Smoothing Problem in 2D Images with Tissue-like P Systems and Parallel Implementation

Francisco Peña-Cantillana[1], Daniel Díaz-Pernil[2], Hepzibah A. Christinal[2,3], Miguel A. Gutiérrez-Naranjo[1]

[1] Research Group on Natural Computing
    Department of Computer Science and Artificial Intelligence
    University of Sevilla, Spain
    `frapencan@gmail.com,magutier@us.es`
[2] Computational Algebraic Topology and Applied Mathematics Research Group
    Department of Applied Mathematics I
    University of Sevilla, Spain
    `sbdani@us.es`
[3] Karunya University, Coimbatore, Tamilnadu, India
    `hepzi@yahoo.com`

Summary. Smoothing is often used in Digital Imagery to reduce noise within an image. In this paper we present a Membrane Computing algorithm for smoothing 2D images in the framework of tissue-like P systems. The algorithm has been implemented by using a novel device architecture called CUDA™, (Compute Unified Device Architecture). We present some examples, compare the obtained time and present some research lines for the future.

1 Introduction

The study of digital images [22] has seen a large progress over the last decades. The aim of dealing with an image in its digital form is improving its quality, in some sense, or to simply achieving some artistic effect. The physical properties of camera technology are inherently linked to different sources of noise, so the application of smoothing algorithm are necessary for an appropriate use of the images. Smoothing is often used to reduce such noise within an image.

In this paper we use Membrane Computing[4] techniques for smoothing 2D images in the framework of tissue-like P systems. The algorithm has been implemented by using a novel device architecture called CUDA™, (Compute Unified Device Architecture) [16, 23]. CUDA™ is a general purpose parallel computing architecture that allows the parallel NVIDIA Graphics Processors Units (GPUs)

---

[4] We refer to [19] for basic information in this area, to [20] for a comprehensive presentation and the web site [24] for the up-to-date information.

to solve many complex computational problems[5] in a more efficient way than on a CPU . This parallel architecture has been previously used in Membrane Computing [1, 2, 3] but, to the best of our knowledge, this is the first time that it is used for implementing smoothing algorithms with tissue-like P systems.

Dealing with Digital Imagery has several features which make it suitable for techniques inspired by nature. One of them is that it can be parallelized and locally solved. Regardless how large the picture is, the segmentation process can be performed in parallel in different local areas. Another interesting feature is that the basic necessary information can be easily encoded by bio-inspired representations.

In the literature, one can find several attempts for bridging problems from Digital Imagery with Natural Computing as the works by K.G. Subramanian *et al.* [4, 5] or the work by Chao and Nakayama where Natural Computing and Algebraic Topology are linked by using Neural Networks [6] (extended Kohonen mapping). Recently, new approaches have been presented in the framework of Membrane Computing [7, 11]. In [8, 9, 10], Christinal *et al.* started a new bio-inspired research line where the power and efficiency of tissue-like P systems [12, 13] were applied to topological processes for 2D and 3D digital images.

The paper is organised as follows: Firstly, we recall some basics of tissue-like P systems and the foundations of Digital Imagery. Next we present our P systems family and an easy example showing different results by using different thresholds. In Section 3 we present the implementation in CUDA™ of the algorithm and show an illustrative example, including a comparative of the time obtained in the different variants. The paper finishes with some final remarks and hints for future work.

2 Preliminaries

In this section we provide some basics on the used P system model, tissue-like P systems, and on the foundation of Digital Imagery.

Tissue-like P systems [14, 15] have two biological inspirations: intercellular communication and cooperation between neurons. The common mathematical model of these two mechanisms is a network of processors dealing with symbols and communicating these symbols along channels specified in advance.

