Computational Materials Science 148 (2018) 336–342

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Material phase classification by means of Support Vector Machines

Jaime Ortegon [a,*], Rene Ledesma-Alonso [b,c], Romeli Barbosa [a], Javier Vázquez Castillo [a], Alejandro Castillo Atoche [d]

[a] Universidad de Quintana Roo, Boulevar Bahía s/n, Chetumal 77019, Quintana Roo, Mexico
[b] CONACYT-Universidad de Quintana Roo, Boulevar Bahía s/n, Chetumal 77019, Quintana Roo, Mexico
[c] Universidad de las Aḿericas Puebla, Sta. Catarina Mártir, San Andrés Cholula, 72810 Puebla, Mexico
[d] Universidad Autónoma de Yucatán, Av. Industrias no contaminantes s/n, Mérida 150, Yucatán, Mexico

ARTICLE INFO

Article history:
Received 30 November 2017
Received in revised form 23 February 2018
Accepted 25 February 2018
Available online 23 March 2018

Keywords:
Phase classification
Support Vector Machines
Random heterogeneous materials

ABSTRACT

The pixel's classification of images obtained from random heterogeneous materials (RHM) is a relevant step for 3D stochastic reconstruction and to compute their physical properties, like Effective Transport Coefficients (ETC). A bad classification will impact on the computed properties. However, the literature on the topic discusses mainly the correlation functions or the properties formulae, giving little or no attention to the classification; authors mention either the use of a threshold or, in few cases, the use of Otsu's method. This paper presents a classification approach based on Support Vector Machines (SVM) and a comparison with the Otsu-based approach, based on accuracy, precision and recall. The data used for the SVM training are the key for a better classification; these data are the grayscale value, the magnitude and direction of pixels gradient. For the validation cases, the recall of the solid phase is significantly better, whilst improving the accuracy for the SVM method. Finally, a discussion about the impact on the correlation functions is presented in order to show the benefits of the proposal.

1. Introduction

The study of random heterogeneous materials is common in the literature, and it is useful to determine its Effective Transport Coefficients (ETC), electromagnetic and mechanical properties. Digital images are used to perform a characterization of materials based on morphological properties. Coker and Torquato [1] established a methodology to compute correlation functions with the purpose of describing the morphology of a digitized material. Recent works are devoted to reconstruct three-dimensional materials based on the correlation functions using different methods [2–7].

The authors' interests are the reconstruction and determination of ETC of random heterogeneous materials (RHM) on fuel cells, particularly, porous carbon electrodes. In synthesis, manufacturing and processing of electrodes in experimental studies, Scanning Electron Microscopy (SEM) is useful for detailed study of a surface. SEM use a high-energy electron beam, which scans across the surface in order to form an image [8]. A SEM image is a map of the signal produced by the interaction of a sample and the electron beam. In this sense, the most important used signal are backscattered electrons (BSE) and secondary electrons (SE). SE is the best signal that characterizes the microstructure morphology [9]. The intensity of the SE signal depends on several parameters, including the height above the surface in the intrinsic topography of the sample and the atomic number of the sample. Samples with high atomic numbers allow the beam to penetrate the sample with less depth, in such a way that the signal is more intense, and as a result, a brighter SEM image. Other effects such as charge accumulation in non-conductive materials are also important in RHMs.

A common practice to enhance SEM images is the use of smoothing and denoising methods before converting gray-scale to binary images. However, most works use a threshold for the binarization. The threshold may be fixed [6,10], automatically selected with Otsu's method [2,3,5], Maximum Likelihood [7] or Maximum Entropy [11]. Recently, Sabharwal et al. [12] use Sauvola's method to enhance the binarization [13]; however, they barely comment that there is a insignificant difference with the results obtained with the Otsu's method. Likewise, Joos et al. [5] used Otsu's method and compared its performance with that of a mean-algorithm and "intuitive" inspection, exhibiting the importance of image processing and binarization.

Support Vector Machines (SVM) are applied to various pattern classification problems [14]. SVM are based on hyperplane classifiers [15] that seek for a separating hyperplane using Statistical Learning Theory. There are few examples of the employment of

* Corresponding author.
E-mail address: jortegon@uqroo.edu.mx (J. Ortegon).

https://doi.org/10.1016/j.commatsci.2018.02.054
0927-0256/

SVMs in materials science. Sundararaghavan and Zabaras [16,17] used SVM to classify 3D microstructures obtained computationally using a Monte-Carlo Potts grain growth model. Likewise, SVM are useful for color image segmentation [18,19], where the authors use color and texture features as elements for the pixel classification (segmentation); those papers take advantage of SVM's capacity to consider data from different contexts. Li et al. [20] use a SVM for the segmentation of white blood cells microscopic images. Recently, Ortegon et al. [21] have presented the idea of using SVM to binarize electrode images; however, neither the presented procedure was validated nor the effects on the correlation functions were studied.

Next Section presents the image processing algorithms used. Section 3 gives a brief on SVM, while the classification procedure is presented on Section 4. In order to observe the effects on the correlation functions, they are described in Sections 5 and 6 is devoted to the corresponding results and discussion. Finally, conclusions are given in Section 7.

