We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Robust Classification of Texture Images using Distributional-based Multivariate Analysis

Vasileios K. Pothos, Christos Theoharatos, George Economou
and Spiros Fotopoulos
Electronics laboratory, Dept. of Physics, University of Patras
Greece

1. Introduction

Classification of texture images has been recognized as an important task in the field of image analysis and computer vision through the last few decades. A plethora of research papers have appeared in the literature trying to cope with effective ways to extract faithful distributions that accurately represent the inner content and attributes of texture images. An issue of great importance is, also, the incorporation of a valid similarity measure that can successfully estimate how close these distributions are with respect to some pre-classified texture categories. The basic operations that need to be carried out in order to estimate the similarity between texture images and thereafter assess the classification problem are (a) choose an appropriate feature space for texture representation, (b) construct a theoretically valid distribution in the texture feature space, i.e. the texture signature, which provide a representation of the texture image in a multivariate feature space, (c) perform pairwise comparisons between corresponding texture signatures that constitute the consequent content distributions of the texture images and (d) choose an experimentally valid classifier for the subsequent evaluation.

The scope of this chapter is the survey of a recently introduced methodology for distributional-based classification of texture images (Pothos et al., 2007), its enhancement via the incorporation of a self-organizing module and its adaptation so as to work in multivariate feature spaces. The original approach is based on an efficient strategy for analyzing texture patterns within a distributional framework and the use of a statistical distributional measure for comparing multivariate data, also known as the multivariate Wald-Wolfowitz test (WW-test) (Friedman & Rafsky, 1979). By combining the flexible character of the original methodology with the learning abilities of neural networks, we build a general-purpose platform for the efficient information management and classification of texture patches without any restriction regarding the exact image content.

Here, we will first describe the enrollment of standard feature extraction techniques for summarizing texture information and structuring multivariate texture spaces. These techniques include wavelet analysis, discrete cosine transform (DCT), Gabor filters and edge histogram descriptor. The above methods have been considered as golden standards for extracting appropriate distributions from texture images. In the following stage, we will test the applicability of some multivariate distributional-based measures for estimating the

classification accuracy over widely available databases and outline the different alternative facets under which a texture database can be accessed within the introduced platform. In this way, the multivariate distributions representing the individual images will be compared via the standard WW-test and the Kantorovich-Wasserstein (KWass) distance, building a content-based texture classification scheme. The proposed distributional measures will be revealed to handle efficiently the texture-space dimensionality and the limited sample size drawn from a given image.

In order to further boost the classification accuracy of the entire scheme, we will finally introduce a computational intelligent module for content representation based on a self-organizing neural network (SONN), the Neural-Gas algorithm (Martinez et al., 1993). The incorporated feature-extraction unit will be responsible for generating a parsimonious description of the texture distribution of each image. The resulting performance will be used to evaluate the four utilized approaches for texture representation. Emphasis will be given to the study of the two above subject and not in the design of the classifier.

2. Background and related work

Besides color and shape, texture plays an important role in the human visual system to recognize and categorize objects and properties in several kinds of images, from natural and artificial color images, to medical, remote sensing and quality control ones. It is proven to be an important visual property of the materials, encountered in many low-level image analysis and computer vision tasks. The study of texture is recognized to be a difficult subject in image science, while texture classification is a central research direction that has a wide variety of applications. In order to build a texture-based classification system, the basic building elements are first robust texture representation, then the design of a (dis)similarity measure between textures and finally the choice of the classifier.

Texture representation is a difficult problem due to the high – and usually unknown – true dimensionality of the feature space required to represent textures. Texture is defined as a homogeneous and coherent field of an image, characterized by features like roughness, variability, repeatability, directionality etc., which are characterized over a certain spatial extend. The preferred approach for texture feature extraction by the majority of researchers is based on image decomposition, by filtering with a subband or wavelet filter bank (Laine & Fan, 1993; Randen & Husoy, 1999; Leung & Malik, 2001; Do & Vetterli, 2002). The image is decomposed into several images for separate processing. The goal is to concentrate the involved energy in a few features and reduce the correlation. An analogous move is to apply a linear transformation by using Fourier or DCT. Older techniques that use direct image domain representation or express co-occurrence properties of image pixels are not commonly used lately. The most suitable representation for summarizing a nonparametric estimate of texture distribution is histogram, since texture is considered to describe the appearance of a region by the distribution of features rather than by some individual feature vectors (Ojala et al., 1996; Rubner et al., 2001). However, histogram is bound with the well-known binning problem which is a difficult one to solve. Regular binning of high-dimensional feature spaces results poor performance, coarse binning affects resolving power, while fine binning leads to large fluctuations due to the statistically insignificant sample sizes for most bins. In the literature, adaptive binning has been proposed to tackle the binning effect (Leow & Li, 2004), as well as binning induced by a set of prototypes.

The other important issue of a texture-based classification system is the design of an appropriate distance measure between textures. Towards this objective, several methods have been proposed based on histogram comparison. An alternative approach to measure texture resemblance is by means of non-parametric statistical tests that make no a-priori assumptions about the underlying sample distribution. This guarantees the similarities to be assessable in terms of statistical significance, but avoids direct statistical parameter estimation. Non-parametric and distributional based classification methods share a common characteristic; they require the availability of a number of independent – identically distributed – samples from the underlying distribution to operate on. These samples are extracted from the involved data during the texture representation stage. In the core of the proposed technique lies a non-parametric test dealing with the "Multivariate Two-Sample Problem", which has been adopted here for expressing texture image similarity. The specific test is a multivariate extension of the classical Wald-Wolfowitz test (WW-test) and compares two different vectorial samples by checking whether they form different branches in the overall minimal-spanning-tree (MST) (Friedman & Rafsky, 1979). The output of this test can be expressed as the probability that the two point-samples are coming from the same distribution. Its great advantage is that no a-priori assumption about the distribution of points in the two samples is a prerequisite. In order to enrich the evaluation results of texture image classification, another distributional distance is also utilized in this work to measure the dissimilarity between the extracted feature distributions, namely the Kantorovich-Wasserstein (KWass) distance (Gibbs & Su, 2002). It should be noticed that the KWass distance is equivalent to the well known Earth Movers Distance (EMD) (Rubner et al., 2001), which is an optimized solution to the transportation problem, when the later is applied on distributions with equal masses signatures.

Before examining in some detail the application of the above described methodology on a texture classification problem, it is interesting to discuss the importance of features' dimensions in an image classification problem. In general, individual image pixels are characterized by the grayscale value, which overall describe the image in a low-dimensionality feature space and can classify a set of images based on the first-order distribution of pixels' intensities. An image has many pixels and there are a large number of samples available to estimate the distribution. The texture case is quite different. For grayscale textures, instead of single pixels, a neighborhood needs to be considered as the basic texture distribution element, so as to account for pixel intensity correlations inside it. The size of the neighborhood is application dependent and should be large enough to encompass significant texture variation. Correlations among pixels can be accounted by increasing the space dimension required for sample representation, to equal the number of included pixels. Two are the basic implications of the above discussion; the expansion in space dimension that increases computational complexity and the limited sample size which in turn influences the classification error.

The success of the previously described methodologies for feature extraction and distributional similarity estimation were tested on a part of the OUTex and the Photometric texture database. The classification problem is stated as follows; given a new texture sample, assign it to the most similar one of a predetermined set of texture classes. Results should be in accordance with the intuitive notion of visual similarity of the different textures. Special effort is taken to judge all techniques under equal terms and use the available databases in an optimal way. Regarding texture feature extraction, four different methodologies were

incorporated in this study which are considered as golden standard in the scientific community: wavelet transform, DCT, Gabor filters and edge histogram approach proposed in the MPEG-7 standard. These methodologies are shortly analyzed in the following section.

