$ ) is the correlation threshold parameter. The peaks with correlations smaller than  $\tau$  are discarded. This parameter must be set iteratively by the

user, because there is no way for the computer to decide if a faint object is to be discarded or not.

The second parameter  $\gamma$  ( $0 \leq \gamma \leq 1$ ) controls the allowed amount of intersection between the grains. If  $\gamma=0$ , the grains cannot intersect each other in the slightest. If  $\gamma=1$ , a grain can lie completely inside another grain. Our algorithm sorts and scans the correlation peaks in decreasing correlation order, discarding an object at pixel  $p_1$  with radius  $r_1$  and correlation  $c_1$  if there is some other object at pixel  $p_2$  with radius  $r_2$  and correlation  $c_2$ , satisfying:

$$\begin{cases} c_1 < c_2 \\ \text{distance}(p_1, p_2) < r_1 + r_2 - \gamma r_1 \end{cases} \quad (8)$$

Figure 2 depicts the outputs of our algorithm and compares it with the outputs of thresholding (the usual preliminary step of binary morphology-based granulometry) and Canny edge detection (the usual preliminary step of edge-based granulometry). In figure 2a, there are 21 pores that do not intersect with the image canvas, where 18 of them are clearly visible and 3 are faintly visible. Our algorithm successfully detected all the 21 pores or only the 18 clearly visible ones, depending on the chosen threshold  $\tau$ . It is not possible to detect all the pores after the thresholding (figures 2d and 2e), because if the faint pores become visible (threshold=70) then the clear pores get merged. Similarly, it is not possible to detect the pores after edge detection (figures 2f and 2g) because edge detector fails to clearly delimit all the pores. Moreover, the algorithm yielded double edges, due to the characteristics of the SEM images. Figure 3 illustrates the ability of our algorithm to detect intersecting pores.

This algorithm can also be used to detect circle-like objects, like equilateral triangles, squares, pentagons and hexagons (although it was not primarily designed for this task). After detecting the objects, other techniques can be used to classify their shapes. However, this algorithm is not appropriate to detect, for example, ellipse or rectangle with high eccentricity. In these cases, this algorithm will likely detect ellipse or rectangle as two or more circles.

### ***Correlation-Based Granulometry with Multi-Shaped Kernels***

In this subsection, we extend our circular-kernel granulometry to detect also non-circular grains (besides the circular ones). As examples, figures 4, 5, 6 and 7 depict porous silicon materials with circular and rectangular pores. We want to count each pore together with its shape, eccentricity, size and orientation. A rotation and scale-invariant template matching would be a perfect technique to solve this problem [11-12, 16]. Unfortunately, these techniques are not adequate to find very simple and small shapes that may appear faintly in the image. In granulometry, we are highly concerned with the accuracy, and not deeply concerned with the computer performance. There is no problem if the computer takes, say, one minute to analyze an image. So, we propose to use the most simple but highly accurate idea: compute the correlations with kernels that represent all the shapes, eccentricities, orientations and sizes (figure 1). Even using this simple idea, the extension of the granulometry to non-circular kernels is not as straightforward as it may seem at the first sight, because there are some subtleties.

The first subtlety is that the adequate correlation threshold  $\tau$  may be different for different shapes. Indeed, we had to use two different threshold values to process the image

in figure 4: a small value for the circles (to detect faint circles) and a large value for the rectangles (to not detect false faint rectangles, because all rectangle pores have high contrast). This leads to the necessity of making an unsuspected modification in the algorithm. The second subtlety is how to compute the intersection between different shapes. In the last subsection, we estimated the intersection between circles using the distance between their centers. Clearly, this approach cannot be used for non-circular shapes and this leads to another unsuspected difficulty.

Let us say that our application has  $m$  shapes with  $m$  different correlation threshold values  $\tau^i$ ,  $1 \leq i \leq m$ . Our granulometry computes cross-correlations of  $A$  with all kernels  $T_j^i$  with shape  $i$  and index  $j$  within the shape. Let us denote the images resulting from the cross correlation between  $A$  and  $T_j^i$  as  $C_j^i$ :

$$C_j^i(x, y) = T_j^i(x, y) \circ A(x, y), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n^i. \quad (9)$$

Let us define the maximal correlation image as:

$$C(x, y) = \underset{i,j}{MAX} [C_j^i(x, y)]. \quad (10)$$

Surprisingly, we miss many objects if we detect the peaks in the maximal correlation image  $C$ . Figure 4a depicts the pores detected using this strategy. Compare it with figure 4b obtained using our ultimate algorithm with the same parameters. Some important circles are missing in figure 4a because we are using a small correlation threshold  $\tau_c$  for the circles and a large threshold  $\tau_r$  for the rectangles. It may happen that a false rectangle with correlation  $c_r$  is located in the neighborhood of a faint real circle with correlation  $c_c$ , with  $c_r > c_c$ . Suppose also that  $c_r < \tau_r$  and  $c_c > \tau_c$ . In this case, the rectangle must be discarded and the circle must be detected. However, the circle will not be detected as a peak because there is another higher peak in its neighborhood. To avoid this problem, we suggest computing a maximal correlation image for each shape:

$$C^i(x, y) = \underset{j}{MAX} [C_j^i(x, y)], \quad 1 \leq i \leq m. \quad (11)$$

A pixel  $(x, y)$  is a peak in  $C^i$  if it is greater than or equal to its eight spatial neighbor pixels. Each peak  $C^i(x, y)$  corresponds to an object/grain of shape  $i$  in  $A$ . The scale, eccentricity and the orientation of the detected object is given by the kernel with shape  $i$  and index  $G^i(x, y)$ :

$$G^i(x, y) = \underset{j}{ARGMAX} [C_j^i(x, y)], \quad 1 \leq i \leq m. \quad (12)$$

The set  $V$  of all peaks in  $A$  is the union of peaks detected for each shape.