Formally, a *tissue-like P system* with input of degree $q \geq 1$ is a tuple

$$\Pi = (\Gamma, \Sigma, \mathcal{E}, w_1, \dots, w_q, \mathcal{R}, i_\Pi, o_\Pi),$$

where
1. Γ is a finite *alphabet*, whose symbols will be called *objects*;
2. $\Sigma(\subset \Gamma)$ is the input alphabet;
3. $\mathcal{E} \subseteq \Gamma$ is the alphabet of objects in the environment;
4. $w_1, \dots, w_q$ are strings over Γ representing the multisets of objects associated with the cells at the initial configuration;

---

[5] For a good overview, the reader can refer to [17, 18].

5. $\mathcal{R}$ is a finite set of communication rules of the following form:

$$(i, u/v, j)$$

for $i, j \in \{0, 1, 2, \ldots, q\}, i \neq j, u, v \in \Gamma^*$;
6. $i_\Pi \in \{1, 2, \ldots, q\}$ is the input cell;
7. $o_\Pi \in \{0, 1, 2, \ldots, q\}$ is the output cells

A tissue-like P system of degree $q \geq 1$ can be seen as a set of q cells (each one consisting of an elementary membrane) labelled by $1, 2, \ldots, q$. We will use 0 to refer to the label of the environment, i_Π denotes the input region and o_Π denotes the output region (which can be the region inside a cell or the environment).

The strings $w_1, \ldots, w_q$ describe the multisets of objects placed in the q cells of the P system. We interpret that $\mathcal{E} \subseteq \Gamma$ is the set of objects placed in the environment, each one of them available in an arbitrary large amount of copies.

The communication rule $(i, u/v, j)$ can be applied over two cells labelled by i and j such that u is contained in cell i and v is contained in cell j. The application of this rule means that the objects of the multisets represented by u and v are interchanged between the two cells. Note that if either $i = 0$ or $j = 0$ then the objects are interchanged between a cell and the environment.

Rules are used as usual in the framework of membrane computing, that is, in a maximally parallel way (a universal clock is considered). In one step, each object in a membrane can only be used for one rule (non-deterministically chosen when there are several possibilities), but any object which can participate in a rule of any form must do it, i.e., in each step we apply a maximal set of rules.

A *configuration* is an instantaneous description of the P system Π. Given a configuration, we can perform a computation step and obtain a new configuration by applying the rules in a parallel manner as it is shown above. A sequence of computation steps is called a *computation*. A configuration is *halting* when no rules can be applied to it. Then, a computation halts when the P system reaches a halting configuration.

Next we recall some basics on Digital Imagery[6].

A *point set* is simply a topological space consisting of a collection of objects called points and a topology which provides for such notions as *nearness* of two points, the *connectivity* of a subset of the point set, the *neighborhood* of a point, *boundary points*, and *curves* and *arcs*. For a point set X in Z, a *neighborhood function* from X in Z, is a function $N : X \to 2^Z$. For each point $x \in X, N(x) \subseteq Z$. The set $N(x)$ is called a *neighborhood* for x.

There are two neighborhood function on subsets of $\mathbb{Z}^2$ which are of particular importance in image processing, the *von Neumann* neighborhood and the *Moore* neighborhood. The first one $N : X \to 2^{\mathbb{Z}^2}$ is defined by $N(x) = \{y : y = (x_1 \pm j, x_2) \text{ or } y = (x_1, x_2 \pm k), j, k \in \{0, 1\}\}$, where $x = (x_1, x_2) \in X \subset \mathbb{Z}^2$. While the Moore neighborhood $M : X \to 2^{\mathbb{Z}^2}$ is defined by $M(x) = \{y : y = (x_1 \pm j, x_2 \pm k), j, k \in \{0, 1\}\}$, where $x = (x_1, x_2) \in X \subset \mathbb{Z}^2$. The von Neumann

---

[6] We refer the interested reader to [21] for a detailed introduction.

and Moore neighborhood are also called the *four neighborhood* (4-adjacency) and *eight neighborhood* (8-adjacency), respectively.

