A Generic Scheme for Color Image Retrieval Based on the Multivariate Wald-Wolfowitz Test

Christos Theoharatos, *Member*, *IEEE*, Nikolaos A. Laskaris, George Economou, and Spiros Fotopoulos

Abstract—In this study, a conceptually simple, yet flexible and extendable strategy to contrast two different color images is introduced. The proposed approach is based on the multivariate Wald-Wolfowitz test, a nonparametric test that assesses the commonality between two different sets of multivariate observations. It provides an aggregate gauge of the match between color images, taking into consideration all the (selected) low-level characteristics, while alleviating correspondence issues. We show that a powerful measure of similarity between two color images can emerge from the statistical comparison of their representations in a properly formed feature space. For the sake of simplicity, the RGB-space is selected as the feature space, while we are experimenting with different ways to represent the images within this space. By altering the feature-extraction implementation, complementary ways to portray the image content appear. The reported results, from the application on a diverse collection of images, clearly demonstrate the effectiveness of our method, its superiority over previous methods, and suggest that even further improvements can be achieved along the same line of research. It is not only the unifying character that makes our strategy appealing, but also the fact that the retrieval performance does not increase continuously with the amount of details in the image representation. The latter sets an upper limit to the computational demands and reminds of performance plateaus reached by novel approaches in information retrieval.

Index Terms—Image retrieval, multivariate statistics, sampling, graph-theoretic methods, similarity measures, multivariate visualization.

1 Introduction

During the last years, large collections of digital libraries are being created due to the low cost of digital storage and the rapid growth of computational power. The increasing number of images has opened the need for developing powerful tools for searching through such image databases. Automatic retrieval of an image from a whole data set using representative text-annotation comes from the early 1970s. However, the limitations of this conventional method forced the development of image-based retrieval systems such as IBM's QBIC project [1], PhotoBook system [2], VisualSEEK [3], etc. These systems make use of low-level features such as color, shape, and texture to represent the image content. In addition, region-based retrieval systems such as NETRA toolbox [4], Blobworld system [5], ImageMap [6], Schema project [7], and WALRUS [8], make use of subimage region extraction and corresponding features to locate relevant images. Among others, MPEG-7 standard [9] defines a set of visual descriptors for image content representation designed to meet the requirements of different application domains. Currently, query-by-example search engines are becoming very popular because they provide a convenient way to formulate the inquiry targets [10]. This can be probably set only loosely initially and refined later through relevance-feedback [11] and active learning procedures [12]. For a comprehensive comparative study of content-based retrieval systems, the interested reader is referred to [13].

Considerable research has been carried out on the basis of color content [14]. The most popular representation of color information is global histogram. Statistically, it denotes the joint probability of intensities of the three-color channels, thus describing the global color distribution in an image. In general, the color histogram provides useful clues for the subsequent expression of similarity between images, due to its robustness to background complications and object distortion. Moreover, it is translation, scale, and rotation invariant, very simple to implement and systems encountering histograms exhibit a fast retrieval response that makes real-time implementation easier. A lot of (dis)similarity measures have been proposed for computing the distance between histograms from two different images. These measures can be distinguished into four categories [15]: heuristic histogram distances, nonparametric test statistics, information-theory divergences, and ground distance ones. The innovative work by Swain and Ballard [16], who proposed a *color-indexing* algorithm to identify similar color images using histogram intersection (HI), remains a golden standard.

The main weakness of extracting color histograms globally is that it does not take into account the spatial distribution of color across different areas of the image. A number of methods have been developed for integrating color and spatial information for content-based queries [13], [14]. Since in most applications, complete segmentation implies a great deal of user interaction during database

- *Ch. Theoharatos, G. Economou, and S. Fotopoulos are with the Electronics Laboratory, Department of Physics, University of Patras, Patras 26500, Greece.*
  E-mail: htheohar@upatras.gr, {economou, spiros}@physics.upatras.gr.
- *N. Laskaris is with the Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Artistole University of Thessaloniki, Thessaloniki 54124, Greece. E-mail: laskaris@zeus.csd.auth.gr.*

Manuscript received 19 July 2004; revised 28 Oct. 2004; accepted 10 Feb. 2005; published online 20 Apr. 2005.
For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org, and reference IEEECS Log Number TKDE-0249-0704.

acquisition; this approach is not feasible for large image databases. The standard approach to provide spatial information is to divide the image into multiple regions and extract color descriptors from each of them [17]. Corresponding region-based color descriptors are compared in order to assess the similarity between two images. Gong et al. [18], for example, partitioned the image into nine equal subimages and represented each subregion by a color histogram. In addition, Stricker and Dimai [19] developed a method for splitting each image into five nonoverlapping spatial regions and employ, afterward, color descriptors for each region in order to establish matching procedures. Alternatively, Smith and Chang [20] used color sets to approximate histograms, which correspond to salient image regions and are represented by binary vectors to allow a more rapid search. A totally different approach was introduced by Pass et. al. [21], who partitioned the histogram bins by the spatial coherence of pixels, classifying in this way each pixel of the quantized image based on the associate color coherence vector (CCV). In this direction, Huang et. al. [22] proposed new color features called color correlograms, which include the local color spatial correlation as well as the global distribution of this spatial correlation. Recently, several other techniques have been introduced as well in which different color-related features are used as descriptors, including chromaticity moments based on cumulated [23] and regular [24] histograms, reference color table methods [25], color feature hashing techniques [26], color adjacency graphs [27], and pure chrominance measurements [28]. The plethora of all these techniques has been motivated by the fact that depending on the application at hand and the user's priorities, sometimes we want to optimize performance regardless of speed and/or storage requirements, while some other times suboptimum performance can be tolerated if there are strict time and space limitations.

It was the scope of this work to develop a generic approach to color image retrieval, which could incorporate inquiries at different levels of information complexity (e.g., ranging from the color of individual pixels to the dominant color-variation within individual image-patches). In order to avoid the limitations with the recent conventional methods (histogram binning problems, complicated image representations, and different techniques for spatial color distribution matching), we tried to devise and experimentally justify a potential alternative approach that would represent the color information of each image in a trustworthy format and employ statistical comparisons between color distributions. Among the main objectives was also the achievement of high performance without resorting to complicated image representations, the compatibility with the conventional notion of a color image as a collection of RGB-pixels and the foundation on a mathematically sound theoretical framework. Overall, it was within our intentions to provide a conceptually simple method with flexible/extendable character (that would not limit the application of the method within a specific color space or tight it with a particular prefiltering technique), while keeping the computational load low.

In our study, we followed a hybrid, pattern-analytic, and graph-theoretic approach. The visual content of each color image is described by means of a vectorial distribution in a properly defined feature space. The comparison between two such images incorporates the computation of a distributional difference and, therefore, shares the invariant characteristics of the histogram-like methods. In the core of the proposed method lies a nonparametric test dealing with the "*Multivariate Two-Sample Problem*" [29], which has been adopted here for expressing color image similarity. The specific test is a multivariate extension of the classical *Wald-Wolfowitz test* (WW-test) and compares two different samples of vectorial observations (i.e., two sets of points in $\mathbf{R}^{\mathbf{P}}$) by checking whether they form different branches in the overall minimal-spanning-tree (MST) [29]. 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. The bottleneck of our approach is the feature-extraction step, i.e., how to represent the content of each image in a properly constructed feature space. Without loss of generality and in order to gain some insights, the feature space was selected in all cases to coincide with the trichromatic RGB-space. However, we varied the procedure with which a vectorial distribution, representative of the image content, is formed in the RGB-space. In the simplest case, a few RGB-vectors from randomly selected pixel-locations suffice for fast, but medium-performance retrieval. For a better performance, each image is partitioned regularly into blocks and representative RGB-vectors are computed from these blocks. Other more complicated variants of the feature-extraction step, though straightforward, are left for future consideration.

