

Texture Feature Neural Classifier for Remote Sensing Image Retrieval Systems

MAURICIO POZZOBON MARTINS¹, LAMARTINE N. FRUTUOSO GUIMARÃES², LEILA MARIA GARCIA FONSECA³

¹INPE - Instituto Nacional de Pesquisas Espaciais, Laboratório Associado de Computação e Matemática Aplicada, Av. dos Astronautas, 1.758 - Jd. Granja - CEP 12227-010 São José dos Campos, SP, Brasil.

mauriciopoz@directnet.com.br

²IEAv - Instituto de Estudos Avançados, CTA – Centro Técnico Aeroespacial, Rod. dos Tamoios km 5,5, 12228-840, São José dos Campos, SP, Brasil.

guimarae@ieav.cta.br

³INPE - Instituto Nacional de Pesquisas Espaciais, Divisão de Processamento de Imagens, Av. dos Astronautas, 1.758 - Jd. Granja - CEP 12227-010 São José dos Campos, SP, Brasil.

leila@dpi.inpe.br

Abstract. Texture information is useful for image data browsing and retrieval. The goal of this paper is to present a texture classification system for remote sensing images addressed to the administration of great collections of those images. The proposed classifier is a hybrid system composed by an unsupervised neural network and a supervised one. Starting from a small portion of the image (pattern) the system should recognize the most similar class to a pattern in a database as well as to identify the images that contain similar patterns. The texture feature vectors used to characterize the patterns are obtained from the images processed by a bank of Gabor Filters. Some experimental results using textures of the Brodatz album, multi-spectral and radar images are presented here.

1. Introduction

In the last few years a large amount of digital image data has been collected every day. Therefore, it is important the existence of an easy way to store, to organize and to retrieve those image data. Systems developed to organize and to retrieve large digital image collections in a database are called image retrieval systems. Research in this subject can be divided in two major groups: Database Management and Computer Vision [1]. They have studied image data retrieval from two different points of view: text-based and visual-based systems.

Text-based image retrieval system use image content annotation and Database Management System (DBMS) to accomplish the image search and recovery through keywords. The data processed by these systems are not the image itself but information related to the image content. This information is represented by concatenated alphanumeric characters (strings) and the queries are performed by traditional languages like SQL (Sequence Query Language). The main drawback is that the manual image annotation and the human perception subjectivity can affect the annotation precision. Two comprehensive surveys on this topic are presented in [2] and [3].

Computer vision researchers have attempted to overcome the difficulties found in the text-based image retrieval process [4, 5, 6]. Instead of using text-based

keywords, the images are indexed by their own visual content data such as color, texture, shape etc. [1].

Aerial and satellite images have become very important for public and private institutions due their various remote sensing applications. One may wish to locate, for instance, all images that contain a certain region such as parking lots, agriculture, train stations, popular house set, highway crossings, bridges, clandestine mines, factories etc. Few works have been focused on remote sensing image retrieval systems [7, 8, 9, 10]. It is evident that there are many difficulties or even the impossibility to perform the manual image browsing and retrieval in a reasonably short time acceptable by the user.

This work proposes the development of an image texture classification system addressed to the administration of great remote sensing image collections. The classification system identifies similar objects that belong to the same class that is characterized by texture features. Each object has its specific texture pattern that characterizes the class that it belongs. Besides the texture, other features such as color and shape can also be used to characterize the classes. In this work a texture descriptor based on a multiresolution decomposition using Gabor wavelets has been proposed.

Several works have considered the use of texture information for image data retrieval [10, 11, 12, 13, 14]. Manjunath and Ma [11, 12] have performed an evaluation using Gabor filters for texture analysis, but they performed

tests only with aerial photographs. In this work tests with multispectral and radar images, besides aerial photographs, were performed to verify the system robustness in providing suitable texture feature for different kinds of remote sensing images. A brief description of Gabor filters is presented in Section 2.