As before, this process detects many non-existing objects together with the real ones. It is very easy to filter out faint objects in  $V$  characterized by low peaks with correlation smaller than  $\tau^i$ . However, it is not so easy to filter out objects in  $V$  that have large intersections with other objects. In multi-shape granulometry where the sizes of objects can vary greatly a simple definition like equation (8) cannot be used. We had to use a more elaborated definition. Our algorithm discards from  $V$  an object  $O_j^i$  with shape  $i$  and correlation  $c_j^i$  in two cases:

(1) If the set of all the objects in  $V$  with correlations greater than  $c_j^i$  hides more than  $\gamma^i$  of the area of  $O_j^i$ :

$$\text{area}[O_j^i \cap W(c_j^i)] > \gamma^i \text{area}[O_j^i] \quad (13)$$

where  $W(c_j^i)$  is the set of all objects in  $V$  with correlation greater than  $c_j^i$ .

(2) If there is an object  $O_i^k$  in  $V$  with shape  $k$  and correlation  $c_i^k$  such that:

$$\begin{cases} c_j^i < c_i^k \\ \text{area}[O_j^i \cap O_i^k] > \gamma^k \text{area}[O_i^k] \end{cases} \quad (14)$$

Computationally, we measure the area of an object by drawing it in a temporary image and counting the number of pixels. Similarly, we count the intersection between two or more objects by drawing them in a temporary image and counting the number of intersecting pixels. For the sake of simplicity, we will not describe the function that filters the peaks in  $V$  in a pseudo computer language. We only note that the key structure for the fast implementation of equations (13) and (14) is a matrix  $I$  (of the same size as the original image  $A$ ) where each element  $I(x, y)$  is a list of objects of  $V$  that intersect that pixel.

### ***Pyramid Structure***

Some very small kernels may not have a precisely defined shape: it may be impossible to say if a very small kernel is a circle or a rotated square. For this reason, we assume that all small pores are circular when analyzing porous silicon SEM images. On the other hand, medium-sized pores have definite shapes. The user must be aware that very small grains/objects can be correctly detected but their shapes may be undetermined.

For finding large objects, it is computationally more efficient to reduce the size of both kernel and image  $A$  and to find the reduced kernel in the reduced  $A$ . However, the size reduction should not be too much to generate imprecise or ambiguous kernels.

For finding objects/grains that can greatly vary its size, we use pyramid structure. We construct a pyramid for image  $A$ , where the basis (or the 0-th floor) is the original image  $A$  with scale 1, the first floor is the image  $A$  at scale 0.5 and  $n$ -th floor is the image  $A$  at scale  $2^{-n}$ . If we find, say, a circle with diameter  $d$  at the first floor of the pyramid (scale 0.5), actually there is a circle with diameter  $2d$  in the basis.

### ***Implementations***

We wrote two C++ programs named *Granul* (for circular shapes only granulometry) and *MGranul* (for granulometry with multiple shapes). The two programs with executable and source files are available in our web site<sup>1</sup>.

Program *Granul* contains the proposed granulometry for the circular shapes and some ad hoc adaptations to analyze images of macroporous silicon layers with circular and rectangular pores. This program consists of four sub-programs:

---

<sup>1</sup> [www.lps.usp.br/~hae/granul](http://www.lps.usp.br/~hae/granul)

1. Corrcirc: Implements the proposed correlation-based granulometry to detect circular pores. This program also detects correctly the square pores, even it was not especially programmed for this task.
2. Classify: Classifies automatically the detected pores in circle or square, based on the size of the pore and on the “influence zone”, that is, the pore’s neighbor area where there is no other pore.
3. Edit: Allows to rectify manually eventual errors made by the automatic detection and classification.
4. Relat: Computes the area of square pores by the seed growing algorithm and generates the pattern spectrum.

Program *MGranul* is the implementation of the proposed multi-shape granulometry. It contains many sub-programs, where the most important are:

1. Correla: Computes the local maxima of the correlations with multi-shaped kernels. This program takes about 50s to process a 710x420 image with 600 kernels in a 2GHz computer. Until now, we have implemented circle, ellipsis, square and rectangle, but other shapes can be easily added.
2. Filtra: Filters the local maxima computed by the previous sub-program using the parameters  $\tau^i$  and  $\gamma^i$  defined by the user. This program takes less than 1s, so the user can try different parameters and verify if all pores were correctly detected, without waiting too long.