An Z-*valued image* on X is any element of Z^X. Given an Z-valued image $I \in Z^X$, i.e. $I : X \to Z$, then Z is called the set of possible range values of I and X the spatial domain of I. The graph of an image is also referred to as the *data structure representation* of the image. Given the data structure representation $I = \{(x, I(x)) : x \in X\}$, then an element $(x, I(x))$ is called a *picture element* or *pixel*. The first coordinate x of a pixel is called the *pixel location* or *image point*, and the second coordinate $I(x)$ is called the *pixel value* of I at location x.

For example, X could be a subset of $\mathbb{Z}^2$ where $x = (i, j)$ denotes spatial location, and Z could be a subset of $\mathbb{N}$ or $\mathbb{N}^3$, etc. So, given an image $I \in Z^{\mathbb{Z}^2}$, a pixel of I is the form $((i, j), I(x))$, which be denoted by $I(x)_{ij}$. We call the *set of colors* or *alphabet of colors of I*, $\mathcal{C}_I \subseteq Z$, to the image set of the function I with domain X and the image point of each pixel is called *associated color*. We can consider an order in this set. Usually, we consider in digital image a predefined alphabet of colors $\mathcal{C} \subseteq Z$. We define $h = |\mathcal{C}|$ as the size (number of colors) of $\mathcal{C}$. In this paper, we work with images in grey scale, then $\mathcal{C} = \{0, \dots, 255\}$, where 0 codify the black color and 255 the white color.

A *region* could be defined by a subset of the domain of I whose points are all mapped to the same (or similar) pixel value by I. So, we can consider the region R_i as the set $\{x \in X : I(x) = i\}$ but we prefer to consider a region r as a maximal connected subset of a set like R_i. We say two regions r_1, r_2 are adjacent when at less a pair of pixel $x_1 \in r_1$ and $x_2 \in r_2$ are adjacent. We say x_1 and x_2 are *border pixels*. If $I(x_1) < I(x_2)$ we say x_1 is an *edge pixel*. The set of connected edge pixels with the same pixel value is called a *boundary* between two regions.

The purpose of image enhancement is to improve the visual appearance of an image, or to transform an image into a form that is better suited for human interpretation or machine analysis. There exists a multitude of image enhancement techniques, as are averaging of multiple images, local averaging, Gaussian smoothing, max-min sharpening transform, etc.

One of the form to enhancement an image could be eliminate non important regions of an image, i.e., remove regions which do not provide relevant information. This technique is known as smoothing.

2.1 A Family of Tissue-like P Systems

Given a digital image with n^2 pixels and $n \in \mathbb{N}$ we define a tissue-like P system whose input is given by the pixels of the image encoded by the objects a_{ij}, where $1 \le i, j \le n$ and $a \in \mathcal{C}$. Next, we shall give some outlines how to prove that our *smoothing problem* can be solved in a logarithmic number of steps using a family of tissue-like P systems Π.

We define a family of tissue-like P systems to do an smoothing of a 2D image. For each image of size n^2 with $n \in \mathbb{N}$, we consider the tissue-like P system with input of degree 1:

$$\mathbf{\Pi}(r,n) = (\Gamma, \Sigma, \mathcal{E}, w_1, \mathcal{R}, i_\Pi, o_\Pi),$$

where

- $\Gamma = \Sigma \cup \mathcal{E}$,
- $\Sigma = \{a_{ij} : a \in \mathcal{C}, \ 1 \le i, j \le n\}$,
- $\mathcal{E} = \{a_{ij} : 1 \le i, j \le n, \ a \in \mathcal{C}\}$,
- $w_1 = \emptyset$,
- R is the following set of communication rules:
  - $(1, a_{ij}b_{kl}/a_{ij}a_{kl}, 0)$,
    - for $1 \le i, j \le n$,
    - $a, b \in \mathcal{C}$, $a < b$ and $d(a,b) \le r$,

    These rules are used to simplify the image. If we have two colors whose distance in the alphabet of colors of the image is small, then we change the high color by the small one. Of this manner, we change the regions structure.
- $i_\Pi = o_\Pi = 1$.