In order to facilitate the pattern retrieval task we propose to use a neural networks that learns how to recognize the similarities among the objects in the feature space. The proposed system is composed by two processing steps: (1) an unsupervised neural network (Kohonen's Self-Organizing Map – SOM and (2) a supervised neural network (Learning Vector Quantization - LVQ). The former generates the initial classification map, while the later one refines the initial classification accuracy.

The Kohonen's Self-Organizing Map divides the feature space in different regions. Each region belongs to a specific class that is characterized by a texture feature vector. SOM algorithm is a kind of code and compression system [15] similar to a vector quantization process [16].

Section 3 describes the neural network operation for the proposed system. The network training phase and the region-based feature space representation named "contextual map" are also described in Section 3. The method has been tested on Brodatz's textures album [17]. Section 4 presents some preliminary results obtained for remote sensing images.

2. Gabor Filter Bank

Edge detection is an important step in many computer vision process because edges contain the bulk of information within an image [18]. Another important aspect is that the most interesting features in the images are usually present in many different sizes and orientations. An efficient way of analyzing such features is using multiresolution decomposition based on Gabor wavelets.

Wavelets are families of basis functions obtained through dilations and translations of a *basic wavelet*. This decomposition provides a compact data structure for image representation [19]. The image representation obtained by convolving the original image with Gabor filters has been shown to be an optimum solution in the sense of minimizing the joint two-dimensional uncertainty in the space and frequency [20]. These filters can be used to detect lines and edges of different sizes and orientations. Besides, statistic measures obtained from filtered images can also be used to characterize the underlying texture information.

2.1 Gabor Functions

Gabor function describes a complex sinusoid with frequency W modulated by a Gaussian envelope with duration σ . A two-dimensional Gabor function and its Fourier transform can be written as [20]:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right] \quad (1)$$

$$G(u, v) = \exp \left\{ -\frac{1}{2} \left[\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\} \quad (2)$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$ and $j=(-1)^{1/2}$. A self-similar filter dictionary can be obtained by appropriate dilations and translations of $g(x,y)$ through the generation function [11]:

$$g_{mn}(x, y) = a^{-m} g(x', y') \quad a > 1, \quad m, n = \text{integer}$$

$$x' = a^{-m} (x \cos \theta + y \sin \theta),$$

$$y' = a^{-m} (-x \sin \theta + y \cos \theta), \quad (3)$$

where $\theta = n\pi/K$ and K is the total number of orientations. The scale factor a^{-m} in (3) is meant to ensure that the energy is independent of m . This set of functions form a non-orthogonal basis of functions for the multiresolution decomposition [20].

Figure 1 shows an example of the real part in the spatial domain of a filter in a certain scale and orientation (eq. 1). The elements of the filter are presented in gray levels to facilitate the wavelet form identification. Figure 2 shows the 3D profile of the same Gabor filter. The *amplitude* axis indicates the value of the elements.

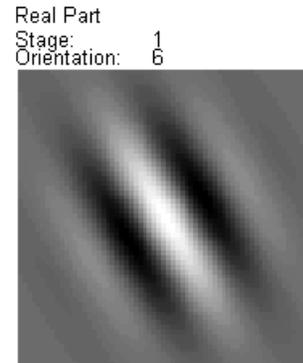


Figure 1 - The real part of a 64x64 pixels Gabor filter.

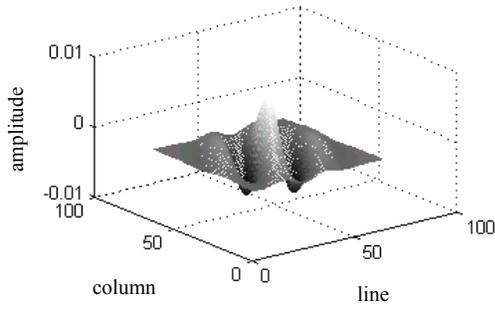


Figure 2 - The Gabor filter 3D profile.

