Juan Pavón
Néstor D. Duque-Méndez
Rubén Fuentes-Fernández (Eds.)

LNAI 7637

Advances in Artificial Intelligence – IBERAMIA 2012

13th Ibero-American Conference on AI
Cartagena de Indias, Colombia, November 2012
Proceedings

Springer

Using SOM Maps for Clustering and Visualization of Diamond Films Deposited by HFCVD Process

Leandro A. Pasa[1], José Alfredo F. Costa[2],
Marcelo C. Tosin[3], and Fábio A. Procópio de Paiva[4]

[1] UTFPR, Medianeira/PR, Brasil
pasa@utfpr.edu.br
[2] UFRN, Departamento de Engenharia Elétrica, Natal/RN, Brasil
jafcosta@gmail.com
[3] UEL, Departamento de Engenharia Elétrica, Londrina/PR, Brasil
tosin@uel.br
[4] IFRN, Campus Zona Norte, Natal/RN, Brasil
fabio.procopio@ifrn.edu.br

Abstract. Diamond is a material with unique properties to be exploited in various applications. The deposition of diamond thin films on surfaces enables its use in mechanics, electronics and optics, among others. To ensure the quality of crystals is important to control the parameters involved in the deposition process. In this study we used the SOM algorithm for clustering and visualization of the parameters of a reactor for deposition of diamond thin films, which together with scanning electron microscopy and Raman spectroscopy, collaborated in reactor calibration. The results show the importance of temperature in this process.

Keywords: SOM, hot filament chemical vapour deposition, diamond thin films.

1 Introduction

The interest in synthesizing diamonds is due to its physical and chemical stable properties, enabling unique technological applications in different fields like electronics, mechanics and optics. There are basically two methods for synthesizing diamonds. The first one consists in submitting graphite to high pressure and high temperature, reaching the region where the diamond is the stable phase. This method is known as HPHT - High Pressure and High Temperature. The second method is the synthesis of diamond at low pressures in the metastable region. Among these low pressure processes are the CVD - Chemical Vapor Deposition. It is a process that involves depositing a solid material on a substrate by activating the precursors in gaseous phase and making them react chemically. The hydrocarbon is dissociated into atomic hydrogen and active carbon. The carbon is deposited on a substrate in the tetrahedral form (diamond). This is a slow process because the atoms are deposited one by one to form a solid film adhered to the substrate.

J. Pavón et al. (Eds.): IBERAMIA 2012, LNAI 7637, pp. 140–148, 2012.

The major advantage of this kind of growth processes is that the diamond can be deposited as thin film on substrates of different geometries and thicknesses, allowing its use in components for use in several areas.

Carbon, in its pure form, appears in several allotropic, crystalline and amorphous forms. Two types of crystalline carbon are diamond and graphite, although there are other elements whose links form a crystal lattice. Among the amorphous forms, we may cite carbon fiber, amorphous carbon, coal and DLC - Diamond Like Carbon, which is an amorphous form and maintains some of the characteristics of diamond [1]. Besides these, there are several other elements formed from carbon atoms. For this reason, it is necessary to maintain tight control over the deposition parameters, to ensure that there is, mainly, the deposition of diamond crystals.

In this study, the Self Organizing Maps (SOM) algorithm was used for clustering and visualization of the parameters involved in a HFCVD process - Hot Filament Chemical Vapor Deposition, detailed in the next section. The parameters of each film deposition were purposely varied to calibrate the reactor to reach the best depositions.

The paper is organized as follows: Section 2 presents the key concepts about the HFCVD process; Sect. 3 describes a brief introduction to SOM maps; Sect. 4 shows the methodology used in the films deposition; Sect. 5 shows the details about the clustering using the SOM. And in the last section, the conclusions are presented.

2 The HFCVD Process

In the HFCVD process (Hot Filament Chemical Vapor Deposition), the reaction occurs in a chamber, as shown in Fig. 1. There is a metallic filament, generally made of tungsten, positioned between 3 and 10 mm above the substrate [2]. Usually, the substrate used is monocrystalline silicon, because it has high melting point, crystal structure and thermal expansion coefficient similar to diamond and has a relatively low cost [3], although other substrates may also be used, such as molybdenum, titanium, and copper, among others [4].