The rest of this paper is organized as follows: The graph-theoretic framework of MST and the multivariate WW-test is provided in Section 2. The feature extraction process is described in Section 3. Extended results including a short discussion of the experimental observations are presented in Section 4. Finally, conclusions are drawn in Section 5, along with an outline of our future research objectives.

2 Theoretical Framework

2.1 Graph-Theoretic Description of the Image Content and MST

Given the establishment of a systematic procedure for extracting homologous, low-level characteristics from a color image (e.g., the chrominance of individual pixels) that are individually represented as vectors in a predetermined space, we can rely on graph theory to provide a collective perspective that captures the essence of the visual content of the image under study. Graph theory, by putting emphasis on the structural relationships between the extracted characteristics, provides robust descriptions to noise and simple transformations like image scaling. Specifically, the MST-graph appears as an extremely useful condensation of the bulk of information conveyed within the ensemble of image characteristics.

Graph theory sketches the MST structure with the following definitions [30]. A *graph* is a structure for

representing pairwise relationships among data. It consists of a set of points called *nodes* $\mathrm{V} = \{\mathrm{V_i}\}_{\mathrm{i=1:N}}$ and a set of links $\mathrm{E} = \{\mathrm{E_{ij}}\}_{\mathrm{i \neq j}}$ between nodes called *edges*. The *degree* $\mathrm{d_i}$ of a node is the number of edges incident to it. When a weight $\mathrm{e_{ij}}$ is assigned to each link, a weighted-graph is formed and, in the particular case that $\mathrm{e_{ij}} = \mathrm{e_{ji}}$, this graph is called *undirected weighted graph*. A *tree* is a connected graph with no cycles. A *spanning tree T* of a (connected) weighted graph $\mathrm{G(V, E)}$ is a connected subgraph of $\mathrm{G(V, E)}$ such that: 1) it contains every node of $\mathrm{G(V, E)}$ and 2) it does not contain any cycle. The MST is a spanning tree containing exactly $(\mathrm{N} - 1)$ edges, for which the sum of edge weights is minimum.

The MST provides a compact description of a point set. It contains the "*nearest-neighbor*" information about each point and the shortest linkage information about subsets of points [31]. Another advantage of MST is *determinacy*, meaning that the results from the application of a method working with MST-graph do not depend on random choices or the order in which points are scrutinized, but are affected solely by the point set provided as input [30]. In general, the MST structure is unchanged under transformations like translation, rotation, and nonlinear ones preserving the ordering of edge lengths. Finally, the MST is *relatively insensitive to small amounts* of noise widely and randomly spread over the field [30].

To conceptualize the previous abstract notions, let us consider the simplest possible feature-extraction procedure (postponing a thorough treatment until Section 3) aiming at mining color information from an image. Suppose that N pixels are selected at random and the corresponding triplets of RGB-values constitute the ensemble of low-level characteristics representing the image at hand. This sample of pixels is represented (based on the RGB-vectors) as a set of points in $\mathbf{R^3}$ space. The specific points are then used as the nodes of the original fully-connected graph, while the interpoint Euclidean distances as the weights of the corresponding edges. Finally, using a standard algorithm [31], the MST is delineated from the original graph. Provided that the number of selected pixels is sufficiently high, the MST offers a parsimonious description of the color variation in the image, with the additional advantage of being inherently insensitive to the pixels location. Given a second image, the color content of which is to be compared with the content of the first one, we can proceed with the selection of pixels (and the construction of MST in RGB-space) as previously and transform the comparison between color-contents into a comparison between the corresponding MST-graphs. To perform such a comparison, a well-defined statistical test is available in the literature of multivariate statistics. A short description of this test is provided in the sequel.

2.2 The Multivariate Wald-Wolfowitz Test (WW-Test)

Given two multidimensional point samples $\{X_i\}_{i=1:m}$ and $\{Y_i\}_{i=1:n}$, the hypothesis H_o to be tested is whether they are coming from the same multivariate distribution. At first, the sample identity of each point is not encountered and the MST of the overall sample is constructed. Then, based on the sample identities of the points, a test statistic R is computed. R is the total number of *runs*, while a *run* is defined as a consecutive sequence of identical sample identities. Rejection of H_o is for small values of R. The null distribution of the test statistic is derived, based on combinatorial analysis [29].

Consider samples of size m and n, respectively, from distributions F_x and F_y, both defined in $\mathbf{R^P}$. Let $N = m + n$, C be the number of edge pairs of MST sharing a common node, and d_i be the degree of the ith node. Then, $C = \frac{1}{2}\sum_{i=1}^{N} d_i(d_i - 1)$.

Number the $N - 1$ edges of the MST arbitrarily and define $Z_i, 1 \leq i \leq N - 1$, as:

$$Z_i = \begin{cases} 1 & \text{if the } i\text{th edge links nodes from different samples} \\ 0 & \text{otherwise.} \end{cases}$$

Then, $R = \sum_{i=1}^{N-1} Z_i + 1$. Under H_o, the mean and variance of R can be calculated as follows [29]:

$$E[R] = \frac{2mn}{N} + 1, Var[R|C] = \frac{2mn}{N(N-1)} \times \left\{ \frac{2mn - N}{N} + \frac{C - N + 2}{(N-2)(N-3)}[N(N-1) - 4mn + 2] \right\}.$$

It has been shown that the quantity:

$$W = \frac{R - E[R]}{\sqrt{Var[R]}}$$

approaches (asymptotically) the standard normal distribution while $\mathrm{E[R]}$ and $\mathrm{Var[R]}$ are given in closed form based on the size of the two samples [29]. This enables the computation of the *significance level* (and *p-value*) for the acceptance of the hypothesis H_o.

In our case, the above test is utilized as follows: With the feature extraction step, a representative point-sample is selected for each of the two color images that are to be compared. W is then computed and used as a similarity measure in a way that the more positive its value is, the more similar the two images are [29]. In the past, a few other statistical indices have been proposed, as well, as means of measuring similarity between color distributions. These distances (for instance, the *Kolmogorov-Smirnov distance* (KS), the *chi-square test* (χ^2-statistic), etc. [15]), measure how unlikely it is that one distribution is drawn from the population represented by the other. Under this perspective, the WW-test can be directly incorporated in retrieval processes from large image libraries, with the great advantage of being, by design, suitable for dealing with multivariate distributions.

2.3 General Comments

An image-retrieval system (e.g., a query-by-example scheme) based on the WW-test presents a flexible character. It is a combined consequence of representing each image by multiple feature vectors and, subsequently, performing statistical comparisons between the formed distributions of feature vectors that are model-free and not limited to equally sized samples. The quantity W computed between pairs of images plays the role of a "distributional distance" and, therefore, inherits interesting invariant-characteristics. On the other hand, such a retrieval system presents a very

generic character, since by altering the nature of extracted feature-vectors, we can represent information of different kinds and, thus, perform several content-dependent comparisons. Many image analysis techniques could have been utilized in this step, including local (e.g., salient-point methods [32]) or regional approaches [33]. Since clustering and segmentation techniques are time consuming, difficult to implement, and still an open issue in the field of image processing, we employed much simpler and generally applicable procedures for extracting representative vectors.

The only seemingly weak point of the proposed scheme is that it relies on the formation of MST, which is known to be a computationally demanding procedure. To provide some insight about the computational complexity of the methodology, the MST construction requires computational time $O(N^2)$ using Prim's algorithm [31], while the test statistic can be evaluated in time $O(N)$, where N is the number of data points in both cases. The employed feature-extraction procedures showed that satisfactory performance could be achieved even when small-sized samples of representative vectors are used.