2.2 Texture Feature Vector Extraction

Let $I(x, y)$ and $g_{mn}(x, y)$ be the image and the filter, respectively. The filtered image is [20]:

$$W_{mn}(x, y) = \int I(x, y) \cdot g_{mn}^*(x - x_1, y - y_1) dx_1 dy_1, \quad (4)$$

where * indicates the complex conjugate. The statistic measures used to generate the texture feature vectors are the mean (μ_{mn}) and the standard deviation (σ_{mn}) of the Gabor transform coefficient values [11].

Four scales and six orientations have been used in all experiments. Therefore, 24 Gabor filters and feature vectors with 48 elements each one are created as following:

$$\vec{f} = [\mu_{11}, \sigma_{11}, \mu_{12}, \sigma_{12}, \dots, \mu_{46}, \sigma_{46}] \quad (5)$$

The subscript of each vector component in (5) represents the scale and the orientation numbers.

3. Neural Network and Texture Pattern Recognition

According to Haykin [19], artificial neural networks were conceived based on the knowledge that the brain process information in a different way that a digital computer does. The main characteristic of an artificial neural network is its generalization ability. Generalization refers to the fact that the neural network produces appropriate exits for entrances that were not in the training phase (or learning). This makes the neural network an efficient procedure to solve complex problems that, in some cases, are untreatable by conventional methods.

Kohonen's Self-Organizing Map (SOM) and Learning Vector Quantization (LVQ) compose the proposed classification system. Texture features extracted by Gabor filters are input to the SOM algorithm that performs an

initial partition of the feature space into clusters of similar patterns. No previous knowledge of similar pattern groups is necessary. Afterwards the neural network weight vectors are submitted to the LVQ process for a fine adjustment. Figure 3 shows the schematic diagram of the classification procedure.

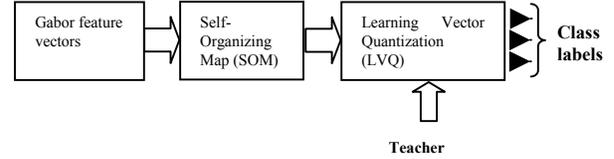


Figure 3- Texture Neural Classifier.

A “conscience factor” is used in the training phase in order to perform a correct representation of neurons in the feature map concerning to the statistical distribution of the classes into the original feature space.

3.1 Kohonen's Self-Organizing Map

The Kohonen's Self-Organizing Map (SOM) [21] is a special class of unsupervised neural network. These networks are formed by one or two-dimensional neuron grids (larger dimensions are possible but it is not common) that modify their synaptic weights in a competitive learning process. This work proposes a two-layer network to transform high-dimensional input features into a two-dimensional discrete map. The topological relationship between different features is preserved. The first neuron layer receives the input features and the second one presents the decision results in a 2-D map.

This network can adaptively separate clusters in the feature space. The output neurons are topologically ranked so that neighboring neurons correspond to similar clusters in the high-dimensional feature space [12].

In the competitive learning process, the neurons calculate their respective Euclidean distance values. The closest neuron to the input pattern is declared winner. Starting from the winner neuron, it is established a neuron neighborhood that will receive reinforcement in its synaptic connections. That will build a cooperation base among the neighbors to correlate the input pattern with an area of the neuron grid. Excited neurons modify their individual values through appropriate adjustments. The winner neuron response (and its neighborhood) gets better whenever a similar pattern to the one that previously excited it is applied to the net.

A lateral interaction in the winner neuron neighborhood is necessary to influence the neighboring neurons and therefore to recognize similar patterns. This

excitation area is function of the distance between the winner neuron and others: as the larger distance the weaker adjustments are applied [21]. It is also important that this lateral interaction includes the entire neuron grid in the initial learning stage (ranking phase) and only a few or just one neuron is included in the final processing (converging) phase [15]. A Gaussian function has been used to perform the lateral interaction process. Figure 4 shows a Gaussian neighborhood function at three different learning stages, from the beginning (left) to the end (right).

