One-Dimensional Kohonen Networks and Their Application to Automatic Classification of Images

Ricardo Pérez-Aguila, Pilar Gómez-Gil, and Antonio Aguilera

***Abstract* – This paper analyses the results obtained when different topologies of 1-Dimensional Kohonen Networks where used to classify color pictures taken to Popocatépetl Volcano (located in the State of Puebla, México, active and monitored since 1997). Due to the fact that volcano images needed to be correlated with other experiments in the research center where this was carried out, it was required to find out if classification was based on the intensity of the images or the topology of the pixels, that is, their connectivity. To do so, we addressed two approaches: one analyzing the results obtained when pixels in the images were permuted; the other using in the clusters generated by the network a novel metric previously tested to be invariant to topology of pixels, and comparing the results to the obtained when using the Euclidean distance.**

Index Terms – **1-Dimensional Kohonen Networks, non-supervised Image Classification, Metrics on Euclidean Spaces, Non-Supervised Classification.**

INTRODUCTION

It is well known the use of 1-Dimensional Kohonen Networks for non-supervised classification when a high redundancy is present [5]. With non-supervised classification, images presenting similar features are grouped in classes. Many processing tasks (as description, object recognition or indexing) are based on such preprocessing [9]. In this paper, we use Kohonen Networks to classify color images, and analyze the criteria followed by the networks to create the clusters. The paper is organized as follows: Section I describes basic concepts of 1-Dimensional Kohonen Networks; Section II describes the pre-processing applied to images for avoiding training bias; Section III describes the methods and results of classification of 2D color images with Kohonen networks and presents some analysis over the classification criteria. Section IV presents conclusions and future work.

Manuscript received October 31, 2006.

R. Pérez-Aguila, P. Gómez-Gil and A. Aguilera are with the Department of Computer Science, Electronics, Physics and Innovation, and the Research Center for Information and Automation Technologies (CENTIA), Universidad de las Américas – Puebla (UDLAP), Ex-Hacienda Santa Catarina Mártir, Cholula, Puebla, México 72820 (e-mail: is104378@prodigy.net.mx, pgomezexttu@gmail.com, antonio.aguilera@udlap.mx).

I. FUNDAMENTALS OF THE 1-DIMENSIONAL KOHONEN NETWORKS

A. Classifying Points in a n-Dimensional Space Through a 1-Dimensional Kohonen Network

A Kohonen Network with two layers showing n input neurons and m output neurons may be used to classify points embedded in a n-dimensional space into m categories ([4], [8]). Input points have the form $(x_1, \ldots, x_i, \ldots, x_n)$. The total number of connections from input layer to output layer is $n\times m$ (See Figure 1). Each neuron j in the output layer 1, $\leq j \leq m$, will have associated a n-dimensional weight vector which describes a representation of class C_j. All these vectors have the form:

$$\text{Output neuron 1: } W_1 = (w_{1,1}, \ldots, w_{1,n})$$
$$\vdots$$
$$\text{Output neuron } m\text{: } W_m = (w_{m,1}, \ldots, w_{m,n})$$

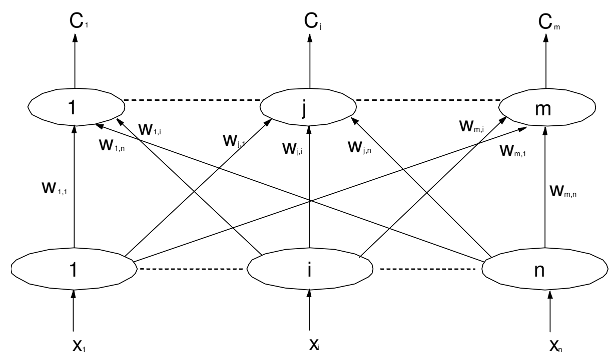


Fig 1. Topology of a 1-dimensional Kohonen Network [5].