3 Image Content Representation

3.1 Random Selection of Pixels

Considering all the RGB-vectors that comprise each image and utilizing them in forming detailed distributions that will be then compared in pairs with the WW-test, is apparently a computationally prohibitive strategy to represent the content of each color image. For that reason, the simplest and most straightforward alternative method of *random sampling* [34] is encountered here. With random sampling, the goal is to choose a representative set of cases from the full population under consideration. In statistical terms, *uniform random sampling* is the basic sampling technique where we select a group of items (a sample) from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample [34]. In the case of a color image, via random sampling a restricted number of pixels are selected. The corresponding RGB-vectors constitute the set of representative vectors. The procedure is repeated for each image (including the query) independently, each time a pairwise similarity computation is performed. Based on these selected sets, relative matches are pursued. Fig. 1 illustrates the performance of WW-test for a pair of similar and a pair of dissimilar RGB-images. For ease in the visualization, only the Red and Green components were taken into account. In both panels, different labels have been associated with each of the two images to be compared. After random selection of 15 pixels from each image, the corresponding RG-vectors were used to locate the selected pixels in the RG-plane. Based on the image they were coming from, the two-dimensional points were labeled accordingly. By contrasting the two overall MSTs, it becomes evident that in the case of similar images (top panel), there are many edges having differently labeled nodes as endpoints while only a few in the case of dissimilar images (bottom panel).

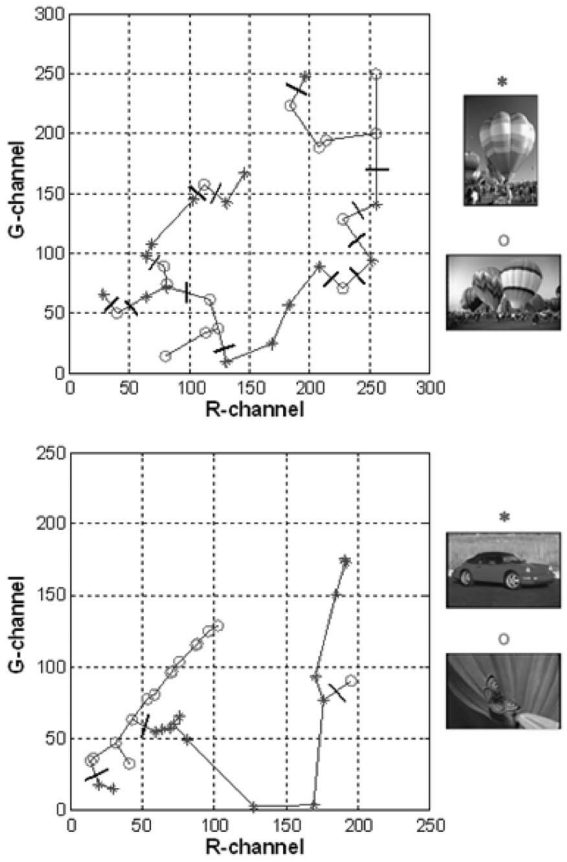


Fig. 1. WW-test for a pair of similar images (top) and dissimilar images (bottom), based on 15 randomly selected pixels from each image ("o" and "*" labels are indicating pixels from different images). In the top panel, there are 13 edges denoted by transverse lines, having differently labeled nodes as endpoints and, therefore, splitting the overall MST into 14 subgraphs, thus $\mathrm{R} = 14\ (\mathrm{W} = -0.7502)$. On the contrary there are only three such edges in the bottom panel, thus $\mathrm{R} = 4\ (\mathrm{W} = -4.4891)$.

Regarding the number of pixels that need to be sampled from each image, it can be intuitively understood that this is related to the nature of each image (i.e., how rich is the image in color information [35]). While in the case of a relatively uniform image a few pixels might suffice, in the case, e.g., of a small object on a uniform background many more RGB-vectors should be drawn so as the small object is represented by a few of them. To provide a practical answer to that question we studied the effect of the number of sampled pixels on the average retrieval performance (see Section 4). This was motivated by the signal processing practice, according to which, in order to define the optimum sampling frequency for an unknown signal, one keeps increasing the sampling rate till the frequency spectrum remains unaltered.

3.2 Segmenting the Image into Nonoverlapping Blocks

The image content representation in the feature space obtained via random pixel selection is such that it does not contain any information about the spatial domain, i.e., the particular color layout on the image plane. To better understand this, consider that any two RGB-vectors can correspond to neighboring or distant pixels as well. As a means of incorporating information about the spatial

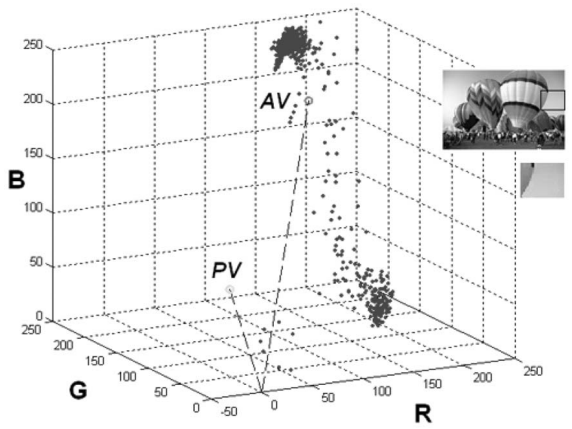


Fig. 2. Representing an image-block by the average vector (AV) and principal vector (PV) of the included set of RGB-vectors.

distribution of color, we followed the standard tactic to partition the image into nonoverlapping equal rectangular blocks [17] and represent the color content within each one by a properly defined RGB-vector. In this way, an image-content representation is developed in the RGB-space in which the representative vectors can be associated with regular pixel locations, that is, the locus of the corresponding block. We experimented with two methods for representing the color-content of each block. The first method is based on the computation of the average vector (AV) and it is expected to perform well in the case of uniform blocks. The second method is based on the computation of the first principal vector (PV) corresponding to the distribution of RGB-vectors comprising each image block. This is achieved by computing first the $[3 \times 3]$ covariance matrix, estimating afterwards the characteristic vector $u_{r_{\max}}$ associated with the maximum eigenvalue $r_{\max}$ and, finally, scaling by $(r_{\max})^{1/2}$). The first PV represents the scale and variation of dominant color variation and is therefore expected to perform well in the case of nonuniform blocks. Our experimentation showed that, in general, these two vectors contain complementary information for each image block and, therefore, can be used simultaneously. A characteristic example is provided in Fig. 2, where from the RGB image shown on the right a $[32 \times 39]$ pixels block has been delineated and the set of all the 1,248 RGB-vectors of the included pixels has been plotted as a point-diagram in which both the AV and the PV have been appended.

Fig. 3 illustrates the performance of WW-test for a pair of similar and a pair of dissimilar RGB-images, when both the set of average vectors (AVs) and the set of principal vectors (PVs) corresponding to the depicted blocks are simultaneously utilized as representative vectors. In each panel, the dot- and star-symbols denote the endpoints of each PV and AV, respectively, while the color is used to distinguish the vectors coming from blocks of different images. For the definition of "optimum" number of blocks, in which each image has to be split so as the representation of its content is as accurate as it is actually necessary, the obvious compromise between a detailed description and computational economy has to be made. Following a procedure similar to the previous one for the definition of the number of selected