The carbon used in the process comes from the ethyl alcohol, methane or other hydrocarbon [5]. The hydrocarbon is dissolved in hydrogen and injected into the chamber at a constant rate, with a ratio of around 0.5%.

The filament is connected to a DC power source and it reaches an average temperature of 2000 °C while warming the substrate to a temperature of about 800 °C. At this temperature, tungsten catalyzes the formation of atomic hydrogen (H^0) which, together with the thermal process, decomposes the hydrocarbon dissolved in activated carboxyl, for example CH_3, CH_2, CH, CHOH, among others. The H^0 and these radicals reach the substrate, where it begins to emerge diamond deposits.

The process begins with the formation of seeds and carbon atoms will be deposited around them. The nucleation process can be accelerated with a pretreatment of the substrate, also called seeding. It consists, for example, directly sprinkling diamond powder on the substrate or a diamond slurry sonication treatment.

The deposited films can be identified by their crystal morphology analyzed by optical microscopy and the Raman spectrum typical for crystalline diamond [6]. The Raman spectroscopy identifies the different phases of carbon, like diamond, nanocrystalline diamond, graphite, amorphous carbon, DLC (Diamond Like Carbon) and

hydrogenated carbon [7-8]. For the diamond, the typical wavelength peak occurs in 1333 cm^{-1}. For graphite, the peak occurs in the Raman spectrum from 1500 cm^{-1} and 1600 cm^{-1}.

It is important to determine and control the parameters involved in the process to define the conditions under which there is diamond films growth and to prevent defects occurring in the crystal formed.

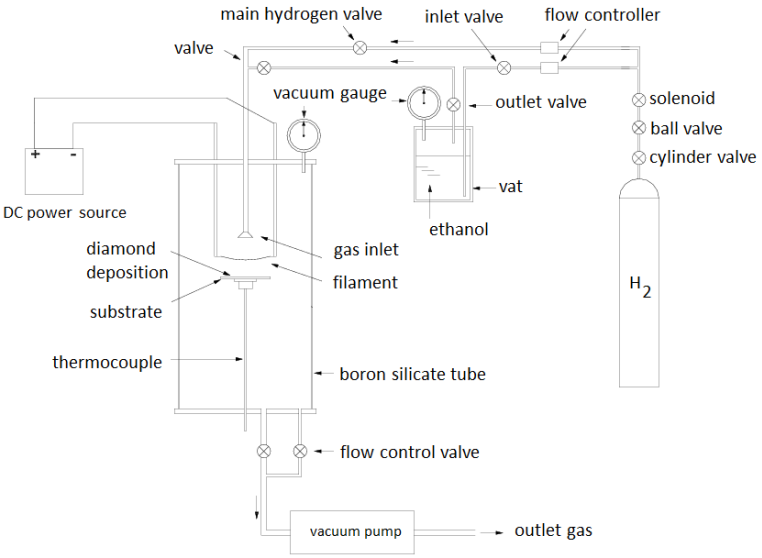

Fig. 1. Schematic view of the HFCVD chamber

3 Self-Organizing Maps (SOM)

The SOM is one of the main models of neural networks at present and is used in countless applications. Unlike other neural network approaches, the SOM is a type of neural net based on competitive and unsupervised learning [9]. The network essentially consists of two layers: an input layer I and an output layer U with neurons generally organized in a 2-dimensional topological array. The input to the net corresponds to a *p*-dimensional vector, *x*, generally in the space R^p. All of the *p* components of the input vector feed each of the neurons on the map. Each neuron *i* can be represented by a synaptic weight vector $w_i = [w_{i1}, w_{i2},..., w_{ip}]T$, also in the *p*-dimensional space. For each input pattern *x* a winner neuron, *c*, is chosen, using the criterion of greatest similarity:

$$\|x - w_c\| = \min_i\{\|x - w_i\|\} \tag{1}$$

where $\|.\|$ represents the Euclidian distance. The winner neuron weights, together with the weights of the neighboring neurons, are adjusted according to the following equation:

$$w_i(t+1) = w_i(t) + h_{ci}(t)[x(t) - w_i(t)] \tag{2}$$

where t indicates the iteration of the training process, $\boldsymbol{x}(t)$ is the input pattern and $h_{ci}(t)$ is the nucleus of neighborhood around the winner neuron c.