## **ANALYZING POROUS SILICON WITH THE PROPOSED ALGORITHM**

### ***Porous Silicon Fabrication***

The macroporous silicon layers [17] with circular and square pores were obtained by electrochemical anodization process of (100) p-type silicon wafer with resistivity of c.a. 10  $\Omega\text{cm}$ . The anodization process was carried out in the single electrochemical cell using HF (48%): DMF mixture where HF corresponds to 12% of total volume of the solution. The cell was anodically polarized at galvanostatic condition fixing the current density at 12  $\text{mA}/\text{cm}^2$  for 20 minutes [18]. The porous silicon structure depends strongly on the previous metallization process of silicon wafer with Al metal. So, we obtained three sets of porous silicon samples. The samples of the first set were obtained after silicon wafer backside metallization and annealing at different temperatures (Fig. 5). The samples of the second set were obtained after backside and front side metallization and annealing at different temperatures (Fig. 6). The samples of the third set were obtained after front side metallization and annealing at different temperatures (Fig. 7).

After Al metallization, all annealing processes were made in  $\text{N}_2$  environment. The porous silicon obtained metalizing the front side before the anodization process present circular and square pores (Figs. 6 and 7). The density of square pores is important for silicon macrotubes formation. So, in the present work we analyze the images obtained by Scanning Electron Microscopy (SEM) in order to control the experimental parameters for adequate macroporous silicon layer formation. The images were obtained using NanoSem 400 microscopy and electron beams from 10 to 30 kV with the secondary electron imaging mode.

### ***Image Processing Results***

The SEM images of macroporous silicon layers annealed for different times are depicted in Figures 5 to 7. These images reveal that the distributions of square and circular pores depend on the thermal annealing process. We applied the proposed algorithm to obtain the quantitative data of the pores structural features.

Figure 8 depicts the histogram of circular pores of samples that were obtained with and without Al metallization at back-side surface of silicon wafer. The sample obtained without metallization shows pores with area between  $0.7$  and  $1.5 \mu\text{m}^2$  and the area of the predominant pores is  $1 \mu\text{m}^2$ . Samples with Al metalized back-side and annealed for  $0.5$  or  $1.5$  hours show pores with area between  $0.7$  and  $1.0 \mu\text{m}^2$  and the area of the predominant pores is  $0.7 \mu\text{m}^2$ . The pores in the sample without metallization have the largest area, so the porosity of this sample is the highest and the pores of this sample are the shallowest. In other words, samples that have pores with smaller areas have deeper pores. These results suggest that the depth propagation rate of the pores is higher in samples with Al metal film than in sample without metallization. Since the etching rate is proportional to the anodization current density, the back-side metallization of silicon wafer promote a local current density enhancement due to local electric field enhancement. This local field enhancement could be due to Al pitches formation [19-21] at silicon back-side region used for electrical contact in the electrochemical process.

The other type of porous silicon structures were obtained after metallization of both front and back sides with Al (Figure 6). In these samples, the porous structures present circular and square pores. The size distribution and depth of pores depend on the thermal annealing process of metalized samples before anodization process. Figures 9(a) and (b) depict the histogram of circular and square pores distribution respectively as a function of annealing time. These Figures show that circular pores density decreases and square pores density increases as the annealing time increases. In these samples, the average area of circular pores is between  $0.3$  and  $0.9 \mu\text{m}^2$  and the average area of square pores is between  $0.9$  and  $1.7 \mu\text{m}^2$ . As the annealing time increases, square pores with area of  $0.9 \mu\text{m}^2$  become predominant.

The silicon microtubes (not shown in this work) are obtained from the porous silicon structure. The appropriate structures for this propose is one that has high density of both circular and square pores. The samples metalized in both surfaces present high square pores density but low circular pores density. The sample obtained without metallization shows many circular pores but no square pore (Figure 5). In order to obtain high density of both circular and square pores, we metalized only the front surface leaving the back side free for a direct external contact (Figure 7). Figure 10(a) depicts the histogram of circular pores of these samples. The circular pores distributions are similar for all samples and the area of the predominant pores is  $0.48 \mu\text{m}^2$ . Figure 10(b) shows the histograms of square pores. The samples annealed for one or two hours present low square pores density with similar area distribution ranging from  $0.48$  to  $12.86 \mu\text{m}^2$ . The sample annealed for 4 hours presents the highest square pores density and the area of the predominant square pores is  $3.72 \mu\text{m}^2$ . The square pore density decreases significantly in the sample annealed for 8 hours and the area of the predominant pores is  $9.38 \mu\text{m}^2$ .

These quantitative results give us information to understand the mechanism of the square pores formation. The square pores are important for the silicon macrotubes formation. The annealing time of Al metalized sample has increased the sheet resistance of silicon wafer (Figure 11) and this is well-correlated with the square pores density increasing of samples annealed for 1 to 4 hours. However, in the sample annealed for 8 hours, in

spite of the increased sheet resistance, the square pores density decreased. These results suggest that the sheet resistance of silicon surface defines the initial pitch distribution during anodization. Since the pitch formation depends on the electric field intensity, the field distribution at initial anodization stage depends on the sheet resistance. So, the initial pitches are originated from the breakdown-like process. In this sense, there is a critical annealing time when the square pores density begins to decrease. Since this critical field has high intensity, the area of pores increases consequently. Additionally, the long annealing time may have induced Al metal to cluster preferentially at some regions where the initial pitches have preferentially formed.