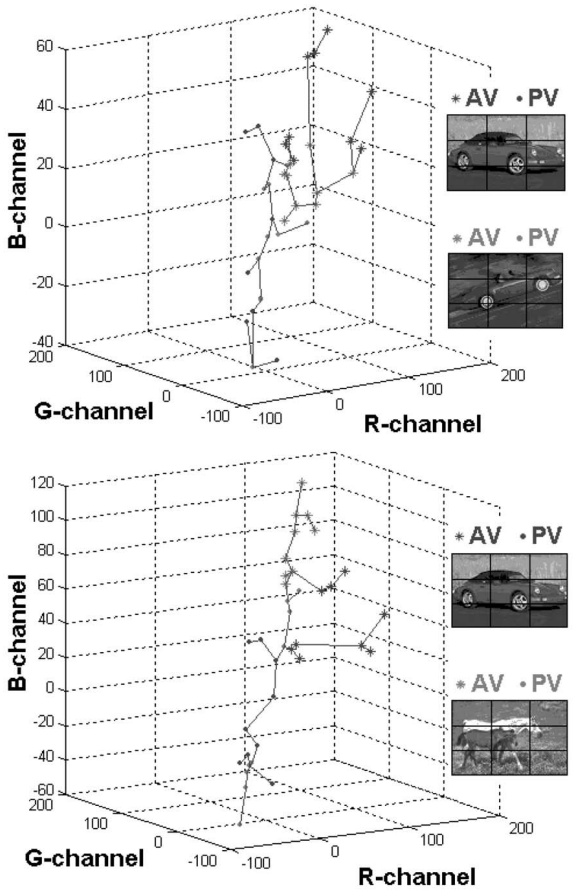


Fig. 3. WW-test for a pair of similar images (top) and dissimilar images (bottom), based on the average vectors (AVs) and the principal vectors (PVs) from the depicted nine blocks. In both panels, the overall MST is comprised from 18 representative vectors from each image. In the top panel $\mathrm{R} = 17\ (\mathrm{W} = -0.6807)$, while in the bottom panel $\mathrm{R} = 7\ (\mathrm{W} = -4.4403)$.

pixels, we measured the retrieval performance as a function of the numbers of blocks. Such a curve not only provided an estimate of the highest performance, but also helped with the settlement of an upper limit beyond of which the increase in computational cost is not counterbalanced by the improvement in retrieval performance.

4 Experimental Study

4.1 Data Set

The ensemble of images utilized to produce the experimental results presented in this paper for demonstrating the introduced methodology is part of the Corel image collection [36]. The specific data set is a heterogeneous subset from the Corel gallery [36], including $D = 1,000$ still color images of 24bpp each, given in portable pixel map format of sizes $[192 \times 128]$ or $[128 \times 192]$ pixels. The entire Corel database contains a wide variety of images, from animals and plants to views and natural images. The utilized data set was formed by preassigning the images into 20 distinct classes (e.g., cars, eagles, flowers, etc.), as introduced in the Schema Reference System [7]. By incorporating the coherent opinion of many individuals, 53 perceptually similar images were kept for each category. Fifty among these images were included in the image

database (which, in total, contained $D = 50 \times 20 = 1,000$ images), while the other remaining three were included in the data set of query images ($Q = 3 \times 20 = 60$ images).

The categorization was employed only for evaluating the performance of the reference system. This classification of the images into 20 perceptually given classes was incorporated in the name of the images, by including in the original numbering a prefix denoting the class (e.g., eagles135074.ppm, fireworks73072.ppm, etc). This served as the ground truth (based on high-level semantics) with which we could compare the behavior of the proposed retrieval scheme. In the experimental results reported in this section, the retrieval procedure had been first evaluated for each of the 60 query images and then the performance score was averaged across all queries.

4.2 Performance Evaluation

The evaluation of an image retrieval system is usually accomplished via empirical approaches. Such an approach typically consists of realizing a set of queries for which the correct answer has been determined beforehand. The most widely used performance measures for the efficiency and accuracy in information retrieval are the two quantities $precision(Pr)$ and $recall(Re)$ [17], which in our case were used as follows: For a given query image, let S be the number of similar images in the database coming from the same class ($S = 50$ in our case) and T be the set of total retrieved results (a user-defined number controlling the size of the selected list of images returned after a query-by-example, e.g., the top 10 images). If R is the number of relevant images retrieved (i.e., correct results) among the T images in the selected list, then precision and recall for the specific query image are defined as [17]:

$$\mathrm{Pr} = \frac{\#\text{ of relevant images retrieved}}{\#\text{ of total retrieved results}} = \frac{R}{T}$$
$$\mathrm{Re} = \frac{\#\text{ of relevent images retrieved}}{\#\text{ of similar images in the database}} = \frac{R}{S}.$$

From the definitions stated above, Pr will decrease when Re increases. The two quantities become equal when the size of the selected list, T, is chosen as the number of similar images in the database. In what follows, the above indices were used for comparing our approach, which is based on the use of W-index as similarity measure, with several other dissimilarity measures for comparing color histograms [15], [37]. In all cases, the reported performances are averaged scores, referring to the same set of query images. Specifically, for the introduced methodology, the results are coming from both basic variants and different settings of the involved parameters, which is basically the number of representative vectors. On the contrary, the only free parameter for the color histogram is the number of bins used, which was set to 20 for each RGB-channel. By experimenting with a wide range of values we noticed no significant changes in recall performance.

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 a query image H and a database image K, respectively, each containing n bins.

Histogram Intersection (HI): It was originally proposed by Swain and Ballard [16] for color image retrieval in the spatial domain and is found to be attractive due to its ability to handle partial matches [37]. It is shown that, when the areas of the two histograms are equal, the HI is equivalent to the L_1 distance. The HI-measure is given by:

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

Kullback-Leibler Divergence (KL): It measures how inefficient, on average, it would be to code one histogram using the other as the code-book [15], [37]:

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

Jeffrey Divergence (JD): This is a modification of the K-L Divergence that is symmetric, numerical stable, and robust with respect to noise and size of histogram bins [37], given by:

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

where $m_i = \frac{h_i + k_i}{2}$.

χ^2 Statistics (χ^2-test): It is a statistical index that measures how unlikely it is that one distribution is drawn from the population represented by the other [15], [37], given by:

$$d_{\chi^2}(H,K) = \sum_{i=1}^{n} \frac{(h_i - m_i)}{m_i}.$$

Earth Movers Distance (EMD): It is based on the solution of the well-known transportation problem [37]. The distance between two distributions is given as the minimum amount of work needed in order to transform one distribution into the other, where the total cost is the sum of the costs needed to move the individual features:

$$d_{EMD}(H,K) = \frac{\sum_{i,j} f_{ij} d_{ij}}{\sum_{i,j} f_{ij}},$$

where d_{ij} denotes the dissimilarity between bins i and j and $f_{ij} \geq 0$ is the optimal flow between the two distributions such that the total cost $\sum_{i,j} f_{ij} d_{ij}$ is minimized, subject to some constraints [37].

4.3 Pixelwise Random Sampling

Regarding the random selection of pixels as a means of mining information about the color content of an image, the performance of our proposal is systematically studied by varying the number of extracted RGB-vectors N from 10 to 100. Since the set of representative vectors is randomly selected each time a pairwise similarity computation is performed, experiments were conducted 10 times and averaged scores of the precision index are plotted in Fig. 4. In this way, performance results are more precise, especially when a small number of pixels are encountered for the WW-based methodology. Fig. 4 includes a plot of this curve for the $T = 10$ top retrieved images, along with the corresponding level of precision for each dissimilarity measure presented previously.

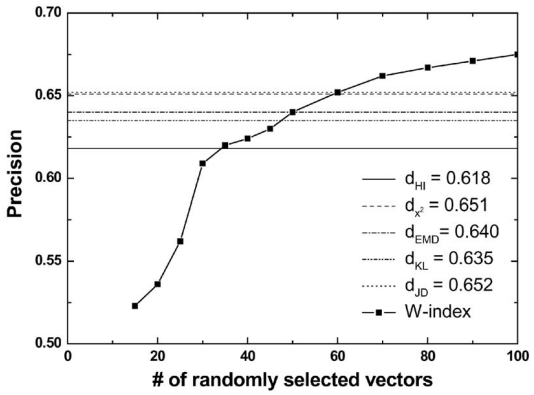


Fig. 4. Precision as a function of the number of randomly selected pixels from each query and database image, for the 10 most similar images ($\mathrm{T} = 10$). Measurements on the curved line are referring to the WW-based system using the W-index as the similarity measure. The precisions of the histogram-related dissimilarity measures are depicted as horizontal lines.