B. Training the 1-dimensional Kohonen Network

A set of training points are presented to the network T times. According to [5], all values of weight vectors can be initialized with random values. The neuron whose weight vector W_j, $1 \leq j \leq m$, is the most similar to the input point P^k is chosen as winner neuron, for each t, $0 < t < T$. In the model proposed by Kohonen, such selection is based on the squared Euclidean distance. The selected neuron will be that with the minimal distance between its weight vector and the input point P^k:

$$d_j = \sum_{i=1}^{n} \left(P_i^k - w_{j,i} \right)^2 \qquad 1 \leq j \leq m \qquad (1)$$

Once the j-th winner neuron in the t-th presentation has been identified, its weights are updated according to:

$$w_{j,i}(t+1) = w_{j,i}(t) + \frac{1}{t+1}\left[P_i^k - w_{j,i}(t)\right] \qquad 1 \le i \le n \qquad (2)$$

When the T presentations have been achieved, values of the weights vectors correspond to coordinates of the 'gravity centers' of the points, or clusters of the m classes.

II. Redistribution in the n-Dimensional Space of Kohonen Network's Training Set

To avoid training bias, the training data was redistributed because it presented a non-uniform distribution over its space. This situation was similar to the next example:
Consider a set of points distributed in a 2D subspace defined by rectangle [0,1] × [0,1]. Moreover, this set of points is embedded in a sub-region delimited, for example, by rectangle [0.3,0.6] × [0.3,0.6] (Figure 2).

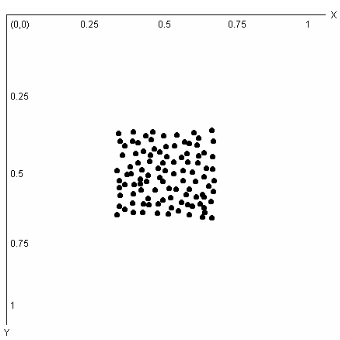


Fig 2. A set of points embedded in $[0.3,0.6] \times [0.3,0.6] \subset [0,1] \times [0,1]$.

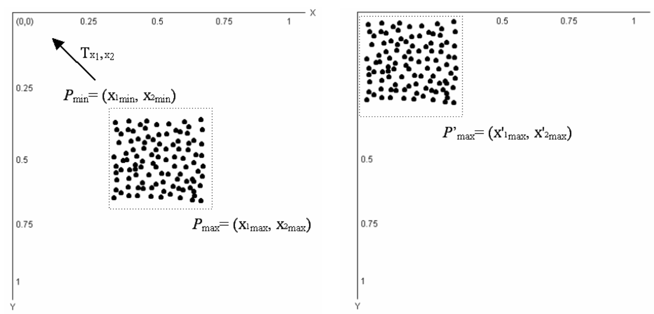
As points are not uniformly distributed in 2D space, we can expect important side effects during their classification process. For example, for a number of classes, we can obtain some clusters that coincide with other clusters or classes without associated training points.

Next we present a simple method to distribute uniformly the points of a training set for the general case of a n-dimensional space.

Consider a unit n-dimensional hypercube H where points are embedded in their corresponding minimal orthogonal bounding *hyper-box* h such that $h \subseteq H$. The point with the minimal coordinates $P_{\min} = (x_{1_{\min}}, x_{2_{\min}}, ..., x_{n-1_{\min}}, x_{n_{\min}})$ and the point with the maximal coordinates $P_{\max} = (x_{1_{\max}}, x_{2_{\max}}, ..., x_{n-1_{\max}}, x_{n_{\max}})$ will describe the main diagonal of h. We proceed to apply to each point $P = (x_1, x_2, ..., x_{n-1}, x_n)$ in the training set, including $P_{\min}$ and $P_{\max}$, the geometric transformation of translation given by:

$$x_i' = x_i - x_{i_{\min}} \qquad 1 \le i \le n \qquad (3)$$

By this way, we will get a new *hyper-box* h' and the points that describe the main diagonal of h' will be $P'_{\min} = (\underbrace{0,...,0}_{n})$ and $P'_{\max} = (x'_{1_{\max}}, x'_{2_{\max}}, ..., x'_{n-1_{\max}}, x'_{n_{\max}})$. (See Figure 3).



a) b)

Fig. 3. a) A training set and its minimal orthogonal bounding hyper-box h. b) Translation of h and the training points such that P'_{min} is the origin of the 2D space.