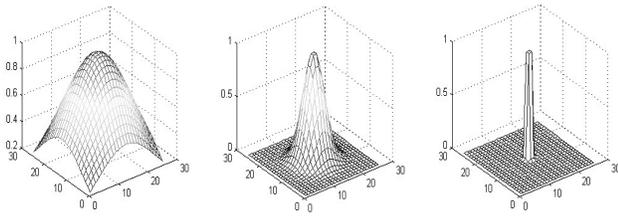


Figure 4- Gaussian neighborhood function in 3 stages.

3.2 Conscience Factor on Kohonen's Map

The Kohonen's map fails in representing the feature space statistical distribution. Computational simulations have demonstrated that the SOM algorithm represents low-density input features excessively whereas high-density input features are poorly represented [15]. The conscience factor in Kohonen net aims to get a probability about $1/N$ of a neuron in the output layer to be chosen to represent an input pattern [22]. In other words, the conscience algorithm is a modification of the SOM algorithm that forces the density matching to be exact [23]. The conscience algorithm registers how many times each neuron wins the competition [15]. If a neuron wins frequently it "feels guilty" and leaves the competition preventing poor representations of high-density input features and favoring the better proportion between the feature space and its representation map.

3.3 Learning Vector Quantization

The SOM algorithm provides an unsupervised initial partition of the feature space into a number of distinct clusters. The patterns belonging to each cluster are topologically similar. After this first phase, the output neurons are labeled using the training feature vectors. The output neurons are assigned to different classes by majority voting. The feature map aims to approximate the input feature vectors or their probability density function [8]. A network fine-tuning phase is necessary to improve the classification accuracy.

A fine-tuning process is performed in a second phase using a Learning Vector Quantization (LVQ) algorithm [20]. The LVQ improves the feature space partition by "moving" the weight vectors away from the decision borders. Kohonen has developed three algorithms named LVQ1, LVQ2 and LVQ3 [21]. The later one has presented better results than the others have and it was chosen to be implemented [8, 20].

4. Results Discussion

Initial tests were performed to evaluate if the feature vectors can efficiently represent their respective texture patterns for the Euclidean distance-based classification. This method has been tested on the Brodatz's album [17] as well as on remote sensing images. The neural classification system was tested only for the Brodatz's album (Figure 5).

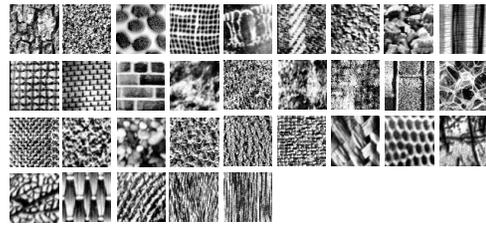


Figure 5- Textures classes from Brodatz's album.

4.1 Feature Vector Validation

Some experiments were performed using the Brodatz's album with 32 classes [17]. The idea is to evaluate the suitability of the texture feature for the classification process. The distance between the feature vectors that belong to the same class should present the smallest value. Each class has 16 patterns. Euclidean distances between each two different patterns were calculated. The retrieval system evaluation has been accomplished by verifying how many feature vectors belonging to the same class are among the first 16 retrievals. If all of the 16 first retrievals belong to the pattern class one has 100% of success. Figure 6 presents the success percentages for each one of the 32 classes using texture features and Euclidean distance normalized by the standard deviation.

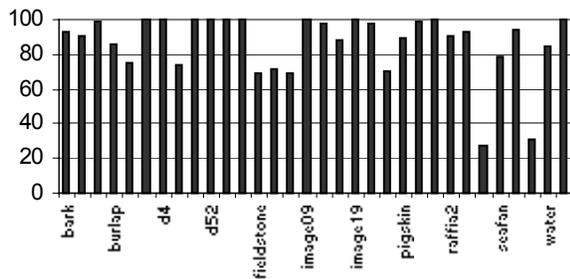


Figure 6- Success percentage in classification.