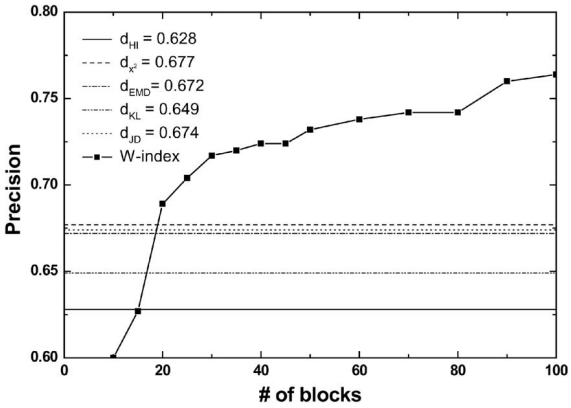


Fig. 5. Precision for a representation using AVs and PVs as a function of the number of blocks in which each query and database image was split, for the 10 most similar images ($\mathrm{T} = 10$). Measurements on the curved line are referring to the WW-based system using the W-index as the similarity measure. The precisions of the histogram-based dissimilarity measures are depicted as horizontal lines.

As it can be seen in Fig. 4, the proposed system performs satisfactorily even for a small number of randomly selected pixels. Describing the general trend of the depicted curve, we could say that the increase in the number of extracted vectors is followed by an increase in the precision. However, this performance enhancement is concurrent with the increase in sampling density only at the leftmost part of the plot. Beyond ~ 60 pixels, the performance improves very slowly with the number of extracted vectors. This observation is very important for finding the best trade-off between effectiveness and efficiency when applying our algorithmic procedure. Considering that in applications where high speed in retrieval is the basic specification, the gain in precision cannot counterbalance the computational load of constructing lengthy MSTs. Such a curve can be suggestive for the minimum number of selected pixels so as to achieve a certain level of performance at a tolerable computational time. To provide some indication regarding effectiveness, we can mention that the precision of the proposed algorithmic procedure for approximately $\mathrm{N} = 30$ vectors is almost equal to the precision of the HI-method, while for $\mathrm{N} \geq 60$ vectors it outperforms all other measures. To provide some indication regarding computational efficiency, the execution time of the proposed procedure was measured and expressed relatively to the execution time of the HI-method denoted as t_{HI}, which is the fastest computational approach among the included in the study. Retrieval time for one query is, on average, $[1.35 \times t_{HI}]$ for $\mathrm{N} = 30$ and $[2.02 \times t_{HI}]$ for $\mathrm{N} = 60$.

4.4 Blockwise Sampling

Regarding the segmentation of images into nonoverlapping rectangular blocks and the subsequent computation of representative vectors as a means of incorporating information about the spatial layout of color, we studied the performance of our proposal as a function of blocks number. In our initial approach, the average vector (AV) was computed within each block and selected as the representative vector. By studying the blockwise AV sampling, a curve having similar trend as in Fig. 4 was obtained (not shown here due to space limitations). The same plateau in performance was observed, after approximately partitioning the images into 20 blocks. However, by increasing the number of blocks, the performance could not be improved beyond the level that was reached by selecting random pixels.

This barrier in performance, motivated the search for alternative representations that could carry more (or complementary) content-related information. The principal vector (PV) expresses a theoretically different concept for vectorial distributions than the average vector (AV) since the AV and PV relate with the first and second order statistics, respectively. Moreover, the PV has a straightforward physical interpretation that carries the information of dominant color variation. Thus, the PV appeared a logical candidate for representing the content of the blocks and proved to actually behave differently when incorporated in a WW-based query engine. Experiments had proven that images belonging to the same class are retrieved at different ranking positions when using the representative AVs and PVs from the delineated blocks separately. This complementary information could be exploited by building an overall system that employs both the AV and PV descriptors. As a consequence, more images belonging to the same class to the query would be retrieved among the first ranking positions. In addition, the common images retrieved among the selected list having high similarity index value, would be enforced and retrieved in higher-ranking positions. Finally, false alarms (irrelevant images belonging to a different class than the query image) would be fairly rejected in the final retrieval choice. By using the two descriptors independently and performing detailed analysis for the entire query-set, it was found that among the different tens of images retrieved by the two corresponding search engines, only a small portion (38 percent) was in common. This motivated the introduced algorithmic practice that was followed as described below.

After extracting the AV and PV as a pair of representative vectors from each block, a single WW-based search engine was built (see Section 3.2). The corresponding precision curve is provided in Fig. 5 in exactly the same

format as in Fig. 4, for the top $T = 10$ retrieved images, versus the number of blocks that each image was split into. In addition, the corresponding level of precision for each dissimilarity measure applied on blocks of the image is also provided in this plot. In this case, the color-spatial information of an image has been modeled by splitting it into a set of equal subimages and representing each subimage by a distinct color histogram [18]. After proper integration of all partial dissimilarity measures for the block-related histograms into an aggregate dissimilarity measure, we built a search engine and performed the corresponding precision-index measurements for different number of blocks. This experimentation showed that the precision-index started to degrade when the images were split into more than 20 blocks. The horizontal lines for each aggregate dissimilarity measure shown in Fig. 5 correspond to the highest precision-index that was identified. This was calculated by testing for different pairs of values $H\&V$, defining the number of horizontal (H) and vertical (V) blocks that each image was split into, when varying H and V from 2 to 5. In this way, the spatial distribution of color is equally considered in all related experiments, making fair comparisons between the proposed WW-based similarity and the predefined histogram-based ones.

From what we can perceive in Fig. 5, once again a plateau emerges in the precision curve, but in this case, its upper limit is much higher (compared to Fig. 4). The increase in retrieval effectiveness is striking; 10-15 percent with respect to all methods and 7-8 percent with respect to the previous implementations of WW-based engine. By measuring the performance in the case that only the PVs were used as representative vectors from the blocks, we verified that the relative high performance seen in Fig. 5 is a result of the combined use of both the AV and the PV of each block. To provide some indication regarding computational efficiency, the execution time was measured for the WW-based approach of blockwise AV and PV sampling and expressed it once again relatively to the execution time of the HI-method t_{HI}. The time for one query is, on average, $[1.30 \times t_{HI}]$ for $\mathrm{N} = 20$ vectors (i.e., after partitioning the image into 10 nonoverlapping blocks) and $[2.20 \times t_{HI}]$ for $\mathrm{N} = 60$ (i.e., after partitioning the image into 30 nonoverlapping blocks). Bearing in mind that the color-histogram technique is a very fast computation, and taking into account the improved performance of our method, the slight increase in execution time should be considered insignificant.

In order to demonstrate more vividly the benefit of efficient information handling from the use of the similarity measure that is based on the application of WW-test on AVs and PVs from 30 nonoverlapping image blocks, we have included in Fig. 6 visualization of the image database structure provided by the dimensionality reduction technique of Multidimensional Scaling (MDS) [37], [38]. This is a standard approach in multivariate statistics, appearing in many algorithmic variants though. Given all the $M(M-1)/2$ pairwise distances between M objects, a distance preserving mapping of these objects can be performed in a low-dimensional space (usually, a plane) and in such a way that the objects are mapped onto points with interpoint distances which are equal to the

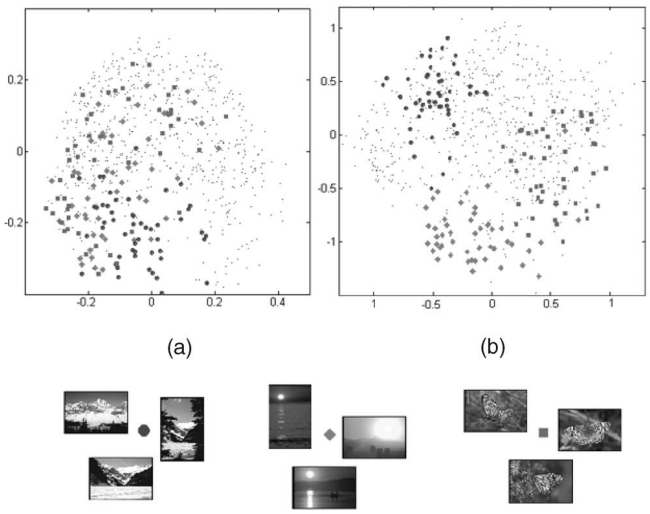


Fig. 6. MDS-maps of the entire database using color histogram (a) and WW-test (b), along with the members from three different image categories denoted using the specific symbols shown at the bottom.