3. Texture feature extraction techniques

3.1 Wavelet transform

A compact representation of image texture needs to be derived in the transform domain for classification (Sebe & Lew, 2000). In the general case, the wavelet transform is applied to a given image in N decomposition levels, decomposing each level in four independent and spatially oriented channels, producing in this way the subbands LL, LH, HL and HH. For the texture feature extraction, the image is partitioned into M non-overlapping square blocks of specific size, and the wavelet transform is applied to each block. The subbands LH and HL are mixed via the type:

$$LHHL_n = \sqrt{LH_n^2 + HL_n^2} \ , \ n = 1, 2, ..., N \tag{1}$$

producing the $LHHL$ subband, for each level of decomposition n. In our study, we used the ordinary number of $N = 3$ decomposition levels, which has prevailed as a kind of standard practice in the literature. The block's dimensions are relative to the size of the texture pattern embedded in a given image and can take typical values of 8×8 or 16×16.
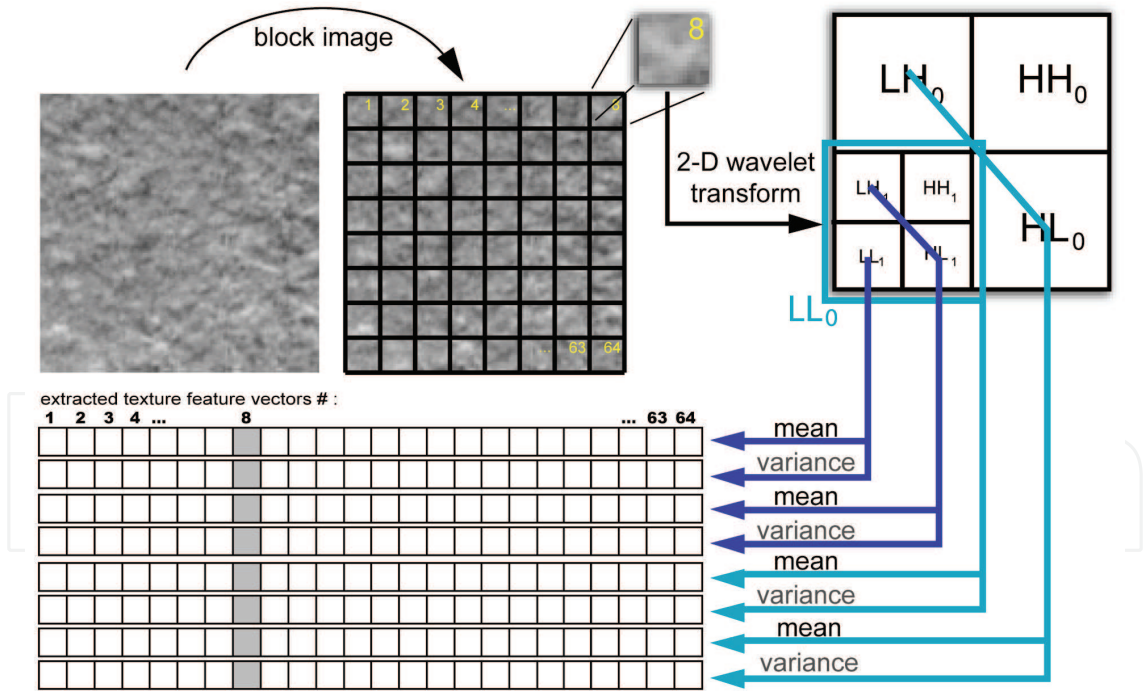

Fig. 1. Texture feature extraction using wavelet transform. Each block of the texture image is decomposed and the mean and variance values are calculated over the LL_n and $LHHL_n$ subbands.

For each square patch, the mean $\mu_{[k,m]}$ and the variance $\sigma_{[k,m]}$ of the energy distribution of the transform coefficients are calculated, as presented in (Pothos et al., 2007). Grouping

together the mean and variance values of the same $m-$th block from all the subbands, a vector of $2 \times (2 \times N)$ coefficients is produced that describes the local texture information of the specific block. In general we will have a set of M such vectors that best describe the texture information of the entire image. A representative scheme that sketches the above procedure is illustrated in Fig. 1.

3.2 Discrete Cosine Transform (DCT)

The Discrete Cosine Transform (DCT) has been widely used in the literature for efficient texture feature selection. It uses cosines of varying spatial frequencies as basis functions and is commonly known for its implementation in the JPEG compression standard (Bhaskaran & Konstantinides, 1995). The DCT coefficients are obtained covering different spectral bands. For texture images, much of the signal energy lies at low-frequency components, which appear in the upper left corner of the DCT. Knowing that DCT converts the spatial information into the frequency domain, texture features can be defined as the spectrum energies in different localizations of a local block. Since the DC coefficient represents (almost) the average grayscale value of each $N \times N$ block, it is not considered to carry any texture information. The remaining AC coefficients capture the details - or frequency and directionality properties - within the pixel-block and therefore can be considered to characterize image texture and be utilized as texture features.
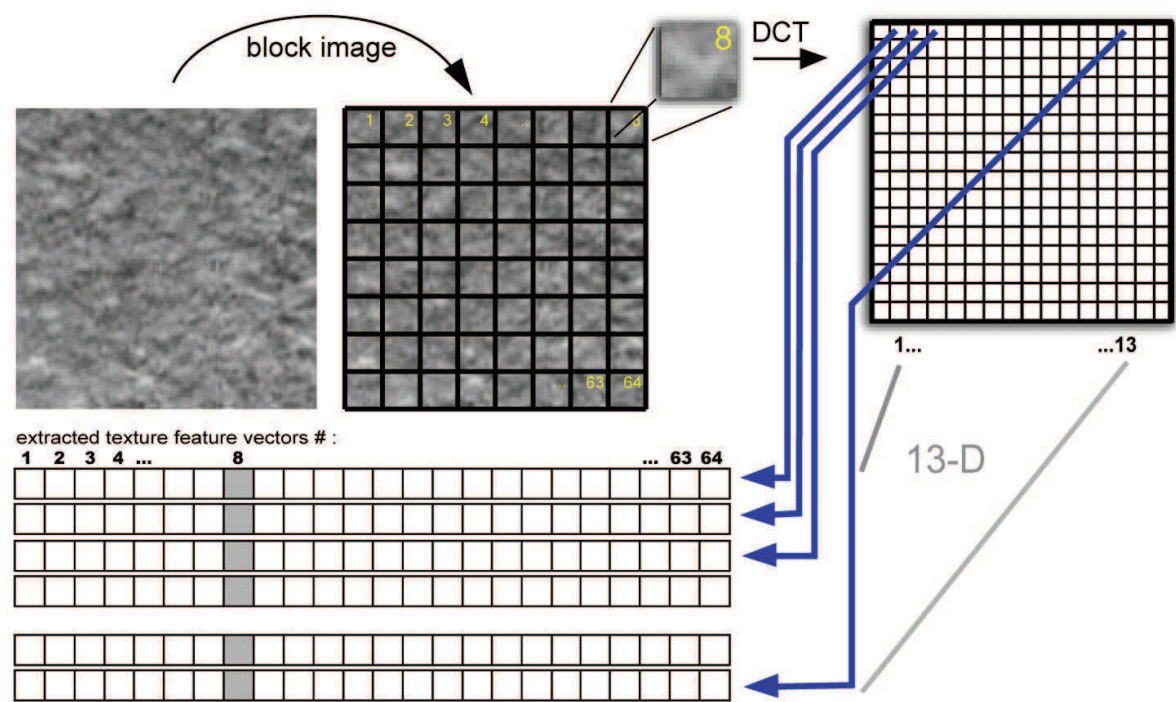

Fig. 2. Texture features extraction using DCT. DCT is applied on each block of the texture image and a feature vector is creating via summing up the square values of the coefficients in the diagonals of the block.