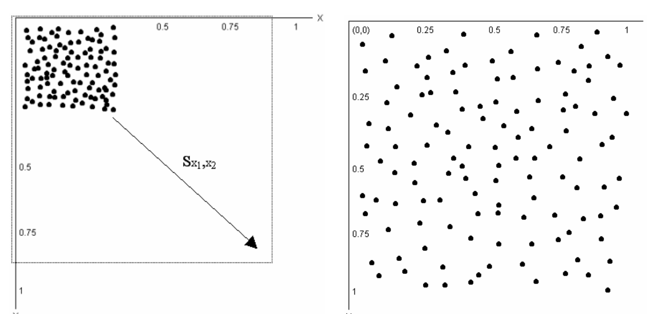
The second part of the distribution procedure consists on the expanding of the current *hyper-box h*' to the whole n-dimensional hypercube H. The scaling of a point $P = (x_1, x_2, ..., x_{n-1}, x_n)$ is given by multiplying their coordinates by factors $S_1, S_2, \ldots, S_n$ each one related with $x_1, x_2, \ldots, x_n$ respectively, in order to produce the new scaled coordinates $x_1', x_2', \ldots, x_n'$ [6]. The goal is to extend the bounding *hyper-box h'* and translated training points to the whole unit hypercube H, by translating point:
$P'_{\max} = (x'_{1_{\max}}, x'_{2_{\max}}, ..., x'_{n-1_{\max}}, x'_{n_{\max}})$ to $(\underbrace{1,...,1}_{n})$.

To do so, we define the set of n equations:

$$1 = x'_{i_{\max}} \cdot S_i \qquad 1 \le i \le n \qquad (4)$$

From equation (4) the scaling factors to apply to all points in *hyper-box h'* are obtained (see Figure 4):

$$S_i = \frac{1}{x'_{i_{\max}}} \qquad 1 \le i \le n \qquad (5)$$



a) b)

Fig. 4. An example of preprocessing. (a) original points (b) translated points.

Finally, each one of the coordinates in the original points of the training set must be transformed in order to be redistributed in the whole unit n-dimensional hypercube $[0,1]^n$ through:

$$x'_i = (x_i - x_{i_{\min}}) \cdot \left(\frac{1}{x'_{i_{\max}}}\right) \qquad 1 \le i \le n \qquad (6)$$

III. Image Classification through 1-Dimensional Kohonen Networks

A. Representing Images through Vectors in $\Re^n$

Let m_1 (rows) and m_2 (columns) be the dimensions of a two-dimensional image. Each pixel in the image, $n = m_1 \cdot m_2$, has associated a 3-dimensional point (x_i, y_i, RGB_i), $RGB_i \in [0, 16{,}777{,}216]$, $1 \le i \le n$, representing the color value of the i-th pixel (assuming that the color of pixels are based in the color model RGB). The color values of the pixels are normalized such that they are in [0.0, 1.0], through the transformation:

$$normalized\_RGB_i = \frac{RGB_i}{16777216} \qquad (7)$$

We define a vector in the n-Dimensional space by concatenating the m_1 rows in the image considering for each pixel its normalized color RGB value. By this way each image is associated to a vector in the n-dimensional Euclidean space. Due to this color normalization, the scalar values are in [0,1], making the training set embedded in a unit n-Dimensional hypercube.

B. Classifications Results

Our training set contains 148 images selected from *CENAPRED* [3] database. These images contain some of the Popocatépetl volcano fumaroles occurred during 2003. The selected images have a resolution of 640×480 pixels and 24-bits color, using format compression JPG.

We present the results obtained by three 1-Dimensional Kohonen Networks with different topologies:

- Network Topology τ_0:
  - Images Resolution: 112×64
  - Input Neurons: n = 112×64 = 7,168
  - Output Neurons (classes): m = 20
  - Presentations: T = 10
- Network Topology τ_1:
  - Images Resolution: 56×32
  - Input Neurons: n = 56×32 = 1,792
  - Output Neurons (classes): m = 30
  - Presentations: T = 1,000
- Network Topology τ_2:
  - Images Resolution: 260×180
  - Input Neurons: n = 260×180 = 46800
  - Output Neurons (classes): m = 25
  - Presentations: T = 500

The set of 148 training points (images) were presented a number of times according to the corresponding topologies. The training procedures were applied as described at Section II. All the weights vectors were initialized to 0.5.

