# Color-Based Descriptors for Image Fingerprinting

Marios A. Gavrielides, Elena Šikudová, and Ioannis Pitas, *Senior Member, IEEE*

***Abstract*—Typically, content-based image retrieval (CBIR) systems receive an image or an image description as input and retrieve images from a database that are similar to the query image in regard to properties such as color, texture, shape, or layout. A kind of system that did not receive much attention compared to CBIR systems, is one that searches for images that are not similar but exact copies of the same image that have undergone some transformation. In this paper, we present such a system referred to as an image fingerprinting system, since it aims to extract unique and robust image descriptors (in analogy to human fingerprints). We examine the use of color-based descriptors and provide comparisons for different quantization methods, histograms calculated using color-only and/or spatial-color information with different similarity measures. The system was evaluated with receiver operating characteristic (ROC) analysis on a large database of 919 original images consisting of randomly drawn art images and similar images from specific categories, along with 30 transformed images for each original, totaling 27570 images. The transformed images were produced with attacks that typically occur during digital image distribution, including different degrees of scaling, rotation, cropping, smoothing, additive noise and compression, as well as illumination contrast changes. Results showed a sensitivity of 96% at the small false positive fraction of 4% and a reduced sensitivity of 88% when 13% of all transformations involved changing the illuminance of the images. The overall performance of the system is encouraging for the use of color, and particularly spatial chromatic descriptors for image fingerprinting.**

***Index Terms*—Color histogram, color quantization, image fingerprinting, image indexing, image representation, retrieval, spatial chromatic histogram.**

## I. INTRODUCTION

RECENT advances in multimedia technology and the growing worldwide access to fast Internet connections have enabled the wide spread distribution of digital images. This trend is amplified by the availability of digital cameras and scanners in an increasing number of households. The enormous amount of the produced media information created the need for intelligent systems for browsing, retrieval, and storage of images. In particular, significant research effort has been devoted to content-based information retrieval systems that enable users to search and retrieve images based on the users' needs and interests. Content-based image retrieval systems (CBIR) aim to provide an automatic way to extract information from images that depends only on the content of the image. Typically, CBIR systems receive an image or an image description as input and retrieve images from a database that are similar to the query image in regard to properties such as color, texture, shape, or layout. A comprehensive survey of CBIR systems can be found in [1].

A kind of system that did not receive much attention compared to CBIR systems, is one that searches for images that are not only similar to but versions (possibly modified) of the same image. The internet is flooded with copies of certain images that have undergone a number of transformations which typically occur during image distribution, such as resizing, cropping, compression etc. Such transformations are essentially "attacks" from the viewpoint of intellectual property protection or digital rights management. A copyright owner of images would find useful a system that would detect only those images that are transformed versions of his/her own images. Detecting transformed versions of images could be used to fight piracy of such material and additionally, it could be used to trace the distribution of their products in the Internet for marketing purposes. Moreover, users would be able to search for different versions of their image of choice without getting back a large number of similar images.

In order for such a system to be successful, it would need to fulfill the following characteristics:

1) Robustness against a number of frequent attacks, including scaling, rotation, cropping, compression, additive noise, and smoothing, which can be easily performed by nonexperts due to the abundance of image manipulation software in the market.
2) Good discriminating ability to avoid the retrieval of false alarms.
3) Efficient storage of extracted image descriptors that would be used for image matching.
4) Efficient search mechanism that would compare the image description of the query image to those in a database of image descriptors.

Research related to the retrieval of different versions of the same image within a database has been given a number of names depending on the application. More specifically, the term *image fingerprinting* has been used, as the extraction of a unique description of an image that would be resilient to transformations, in an analogous manner to human fingerprints. Fingerprinting

Manuscript received May 20, 2004; revised October 19, 2005. This work was supported by the RTN Project MOUMIR (Models for Unified Multimedia Information Retrieval, HPRCN-1999-00108) and in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alexander Loui.

M. A. Gavrielides was with the Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece. He is now with the Division of Imaging and Applied Mathematics, Office of Science and Engineering Labs, Center for Devices and Radiological Health, Food and Drug Administration, and the National Institute of Biomedical Imaging and Bioengineering, Rockville, MD 20852 USA (e-mail: marios.gavrielides@fda.hhs.gov.).

E. Šikudová was with the Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece. She is now with the Department of Applied Informatics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia (e-mail: sikudova@fmph.uniba.sk).

I. Pitas is with the Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece (e-mail: pitas@zeus.csd.auth.gr).

Digital Object Identifier 10.1109/TMM.2006.876290

had been originally proposed as a tool to track music and video broadcasting in order to perform copyright infringement monitoring or to provide input to Digital Rights Management systems for revenue distribution. Image fingerprinting differs from the body of work falling under image *watermarking* in the sense that watermarking involves embedding information into the image that is recovered to trace the image, whereas fingerprinting, as defined here, involves descriptors extracted from the image content. Watermarking changes the content of an image (in a varying degree, usually related to the effectiveness of the watermark) whereas fingerprinting is a passive technique. Moreover, fingerprinting can be used to search for those copies of an image that have already been circulating in the internet with no watermarks embedded on them.