In order to extract textural attributes, the images are initially partitioned into $N \times N$ pixel-blocks, with $N = 16$ in our case. The block size was selected in order to reduce the number of extracted feature vectors and also try to effectively capture the texture information using

a larger image patch. In addition, it was experimentally verified to produce enhanced classification results compared to a smaller pixel-block (e.g., $N = 8$). Then, the DCT is applied to each distinct block, as illustrated in Fig. 2. From each DCT block, texture can be now represented by a feature vector V_m, with $m \in [1, 2N - 2]$, the elements of which are the square sums of coefficients of the corresponding diagonals (i.e., zig-zag traversal lines). The vector resulting from the zig-zag ordering contain all the AC coefficients starting from the upper left location (i.e., $(0, 1)$) to the bottom right (i.e., $(N - 1, N - 1)$). Assuming that a given image is initially divided into M blocks of 16×16 pixels, then a set of M feature vectors can be extracted that best describes the texture image content of the particular image. The specific indexing scheme was found to be robust, when similarity-based image rotation is considered (Theoharatos et al., 2006).

3.3 Gabor filters

The relation between the human vision system and the Gabor filters is a strong motive to test Gabor filtering for texture feature extraction. Spatially, a Gabor function is a Gaussian modulated sinusoid. In his work (Daugman, 1985), Daugman generalized the Gabor function to the model of the 2-D form:

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\pi\left[\frac{(x-x_0)^2}{\sigma_x^2} + \frac{(y-y_0)^2}{\sigma_y^2}\right]} e^{i(\xi_0 x + v_0 y)} \tag{2}$$

where (x_0, y_0) is the center of the field in the spatial domain and (ξ_0, v_0) is the optimal spatial frequency of the filter in frequency domain. σ_x and σ_y are the standard deviations of the elliptical Gaussian for axis x and axis y respectively.

In order to extract texture information, we first partitioned the images into M non-overlapping blocks. Then, the Gabor filters were applied using 4 scales and 6 orientations, creating $N = 24$ filtered subimages. These subimages are obtained by computing the magnitude from the real $G_{\Re_n}$ and imaginary $G_{\Im_n}$ parts of each n subbands:

$$G_n = \sqrt{G_{\Re_n}^2 + G_{\Im_n}^2}, \; n = 1, 2, ..., N \tag{3}$$

Mean and variance are calculated by G_n for each one of the N-filtered subimages. In (Zhang et al., 2000), a $2 \times N - D$ multidimensional vector is constructed such that to be used for similarity matching using a valid (dis)similarity measure (i.e., the sum of Euclidean distances). In this study, a 2×24 dimensional feature vector is built for the description of the texture information, corresponding to the mean and variance values per filtered subimages that are contained in each corresponding block. In the final stage, a total number of M feature vectors of $48 - D$ is constructed for the description of the texture information of all database images. A representative scheme of the previously reported technique is illustrated in Fig. 3, clearly sketching the overall procedure.
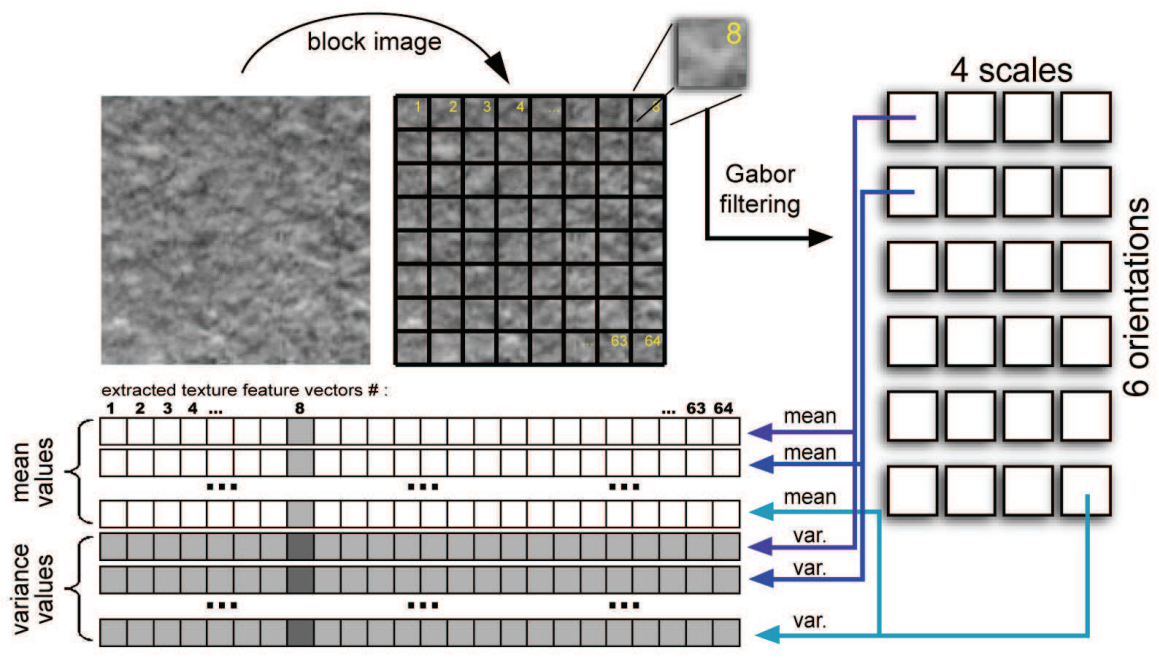
Fig. 3. Texture features extraction using Gabor filters. Each image is blocked into M blocks, and from each block 48 values are extracted, after applying 24 Gabor filters.

3.4 Edge Histogram descriptor

The Edge Histogram is a useful texture descriptor that captures the spatial distribution of edges in the image and is defined in MPEG-7 for similarity search and retrieval practices (Manjunath et al., 2002). The local distribution of edges is considered a reliable candidate attribute and useful for image to image matching, even when the underlying texture structure is not homogeneous in texture properties. The extraction of textural features is estimated as follows. In the first step, the image is divided into 4×4 subimages and the local-edge distribution for each subimage is represented by a histogram. To do this, edges in the subimages are categorized into five types: vertical, horizontal, 45° diagonal, 135° diagonal and nondirectional (or isotropic). Therefore, each image is comprised of 16 subimages with five bins each (corresponding to the above five category-types). The overall histogram is composed of $16 \times 5 = 80$ bins.

Edge detection and classification in each subimage can be done by further dividing those subimages into non-overlapping square blocks (e.g., into a number of 2×2 pixel-images) and afterwards apply appropriate oriented edge detectors (including four directional selective detectors and one isotropic operator (Manjunath et al., 2001)) to compute the corresponding edge strengths. If the maximum edge strength of these oriented edge detectors is found above a given threshold, then the corresponding edge orientation is associated with the image-block which is considered to be an edge-block. If not, then the image-block is not classified as edge-block. Finally, the edge blocks that result contribute to the appropriate binning procedure of the histogram descriptor, with each bin value normalized to the total number of image-blocks in the subimage (i.e., $[0, 1]$). The individual algorithmic steps of the local-edge histogram descriptor are summarized in Fig. 4, with $h(i)$ denoting the overall histogram comprised of 80 bins.
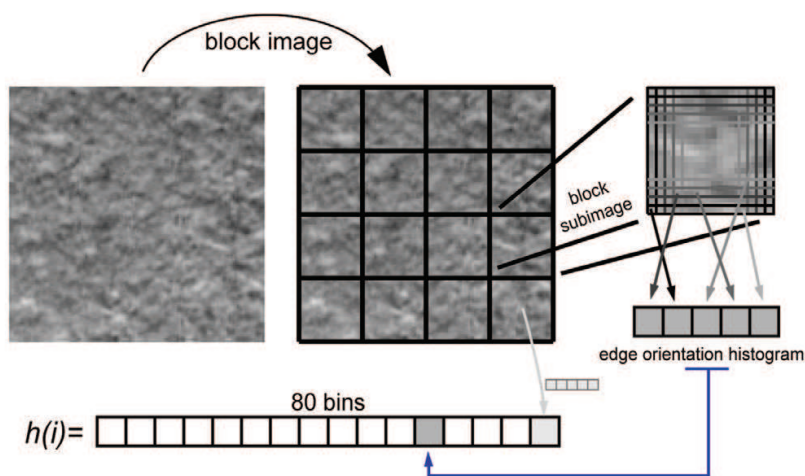
Fig. 4. Construction of local-edge histogram $h(i)$.