The SEM images analysis by the proposed granulometry algorithm yielded quantitative information of the pores size and their distribution in circular and square pores. These results were important to understand the formation mechanism of square macropores.

## CONCLUSIONS

The new correlation-based granulometry demonstrated to be an effective tool for SEM images processing, making it possible to analyze quantitatively the porous materials parameters such as pores radii, the size distribution and pores density. Additionally, the proposed granulometry allowed us to classify the pores as circular or square ones. The developed algorithm can contribute in the porous material research area, allowing rapid analysis of structures and rapid extraction of their geometric parameters. The proposed algorithm showed to be an excellent tool for discussing the square pores formation mechanism in silicon substrate by electrochemical anodization process. The results obtained indicate that the annealing time of the Al metalized silicon wafer before anodization process is an important parameter for controlling the porous density and their sizes. The non-metalized silicon wafer presented high circular pores density. When the porous silicon is obtained from surface metalized on both sides, the circular pores density decreases and the square pores density increases when the annealing time increases. The porous silicon structure with high density of both circular and square pores could be obtained using silicon wafer metalized only in the front side. This kind of porous structure has shown to be an excellent precursor substrate for silicon macrotubes formation.

## ACKNOWLEDGEMENTS

The authors thank to CNPq for the financial support.

## References

- [1] C. P. Bean, I. S. Jacobs, Magnetic Granulometry and Super-Paramagnetism. *J. Appl. Phys.* 27(12), 1448-1452 (1956).
- [2] N. Azema, M. F. Pouet, C. Berho, O. Thomas, Wastewater suspended solids study by optical methods. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 204, 131–140 (2002).
- [3] G. Matheron, *Random Sets and Integral Equation*. (Wiley, New York, 1978).
- [4] E. R. Dougherty, J. T. Newell and J. B. Pelz, Morphological Texture-Based Maximum-Likelihood Pixel Classification Based on Local Granulometric Moments. *Pattern Recognition*, 25(10), 1181-1198 (1992).
- [5] D. S. Raimundo, P. B. Calipe; D. R. Huanca, W. J. Salcedo, Anodic Porous Alumina Structural Characteristics Study Based on SEM Image Processing and Analysis. *Microelectronics Journal*, 40, 844-847 (2009).
- [6] L. Vincent, *Fast Grayscale Granulometry Algorithms*. Proc. Int. Symp. Mathematical Morphology, Fontainebleau (1994).
- [7] Mathworks Image Processing Toolbox 6.4 demo “Granulometry of Snowflakes”, accessed on 2009.  
<http://www.mathworks.com/products/image/demos.html?file=/products/demos/shipping/images/ipexsnow.html>
- [8] N. H. Maerz, T. C. Palangio, J. A. Franklin, WipFrag image based granulometry system. Proc. of the FRAGBLAST, 5 Workshop on Measurement of Blast Fragmentation, Montreal, 91-99 (1996).
- [9] R. C. Gonzalez, R. E. Woods, *Digital Image Processing, Second Edition*. Prentice-Hall, 2002.
- [10] J. P. Lewis, Fast normalized cross-correlation. *Vision Interface*, 120-123 (1995), url = [citeseer.ist.psu.edu/lewis95fast.html](http://citeseer.ist.psu.edu/lewis95fast.html).
- [11] S. A. Araújo, H. Y. Kim, Ciratefi: An RST-Invariant Template Matching with Extension to Color Images. *Integrated Computer-Aided Engineering*, 18(1), 75-90, (2011).
- [12] H. Y. Kim, Rotation-Discriminating Template Matching Based on Fourier Coefficients of Radial Projections with Robustness to Scaling and Partial Occlusion. *Pattern Recognition*, 43(3), 859-872 (2010).
- [13] A. Elmoutaouakkil, L. Salvo, E. Maire, G. Peix, 2D and 3D Characterization of Metal Foams Using X-ray Tomography. *Advanced Engineering Materials*, 4(10), 803-807 (2002).
- [14] E. Maire, N. Gimenez, V. S. Moynot, H. Sautereau, X-ray tomography and three-dimensional image analysis of epoxy-glass syntactic foams. *Phil. Trans. R. Soc. A*, 364, 69-88 (2006).

- [15] E. Maire, P. Colombo, J. Adrien, L. Babout, L. Biasetto, Characterization of the morphology of cellular ceramics by 3D image processing of X-ray tomography. *Journal of the European Ceramic Society*, 27(4), 1973-1981 (2007).
- [16] G. Takacs, V. Chandrasekhar, S. Tsai, D. Chen, R. Grzeszczuk, B. Girod, Unified Real-Time Tracking and Recognition with Rotation-Invariant Fast Features. *IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR)*, 934-941 (2010).
- [17] H. Föll, M. Christophersen, J. Carstensen, G. Hasse, Formation and Application of Porous Silicon. *Materials Science and Engineering: R: Reports*, 39(4), 93-141 (2002).
- [18] D. R. Huanca, J. Ramirez-Fernandez, W. J. Salcedo, Morphological and Structural Effect of Aluminum on Macroporous Silicon Layer. *J. Mat. Sci and Eng.* 4(8), 55-59 (2010).
- [19] S. K. Ghandi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*. Wiley-Interscience Publication, New York, USA, 1983.
- [20] S. Wolf, *Silicon Processing for the VLSI Era. Vol II*, Lattice Press, California, USA, 1990.
- [21] G. L. Schnable, R. S. Keen, Aluminum metallization-advantages and limitations for integrated circuit applications. *Proceedings of the IEEE* 57, 1570-1580 (1969).