Once the SOM training algorithm has converged the computed feature map displays important statistical characteristics of the input space, which can be summarized as follows [10]:

(i) Vector quantization: the basic objective of SOM is to store a large set of input vectors by finding a smaller set of prototypes that provides a good approximation to the input space.

(ii) Topological ordering: the features map computed by SOM is ordered topologically. Similar input vectors are mapped close to each other, while dissimilar ones are mapped far apart.

(iii) Density Matching: the SOM reflects the probability distribution of data in the input space. Regions in the input space in which the input patterns are taken with a high probability of occurrence are mapped onto larger domains of the output space, and thus have better resolution than regions in the output space from which input patterns are taken with a low probability of occurrence.

There are some studies that apply SOM for visualization and analysis of processes in several areas. Pilsung [11] developed a virtual metrology system for an etching process in semiconductor manufacturing based on various data mining techniques that can not only predict the metrology measurement accurately, but also detect possible faulty wafers with a reasonable confidence. Abonyi [12] applied SOM on the analysis of an industrial polyethylene plant and demonstrates that the SOM is very effective in the detection of the typical operating regions related to different product grades, and the model can be used to predict the product quality based on measured process variables. Rasanen [13] presents an approach for dynamic process state monitoring and applied it in a circulating fluidized bed energy plant. That was based on a self-refreshing modification of the self-organizing map where previously learned data was used recursively to avoid catastrophic forgetting. The results of the simulations showed that method is a useful tool for monitoring process states. Sanchez [14] describes a virtual sensor design for coating thickness estimation in a hot dip galvanising line based on local models using SOM. Domínguez [15] proposes a method that defines a new visual exploration tool, called dissimilarity map for the visual comparison of industrial processes.

4 Diamond Depositions

Twenty one 21 diamond films depositions were made, each one consisted of 18 parameters: substrate area, pressure in the chamber, the main stream gas, ethanol flow, the bubbler (vat) valve opening, pressure in the bubbler, main valve opening, the thermocouple voltage, the substrate temperature, electrical current source, voltage source, the ethanol temperature, the ethanol vapor pressure, hydrogen line pressure, filament diameter, distance between filament and the substrate, deposition time, the percentage of carbon. The samples numbers 1 - 5 were deposited on silicon substrates and not passed through the seeding process. Samples 6 - 14 were deposited on silicon substrates and the sowing was carried with tungsten powder with a particle size of

Table 1. Samples separation into groups

Group	Deposition	Sample
1	None	1, 2, 3, 4 e 5
2	No diamond	8, 9, 11, 12, 13, 14 e 15
3	Diamond	6, 7, 10, 16, 17, 18, 19, 20, 21