Another way to evaluate the classification system performance is through the kappa coefficient [24]; the larger the value of Kappa the better the classification system performance ($\text{kappa} \leq 1$). The resulting classification shown in Figure 6 presented a kappa value of 0.8594 (considered excellent).

A test to verify the efficiency of the texture feature vectors for remote sensing images has been performed. Normalized Euclidean distance was used as a similarity measure. An aerial photograph of Recife's airport runway was partitioned in squares of 64x64 pixels (sample). For each sample a feature vector was extracted. The test consisted on choosing a sample (at the pointer) of a significant texture pattern (Figure 7, left) and locating the 10 closest samples by comparing their features through the normalized Euclidean distance. The closest samples retrieved by the system are shown in Figure 7 (right). One may notice that the results are visually similar to the first one.

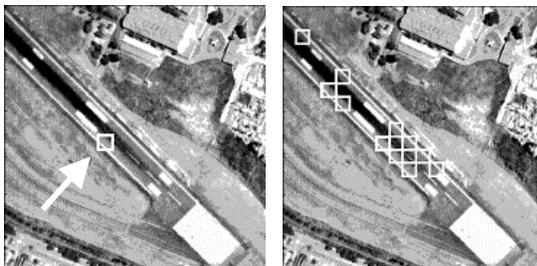


Figure 7- Air photograph. Texture pattern chosen (left). Squares returned (right).

The same test was performed for other types of remote sensing images including radar images. The test radar image, acquired from a SAR- 580 airborne sensor, has seven texture classes that were defined by a remote sensing interpreter. One 32x32 pixels square extracted from a texture class was chosen as the pattern (Figure 8 (left)). In this case, the 20 closest samples were retrieved (Figure 8 (right)). One can observe that samples with similar texture were found.

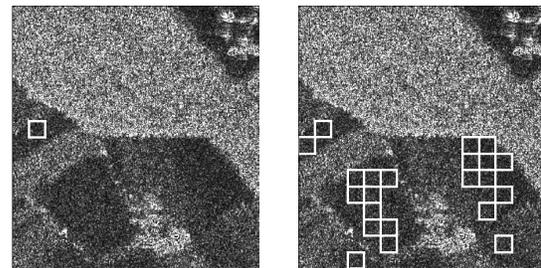


Figure 8- Radar image. Query texture (left). Returned ones (right).

A similar test was performed for a multispectral image acquired from Landsat-5 (bands 3, 4 and 5). Feature vectors were extracted from 64x64 pixels squares using three bands. The feature vectors obtained for each band were concatenated in just one feature vector. The test consisted on evaluating if this new feature could improve the classification when compared to the resulting classification obtained for just one band.

Figure 9 (left) shows the query square (at the pointer) and the retrievals for band 3. Figure 9 (right) shows the results when considering the three bands.

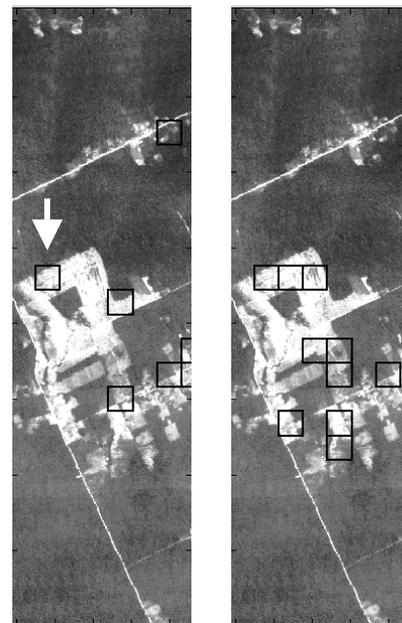


Figure 9- Multispectral image. Query with features from one band (left). Query with associated features (right).

One may observe that in Figure 9 (right) there are more squares visually closer to the query pattern than those in the Figure 9 (left). Apparently the use of a multi-spectral image (3 bands) improves the resulting classification.