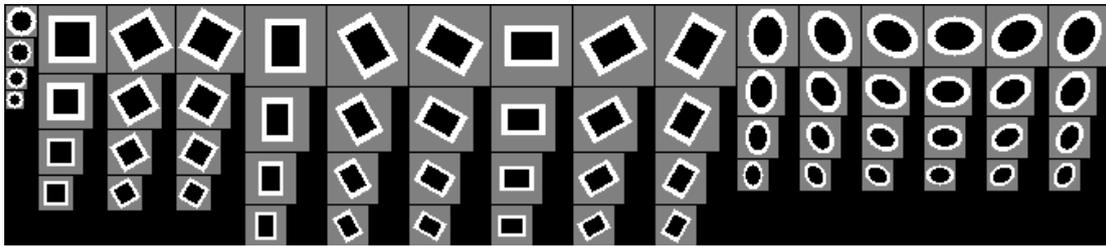


Fig. 1: Some of the kernels used in our granulometry. Black pixels have negative values, white ones are positive, and gray ones are zeros. Correlation-based granulometry with circular kernels use only the kernels in the leftmost column. Multi-shape correlation-based granulometry uses all the kernels.

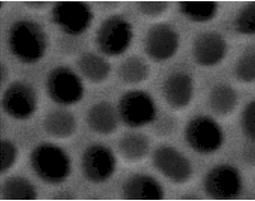
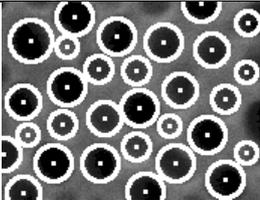
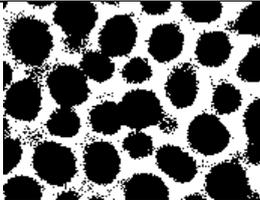
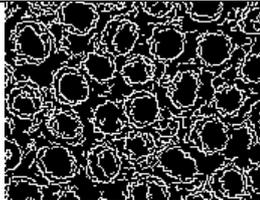
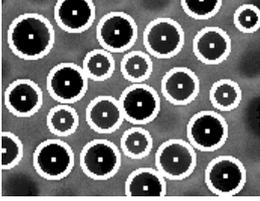
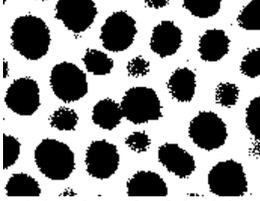
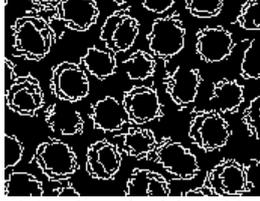
Image of porous material	Correlation-based granulometry	Thresholding	Canny edge detector
 (a) A portion of the original SEM image.	 (b) $\tau=0.01$ .	 (d) Threshold=70.	 (f) Threshold=100.
	 (c) $\tau=0.15$ .	 (e) Threshold=90.	 (g) Threshold=140.

Fig. 2: (a) A sub-image of a porous material image with 21 pores located entirely within the image (18 clear pores and 3 faint pores). (b-c) The output of our algorithm detecting all 21 pores or only 18 clear pores depending on the threshold value. (d-e) Thresholding with different parameters (thresholding is usually the first step of the binary morphology-based granulometry). (f-g) Canny edge detector (edge detection is usually the first step of the edge-base granulometry).

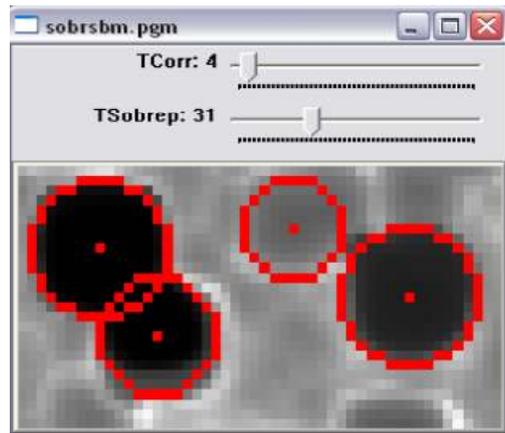


Fig. 3: Our correlation-based granulometry can detect even intersecting grains.