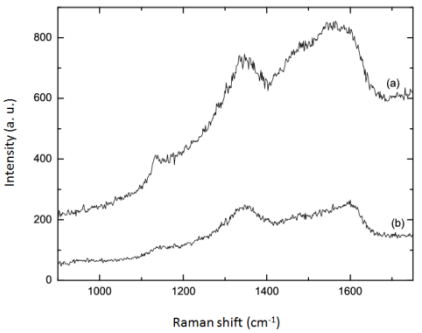

Fig. 2. Raman spectrum for sample 15

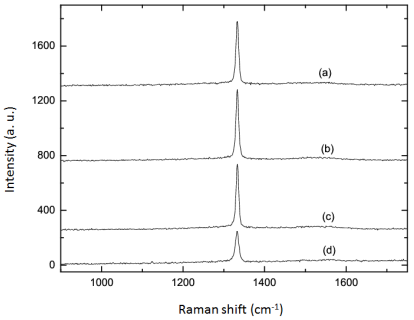

Fig. 3. Raman spectrum for sample 16

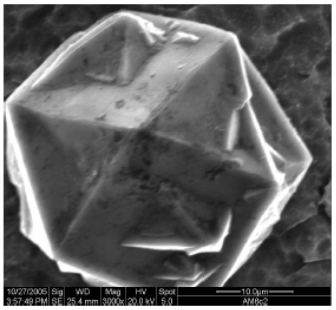

Fig. 4. Scanning microscopy image for sample 6

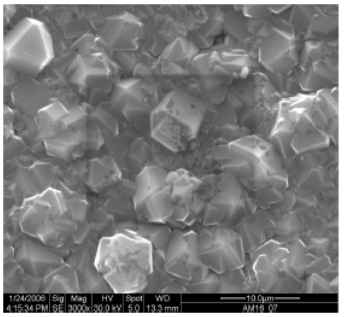

Fig. 5. Scanning microscopy image for sample 16

about 12 microns. The sample numbers 15 - 19 were deposited on silicon substrates, and the sowing was carried out with diamond powder with a particle size of about 1 micron. Samples 20 and 21 were deposited on substrates of pure titanium and the sowing was also performed with the diamond powder with a particle size of about 1 micron.

The analyzes carried out by scanning electron microscopy and Raman spectroscopy show the existence of three groups of samples, as shown in Table 1. Group 1 represents a sample where there was not the deposition of any material. Group 2

represents samples where there was no diamond deposition, but other allotropic carbon forms. Group 3 represents the diamond crystals deposition.

The Raman spectrum of the sample 15, shown in Fig. 2, reveals a prevalence of graphite deposition, since the Raman shift peak is near 1600 cm^{-1}, curve (a) at the periphery and the curve (b) in the center of sample. In sample 16, there were crystals deposition in diamond lattice, with a Raman shift peak around 1330 cm^{-1} as shown in Fig. 3, curves (a) and (b) at the periphery and the curves (c) and (d) in the center of the sample.

Figure 4 shows the scanning electron microscopy image for sample number 6 corresponding to diamond deposition. In Fig. 5, the image shows diamond crystals formation in the sample 16.

5 Clustering and Visualizations

The database used in this paper contains 21 records, divided into three classes, namely: 5 instances for the first class, 7 instances for second class and 9 instances for third class. The first class indicates that there was no material deposition, the second class indicates that there was material deposition, but not diamond crystals and the third class indicates diamond deposition.

Data were normalized, because SOM uses the Euclidean distance to measure distances between vectors. As the ranges of the variables are different, standardization is necessary so that a variable does not stand out compared to other.

The SOM Toolbox is contains function for creation, visualization and analysis of self-Organizing Maps [16]. It is a software package for Matlab computing environment available free of charge from http://www.cis.hut.fi/projects/somtoolbox.

Viewing the results produced by the SOM Toolbox it is necessary to set some parameters for initialization, training and viewing the map. In the initialization it was used the function *som_lininit* function, that initializes a SOM linearly, and for training it was used the *som_batchtrain* function, that trains the SOM with the given data using the batch training algorithm.

The SOM parameters configuration that were used for clustering and data visualization were: the variance is normalized to one, linear initialization of weights, batch training mode, gaussian neighborhood function, 3000 epochs and size of the map 12x12.

After training, the map is displayed in the U-Matrix, with the function *som_show*. In the U-Matrix different colors are used to represent the distances between neurons. A light shade means that they are close, as a darker can be interpreted as a separator clusters. Figure 6 shows the U-matrix color in three different classes, each one representing a cluster of samples. Figure 7 shows the map containing the sample clustering represented by their respective numbers, as follows:

- Red – class 1, no material deposition;
- Yellow – class 2, no diamond deposition, but other allotropic carbon forms;
- Blue – class 3, diamond deposition.