Figure 5 shows the classification of the training images using the three proposed topologies. The figures show the distribution of the 148 training images in each one of the classes.

C. Intensity Based Classification Vs. Classification Based in the Topology of Pixels in the Images

One of the problems to consider is related with the question if our implementations of the Kohonen Networks classify according to intensity color or if they classify according to the connectivity, i.e. the topology between pixels composing an image. In order to give arguments that support our hypothesis that the network classified according to topology, we developed two approaches:

- An approach (section III.C.1) based in a classification of the training images with pixels permuted. If our implementation classifies by color intensity, we may expect a distribution in the classes similar to distributions presented before permutation, as in Figure 5.
- An approach (section III.C.2) based in the distances between the weight vectors associated to each output neuron. The clusters themselves are 2D color images. In this approach we will use an additional metric that guarantee the comparison of images only by their color intensity. According to Kohonen's training algorithm, the clusters (class representatives) have been distributed uniformly in a unit nD hypercube. Such distribution implies that each cluster has itself specific characteristics that distinguish its respective class among other classes. By applying the new proposed metric, we can expect that distances show a considerable proximity between clusters, hence, they have similar color intensities. Moreover, this last result should establish a considerable distinct distribution respect to the distribution indicated by the Euclidean metric. In the case that Kohonen Network classifies only by color intensity, then the clusters distribution reported by both metrics should be similar.

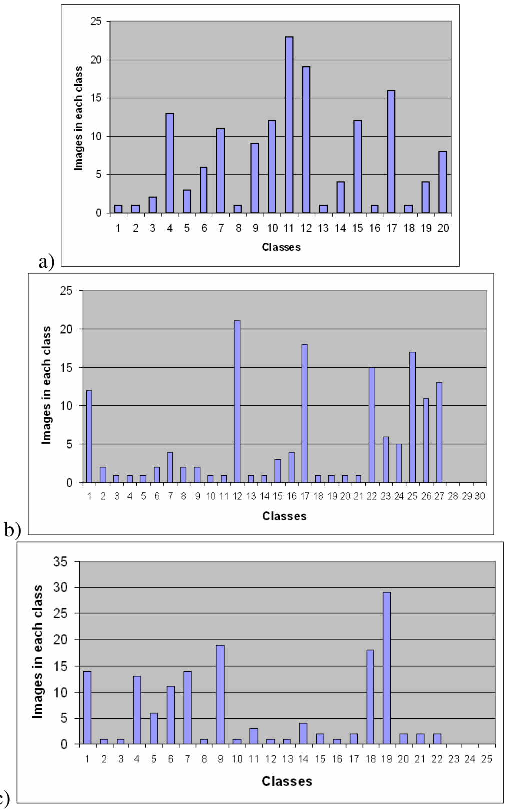


Fig. 5. Classification of the 148 training images according to Network Topology a) τ_0, b) τ_1 and c) τ_2.

C.1 Permutation of Pixels in the Training Images

(See Table 1 for examples of the permutations described here.)

- P_1: Random permutation of all the pixels in the image.
- P_2: Division of the image in 25 rectangular regions and random permutation of the pixels in each region.
- P_3: Division of the image in 25 rectangular regions and random permutation of such regions.
- P_4: Division of the image in 25 rectangular regions, random permutation of the pixels in each region and random permutation of the regions.

Table 1. Permutations of pixels applied to the training images.