Despite the volume of research conducted in the field of CBIR, limited efforts have focused in image fingerprinting. Published work has suffered from limitations regarding the types of attacks (i.e. algorithm addressing only rotation invariance etc) and/or the degree of an attack (i.e., algorithm limited to images scaled by only 10% etc.). In the digital communication era, it is important for image fingerprinting methods to be resistant to distribution "attacks" such as compression and down scaling, which are common in the internet. Moreover, the published methods suffer from the small size of the evaluation database and the lack of quantitative evaluations (including sensitivity and false alarm rates).

Related work includes the method of Schmid *et al.* [2] for matching an image to an image database. They used local gray value invariants detected at corner points and a voting algorithm for the retrieval of correct matches. Their method is limited by scale constraints, due to stability issues of the corner detector. The extracted invariants were tested only for zooming by a factor of two, rotation, viewpoint change and partial visibility, without addressing other common attacks such as down scaling or compression. No complete quantitative evaluation was presented other than recognition rates for a few examples. Datta *et al.* [3] developed a method where they used texture measures to segment an image into regions and extracted a feature vector of invariant moments and region properties. The extracted vector was used for image authentication against images under certain transformations. Transformations used included down scaling to 50%, cropping up to 10% and rotation limited to 90° and 180°. However, they only presented results comparing the distance of the vector of an original image to that of transformed images without demonstrating how such measures would perform in a general database of other images. Moreover, they only used a single image for their evaluation. Johnson *et al.* [4] proposed another system based on salient points. They detected corner points in an image and used normalized cross-correlation between the neighborhoods of points to select salient corner points for image representation. Their method was designed to withstand affine transformations but, since it is based on gradient detectors for corner location, it is sensitive to blurring and compression attacks. No evaluation other than a few examples was presented. Recently, Seo *et al.* [5] developed a fingerprinting method based on the Radon Transform, log mapping and the Fourier transform. They used bit error rate (BER) to evaluate the performance of their algorithm in the presence of several kinds of image degradations. The results for four images showed good performance for compression, smoothing, rotation and scaling. However, the method did not perform well for cropping. The authors did not present BER results in-between the four test images or in a larger database to give some indication of false alarm rates.

Another research area closely linked to this work is the field of hashing, which is more concerned with security issues and image authentication. Schneider *et al.* [6] presented a methodology for designing content-based digital signatures for image authentication but did not provide any experimental results. Venkatesan *et al.* [7] proposed an algorithm using randomized signal processing strategies for a nonreversible compression of images into random binary strings. They evaluated the robustness of their algorithm on a variety of attacks with satisfactory results. However, they used small degrees of scaling and rotation. Johnson *et al.* [8] used compressed dither-based sequences for secure image hashing. They evaluated their method on a small set of images and limited attacks (i.e 1% rotation, 1% cropping). An overview of image security techniques focusing on copyright protection was reported by Wolfgang and Delp [9].

In this manuscript, we examine the use of novel color-based descriptors for an image fingerprinting system and its robustness to most common attacks in degrees beyond which an image would no longer be the same. It is well known that color and more specifically, color histogram is a robust global descriptor as long as there is uniqueness in the color pattern held against the pattern in the rest of the data set [1]. However, there are many factors affecting the performance of color-based descriptors. We identified the following factors and examined the performance of our system in relation to them: a) various quantization methods, b) various color descriptors including color-only and color-spatial information, c) reduced number of colors, d) type of histogram-similarity measures, including measures incorporating perceptual similarity between colors, e) effect of transformations, including scale, cropping, compression, smoothing, rotation, additive noise and luminance change on the performance of a fingerprinting system, and f) effect of data set used, by evaluating color descriptors on a set of random images but also on sets of similar images. The method was evaluated on a relatively large database of 27570 transformed images, produced from 919 original images, from which meaningful conclusions can be deduced regarding the fingerprinting performance when using color descriptors for image representation. The contribution of this work is the thorough analysis and comparison of the factors affecting the performance of color-based descriptors when applied to the specific problem of image fingerprinting and the comprehensive evaluation of such an application on a large data set.

The manuscript is organized as follows. In Section II, the database used for the evaluation, along with a description of the transformations that will be used to examine the robustness of the algorithm are presented. In Sections III and IV, the methods for the extraction of image fingerprints as well as the matching process are described in detail. Section V includes the performance evaluation of the method, along with a discussion of the results. Finally, this study is summarized and conclusions are presented in Section VI.