Each P system works as follows: We take pairs of adjacent pixels and change the color of the pixel with lower color. We do it in a parallel way with all the possible pairs of pixels. In the next step, we will repeat the previous process, but the colors of the pixels have could to be changed. So, we need, in the worst case, a linear number of steps to do all the possible changes to obtain an smoothing of our image.
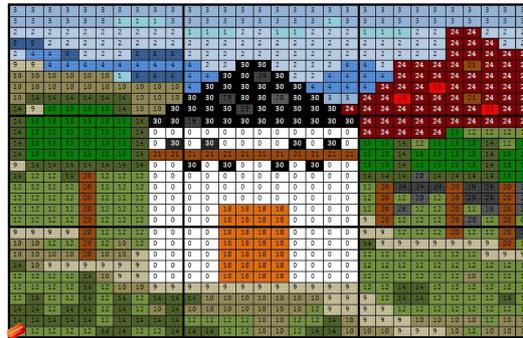

Fig. 1. An example

2.2 An Easy Example

In this section, we show the results obtained by the application of our method. Our input image (of size 30×30) can be seen in Figure 1. In this case, 0 represents

to white color and 30 represents to black color, i.e., the inverse order for $\mathcal{C}$ is considered.
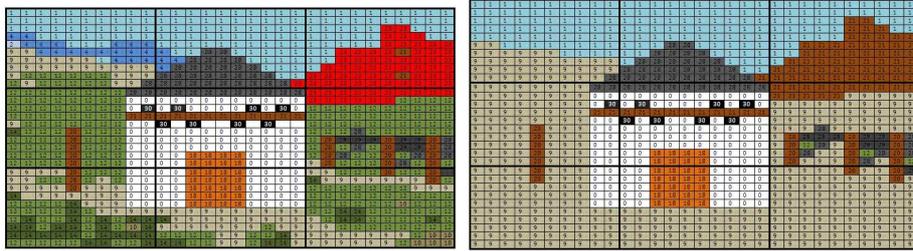

Fig. 2. Theoretical Smoothness of the Figure 1.

Working with different thresholds provides different results as we can observe in Figure 2. If we take $r = 5$, then we get the first image, and when the threshold is $r = 10$ the second image is obtained. By using this method, it is clear that the structure of regions is (more or less) conserved with a threshold of $r = 5$ and we need to a high number to obtain a more simplified image. Moreover, we can observe in both cases the color of regions is similar to the regions of input image in this method. Therefore, we can use this technique for smoothing images, and clarify the structure of a region without eliminate important information.

Bearing in mind the size of the input data is $O(n^2)$, $|\mathcal{C}| = h$ and r is the threshold used with both membrane solutions. The amount of necessary resources for defining the systems of our families and the complexity of our problem is determined in the following table:

Smoothing Problem	
Complexity	Dynamical
Number of steps of a computation	$O(n)$
Necessary Resources	
Size of the alphabet	$n^2 \cdot h$
Initial number of cells	1
Initial number of objects	0
Number of rules	$O(n^2 \cdot h)$
Upper bound for the length of the rules	4

3 Parallel Implementation

GPUs constitute nowadays a solid alternative for high performance computing, and the advent of CUDA™ allow programmers a friendly model to accelerate a broad range of applications. The way GPUs exploit parallelism differ from multi-core

CPUs, which raises new challenges to take advantage of its tremendous computing power. GPU is especially well-suited to address problems that can be expressed as data-parallel computations. GPUs can support several thousand of concurrent threads providing a massively parallel environment. This parallel computation model leads us to look for a highly parallel computational technology where a parallel simulator can run efficiently.

In this paper, we present a parallel software tool based in our membrane solution for smoothing images. It has been developed by using Microsoft Visual Studio 2008 Professional Edition (C++) with the plugging Parallel Nsight (CUDA™) under Microsoft Windows 7 Professional with 32 bits.