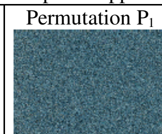
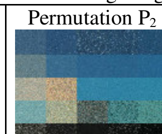
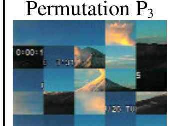
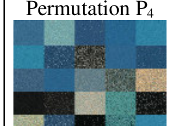
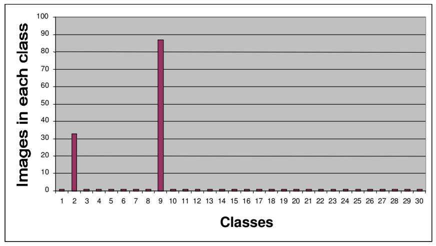
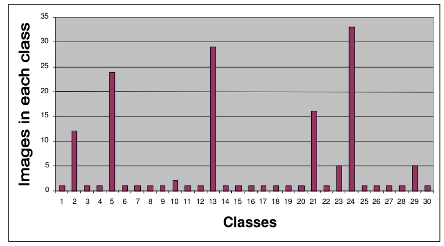
Original Image	Permutation P_1	Permutation P_2
		
Permutation P_3	Permutation P_4	
		

Table 2. Distribution of the training images in the classes of network topology τ_1.

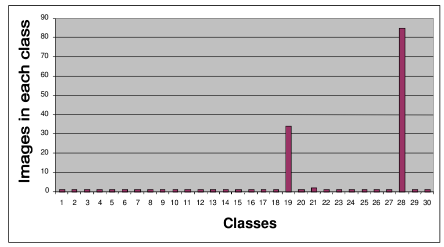


Permutation P_1



Permutation P_2

Permutation P_3



Permutation P_4

Consider network topology τ_1. For permutations P_1, P_3 and P_4, we observe in Table 2 that two classes grouped the 80% of training images.

The results using permutation τ_2 differs from others by the fact that 80% of training images are grouped in seven classes with more than 5 images each one. Permutation P_2 may be considered visually as one that preserves the connectivity of the pixels respect to the original training images. This is because if we increment the number of rectangular regions (more regions than those in permutation P_2) and permute its corresponding pixels, as the number of regions increase the corresponding image will approximate to the original image. In fact, the original images can be seen as images divided in regions with only one pixel each one; obviously, the permutation of the pixel in each region leave to the image in its original state.

C.2 Analysis Based in an Additional Metric over $\Re^+$

Definition 1 ([1] & [2]): Let x, y ∈ $\Re^+$. Let ρ be the function described as

$$\rho(x,y) = \begin{cases} 1 - \dfrac{x}{y} & if \quad x < y \\ 1 - \dfrac{y}{x} & if \quad y < x \\ 0 & if \quad x = y \end{cases} \qquad (8)$$

Such function is in fact a metric over $\Re^+$. See [7] for more details.

Let I be an image. We know that each one of its pixels p_i will have associated a vector (x_i, y_i, RGB_i), $i \in [1,n]$, $RGB_i \in$ [0, 16777216]. Lets assume that the dimensions of each pixel are equal to one. We will define the Total Intensity of I, denoted by T(I), as:

$$T(I) = \sum_{i=1}^{n} RGB_i \qquad (9)$$

Let I_A and I_B be two images with the same geometrical dimensions. Let $T(I_A)$ and $T(I_B)$ be their corresponding Total Intensities. Because $T(I_A)$, $T(I_B) \in \Re^+$ we can calculate its distance through metric ρ.

Now, we will define the similarity between images I_A and I_B according to the value of $\rho(T(I_A), T(I_B))$. Let $0 \le \varepsilon < 1$ be an arbitrary value such that

$$I_A \text{ is similar to } I_B \Leftrightarrow \rho(T(I_A), T(I_B)) < \varepsilon$$

It is shown that a classification based in metric ρ will not take in account the connectivity between the pixels in the images. For example, for the images presented in Figure 6 we have that $\rho(T(I_A), T(I_B)) = 0$.

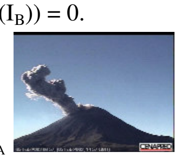
I_A  I_B 

Fig. 6. An example where $\rho(T(I_A), T(I_B)) = 0$.
I_B is image I_A applying permutation P_3.

Kohonen Network uses the Euclidean metric over $\Re^n$. Because the representatives of the classes (clusters) in the network are themselves vectors in $\Re^n$, then we can determine the Euclidean distance between any pair of clusters.