original distances between the corresponding objects. The advantage of such a low-dimensional point-diagram is that by visual inspection, the neighboring relationships between objects and, therefore, possible grouping tendencies in the data are becoming apparent. In our case, we first estimated the pairwise similarity measures for all the 1,000 images in our database (i.e., $1,000 \times (999/2) = 499,500$ W-indices in total) and then, by using Kruskal's algorithm [38], the 2-dimensional MDS-map was computed, as seen in Fig. 6b. For comparison purposes, we include in Fig. 6a the corresponding MDS-map from the color histogram approach. In the MDS-map, the entire image database is shown as a point-swarm. Of great interest is the relative structural difference between the subsets of points belonging to different image classes. To portray this kind of information, light black dots are used to represent all the images in the background of Fig. 6 and three types of symbols to distinct the images from three image classes (consisting of 50 images each). The specific classes were chosen for visual representation between the different approaches, due to their evident perceptual color dissimilarity. The specific implementation was made using 30 AVs and PVs. A few images from these three selected classes are given at the bottom of Fig. 6 along with the associated symbols. The direct comparison between Fig. 6a and 6b reveals enough evidence that the WW-test provides a similarity measure that is superior to histogram-related one when visualization of mutual distances of all database images is encountered. From what we can observe, the three different image classes are clearly separated in our method, forming clusters of high concentration. On the other hand, using the histogram-approach, the same groups of images tend to mix together, i.e., the different classes seem to overlap.

4.5 Comparative Study of the Different Variants of the Proposed Scheme

To enable a direct comparison, in terms of performance, of the different variants of the WW-test based search engine, precision measurements have been included in Table 1 for different numbers of N of representative

TABLE 1
Precision Measurements for a Query-System Returning the $T = 10$ Most Similar Images Based on Different Ways to Express Image Content (Random Sampling, AVs & PVs from Blocks) and Variable Number N of Representative Vectors

	N = 20	N = 40	N = 60	N = 100
Random Sampling + WW-test	0.536	0.624	0.652	0.675
AVs & PVs +WW-test	0.600	0.689	0.717	0.732

vectors. In addition, measurements have been carried out for the corresponding execution times. Provided that t_{HI} denotes the execution time of the HI-method, the execution times of WW-test based methods have been expressed relative to it and tabulated in Table 2.

In order to provide some insight to the complexity of our method, a detailed analytic and experimental cost model is presented in Table 3, listing the proportions of the individual steps of the algorithm (partitioning, feature extraction, MST construction, and WW-test), prior to the overall execution time of the different implementations. From what we can perceive here, the MST-construction is proportionally the most time-consuming operation of the overall system implementation. The ongoing execution time of the WW-test that derives from the growing number of vectors extracted from each RGB image (Table 1) is the outcome of the increase in computational time needed for the construction of the overall MST-graph (Table 3). Despite the apparent considerable computational load that our proposal requires (since, for the execution of WW-test, the construction of the MST is a prerequisite), the practice shows that efficiency might be an advantage as well. It is worth mentioning that MST is also a key part in the *entropic spanning graphs* [39], an important methodological approach in image analysis with many interesting applications. In their case, the authors suggest the use of dedicated hardware. Nowadays, the theory of randomized algorithms [40] provides alternative fast approximations to the MST construction problem with the incorporation of which the efficiency of our work can improve significantly.

Fig. 7 provides an illustration of the performance of different approaches by including a detailed query example. The query image (a sunset) can be seen at the top. Each of the following three rows is devoted to the results obtained by using one of three different techniques (color-histogram method, WW-test applied on 30 randomly selected vectors, WW-test applied on AVs, and PVs from 30 blocks). The returned images, in each row, are represented in decreasing order of similarity to the query image. As we can see, some of the returned images are different from the query image in the case of histogram-approach (the images ranking from six to nine contain a red car with green or dark-green background). On the contrary, WW-test in its random sampling implementation retrieves correctly eight sunrise-images among the eight first returned ones and includes a red car with dark or yellow-dark background at the ranks 9 and 10. The best recall result is provided via the WW-based implementation using both the AVs and PVs blockwise sampling.

TABLE 2
Average Execution Times for a Single Query, Relative to the Execution Time of HI

	N=20	N=40	N=60	N=100
Random Sampling + WW-test	×1.29	×1.56	×2.02	×3.49
AVs & PVs +WW-test	×1.30	×1.65	×2.20	×3.72

Finally, the performance of the different similarity measures as methods for accessing image databases was evaluated following the standard procedure of constructing the Precision versus Recall diagram. Based on Fig. 8, the following observation can be made. The WW-test when acting on randomly selected pixels behaves pretty much the same as the color-histogram approaches. One needs to bear in mind that color is a low-level image characteristic, considered as a sufficient signature of the general "mood" of the picture and there exactly lies the broad applicability and relative success of the histogram methods. By randomly selecting a set of RGB-vectors from each database member and, subsequently, performing pairwise WW-based comparisons, we stay conceptually at the level of color histogram; since each distribution reflects the relative frequency with which the RGB-neighborhoods are populated by the vectors of individual pixels, we are still performing a statistical comparison of two distributions in RGB-space. Surprisingly, the formed distributions in RGB-space, regarding the retrieval performance of the WW-test, need not to be acquired in high details (sampling of only a few pixels provided satisfactory results).

On the other hand, when the WW-test acts on vectors representing both the color content and the dominant color variation within each block, it results in a significantly improved performance. This can be understood by con-

TABLE 3
Execution Times of the Individual Algorithmic Steps Prior to the Overall Execution Time of the Different Implementations (Random Sampling and Blockwise Sampling)

	Algorithmic Steps	N=20	N=60	N=100
Random Sampling	RGB extraction	25%	10%	6%
	MST construction	67%	87%	92%
	WW-test	8%	3%	2%
Blockwise Sampling	Block Partitioning	21%	7%	3%
	AV and PV extraction	22%	9%	4%
	MST construction	52%	83%	92%
	WW-test	5%	1%	1%

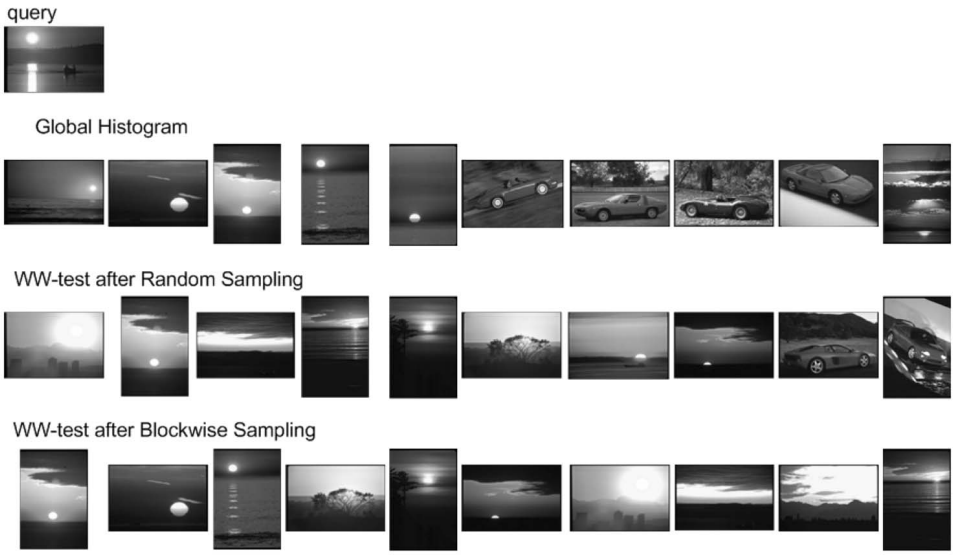


Fig. 7. Retrieval results for color histogram (first row), WW-test after pixelwise random sampling (second row), and WW-test after blockwise sampling of average and principal vectors (third row).

sidering that while the average vectors from the blocks convey information similar in nature (apart from a "low-pass" effect) with the randomly sampled vectors, the principal vectors provide additional information, which is the local color variation within the different image regions. In terms of information, it is exactly this enhanced representation of the image content that constitutes the subsequent WW-test so powerful. It is worth mentioning that the same high performance could not be reached by random selection (even for 100 times more pixels) and applying the WW-test on the corresponding RGB-vectors.