To implement the P systems, CUDA™ C, an extension of C for implementations of executable kernels in parallel with graphical cards NVIDIA has been used. It has been necessary the *nvcc compiler* of CUDA™ Toolkit. Moreover, we use libraries from openCV to the treatment of input and output images. Microsoft Visual Studio 2008 is responsible for calling to the compilers to build the objects, and to link them with the final program. This allows us to deal with images stored in .BMP, .DIB, .JPEG, .JPG, .JPE, .PNG, .PBM, .PGM, .PPM, .SR, .RAS, .TIFF and .TIF formats

The experiments have been performed on a computer with a CPU Intel Pentium 4 650, with support for HT technology which allows to work like two CPUs of 32 bits to 3412 MHz. Computer has 2 MB of L2 cache memory and 1 GB DDR SDRAM of main memory with 64 bits bus wide to 200 MHz. Moreover, it has a hard disc of 160 GB SATA2 with a transfer rate of 300 Mbps in a 8 MB buffer.

The graphical card (GPU) is an NVIDIA Geforce 8600 GT composed by 4 *Stream Processors* with a total of 32 cores to 1300 MHz and executes 512 threads per block as maximum. It has a 512 MB DDR2 main memory, but 499 MB could be used by processing in a 128 bits bus to 700 MHz. So, the transfer rate obtained is by 22.4 Gbps. For constant memory used 64 KB and for shared memory 16 KB (It is not a good data for a good CUDA™ graphical card). Its Compute Capability is 1.1 (from 1.0 to 2.1), then we can obtain a lot of improvements in the efficiency of the algorithms.

If we compare CPU and GPU, we observe that the former has two cores to 3412 MHz and the latter has 32 to 1300Mhz and has larger memory.

We have developed two applications of our P systems. In this case, we consider the natural order in the set of colors $\mathcal{C} = \{0, \ldots, 255\}$. In the first one, we have considered a deterministic implementation, where we work with the Moore neighborhood. So, the system checks if the rules can be applied for eight adjacent pixels. In the second one, we have considered an random selection of an adjacent pixel to work. Then, the system checks only one possibility chosen in an random way. Of this form, we simulate the characteristic non determinism of P systems with random. Moreover, we have decided to stop the system before the halting configuration, because more than an appropriate number of parallel steps of processing could be non-operative. In fact, in the deterministic version, the process could fin-

ish before the pre-fixed number of steps. So, the system needs some time to check this possibility. In the second one, it is not necessary to look at this question.
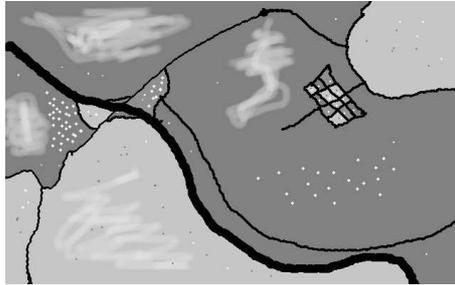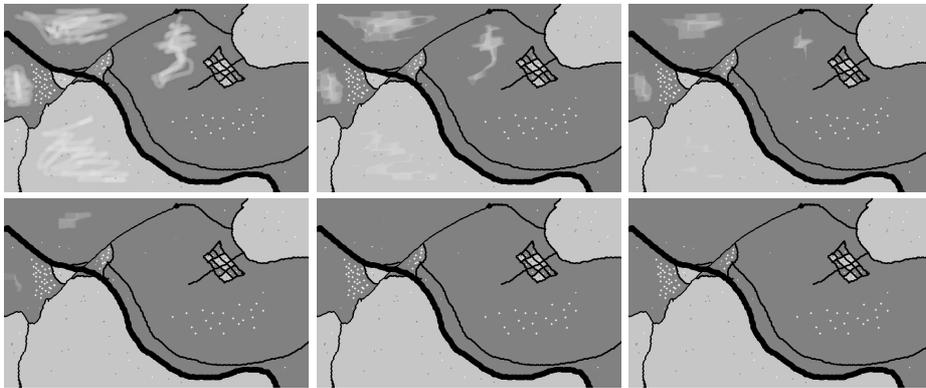