4.2 Neural Classifier

Initial tests were performed for textures in Brodatz's album. Figure 10 shows a 10x10 output neuron grid of a Kohonen's feature map. Five texture classes were used in the training stage and different colors were associated to each class. The training features were introduced to the net and the winner neurons received the respective class color. As can be seen in Figure 10 (left) there is a topological organization of the neuron clusters. They are the most representative neurons to its respective class.

Classification tests of unknown patterns, however, are processed on a modified feature map, called contextual (or semantic) map [15]. Contextual map displays the self-organizing feature map so that each neuron appears labeled according to the class it stored during the learning stage. The neurons receive labeled patterns and the closest pattern determines the label of the neuron. The resulting map for the same five classes is shown in Figure 10 (right). This is a good way to verify the topological map organization; an unknown input pattern is labeled by the neuron it excited on the grid. It is expected that similar patterns, especially those not trained by the net, excite neurons in the same region on the map.

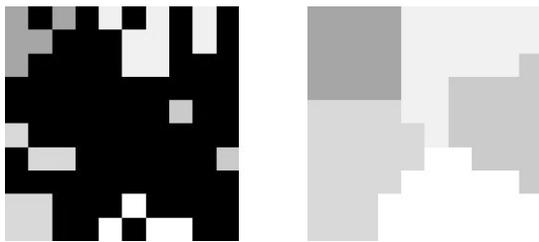


Figure 10- Kohonen's map output (left).
Contextual map (right).

The Brodatz's 32-texture classes were divided into two groups: a small group (1/3 of the features) for training and another large one for test. Feature vectors were extracted from those patterns. The training feature set was submitted to the SOM algorithm in a 25x25 neurons grid with and without conscience factor. The classification accuracy evolution during the training process was calculated at each 5 epochs by the kappa coefficient. The contextual map was generated and the testing features were submitted to the classification system. The final kappa coefficients were 0.6276 and 0.6628 without and with conscience factor, respectively (see Figure 11).

The LVQ3 algorithm was applied in the net considering the conscience factor. The kappa value obtained was the same value as obtained before, 0.6628.

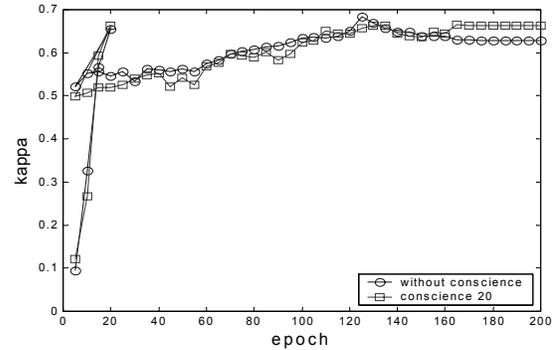


Figure 11- Kappa coefficients in the SOM training process.

5. Conclusions

One has observed that texture features obtained from Gabor filters are a good texture pattern representation even for remote sensing images.

The classification accuracy obtained with the SOM algorithm is considered good although no significant improvement was observed when using conscience factor or LVQ3 algorithm. It is possible that the texture classes chosen in the experiments are very different amongst themselves; it is necessary a larger bank of classes with evident similarity between some of them to make the classification task harder. One must take into account that training phase and the net parameters influence the results.

Most of the researches on texture feature extraction by Gabor filters use only air photographs as remote sensing images. This work showed that this methodology can also be applied on others kinds of images, like radar and multispectral images.

This work is addressed on feature extraction and neural network implementation. Unfortunately, comparisons with others approaches were not performed up to this moment.

For the future, we intend to test the neural classification system with textures obtained from remote sensing images, to study an indexing system to recover images from a database and perform comparisons with others approaches.

6. References

- [1] Rui, Y., Huang, T. S. and Chang, S. "Image Retrieval: Current Techniques, Promising Directions and Open Issues". *Journal of Visual Communication and Image Representation* 10 (1999), 39--62.