However, the local-edge histogram is not sufficient enough for effective image matching. For this reason, global-edge distributions are used in association to local-based ones (Manjunath et al., 2002). In this way, global- and semiglobal-edge histograms are produced respectively, computed directly from the 80 local histogram bins. Regarding the global-edge histogram, the five types of edge distributions in all subimages are accumulated. For semiglobal-edge histograms subsets of subimages are grouped as shown in Fig. 5. The combination of all those distinct histograms, produce a histogram of 150 bins. For matching the extracted features, MPEG-7 defines the $L_1 -$ norm $D(A,B)$ as the distance measure for comparing two image histograms A and B, using the following formula:

$$D(A,B) = \sum_{i=0}^{79} |h_A(i) - h_B(i)| + 5 \times \sum_{i=0}^{4} |h_A^g(i) - h_B^g(i)| + \sum_{i=0}^{64} |h_A^S(i) + h_B^S(i)| \tag{4}$$

where $h_A(i)$ and $h_B(i)$ represent the normalized edge histogram bin values of image A and image B respectively. In the above equation, $h_A^g(i)$ and $h_B^g(i)$, as well as $h_A^S(i)$ and $h_B^S(i)$ represent the normalized bin values for the global-edge and semiglobal-edge histograms respectively, of consecutive images A and B.
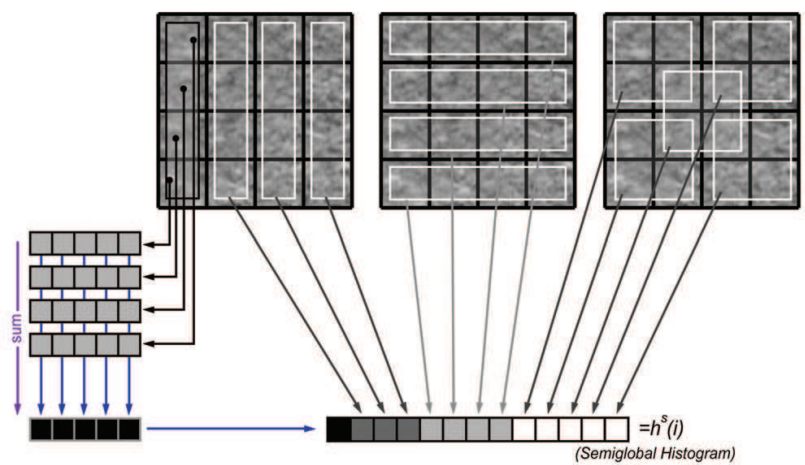

Fig. 5. Construction of semi-global edge histogram $h^S(i)$.

In order to measure the similarity of the different texture images using distributional (dis)similarity measures, the extracted edge histogram data need to be re-arranged into a form of multiple vectors, instead of a histogram. We perform the multivariate nature our distributional measures, by operating in five dimensions as follows. Firstly, we calculate separately each edge histogram (see Fig. 4) from each subimage of the local-edge histogram, resulting in 16 vectors of five dimensions (i.e., $5-D$) each. Additionally, an extra one $5-D$ vector is computed from the global-edge histogram. Finally, we utilize the 13 parts of the semiglobal-edge histogram, in order to create another 13 vectors of $5-D$ each. The ensemble of these practises produce a total of 30 vectors in $5-D$ and thus the distributional (dis)similarity measures can now straightforwardly applied for texture matching over the selected attributes of two separate images. We note here that, each dimension of the formed feature space corresponds to the normalized number of edges that are found to be vertical, horizontal, diagonal, diagonal and nondirectional.

4. Distance between distributions

In order to compare the extracted spectral representations in pairs and therefore estimate the similarity between texture images, we resorted to the field of multivariate statistics for choosing proper distributional-based measures. There exists a rich literature on probability distribution distance measures (Do & Vetterli, 2002; Theoharatos et al., 2006; Rubner et al., 2001; Gibbs & Su, 2002), the choice of which is relatively influenced by a number of inherent parameters of the data distribution. These include the amount of existing data, the dimensionality of the resulted feature space and the distributional structure of the selected multidimensional vectors. In applications like texture classification, face recognition, fingerprint identification etc., intrinsic representations that constitute characteristic manifolds coming from the corresponding high-dimensional data distributions are realized. The adaptation of valid/proper (dis)similarity measures capable for capturing these structures and, moreover, performing pairwise comparisons between pairs of suchlike distributions is the key factor for successfully assessing huge data archives. Several well known distance measures were examined towards achieving the above goal. In our analysis that follows, we present results for the best (and simultaneously more appropriate) two ones, i.e. the multivariate Wald-Wolfowitz test (WW-test) (Friedman & Rafsky, 1979) and the Kantorovich-Wasserstein distance (Gibbs & Su, 2002). In their basic implementation, both measures utilize the L_2-norm for calculating the ground distance between sample vectors. For comparison reasons, histogram-based measures are also utilized in several experimental trials such as Histogram Intersection (HI), Kullback-Leibler Divergence (KL-D) and Jeffrey Divergence (J-D).

4.1 The multivariate Wald-Wolfowitz test (WW-test)

As an important constituent of the introduced framework, a nonparametric test is adopted for estimating texture content similarity in a reliable and convenient way. It is a multivariate extension of the classical Wald-Wolfowitz test, comparing two different sets of points in $\Re^p$ by checking whether they form different branches in the overall MST-graph (Zahn, 1971). The output can be expressed as the probability that the two samples are coming from the same distribution (Friedman & Rafsky, 1979).

In the multivariate case, the graph is built over points in $\Re^p$: a single node corresponds to every given point and the weight associated with every possible edge is the corresponding interpoint norm (i.e., ground distance used in its basic implementation). The edges involved in the construction of MST are the ensemble of straight-line segments connecting all points with minimum total length. WW-test can be used to test the hypothesis $\mathbf{H_o}$, whether any two given multidimensional point samples $\{X_i\}_{i=1:m}$ and $\{Y_i\}_{i=1:n}$ come from the same multivariate distribution. At first, the two data samples of size m and n are considered, respectively, from distributions defined in $\Re^p$. Then, the sample identity of each point is not encountered and the MST of the overall sample is constructed. Based on the sample identities of the points, a test statistic R is computed that is defined as the number of disjoint subtrees that finally result. Rejection of $\mathbf{H_o}$ is for small values of R. It has been shown that the quantity:

$$W = \frac{R - E[R]}{\sqrt{Var[R \,|\, C]}} \tag{5}$$

approaches the standard normal distribution, while the mean $E[R]$ and variance $Var[R\,|\,C]$ of R are given in closed form (Friedman & Rafsky, 1979). Its importance is that using simple formulae, the significance level (and p-value) for the acceptance of the hypothesis $\mathbf{H_o}$ can be readily estimated. WW-test, based on the MST-planning procedure, offers a unique environment for contrasting different signal representations. Thus, it can effectively cope with the understanding and matching of manifold-type structures, which is actually the case of the vectorial spectral representations of the texture images under study.

4.2 Kantorovich-Wasserstein distance

The Kantorovich-Wasserstein (KWass) metric defines a "distance" between two stochastic distributions. It is described by the formula:

$$d_w(\mu, \nu) = \inf_J \{\mathbf{E}[d(X, Y)] : \mathrm{L}(X) = \mu, \mathrm{L}(Y) = \nu\} \tag{6}$$

where the infimum is taken over all joint distributions J with marginals μ and ν (Gibbs & Su, 2002).

For discrete distributions X, Y with samples of the same size $X = (x_1, x_2, \ldots, x_n)$ and $Y = (y_1, y_2, \ldots, y_n)$, the minimum is taken over all permutations. It is common to use the Hungarian algorithm in order to solve the optimal assignment problem (Levina & Bickel, 2001).