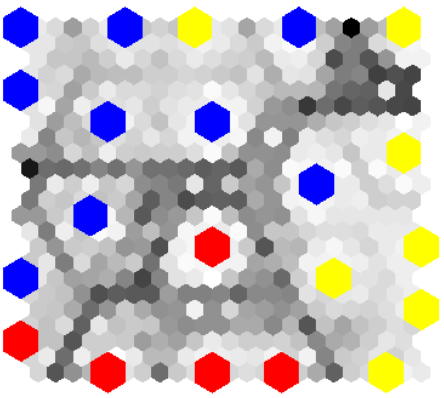

Fig. 6. The U-Matrix

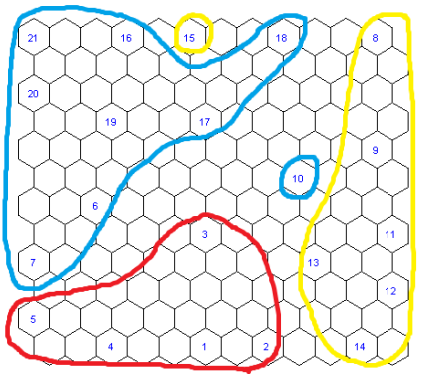

Fig. 7. Clustering map

The sample 10 and sample 15 was intentionally marked to show they are arranged outside the cluster they belong to.

Observing Figs. 6 and 7, it is possible to see that the sample 15 (without diamond deposition) was placed between the samples of group 3, which had diamond deposition and away from the samples that did not result in the deposition of diamond, group 2. It can be explained by the deposition parameters in the sample: as well as the samples of group 3, this one was sowing with diamond powder and the other parameters values result in values similar to group samples, except the substrate temperature, which was 10% below the ideal temperature during deposition. The low temperature did not favor the formation of diamond crystals. The sample number 10 (with diamond deposition) was placed next to samples which showed no diamond deposition. This sample showed parameter values similar to other samples of the second group, but the substrate temperature was good throughout the deposition process, while the other group samples were deposited at a temperature below the minimum required to result in diamond crystals deposition.

Another way to understand the U-Matriz is show in Fig. 8. The valleys represent each cluster and de montains represent the distance between them.

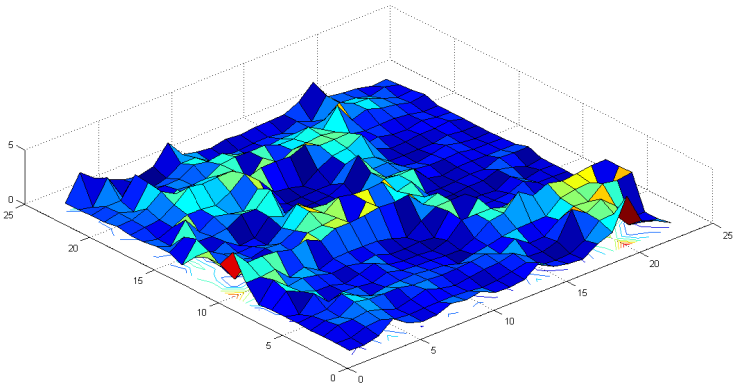

Fig. 8. The U-Matrix surface

Table 2. Quantization error for different map sizes

Map Size	Qe	Te
8x8	1.0066	0
10x8	0.8001	0
10x10	0.6353	0
12x10	0.4300	0
12x12	**0.2431**	0

The quality of the map generated was evaluated with the function *som_quality* through two measurements: Quantization error (Qe) and topographic error (Te). The quantization error represents the average distance of each input vector to the neuron. The topographic error represents the proportion of data vectors in which the first and second winning neurons are not adjacent units. The results obtained for the two metrics are Qe = 0.2431 and Te = 0. It shows, respectively, that the neuron is properly seated on the input vectors and the map represents the data input topology. Table 2 shows the results for the quantization error and topographic error for different map sizes.

6 Conclusions

SOM Toolbox was used to clustering and visualization of the deposition parameters of thin diamond films by HFCVD process. These parameters were used for calibration of a reactor. Analyzing the samples 10 and 15, it is clear how important is the substrate temperature during the process. In the map generated it can be seen that the separation of the samples are consistent with the results of analyzes by scanning electron microscopy and Raman spectroscopy, of the deposited films, showing the algorithm's effectiveness in clustering similar data, helping to identify the optimal parameters for deposition in this reactor.