5 Concluding Remarks and Discussion

We have introduced the use of WW-test as a general strategy for measuring similarity between color images and demonstrated its effectiveness when used for image retrieval. Our approach relies on a dual, segregation-integration algorithmic step. First, each image is represented by a set of low-level characteristics extracted in the form of an ensemble of feature-vectors and then "set-differences" are computed between pairs of image representations. The great advantage of our approach is that since it involves a "distributional distance" acting on samples of image constituents, the emerging similarity measure possesses desirable invariant characteristics, such as rotation and translation invariance. Its generic character stems from the fact that by altering the character of low-level image characteristics, we can modify the flavor of formulated queries. For instance, we could have replaced the utilized pixelwise color information by a textural characterization and retrieved images not of similar color content, but of alike texture. Part of the flexibility of our proposal is due to the statistical nature of the core procedure, the WW-test and, specifically, its multivariate orientation. Not only can different image characteristics, in principle, be combined naturally in one type of query, but also different types of queries can evolve independently and their results can be compared across types. The latter is a direct consequence of the fact that the measured W-index relates directly to significance-level and, therefore, can be used as an absolute measure to rank among the results of different types of query. For instance, the 10 most similar images to a given query-image regarding color representation and the 10 most similar ones regarding textural representation can be immediately ranked from 1-20 using the corresponding W-indices.

In the included experimental study and without loss of generality, we restricted ourselves to image content representations in the RGB-space. Using two simple feature extraction procedures, we show that our methodology outperforms previously related ones, which are considered as classical approaches for color image retrieval. Without the most efficient implementation (since a greedy algorithm was employed for MST [31]), we showed that improved performance could be achieved within a logical execution time.

There are two things from this experimental study that is worth further consideration. The first is the plateau that is

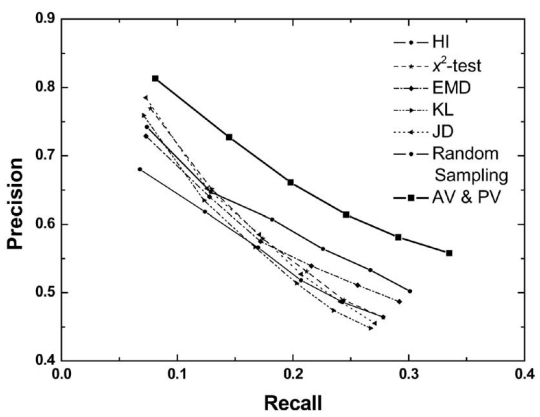


Fig. 8. Comparison of retrieval performance via the Precision versus Recall diagram for the different methodologies.

reached in the precision curves (Figs. 4 and 5) as we increase the number of representative vectors. This is important for either defining the maximum performance at the least computational cost or deciding a reasonable trade-off. While these curves are associated with a specific image database, we believe that a similar behavior can be seen in general. We have to mention that in a (very) preliminary study, we applied the WW-test for color image retrieval on an entirely different color data set (Macmillan) [41], with similar results. Specifically, about the plateau seen in the case of WW-test, it is very interesting to notice that a similar behavior has been reported recently in the work of Bingham and Mannila on random projections for image and text retrieval, providing a related algorithmic paradox that can result in significant computational savings [42]. In this way, by switching from color representation at pixel-level to color representation at block-level and adequately incorporating color-variation characteristics, a significant improvement in performance (~ 15 percent) is realized. This opens the possibility of even higher performance, by experimenting with other higher-order characteristics from image patches.

Finally, we would like to mention the scheduled extensions of this paper and our related feature research objectives:

1. Analytical formulation of ways with which the application of WW-test can achieve more demanding invariance characteristic, like being suitable as a cost for partial matching. For instance, scale invariance can be achieved by incorporating the algorithmic trick of multiple resampled versions of the query image from [43] and exploiting the fact that W-index can act as an absolute measure of statistical significance to rank the similarity across different scales.
2. Experimentation with other feature spaces (e.g., different color spaces) or higher dimensionality features (long vectors from local image patches).
3. The use of alternative ways to extract representative vectors. Of specific interest are neuromorphic signal processing techniques (e.g., [44]) that identify regions of interests from an image by taking into consideration the knowledge from psychophysics about human visual system functionality.

ACKNOWLEDGEMENTS

This work was funded by the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II), under the Program HERAKLEITOS.

REFERENCES

[1] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," *Computer,* vol. 28, no. 9, pp. 23-32, Sept. 1995.

[2] A. Pentland, R.W. Ricard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases," *Int'l J. Computer Vision,* vol. 18, no. 3, pp. 233-255, 1996.

[3] J.R. Smith and S.F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System," *Proc. 1996 ACM Int'l Multimedia Conf.,* pp. 87-98, 1996.

[4] W.Y. Ma and B.S. Manjunath, "Netra: A Toobox for Navigating Large Image Database," *Proc. IEEE Int'l Conf. Image Processing,* vol. 1, pp. 568-571, Oct. 1997.

[5] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Quering," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 8, pp. 1026-1038, Aug. 2002.

[6] E.G.M. Petrakis, C. Faloutsos, and K.-I. Lin, "ImageMap: An Image Indexing Method Based on Spatial Similarity," *IEEE Trans. Knowledge and Data Eng.,* vol. 14, no. 5, pp. 979-987, Sept./Oct. 2002.

[7] V. Mezaris, H. Doulaverakis, R. Medina, S. Herrmann, I. Kompatsiaris, and M.G. Strintzis, "A Test-Bed for Region-Based Image Retrieval Using Multiple Segmentation Algorithms and the MPEG-7 eXperimentation Model: The Schema Reference System," *Proc. Int'l Conf. Image and Video Retrieval (CIVR 2004),* 2004.

[8] A. Natsev, R. Rastogi, and K. Shim, "WALRUS: A Similarity Retrieval Algorithm for Image Databases," *IEEE Trans. Knowledge and Data Eng.,* vol. 16, no. 3, pp. 301-316, Mar. 2004.

[9] B.S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG-7: Multimedia Content Description Interface.* New York: John Wiley & Sons, 2002.

[10] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[11] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 8, no. 5, pp. 644-655, 1998.

[12] E. Chang and S. Tong, "SVMActive—Support Vector Machine Active Learning for Image Retrieval," UCSB Technical Report, Nov. 2001.

[13] R.C. Veltkamp, M. Tanase, and D. Sent, "Features in Content-Based Image Retrieval Systems: A Survey," *State-of-the-Art in Content-Based Image and Video Retrieval,* R.C. Veltkamp, H. Burkhardt, and H.-P. Kriegel, eds., pp. 97-124, Kluwer, 2001.

[14] R. Schettini, G. Ciocca, and S. Zuffi, "A Survey on Methods for Colour Image Indexing and Retrieval in Image Databases," *Color Imaging Science: Exploiting Digital Media,* R. Luo and L. MacDonald eds., J. Wiley, 2001.