4.3 Distances between histograms

In order to provide a short description of the other methods employed here for comparison purposes, let $H = \{h_i\}$ and $K = \{k_i\}$ be histograms from two texture images H and K to be compared respectively, each containing n bins. A plethora of measures have been reported in the literature for calculating the distance between histogram distributions (Rubner et al., 2001). Here, three characteristic measures are utilized that are commonly used by other researchers.

Histogram Intersection (HI)

It was originally proposed for color image retrieval (Swain & Ballard, 1991) in the spatial domain and is found to be attractive due to its ability to handle partial matches (Rubner & Tomasi, 2001). The HI-measure is given by:

$$d_{HI}(H,K) = 1 - \frac{\sum_{i=1}^{n} \min(h_i, k_i)}{\sum_{i=1}^{n} k_i} \tag{7}$$

Kullback-Leibler Divergence (KL-D)

It measures how inefficient on average it would be to code one histogram using the other as the code-book (Rubner et al., 2001):

$$d_{KL}(H,K) = \sum_{i=1}^{n} h_i \log \frac{h_i}{k_i} \tag{8}$$

Jeffrey Divergence (J-D)

This is a modification of the KL-D that is symmetric, numerical stable and robust with respect to noise and size of histogram bins (Rubner & Tomasi, 2001), given by:

$$d_{JD}(H,K) = \sum_{i=1}^{n} \left(h_i \log \frac{h_i}{m_i} + k_i \log \frac{k_i}{m_i} \right), \tag{9}$$

where $m_i = h_i + k_i/2$.

5. Experimental analysis

For our experimental analysis two texture databases were utilized: the OUTex (University of Oulu Texture database) and the Photometric texture database. The first dataset contains 24 distinct texture categories, having 180 grayscale images of similar size in each class and thus resulting in a total number of 4320 texture images. The amount of images comprising a single category is formed by using nine different texture orientation images. From those 4320 images, 216 were selected as queries (i.e., one texture for each orientation and category), while the leave-one-out procedure was used in a $k-\text{NN}$ for the classification procedure. The incorporation of a bigger number of query-images had a very slight impact on the classification results. The second dataset contains 34 different kinds of textures, having 56 images each. For this database, we used an increasing number of images per texture in the database (to train the classifier), while the rest images were utilized as query-images to be classified. Fig. 6 presents characteristic samples from all texture categories for both the OUTex and Photometric texture databases.

In Fig. 7, the classification results are given for the OUTex database using the leave-one-out procedure. The horizontal axis contains the values of k of the simple $k-\text{NN}$ classifier, while the vertical one contains the classification error rate of each method. In all cases, the (distributional-based) multivariate WW-test outperforms significantly all other (histogram-based) measures. The DCT and Gabor feature extraction methods seem to produce the best classification results, while the edge histogram one seems to provide poor outcomes.
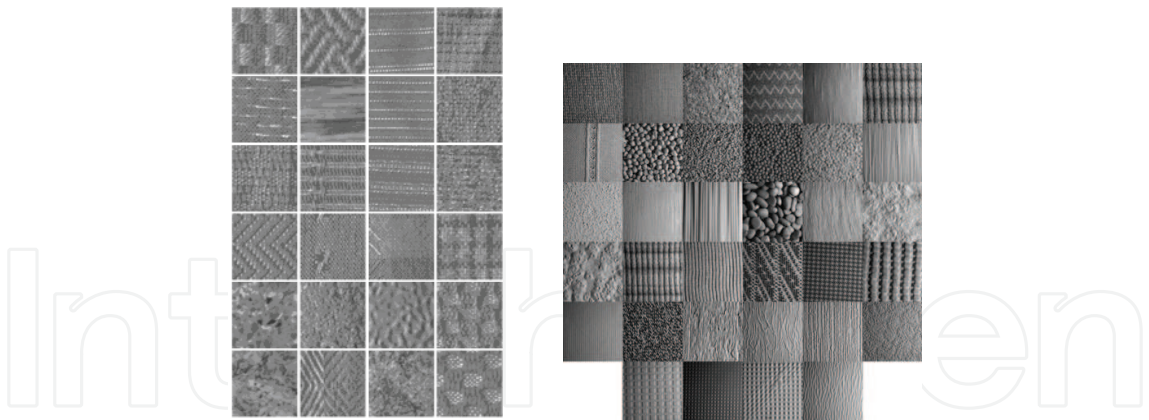
Fig. 6. Sample category images of the utilized texture databases. Left) OUTex database. Right) Photometric texture database.
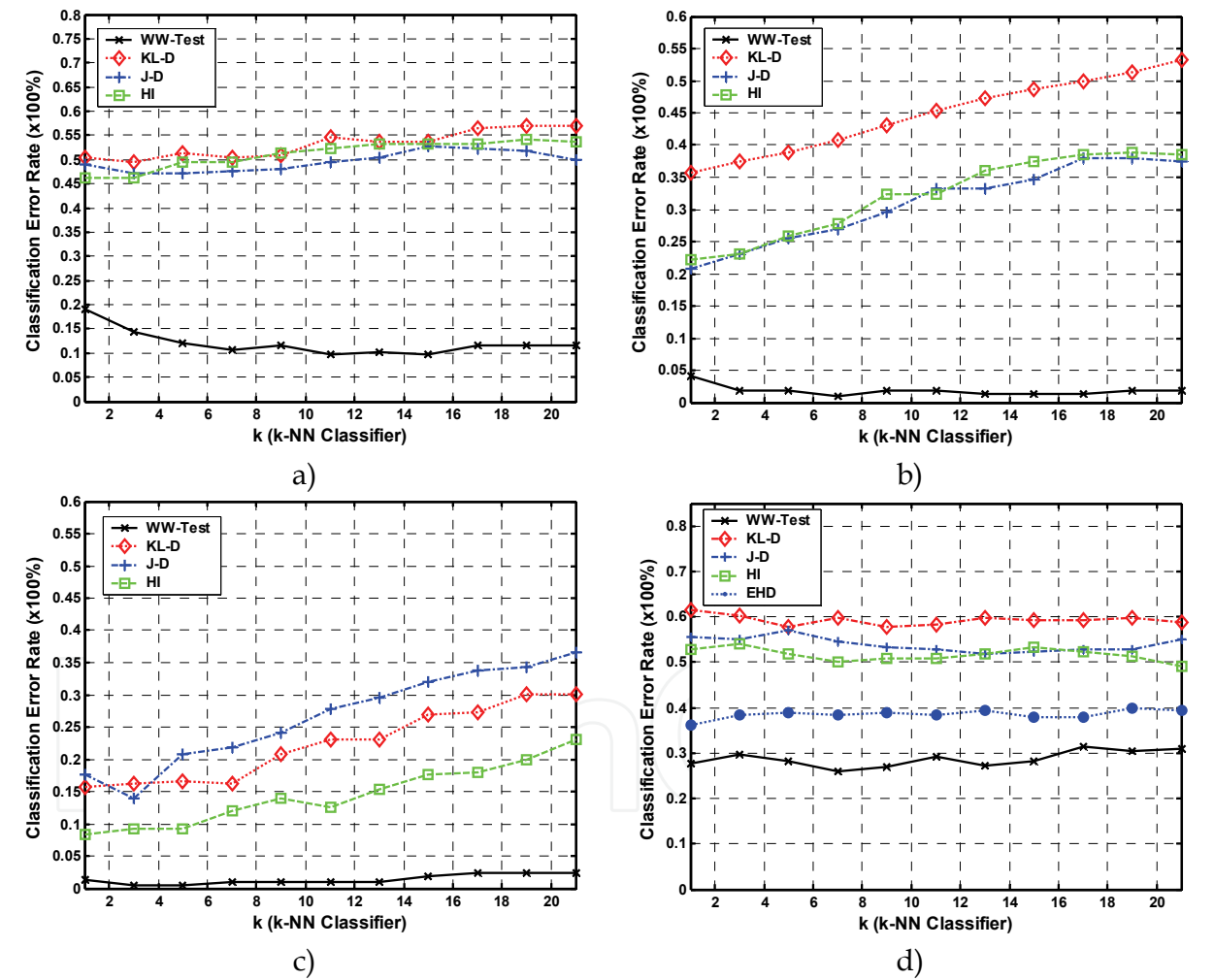

Fig. 7. Classification Error Rate results for a) Wavelets, b) DCT, c) Gabor filters and d) Edge Histogram, for the OUTex database.