We define a false color map that represents the distribution of the clusters in the subspace $[0, 1]^n$. The maximal Euclidean distance between any two clusters will be $d_{\max} = \sqrt{n}$ and the minimal distance will be $d_{min} = 0$. Every Euclidean distance between two clusters will be associated with a color in the grayscale through $\frac{d}{d_{\max}} \cdot 256$. For $d = 0$ black color is associated, for $d = d_{max}$ then white color is associated.

Moreover, we define a false color map that represent the distances between the clusters in the subspace $[0, 1]^n$ under metric ρ. For any clusters a and b, $\rho(a, b)$ will be associated with the grayscale through $\rho(a, b)\cdot 256$. If $\rho(a, b) = 0$ then $a = b$ and therefore such distance will be represented by black color. On the other hand, $\rho(a, b)\cdot 256 \to 256$ while $\rho(a, b) \to 1$.

Consider Network Topology τ_0. The false color maps associated to the distances between the clusters under the Euclidean metric and ρ metric are presented in Table 3. It can be observed in the map under metric ρ that the 47% of the distances between clusters are less than 0.20. This indicates that, according this metric, an important number of clusters are similar with $\varepsilon = 0.20$ (the mean distance in this metric is 0.2542 with variance 0.0373 and standard deviation 0.1933). In the other hand, for topology τ_0, n = 7,168, hence, $d_{\max}=\sqrt{7168}$ = 84.66. Analogously we consider the number of distances whose value is less than the 20% of $d_{\max}$. By this way, the map based in the Euclidean metric reports that only the 19% of the distances between clusters are lower than 16.9328 (the mean distance under Euclidean metric was 24.2119 with variance 94.7531 and standard deviation 9.7341). In conclusion, both metrics report different distributions of the clusters which make visible the differences between a classification based in topology of pixels, by the Kohonen Network, and a classification based in color intensities of the images.

IV. CONCLUSIONS

According to the results provided by the approaches discussed in sections III.C.1 and III.C.2 we can infer that image classification based in a 1-Dimensional Kohonen Network groups an image set according to features based in the connectivity between pixels, i.e., their topology. As part of future work, we will analyze the images contained in each one of the resulting classes and their respective neighborhoods in order to determine some features shared by these images. By identifying these features, in our images domain, we will analyze the possible application of our classifications in the prediction of events of Popocatépetl volcano. Another objective considers the comparison of non-supervised classification with other techniques that allow the automated retrieval and classification of images such as Case Based Reasoning (CBR) and Image Based Reasoning (IBR).

REFERENCES

1. Aguilera, Antonio; Lázzeri Menéndez, Santos Gerardo & Pérez-Aguila, Ricardo. Image Based Reasoning Applied to the Comparison of the Popocatépetl Volcano's Fumaroles. Proceedings of the IX Ibero-American Workshops on Artificial Intelligence, Iberamia 2004, pp. 3-8. November 22 to 23, 2004. National Institute of Astrophysics, Optics and Electronics (INAOE), Puebla, México.
2. Aguilera, Antonio; Lázzeri Menéndez, Santos Gerardo & Pérez-Aguila, Ricardo. A Procedure for Comparing Color 2-Dimensional Images Through their Extrusions to the 5-Dimensional Colorspace. Proceedings of the 15th International Conference on Electronics, Communications, and Computers CONIELECOMP 2005, pp. 300-305. Published by the IEEE Computer Society. February 28 to March 2, 2005. Puebla, México.
3. CENAPRED (Centro Nacional de Prevención de Desastres), México. Web Site: http://www.cenapred.unam.mx (April 2003).
4. Davalo, Eric & Naïm, Patrick. Neural Networks. The Macmillan Press Ltd, 1992.
5. Hilera, José & Martínez, Victor. Redes Neuronales Artificiales. Alfaomega, 2000. México.
6. Pérez Aguila, Ricardo. The Extreme Vertices Model in the 4D space and its Applications in the Visualization and Analysis of Multidimensional Data Under the Context of a Geographical Information System. Thesis for the Master's Degree in Sciences. Universidad de las Américas - Puebla. Puebla, México, 2003.
7. Pérez Aguila, Ricardo; Gómez-Gil, Pilar & Aguilera, Antonio. Non-Supervised Classification of 2D Color Images Using Kohonen Networks and a Novel Metric. Progress in Pattern Recognition, Image Analysis and Applications; 10th Iberoamerican Congress on Pattern Recognition, CIARP 2005; Proceedings. Lecture Notes in Computer Science, Vol. 3773, pp. 271-284. Editors: Lazo, M. & Sanfeliu, A. Springer-Verlag Berlin Heidelberg. November 15 to 18, 2005. La Havana, Cuba.
8. Ritter, Helge; Martinetz, Thomas & Schulten, Klaus. Neural Computation and Self-Organizing Maps, An introduction. Addison-Wesley Publishing Company, 1992.
9. Zerubia, Josiane; Yu, Shan; Kato, Zoltan & Berthod, Mark. Bayesian Image Classification Using Markov Random Fields. Image and Vision Computing, 14:285-295, 1996.