References

1. Grill, A.: Diamond-Like Carbon State of the Art. Diamond and Related Materials 8(2-5), 428–434 (1999)
2. Matsumoto, S.: Development of Diamond Synthesis Techniques at Low Pressures. Thin Solid Films 368(2), 231–236 (2000)
3. Riley, D.J., Alexander, M.S., Latto, M.N., May, P.W., Pastor-Moreno, G.: A Simple Route to Ohmic Contacts on Low Boron-Doped CVD Diamond. Diamond and Related Materials 12(9), 1460–1462 (2003)
4. Ristic, G.S., Bogdanov, Z.D., Zec, S., Romcevic, N., Dohcevic-Mitrovic, Z., Miljanic, S.S.: Effect of the Substrate Material on Diamond CVD Coating Properties. Materials Chemistry and Physics 80(2), 529–536 (2003)
5. May, P.W., Everitt, N.M., Trevor, C.G., Ashfold, M.N.R., Rosser, K.N.: Diamond Deposition in a Hot-Filament Reactor Using Different Hydrocarbon Precursor Gases. Applied Surface Science 68(3), 299–305 (1993)

6. Spear, K.: Diamond (Ceramic Coating of the Future. Journal of the American Ceramic Society 72(2), 171–191 (1989)
7. Breza, J., Kadlečíková, M., Veselý, M., Frgala, Z., Kudrle, V., Janča, V., Janík, J., Buršík, J.: Raman Bands in Microwave Plasma Assisted Chemical Vapour Deposited Films. Microelectronics Journal 34(11), 1075–1077 (2003)
8. Willard, H.H., Merrit Jr., L.L., Dean, J.A., Settle Jr., F.A.: Instrumental Methods of Analysis. Wadsworth Publishing Company (1988)
9. Kohonen, T.: Self-Organizing Maps, 2nd edn. Springer, Berlin (1997)
10. Gonçalves, M., Andrade Netto, M.L., Costa, J.A.F.: Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information. In: Mwasiagi, J.I. (ed.) Self-Organizing Maps Applications and Novel Algorithm Design, vol. 1, pp. 299–322. InTech, Viena (2011)
11. Pilsung, K., Hyoung-joo, L., Sungzoon, C., Dongil, K., Jinwoo, P., Chan-Kyoo, P., Seungyong, D.: A Virtual Metrology System for Semiconductor Manufacturing. Expert Systems with Applications 36(10), 12554–12561 (2009)
12. Abonyi, J., Nemeth, S., Vincze, C., Arva, P.: Process Analysis and Product Quality Estimation by Self-Organizing Maps with an Application to Polyethylene Production. Computers in Industry 52(3), 221–234 (2003)
13. Rasanen, T., Kettunen, A., Niemitalo, E., Hiltunen, Y.: Self-Refreshing SOM for Dynamic Process State Monitoring in a Circulating Fluidized Bed Energy Plant. In: 3rd International IEEE Conference on Intelligent Systems (IS 2006), pp. 344–349. IEEE (2006)
14. Sanchez, A.P., Blanco, I.D., Cuadrado-Vega, A.A., Diez-Gonzalez, A.B., Carrera, F.O., Rubio, V.L.: Virtual Sensor Design for Coating Thickness Estimation in a Hot Dip Galvanising Line based on Interpolated SOM Local Models. In: 28th IEEE Annual Conference of the Industrial Electronics Society (IECON 2002), vol. 2, pp. 1584–1589 (2002)
15. Domínguez, M., Fuertes, J.J., Díaz, I., Prada, M.A., Alonso, S., Morán, A.: Monitoring Industrial Processes with SOM-Based Dissimilarity Maps. Expert Systems with Applications 39(8), 7110–7120 (2012)
16. Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J.: Self-Organizing Map in Matlab: the SOM Toolbox. In: 1999 Matlab DSP Conference, pp. 35–40 (1999)