The OUTex Database downloaded from: http://www.outex.oulu.fi/temp/
The Photometric Database downloaded from: http://www.taurusstudio.net/research/pmtexdb/index.htm

In Fig. 8 the classification results over the Photometric texture database are exemplified. The horizontal axis contains the number of texture images used as database images for the $k - NN$ classifier. In this way, it is possible to understand how the numbers of images used for "training" the classifier affects the results. In practice, the most important outcome would be in the case of using only 5 training images, which is more close to real application implementations. Wavelets and DCT seems to give best results for a few training images, while Gabor filters and edge histogram seems to have high classification error ratio. The use of many images for training the classifier affects the results, producing low classification error rate for all methods, except for the edge histogram that is in accordance with the results produced using the OUTex database. We have to notice here that, in general, the multivariate WW-test performs slightly better than the KWass measure, when distributional-based measures are considered. In addition, histogram-based measures do not perform adequately when compared to distributional-based ones, in all different texture feature extraction methodologies. This internal comparison is in accordance with the results coming from both texture datasets, as also theoretically expected from our initial analysis and the reports coming from the wide literature.
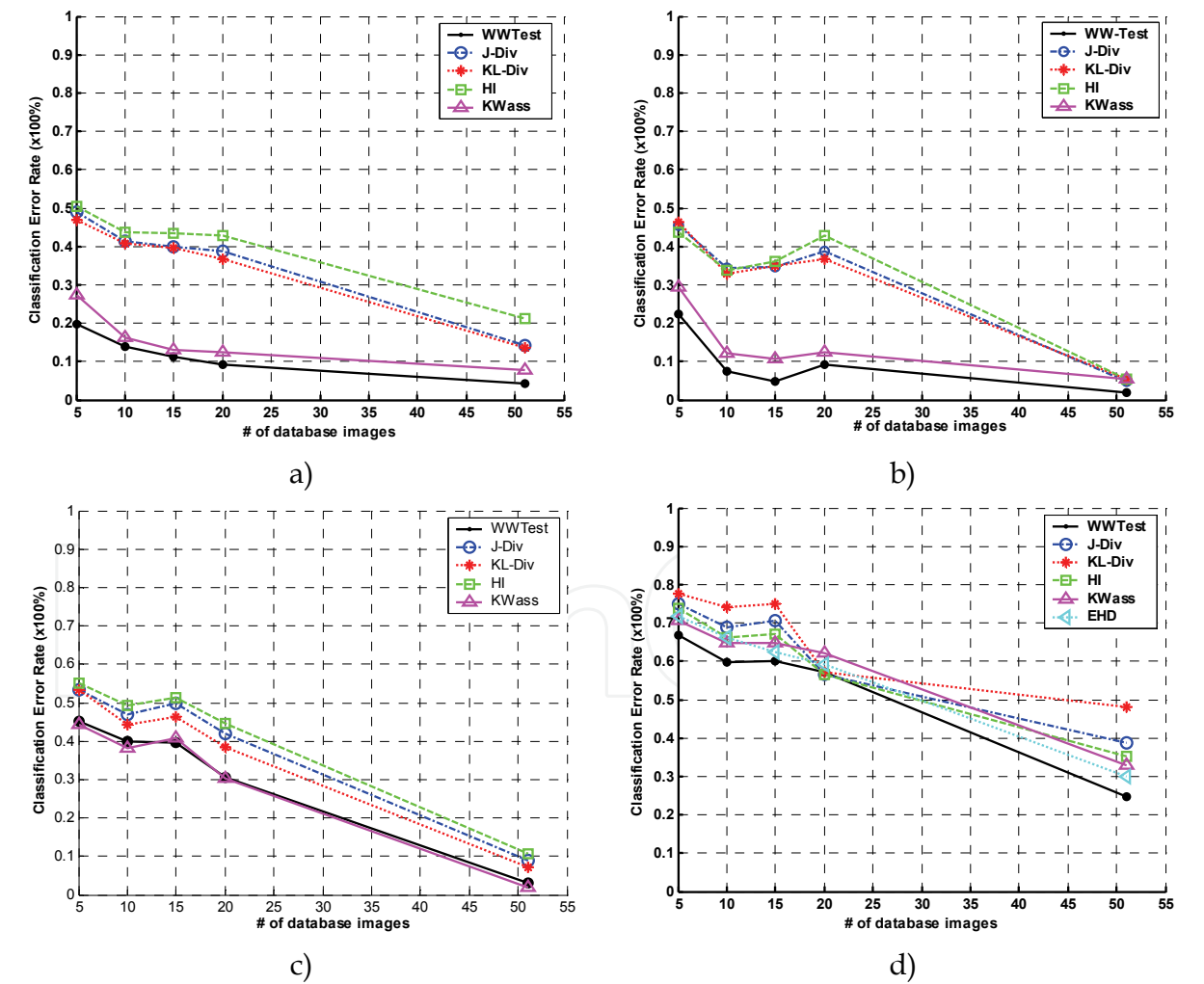

Fig. 8. Classification Error Rate results for a) Wavelets, b) DCT, c) Gabor filters and d) Edge Histogram for different number of training images, for the Photometric database.

6. Boosting the classification performance via Vector Quantization

Vector quantization aims at representing the data with a reduced set of prototype data vectors and thus summarizes the input information while inducing minimal distortion. In the case of texture images, a codebook of k code vectors that summarizes the vectorial representation of the entire spectral information is designed by applying the Neural-Gas algorithm to the data matrix. This algorithm is an artificial neural network model, which converges efficiently to a small, user-defined number of codebook vectors, using a stochastic gradient descent procedure with a "soft-max" adaptation rule that minimizes the average distortion error (Martinez et al., 1993).

The Neural Gas network operates by utilizing first a stochastic sequence of incoming data vectors $X(t)$, which is governed by the distribution $P(\mathbf{X})$ over the manifold $\mathbf{V}$. Then, an adaptation step is performed for adjusting the weights of the k neurons $\{A_j\}_{j=1:k}$:

$$\Delta A_j = \varepsilon \cdot h_\lambda \left(f_j \left(X(t), \{A_i\}_{i=1:k} \right) \right) \left(X(t) - A_j \right), \ j = 1, 2, \ldots, k, \ \forall \, t = 1, \ldots, t_{\max} \tag{10}$$

The function $h_\lambda(y)$ in the above equation has an exponential form $e^{-y/\lambda}$ and $f\left(X, \{A\}_j\right)$ is an indicator function that determines the 'neighborhood-ranking' of the reference vectors according to their distance from the input vector $\mathbf{X}$, while for both parameters ε and λ an exponential decreasing schedule is followed, with $t_{\max}$ being the final number of adaptation steps that can be defined from the data based on simple convergence criteria.

The asymptotic density distribution of the codebook vectors $P(\mathbf{A})$ was proved, mathematically (Martinez et al., 1993), to be proportional to the data density distribution $P(\mathbf{A}) \propto P(\mathbf{X})^{\underline{d}/(\underline{d}+2)}$ where $\underline{d} \leq d$ is the intrinsic dimensionality of the input data. This theoretical evidence along with the accompanying experimental evidence (Martinez & Schulten, 1994) motivated our conjecture that the designed codebook could serve as a faithful representation of the vectorial distribution in color-space and therefore could be utilized in the subsequent comparisons regarding color content. The relationships between filter responses are encoded in the joint multivariate distribution and provide unique information about the textural structure. To reduce the complexity of the classification problem, texture distribution comparisons are carried out in pairwise fashion using the distribution-distance measures presented before to test the efficiency of the sequential algorithmic procedure.