Fig. 3. Original image

Fig. 4. Deterministic version, Threshold 50: 0, 5, 10, 20, 32 and 44 steps, respectively.
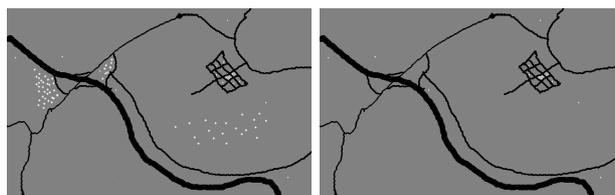

Fig. 5. Deterministic version. First: Threshold 75, step 193. Second: Threshold 125, step 193.

We consider the image of size 640×400 in Figure 3. When we take the deterministic application of our software, we can check that if we use an threshold $r = 50$, our software smooths the original image (see Figure 4) using 44 parallel steps. Nonetheless, when we work with a higher threshold, new important regions are changed, and the output image is different. (See the images of Figure 5).

When we consider the random version of our software, we can check that if we use an threshold $r = 50$ we need 300 steps, but the differences with the resulting image with 150 or 200 steps are minimum, as we can see in Figure 6. When we take higher thresholds, as in the Figure 7, we can check that new regions change of colors.
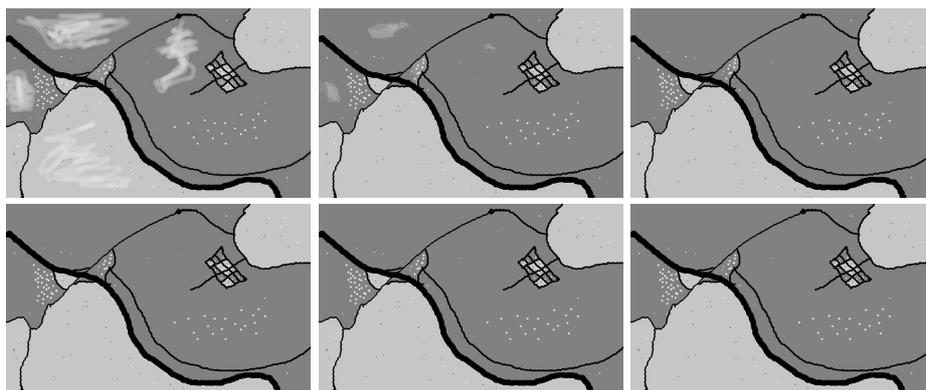

Fig. 6. Random version, Threshold 50: 0, 50, 100, 150, 200 and 300 steps, respectively.
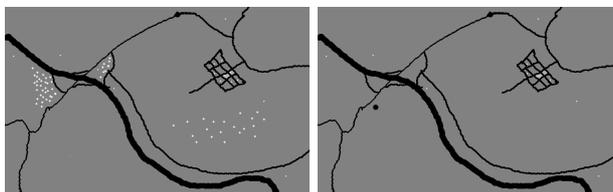

Fig. 7. Random version.First: Threshold 75, step 1000. Second: Threshold 125, step 800.

We present a study the running time of our software for both cases with different thresholds, showed in the above examples, with the next table. We can observe that deterministic version of our software needs less time with respect to the random version. In the first one, it applies eight rules for each pixels while, in the second one, it applies only one rule for each pixel. Moreover, we need a running time to implement the random in each step for each pixel.