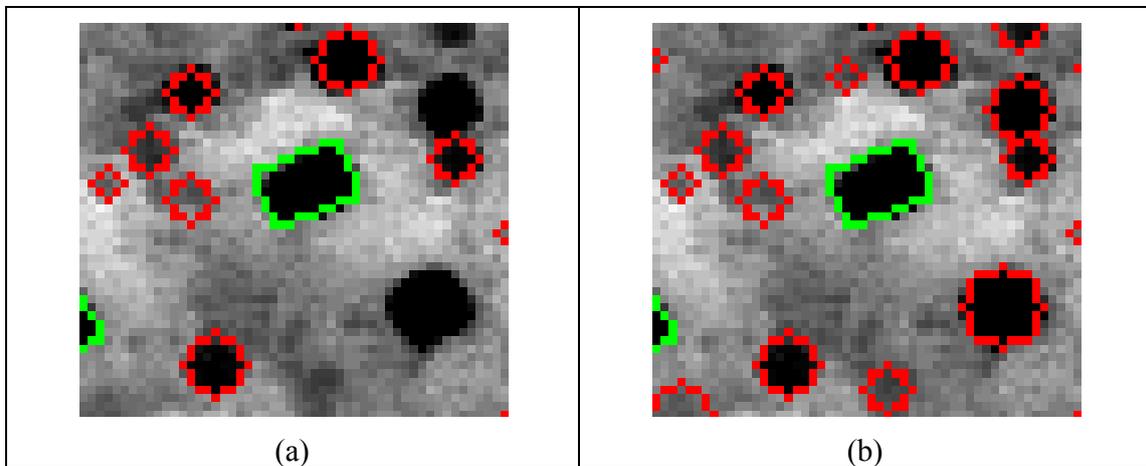


Fig. 4: A porous silicon material with circular and rectangular pores. (a) The multi-shape correlation algorithm misses some pores if it detects peaks in the maximal correlation image  $C$ . (b) If the algorithm detects peaks separately for each shape, it does not miss pores.

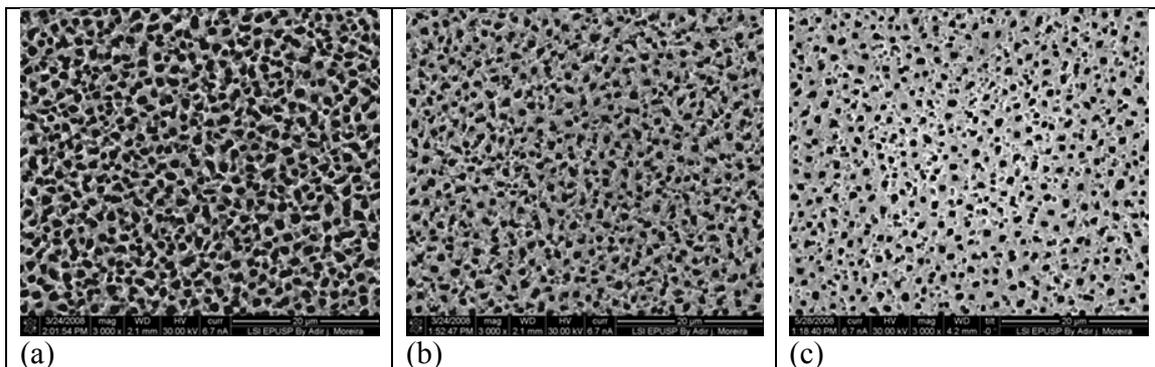


Fig. 5: The SEM images of porous silicon structures obtained from: (a) non-metalized silicon wafer, (b) back-side metalized annealed for 0.5 hours and (c) back-side metalized annealed for 1.5 hours.

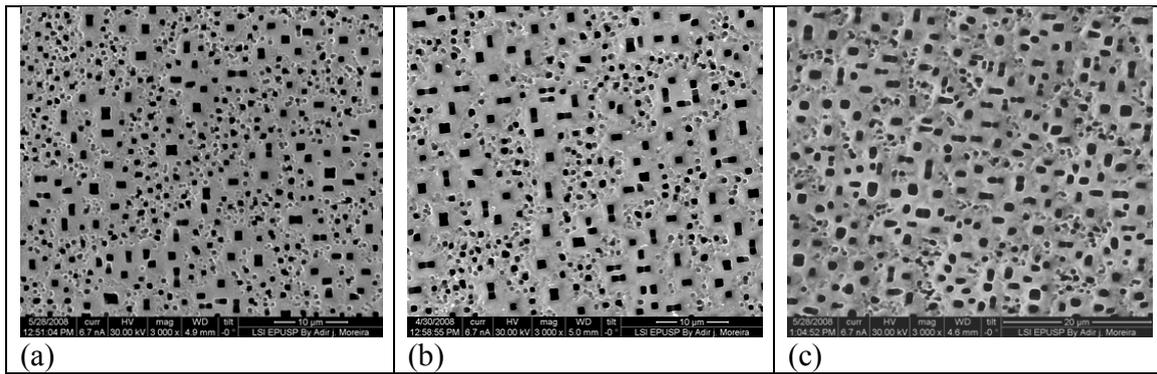


Fig. 6: The SEM images of porous silicon structures obtained from the silicon wafer that was metalized in both sides and annealed for (a) 0.5, (b) 1.5, and (c) 3.0 hours.