In order to evaluate the classification accuracy of our procedure when incorporating the Neural-Gas based vector quantization scheme, the Photometric database was utilized. In our experiments, 4 training sets of images were used for the classifier, with each one containing 5, 20, 35 and 51 database images respectively per class, while the rest ones were used as queries. Experiments took place for Wavelets, DCT and Gabor methods, which were proven to produce the best classification results. Each image was first partitioned into several overlapping blocks. For each block, all three feature extraction methods were applied, as previously explained. Due to the block-overlapping nature of our procedure, the extracted texture signatures are comprised of a large number of feature vectors. The Neural-Gas algorithm is next used to select the most representative ones that best describe the texture feature distribution, boosting in this way the classification results. In Fig. 9 the general scheme of using Neural-Gas is presented.
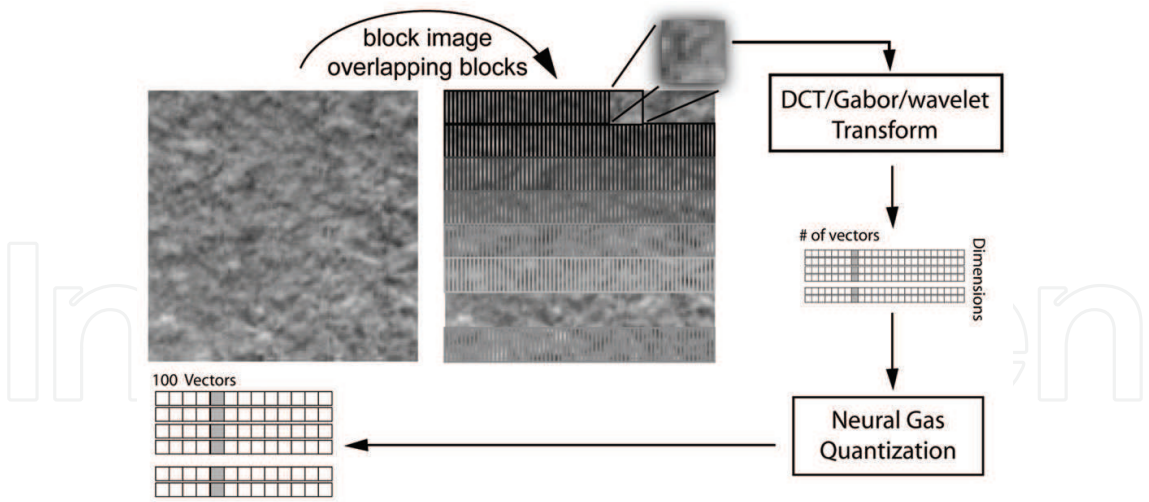
Fig. 9. Overlapping blocks produce a large number of feature vectors. Neural-Gas is used to select the most representative ones that best describe the image texture.
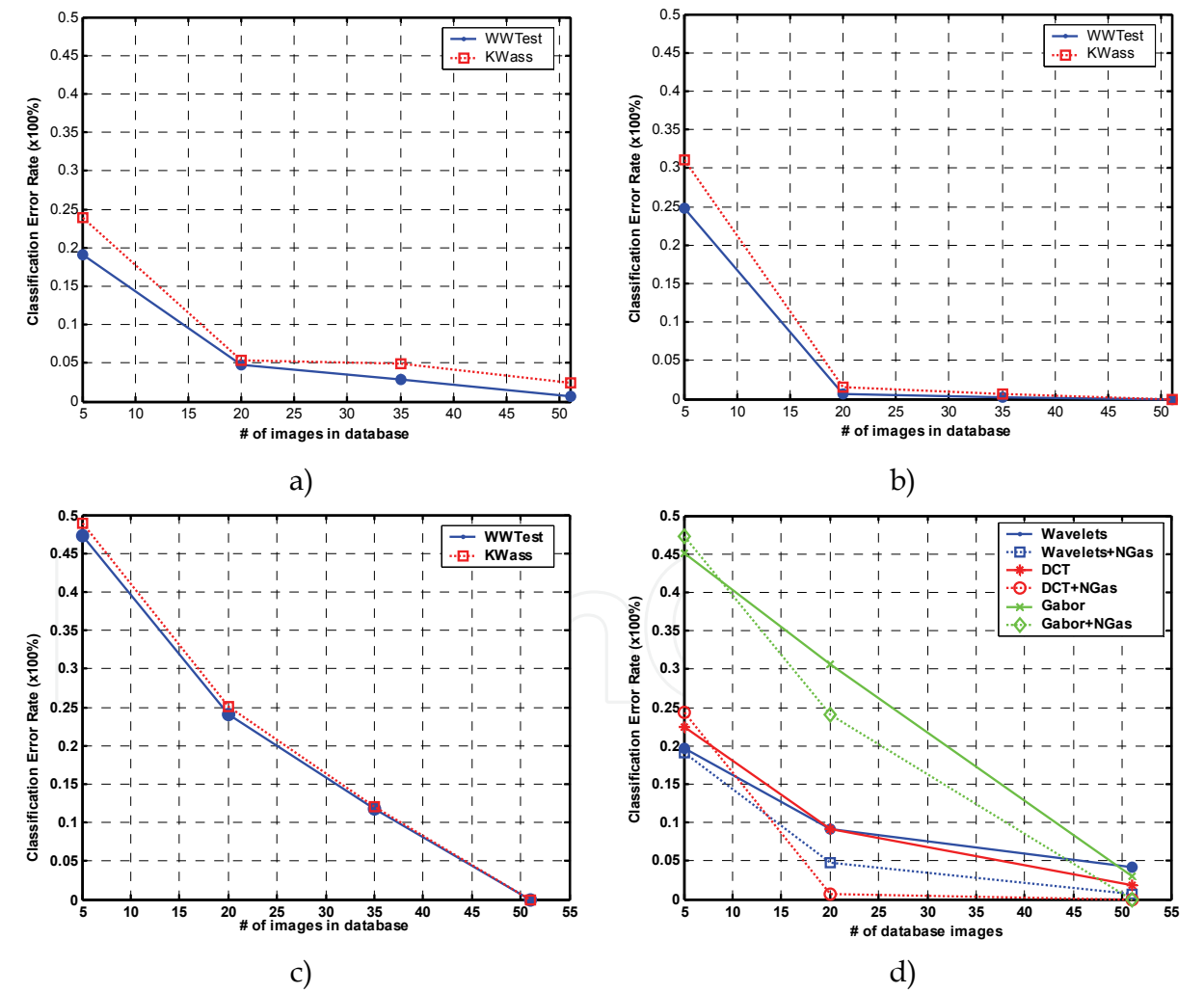

Fig. 10. Experimental results using the Neural-Gas algorithm for a) Wavelets, b) DCT, c) Gabor filters. d) Comparative results of the methods with and without Neural-Gas.

In Fig. 10, the results obtained utilizing the Neural-Gas are presented. When the classifier is trained with a few texture images, the Neural-Gas procedure seems to slightly increase the classification error. In contrast, as the number of database samples is increased, the Neural-Gas procedure boosts the accuracy of the classification quite enough.

7. Conclusion

A robust methodology is presented in this chapter that tries to tackle the texture classification problem. The multivariate WW-test and the KWass distance are used in order to measure how close two distributions are. Their generic character stems from the fact that, by altering the character of texture image characteristics, we can modify the flavour of formulated queries. To avoid problems associated with histograms as empirical estimates of the distribution (e.g., the binning effect), particularly in high-dimensional spaces, dissimilarity between texture distributions is computed using distributional-based multivariate analysis. These multidimensional measures can adequately operate even with a small number of distributional samples and is well suited for texture matching.