Version \ Thresholds	Computation steps	Running Time
Determ. \ 50	44	536.522 ms
Determ. \ 75	193	1582.098 ms
Determ. \ 125	193	1563.660 ms
Random \ 50	300	6823.683 ms
Random \ 75	1000	21537.701 ms
Random \ 100	800	17332.891 ms

Finally, we have done some experiments with our software to know what happened if we work with images of different size. We have checked our software with images until size 512 in both version. The deterministic version needs much time with bigger images, and the random version does not work with those images. This is a physical problem our graphical card, because the shared memory is small.
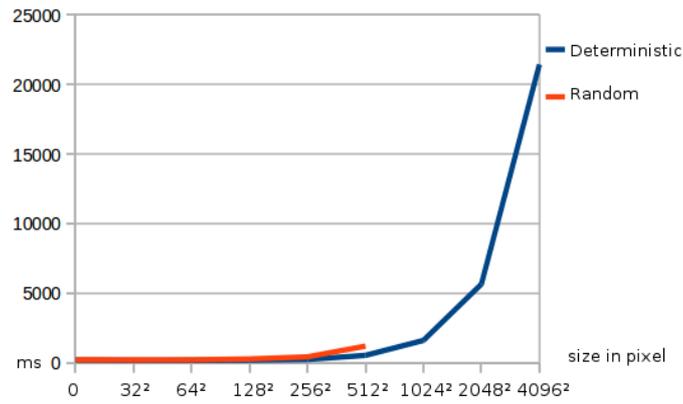

Fig. 8. Comparatives.

4 Conclusions and Future Work

In this paper, three emergent research fields are put together. Firstly, as pointed in [8, 10], Membrane Computing has features as the encapsulation of the information, a simple representation of the knowledge and parallelism, which are appropriate with dealing with digital images. Nonetheless, the use of the intrinsic parallelism of Membrane Computing techniques cannot be implemented in current one-processor computers, so the potential advantages of the theoretical design are lost.

In this paper we show that the drawback of using one-processor computers for implementing Membrane Computing designs can be avoided by using the parallel architecture CUDA™. This new technology provides the hardware needed for a real parallel implementation of Membrane Computing algorithms.

Considering this paper as a starting point, several research lines are open: From Digital Imagery, new parallel algorithms can be proposed adapted to the new technology, from the Membrane Computing side, new design or different P system models can be explored. From the hardware point of view, the advances in the new technology CUDA™ with the new boards Tesla and Fermi open new possibilities for going on with the research.

Acknowledgements

References

1. Cecilia, J.M., García, J.M., Guerrero, G.D., Martínez-del-Amor, M.A., Pérez-Hurtado, I., Pérez-Jiménez, M.J.: Implementing P systems parallelism by means of GPUs. In: Păun, Gh., Pérez-Jiménez, M.J., Riscos-Núñez, A., Rozenberg, G., Salomaa, A. (eds.) Workshop on Membrane Computing. Lecture Notes in Computer Science, vol. 5957, pp. 227–241. Springer (2009)
2. Cecilia, J.M., García, J.M., Guerrero, G.D., Martínez-del-Amor, M.A., Pérez-Hurtado, I., Pérez-Jiménez, M.J.: Simulating a P system based efficient solution to SAT by using GPUs. Journal of Logic and Algebraic Programming 79(6), 317–325 (2010)
3. Cecilia, J.M., García, J.M., Guerrero, G.D., Martínez-del-Amor, M.A., Pérez-Hurtado, I., Pérez-Jiménez, M.J.: Simulation of P systems with active membranes on CUDA. Briefings in Bioinformatics 11(3), 313–322 (2010)
4. Ceterchi, R., Gramatovici, R., Jonoska, N., Subramanian, K.G.: Tissue-like P systems with active membranes for picture generation. Fundamenta Informaticae 56(4), 311–328 (2003)
5. Ceterchi, R., Mutyam, M., Păun, G., Subramanian, K.G.: Array-rewriting P systems. Natural Computing 2(3), 229–249 (2003)
6. Chao, J., Nakayama, J.: Cubical singular simplex model for 3D objects and fast computation of homology groups. In: 13th International Conference on Pattern Recognition (ICPR'96). vol. IV, pp. 190–194. IEEE Computer Society, IEEE Computer Society, Los Alamitos, CA, USA (1996)
7. Christinal, H.A., Díaz-Pernil, D., Gutiérrez-Naranjo, M.A., Pérez-Jiménez, M.J.: Thresholding of 2d images with cell-like P systems. Romanian Journal of Information Science and Technology (ROMJIST) 13(2), 131–140 (2010)