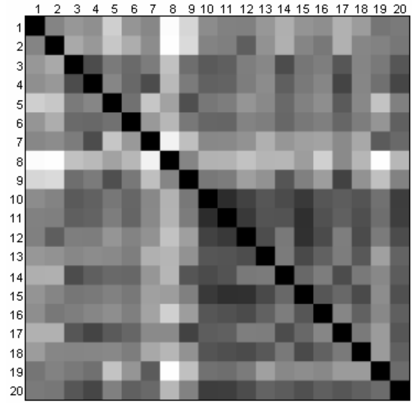
Ricardo Pérez-Aguila is a Ph.D. student in Computer Science at the Universidad de las Américas, Puebla (UDLA) where he has received also the B.Sc. (2001) and M.Sc. (2003) degrees in Computer Science. Since 2003 he has imparted undergraduate courses for the Department of Computing Science, Electronics, Physics and Innovation, and for the Department of Mathematics and Actuarial Sciences, at UDLA. His research interests consider the study of n-Dimensional Polytopes by analyzing their Visualization, Geometry, Topology, and Representation.

Pilar Gomez-Gil received the B.Sc. degree from the Universidad de las Americas (UDLA), at Puebla, Mexico in 1983, and the M.Sc. and Ph.D. degrees from Texas Tech University in 1991 and 1999, respectively, all in Computer Science. Since 1985 she has been with the Department of Computer Science, Electronics, Physics and Innovation at UDLA. Her research interests include Artificial Neural Networks, Pattern Recognition and Forecasting. Currently she is involved with research related to prediction of electrocardiograms, recognition of antique handwritten documents, and identification of HLA sequences. Dr. Gomez-Gil is a senior member of the IEEE, and founder member of the IEEE Computational Intelligence Society chapter Mexico.

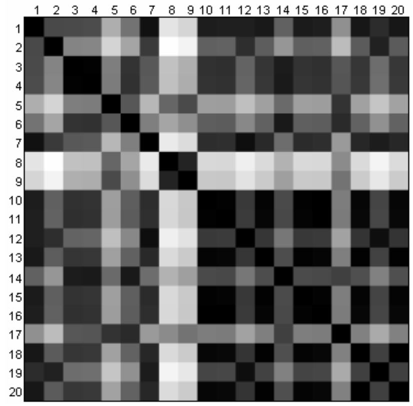
Antonio Aguilera received the B.Sc. degree in Mathematics from the Universidad Autónoma Metropolitana (1985), the M.Sc degree in Information Systems from the Universidad de las Americas, Puebla (UDLA), and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC). Since 1988 he is a full time professor of the Department of Computer Science, Electronics, Physics and Innovation at UDLA. His research interests consider Computer Graphics, Virtual Reality, Numerical Methods, and the study of n-Dimensional Polytopes.

Table 3. False Color Maps that show the distances between clusters in Network Topology τ_0.

Distances according to Euclidean Metric



Distances according to ρ Metric



In both maps:
Maximum Distance
Minimum Distance