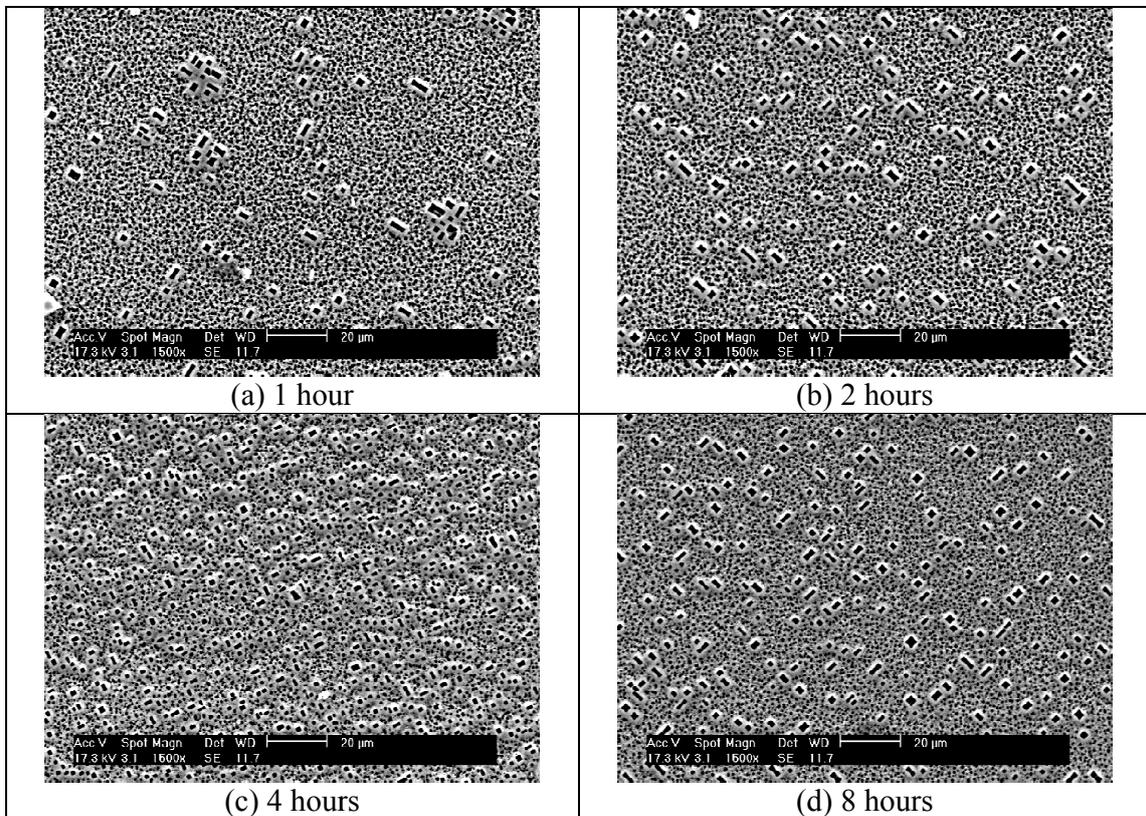


Fig. 7: The SEM images of porous silicon structures obtained from the silicon wafer metalized only on the front side and after thermal annealing for (a) 1, (b) 2, (c) 4 and (d) 8 hours.

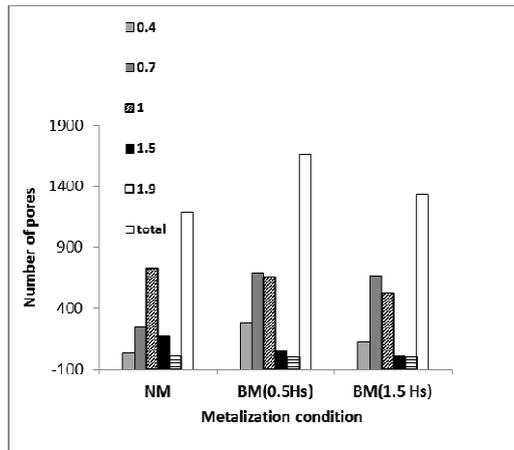


Fig. 8: The histograms of circular pores obtained from the SEM images of Figure 5: non-metalized silicon wafer (NM) and back-side metalized (BM) annealed for 0.5 and 1.5 hours. The inner labels describe the pores' areas (in  $\mu\text{m}^2$ ) corresponding to each bar of the histograms. The white bars represent the total numbers of circular pores.