8. Christinal, H.A., Díaz-Pernil, D., Real, P.: Segmentation in 2D and 3D image using tissue-like P system. In: Bayro-Corrochano, E., Eklundh, J.O. (eds.) CIARP. Lecture Notes in Computer Science, vol. 5856, pp. 169–176. Springer (2009)

9. Christinal, H.A., Díaz-Pernil, D., Real, P.: Using membrane computing for obtaining homology groups of binary 2D digital images. In: Wiederhold, P., Barneva, R.P. (eds.) IWCIA. Lecture Notes in Computer Science, vol. 5852, pp. 383–396. Springer (2009)

10. Christinal, H.A., Díaz-Pernil, D., Real, P.: P systems and computational algebraic topology. Journal of Mathematical and Computer Modelling 52(11-12), 1982 – 1996 (December 2010), the BIC-TA 2009 Special Issue, International Conference on Bio-Inspired Computing: Theory and Applications

11. Díaz-Pernil, D., Gutiérrez-Naranjo, M.A., Molina-Abril, H., Real, P.: A bio-inspired software for segmenting digital images. In: Nagar, A.K., Thamburaj, R., Li, K., Tang, Z., Li, R. (eds.) Proceedings of the 2010 IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications BIC-TA. vol. 2, pp. 1377 – 1381. IEEE Computer Society (2010)

12. Díaz-Pernil, D., Gutiérrez-Naranjo, M.A., Pérez-Jiménez, M.J., Riscos-Núñez, A.: A uniform family of tissue P systems with cell division solving 3-COL in a linear time. Theoretical Computer Science 404(1-2), 76–87 (2008)

13. Díaz-Pernil, D., Pérez-Jiménez, M.J., Romero, A.: Efficient simulation of tissue-like P systems by transition cell-like P systems. Natural Computing 8, 797–806 (2009)

14. Martín-Vide, C., Pazos, J., Păun, Gh., Rodríguez-Patón, A.: A new class of symbolic abstract neural nets: Tissue P systems. In: Ibarra, O.H., Zhang, L. (eds.) COCOON. Lecture Notes in Computer Science, vol. 2387, pp. 290–299. Springer (2002)

15. Martín-Vide, C., Păun, Gh., Pazos, J., Rodríguez-Patón, A.: Tissue P systems. Theoretical Computer Science 296(2), 295–326 (2003)

16. Nickolls, J., Buck, I., Garland, M., Skadron, K.: Scalable parallel programming with cuda. Queue 6, 40–53 (March 2008)

17. Owens, J.D., Houston, M., Luebke, D., Green, S., Stone, J.E., Phillips, J.C.: GPU Computing. Proceedings of the IEEE 96(5), 879–899 (May 2008)

18. Owens, J.D., Luebke, D., Govindaraju, N., Harris, M., Krüger, J., Lefohn, A., Purcell, T.J.: A survey of general-purpose computation on graphics hardware. Computer Graphics Forum 26(1), 80–113 (2007)

19. Păun, Gh.: Membrane Computing. An Introduction. Springer-Verlag, Berlin, Germany (2002)

20. Păun, Gh., Rozenberg, G., Salomaa, A. (eds.): The Oxford Handbook of Membrane Computing. Oxford University Press (2010)

21. Ritter, G.X., Wilson, J.N., Davidson, J.L.: Image algebra: An overview. Computer Vision, Graphics, and Image Processing 49(3), 297–331 (1990)

22. Shapiro, L.G., Stockman, G.C.: Computer Vision. Prentice Hall PTR, Upper Saddle River, NJ, USA (2001)

23. NVIDIA Corporation. NVIDIA CUDA^{tm} Programming Guide.
    `http://www.nvidia.com/object/cuda_home_new.html`

24. P system web page. http://ppage.psystems.eu