- [2] Chang, S and Hsu, A. "Image information systems: Where do we go from here?" *IEEE Transactions on Knowledge and Data Engineering* 4(5) (Oct. 1992).
- [3] Tamura, H. and Yokoya, N. "Image databases systems: A survey". *Pattern Recognition* 17(1) (1984), 29--43.
- [4] Gupta, A., Santini, S. and Jain, R. "In search of information in visual media". *Communications of ACM* 40(12) (Dec. 1997), 35--42.
- [5] Pentland, A. and Picard, R. "Special issues on digital libraries". *IEEE Transactions on Pattern Recognition and Machine Intelligence* 18(7-8) (1996), 673--733.
- [6] Chang, S and Hsu, A. "Image information systems: Where do we go from here?" *IEEE Transactions on Knowledge and Data Engineering* 4(5) (Oct. 1992), 431--442.
- [7] Val Cura, L. M. del, N. J. Leite, Merdeiros, C. B. "An Architecture for Content-Based Retrieval of Remote Sensing Image", *IEEE* (2000).
- [8] Ma, W. Y., and Manjunath, B. S. "A Pattern Thesaurus for Browsing Large Aerial Photographs". *Department of Electrical and Computer Engineering, University of California, ECE Technical Report 96-10* (Jun.1996).
- [9] Kelly, C., and Cannon, M. "Query by image example: The CANDID approach". *Proc. of SPIE Storage and Retrieval for Image and Video Databases III* (1995), 238--248.
- [10] Sheikholeslami, G., Zhang, A., and Bian, L. "A multi-resolution content-based retrieval approach for geographic images". *GEOINFORMATICA* 3(2) (Jun. 1999), 109--139.
- [11] Manjunath, B. S., and Ma, W. Y. "Texture Features for Browsing and Retrieval of Image Data", *IEEE Trans. on Pattern Analysis and Machine Intelligence* 18(8) (Ago. 1996), 837--842.
- [12] Manjunath, B. S., and Ma, W. Y. "A texture thesaurus for browsing large aerial photographs". *Journal of Amer. Soc. Inform.* 49(7) (1998), 633--648.
- [13] Manjunath, B. S., and Ma W. Y. "Browsing Large Satellite and Aerial Photographs", *International Conference on Image Processing*, vol. 3 (Set. 1996).
- [14] Equitz, W. and Niblack, W. "Retrieving images from a database using texture – algorithms from the QBIC system". Technical Report RJ 9805, *Computer Science, IBM Research Report*, (May 1994).
- [15] Haykin, S. "Redes Neurais, Princípios e Prática". 2 ed., *Bookman*, 2001.
- [16] Gersho, A. and Gray, R. M. "Vector Quantization and Signal Compression". *Kluwer Academic Publisher*, 1991.
- [17] Brodatz, P. "Textures: A photographic album for artists and designers". Toronto, Dover, 1966.
- [18] Myler, H. and Weeks, A.R. "Computer Imaging Recipes in C". *P T R Prentice Hall*, 1993.
- [19] Manjunath, B. S. "Perceptual Grouping and Segmentation Using Neural Networks". *Signal and Image Processing Institute, University of Southern California*, 1991.
- [20] Wu, P., Manjunath, B. S., Newsan, S., Shin, H. D. "A texture descriptor for browsing and similarity retrieval". *Signal Processing: Image Communication* (2000), 33-43.
- [21] Kohonen, T. "The Self-Organizing Map". *Proceedings of the IEEE* 78(9) (1990), 1464--1480.
- [22] Costa, R. A., Mendes, R. F. M., "Implementação de Consciência na Rede de Kohonen". *Trabalho de Graduação Interdisciplinar, Universidade Braz Cubas*, 1997.
- [23] Desieno, D. "Adding Conscience to Competitive Learning". *IEEE* (1988), 117--124.
- [24] Cohen, J. "A coefficient of agreement for nominal scales". *Educational and Psychological Meas.* (20) (1960), 37--46.