[15] Y. Rubner, J. Puzicha, C. Tomasi, and J.M. Buhmann, "Empirical Evaluation of Dissimilarity Measures for Color and Texture," *Computer Vision and Image Understanding,* vol. 84, pp. 25-43, 2001.

[16] M.J. Swain and D.H. Ballard, "Color Indexing," *Int'l J. Computer Vision,* vol. 7, no. 1, pp. 11-32, 1991.

[17] V. Castelli and L.D. Bergman, *Image Databases: Search and Retrieval of Digital Imagery.* New York: John Wiley & Sons, 2002.

[18] Y. Gong, C.H. Chuan, and G. Xiaoyi, "Image Indexing and Retrieval Using Color Histograms," *Multimedia Tools and Application,* vol. 2, pp. 133-156, 1996.

[19] M. Stricker and A. Dimai, "Color Indexing with Weak Spatial Constraints," *SPIE Proc.,* no. 2670, pp. 29-40, 1996.

[20] J.R. Smith and S.-F. Chang, "Single Color Extraction and Image Query," *Proc. IEEE Int'l Conf. Image Processing,* Oct. 1995.

[21] G. Pass, R. Zabih, and J. Miller, "Comparing Images Using Color Coherence Vectors," *Proc. Fourth ACM Multimedia Conf.,* 1996.

[22] J. Huang, R. Kumar, S. Mitra, W. Zhu, and R. Zabih, "Image Indexing Using Color Correlogram," *Proc. IEEE Conf. Computer Vision Pattern Recognition,* pp. 762-768, 1997.

[23] M. Stricker and M. Orengo, "Similarity of Color Images," *Proc. SPIE Storage and Retrieval for Image and Video Databases,* pp. 381-392, 1995.

[24] G. Paschos, I. Radev, and N. Prabakar, "Image Content-Based Retrieval Using Chromaticity Moments," *IEEE Trans. Knowledge and Data Eng.,* vol. 15, no. 5, pp. 1069-1072, Sept./Oct. 2003.

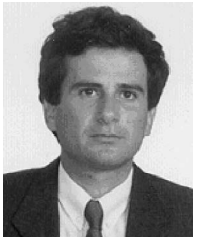
[25] B.M. Mehtre, M.S. Kankanhalli, A.D. Narasimhalu, and G.C. Man, "Color Matching for Image Retrieval," *Pattern Recognition Letters,* vol. 16, no. 3, pp. 325-331, 1995.

[26] T. Gevers and A.W. M. Smeulders, "Content-Based Image Retrieval by Viewpoint-Invariant Color Indexing," *Image and Vision Computing,* vol. 17, no. 7, pp. 475-488, 1999.

[27] I.K. Park, I.D. Yun, and S.U. Lee, "Color Image Retrieval using Hybrid Graph Representation," *Image and Vision Computing,* vol. 17, no. 7, pp. 465-474, 1999.
[28] T. Randen and J.H. Husoy, "Image Content Search by Color and Texture Properties," *Proc. Int'l Conf. Image Processing*, pp. 580-583, 1997.
[29] J.H. Friedman and L.C. Rafsky, "Multivariate Generalizations of the Wald-Wolfowitz and Smirnov Two-Sample Tests," *Annals of Statistics,* vol. 7, no. 4, pp. 697-717, July 1979.
[30] C.T. Zahn, "Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters," *IEEE Trans. Computers,* vol. 20, no. 1, Jan. 1971.
[31] R.C. Prim, "Shortest Connection Networks and Some Generalizations," *Bell Systems Technical J.,* vol. 36, pp. 1389-1401, 1957.
[32] V. Gouet and N. Boujemaa, "Object-Based Queries Using Color Points of Interest," *Proc. IEEE Workshop Content-Based Access of Images and Videos (CBAIVL),* Dec. 2001.
[33] B. Moghaddam, H. Biermann, and D. Margaritis, "Image Retrieval with Local and Spatial Queries," *Proc. IEEE Int'l Conf. Image Processing (ICIP),* pp. 542-545, Sept. 2000.
[34] S.K. Thompson, *Sampling.* New York: John Wiley & Sons, 1992.
[35] J. Costa and A.O. Hero, "Geodesic Entropic Graphs for Dimension and Entropy Estimation in Manifold Learning," *IEEE Trans. Signal Processing,* vol. 52, no. 8, pp. 2210-2221, 2004.
[36] Corel Stock Photo Library, Corel Corp., Ontario, Canada.
[37] Y. Rubner and C. Tomasi, *Perceptual Metrics for Image Database Navigation.* Kluwer Academic Publishers, 2001.
[38] J.B. Kruskal, "Multi-Dimensional Scaling by Optimizing Goodness-of-Fit to a Nonmetric Hypothesis," *Psychometrica,* vol. 29, pp. 1-27, 1964.
[39] A.O. Hero, B. Ma, O.J.J. Michel, and J. Gorman, "Applications of Entropic Spanning Graphs," *IEEE Signal Processing Magazine,* vol. 19, no. 5, pp. 85-95, 2002.
[40] R. Motwani and P. Raghavan, *Randomized Algorithms.* Cambridge Univ. Press, 2000.
[41] Ch. Theoharatos, N. Laskaris, G. Economou, and S. Fotopoulos, "A Similarity Measure for Color Image Retrieval and Indexing Based on the Multivariate Two Sample Problem," *Proc. European-Signal Processing Conf. (EUSIPCO),* 2004.
[42] E. Bingham and H. Mannila, "Random Projection in Dimensionality Reduction: Applications to Image and Text Data," *Proc. Seventh ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD-2001),* pp. 245-250, 2001.
[43] K. Vu, K.A. Hua, and W. Tavanapong, "Image Retrieval Based on Regions of Interest," *IEEE Trans. Knowledge and Data Eng.,* vol. 15, no. 4, pp. 1045-1049, July/Aug. 2003.
[44] N. Laskaris and S. Fotopoulos, "A Novel Training Scheme for Neural-Network Based Vector Quantizers and Its Application in Image Compression," *Neurocomputing,* vol. 61, pp. 421-427, 2004.

Christos Theoharatos received the BSc degree in physics in 1998 and the MSc degree in electronics in 2001, both from the Department of Physics, University of Patras, Greece. He is currently a PhD degree student in image processing at the Electronics Laboratory in the Department of Physics, University of Patras, Greece. His main research interests include pattern recognition, multimedia databases, data mining, and machine vision. He is a member of the IEEE.

Nikolaos A. Laskaris received both the MSc degree in medical physics (1995) and the PhD degree in biomedical signal processing (1998) from the University of Patras, Greece. He joined the Laboratory for Human Brain Dynamics in the Brain Science Institute (BSI) at RIKEN in Japan, for four years (1999-2003), where he worked mostly as a neuroengineer. He received a grant from the Greek GSRT (ENTER-program) for working as an associate researcher in the Electronics Laboratory of the Physics Department at Patras University (2003-2004), where he involved in the development of neuromorphic signal/image processing techniques. Currently, he is a faculty member in the Informatics Department in AUTH. His current research interests include computational intelligence, machine vision, soft computing, data mining, nonlinear dynamics, and their applications.



George Economou received the BSc degree in physics from the University of Patras, Greece, in 1976, the MSc degree in microwaves and modern optics from the University College, London, in 1978, and the PhD degree in fiber optic sensor systems from the University of Patras in 1989. He is currently an assistant professor in the Physics Department at the University of Patras. He has published papers on nonlinear signal and image processing, fuzzy image processing, and fiber optic sensors. His research interests include image processing, computer vision, and optical signal processing.

Spiros Fotopoulos received the BSc degree in physics from the University of Athens in 1974 and the PhD degree from the University of Patras in 1983. He is working in the digital signal and image processing area. His research activity focuses on nonlinear digital filters, multichannel filters, fuzzy image processing, neural networks techniques, graph theoretic approaches, and applications to biomedical signals. Dr. Spiros Fotopoulos is currently an associate professor in the Department of Physics, University of Patras, Greece.