Individual texture samples were extracted from the images of the OUTex and the Photometric texture databases, by partitioning the image into regions of almost homogenous texture content. The intrinsic dimensionality of the texture regions is computed by means of image decomposition implementing some of the well-established techniques. Depending on the technique, the joint spectral distribution is sampled in the form of multivariate vectors. The efficiency in textural feature extraction of the different methods, as well as the competence of the above measures in distributional texture image representations, was tested with quite satisfactory results, which yield future ideas for research and application. In addition, a neural-network based vector quantizer was adopted in order to further boost the classification accuracy of the introduced methodologies. Each texture image distribution was summarized by a vector quantization scheme in order to select a restricted number of prototype code vectors, thus resulting in a sampled representation of the original spectral distribution. These code vectors played the role of a spectral signature for each texture image, capturing its basic structure and providing a sparse-compact representation.

As a scheduled extension of our work, the application of the introduced methodology to color or multispectral images for texture classification can be straightforwardly implemented, by placing other subband's texture information to higher dimensions in the feature space. However, special care has to be taken to the number of extracted feature vectors in the case of multispectral images, due to the "curse of dimensionality" (Costa & Hero, 2004). In addition, more has to be done in order to overcome the problem that arises from the weakness to capture texture patterns that have different plain scales. This might be accomplished by utilizing adaptive scalable blocks, by means of – each time – different sized block patches inside the input images. In this way, the texture information can be efficiently captured in different scales, making use of possible regions of interest. Moreover, texture segmentation algorithms could be incorporated towards this solution, among the plethora that is available in the literature the recent years (Jain & Farrokhnia, 1991; Manjunath & Chellappa, 1991). Finally, the use of alternative ways to extract texture information based on high-level features (i.e., semantic-based texture attributes), so as to correlate better with the human perceptual inspection, is of crucial importance (Liu et al., 2007), as well as the potential application of combining other primitives in the feature extraction methodology, such as color and shape.

8. Acknowledgements

9. References

Bhaskaran, V. & Konstantinides, K. (1995). *Image and video compression standards, algorithms and architectures,* Kluwer Academic Publishers, ISBN: 0-7923-9952-8, Norwell, Massachusetts USA

Costa, J. & Hero, A.O. (2004). Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Trans. on Signal Processing,* Vol. 52, No. 8, pp. 2210-2221

Daugman, J.G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A,* Vol. 2, No. 7, pp. 1160-1169

Do, M.N. & Vetterli, M. (2002). Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. on Image Processing,* Vol. 11, No. 2, pp. 146-158

Friedman, J.H. & Rafsky, L.C. (1979). Multivariate generalizations of the Wald-Wolfowitz and Smirnov two-sample tests. *Annals of Statistics,* Vol. 7, No. 4, pp. 697-717

Gibbs, A. & Su, F. (2002). On choosing and bounding probability metrics. *International Statistical Review,* Vol. 70, No. 3, pp. 419-435

Jain, A.K. & Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition,* Vol. 24, No. 12, pp. 1167–1186

Laine, A. & Fan, J. (1993). Texture classification by wavelet packet signatures. *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 15, No. 11, pp. 1186–1190

Leow, W.K. & Li, R. (2004). The analysis and applications of adaptive-binning color histograms. *Computer Vision and Image Understanding,* Vol. 94, No. 1-3, pp. 67–91

Leung, T. & Malik, J. (2001). Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision,* Vol. 43, No. 1, pp. 29-44

Levina, E. & Bickel, P. (2001). The earth mover's distance is the Mallows distance: some insights from statistics. *Proceedings of IEEE Int. Conf. on Computer Vision,* pp. 251-256, Vancouver, Canada, July 2001

Liu, Y.; Zhang, D.; Lu, G. & Ma, W.Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recognition,* Vol. 40, No. 1, pp. 262-282

Manjunath, B.S. & Chellappa, R. (1991). Unsupervised texture segmentation using Markov random fields models. *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 13, No. 5, pp. 478–482

Manjunath, B.S.; Ohm, J.R.; Vasudevan, V.V. & Yamada, A. (2001). Color and texture descriptors. *IEEE Trans. on Circuits and Systems for Video Technology,* Vol. 11, No. 6, pp. 703-715

Manjunath, B.S.; Salembier, P. & Sikora, T. (2002). *Introduction to MPEG-7: multimedia content description interface,* Wiley, ISBN: 978-0-471-48678-7, New York

Martinez, T.M. & Schulten, K.J. (1994). Topology representing networks. *Neural Networks,* Vol. 7, No. 3, pp. 507-522

Martinez, T.M.; Berkovich, S.G. & Schulten, K.J. (1993). "Neural-gas" network for vector quantization and its application to time-series prediction. *IEEE Trans. on Neural Networks,* Vol. 4, No. 4, pp. 558-569.

Ojala, T.; Pietikainen, M. & Harwood, D. (1996). A comparative study of texture measures with classification based feature distributions. *Pattern Recognition,* Vol. 29, No. 1, pp. 51–59

Pothos, V.K.; Theoharatos, C.; Zygouris, E. & Economou, G. (2007). Distributional-based texture classification using non-parametric statistics. *Pattern Analysis and Applications, Special Issue on Non-parametric Distance-based Classification Techniques and its Applications,* DOI 10.1007/s10044-007-0083-9

Randen, T. & Husoy, J.H. (1999). Filtering for texture classification: a comparative study. *IEEE Trans on Pattern Analysis and Machine Intelligence,* Vol. 21, No. 4, pp. 291-310

Rubner, Y. & Tomasi, C. (2001). *Perceptual metrics for image database navigation,* Kluwer Academic Publishers, ISBN: 0-7923-7219-0, Norwell, Massachusetts, USA

Rubner, Y.; Puzicha, J.; Tomasi, C. & Buhmann, J.M. (2001). Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding,* Vol. 84, No. 1, pp. 25–43

Sebe, N. & Lew, M.S. (2000). Wavelet based texture classification. *Proceedings of IEEE Int. Conference on Pattern Recognition,* pp. 947-950, Barcelona, Spain, September 2000, IEEE Computer Society Washington, DC

Swain, M.J. & Ballard, D.H. (1991). Color indexing. *International Journal of Computer Vision,* Vol. 7, No. 1, pp. 11-32

Theoharatos, C.; Pothos, V.K.; Laskaris, N.A.; Economou, G. & Fotopoulos, S. (2006). Multivariate image similarity in the compressed domain using statistical graph matching. *Pattern Recognition,* Vol. 39, No. 10, pp. 1892-1904

Zhang, D.S.; Wong, A.W.; Indrawan, M. & Lu, G. (2000). Content-based image retrieval using Gabor texture features, *Proceedings of IEEE Pacifin-Rim Int. Conference on Multimedia,* pp. 392-395, University of Sydney, Australia, December 2000
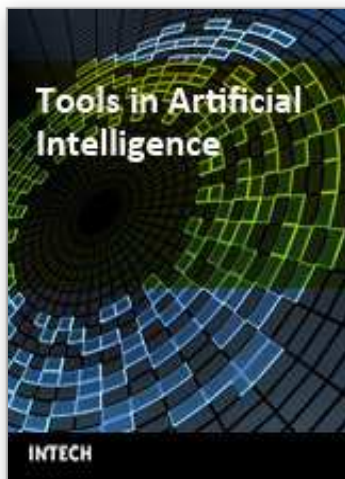
Tools in Artificial Intelligence

Edited by Paula Fritzsche

This book offers in 27 chapters a collection of all the technical aspects of specifying, developing, and evaluating the theoretical underpinnings and applied mechanisms of AI tools. Topics covered include neural networks, fuzzy controls, decision trees, rule-based systems, data mining, genetic algorithm and agent systems, among many others. The goal of this book is to show some potential applications and give a partial picture of the current state-of-the-art of AI. Also, it is useful to inspire some future research ideas by identifying potential research directions. It is dedicated to students, researchers and practitioners in this area or in related fields.

INTECH

open science | open minds