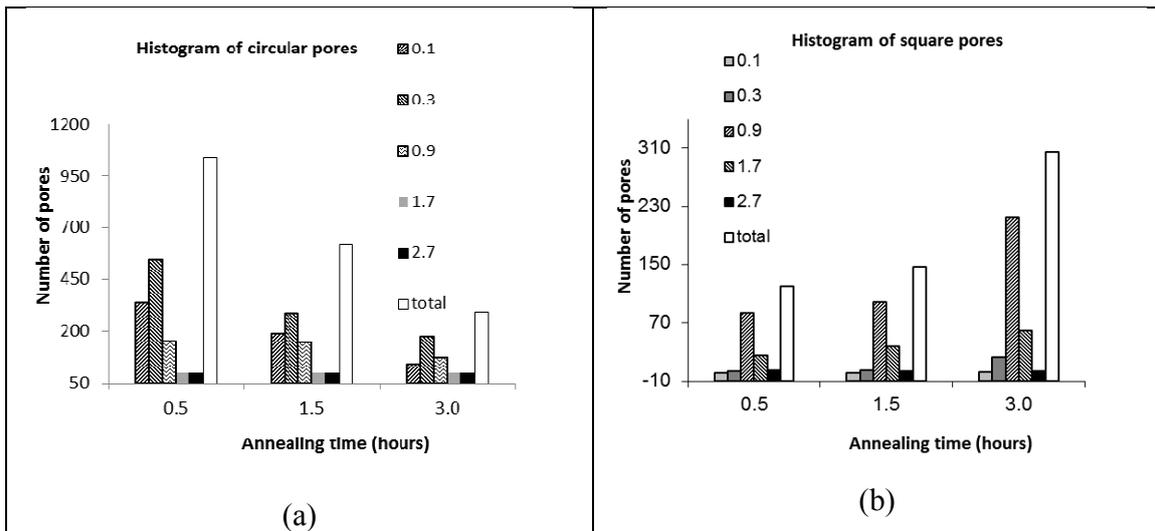


Fig. 9: The histograms of circular (a) and square (b) pores obtained from the SEM images of Figure 6 (silicon wafer that was metalized both front and back sides) annealed for different times. The inner labels describe the pores' areas (in  $\mu\text{m}^2$ ) corresponding to each bar of the histograms. The white bars represent the total numbers of circular or square pores.

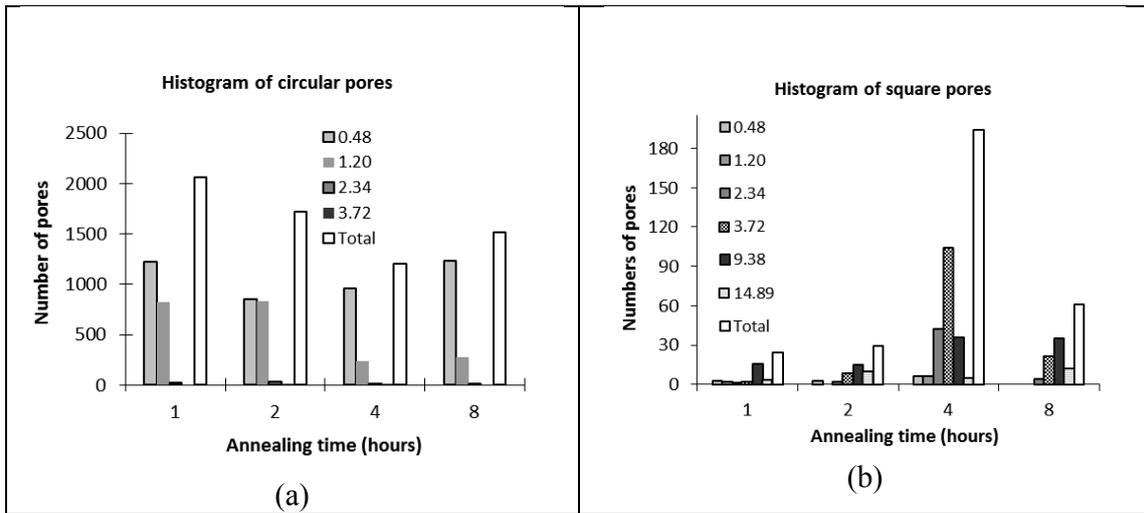


Fig. 10: The histograms of circular (a) and square (b) pores obtained from the SEM images of Figure 7 (metallized only on the front side) and annealed for different times. The inner labels describe the pores' areas (in  $\mu\text{m}^2$ ) corresponding to each bar of the histograms. The white bars represent the total numbers of circular or square pores.

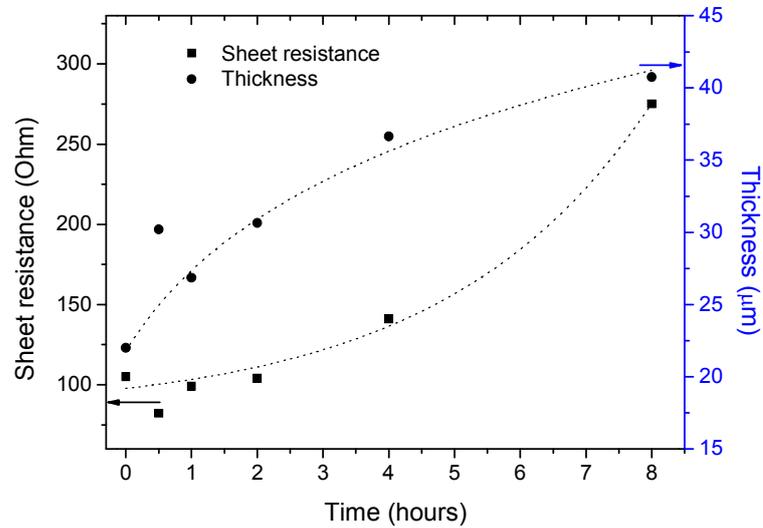


Fig. 11: The sheet resistance of the silicon wafer at the front surface as a function of annealing time. The graph depicts also the thickness of porous silicon layers after anodization process.