2. Image processing

In the present work SEM images are considered as the source of information about the structure of RHM. Particularly, images of the Catalytic Layer (CL) of Proton Exchange Membrane Fuel Cells (PEMFC) are studied. The pixels in each of those images are classified to generate a representation of the configuration of the material phases, from which some microstructure statistical descriptors, i.e., low order correlation functions, are computed. It is illustrated with the sample image shown in Fig. 1.

The images of two-phases RHM are commonly separated in two classes. This represents a binarization of the pixels, where black and white pixels represent void and solid phases of a porous material. It is important to remark, due to SEM imaging process, it is possible that the same gray value corresponds to distinct parts of the material at different depths from the microscope sensor. This situation has been presented in Fig. 2, where pixels at different locations, either in the void phase or in the solid phase, display the same gray value.

It is a common practice to enhance the image with different methods like histogram equalization, normalization and denoising [22]. Equalization is the process to distribute channel values of an image to the whole pixel's domain range, by spreading values and increasing the image contrast. Images with high contrast shows a great distribution of gray-level detail. Since different image formats may use different ranges for pixel's values, a normalization is proposed, in order to convert integers to floating point numbers between 0 and 1, which simplifies the classification. Finally, denoising will remove Gaussian and impulse noise with a Weiner [23] and median filter [22] respectively. Original and processed images, obtained with the aforementioned equalization-normaliza tion-denoising method, are shown in Fig. 3.

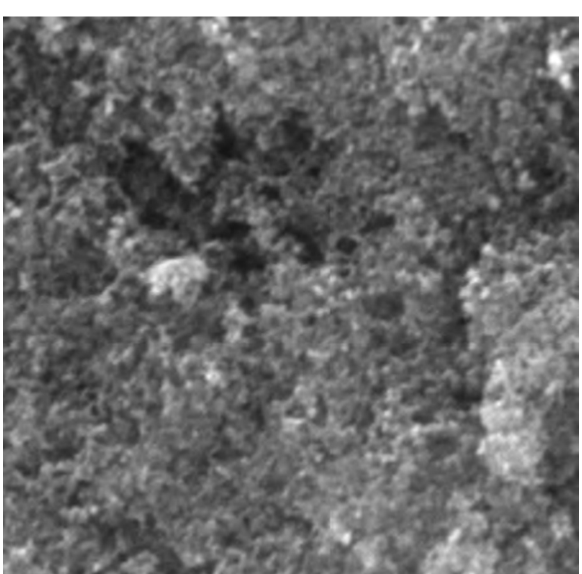


Fig. 1. SEM image of the catalytic layer of a PEMFC.

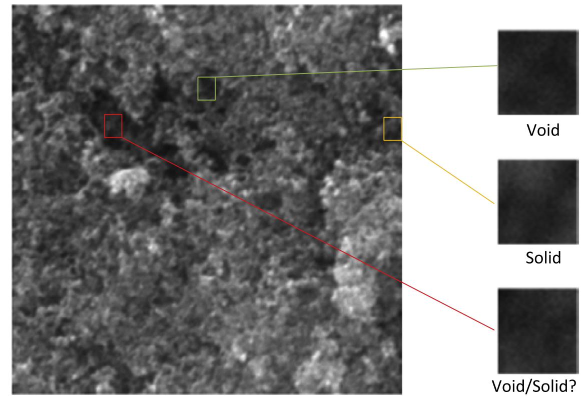
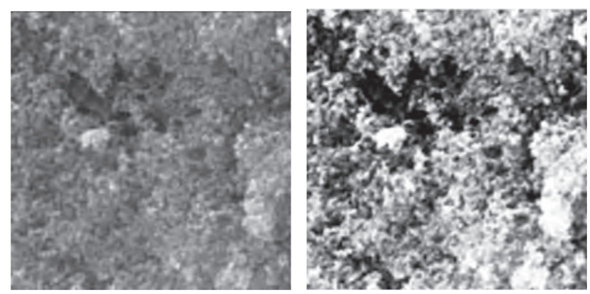


Fig. 2. Close-up to regions with the same gray value corresponding to a different phase of the material.



(a) Original (b) Enhanced

Fig. 3. Catalytic layer of a PEMFC obtained by SEM. (a) Original image and (b) enhanced image by the equalization-normalization-denoising method.

For binarization, there are two useful morphological operators: erosion and dilation [22]. The erosion's goal is to shrink regions of foreground (white) pixels and enlarging holes within them by eroding its boundaries. On the other hand, the dilation enlarges regions of foreground pixels and contracts holes within them by dilating its boundaries. These morphological operators use a structuring element, usually a small pixel mask. To apply the operator, consider each of the pixels of the input image and superimpose the structuring element on top of it. For an erosion, if every pixel in the structuring element coincides with a foreground pixel in the image, then the input pixel is left as foreground, otherwise it is set as background (black). Dilation is the dual of erosion, i.e., dilating foreground pixels is equivalent to eroding background pixels.

3. Support Vector Machines

Support Vector Machines (SVM) are based on hyperplane classifiers, and they are frequently applied to classify samples in different categories called classes. Fig. 4 shows a 2D schematic representation of main elements for an SVM, where the hyperplane is the line separating squares from circles; margin is the maximal