TABLE I
IMAGE DATABASE USED FOR EVALUATION OF COLOR DESCRIPTORS

| Set | Number of images |
|---|---|
| BAL-Original Set | 450 |
| Image categories | |
| -campusinfall | 48 |
| -football | 48 |
| -springflowers | 48 |
| -sanjuans | 48 |
| -cherries | 55 |
| -greenlake | 48 |
| -leaflessflowers | 48 |
| -swissmountains | 30 |
| -yellowstone | 48 |
| -cannonbeach | 48 |
| Total Similar-Original Set | 469 |
| All-Original Set | 919 |
| Transformed-Image Set | (919*30 transformations)= 27570 |

## II. MATERIALS

A database of 919 color images was used to evaluate the method. The database consisted of the following subsets: a) The BAL-Original set, including 450 randomly selected art images provided by the Bridgeman Art Library (BAL), b) The Similar-Original set, including 469 images from ten sets of categories, acquired from http://www.cs.washington.edu/research/image-database/. The categories along with their corresponding number of images are tabulated in Table I. The Similar-Original set was used to examine the performance of color descriptors on similar images. All images were in JPEG format and had variable sizes. The set of 30 transformations shown in Table II was applied to each image using MATLAB (The MathWorks, Inc.) software. The images were resized using nearest neighbor interpolation. For the image rotation, it has to be noted that MATLAB adds a black frame around the image, thus producing an additional source of degradation, except for the case of 90°. We chose not to remove that frame for the calculation of the descriptors since we wanted to treat it as an additional form of attack. Illumination changes were implemented by using nonlinear mapping of the color values for different gamma values, keeping in mind that a gamma value of 1.0 indicates linear mapping. Moreover, contrast adjustment was implemented by stretching the image so that 2% of low and high intensities were saturated. Additive noise was implemented using three types of noise: salt and pepper, gaussian and speckle. All parameters used to create the different transformations are given in Table II, along with the notation for the different subsets. A total of 30 transformed images were created for each of the 919 original images, resulting in the Transformed-Image set of 27 570 images.

## III. COLOR-BASED FINGERPRINT EXTRACTION

An image fingerprinting system consists mainly of two parts: fingerprint extraction and fingerprint matching. In the first part a descriptor is extracted from each image and is used to create an indexed database. In the second part, the index for an image (query image) is compared to the indices of the rest of the database (target images), using some kind of similarity measure to determine close matches between the query image and target images. For this particular application, only transformed versions of the same image are considered good matches and should be retrieved.

The fingerprint extraction procedure involves the quantization of the image colors and the calculation of color histograms based on the resulting colors. As can be seen below, the choices of the quantization method and the colors used have a direct effect on the performance of the fingerprinting algorithm.

### A. Color Quantization

We examined two methods for color quantization. One involves the sampling of a color space using *fibonacci lattices*, as will be explained below. With this method, it was possible to control the number of colors in the resulting palette, and examine the effect of various color palettes on the algorithm performance. The other method uses a predefined color palette, which was designed to match human perception of colors. In this section the two methods for color quantization are described in detail.

*1) Color Quantization Using Fibonacci Lattices:* The first quantization method was an implementation of a palette generation procedure proposed by Mojsilovic *et al.* [10]. They used fibonacci lattices to quantize the $(a, b)$ plane of $CIELab$ space. The technical details of this method can be found elsewhere [10]. Briefly, the method consists of the following procedure.

First, images are transformed from the $RGB$ space to the $CIELab$ space as described in [11]. Chrominance sampling is performed on a lattice defined by a set of points $z_n = n^{\delta} e^{j2\pi n\tau}$, where $\tau,\ \delta \in \mathbb{R},,\ n \in \mathbb{Z}$. Parameters $\tau$ and $\delta$ influence the distribution of its points in the plane. For $\tau = (\sqrt{5}-1)/2$ and $\delta = 1/2$, this lattice is called a *Fibonacci lattice*. For sampling of a color space, the lattice can be rotated by a small angle $\alpha$ for better alignment with the boundaries of the color space. In such case, the lattice points are given by

$$z_n = n^{\delta} e^{j\theta}; \quad \theta = 2\pi n\tau + \alpha; \quad \alpha, \tau, \delta \in \mathbb{R} \quad n \in \mathbb{Z}. \quad (1)$$

To design a palette, the luminance axis is quantized into $N_L$ levels. For each luminance level, $N_p$ points are generated using (1). The lattice is scaled so that it covers the entire $(a, b)$ chrominance plane. The points chosen in this way are uniformly distributed, assuring that, there is neither crowding in the center nor scarcity at the borders. But, since the slice section of the $CIELab$ color space is not circular, some of the generated points can lie outside the valid colors in the $RGB$ cube space. Therefore, all the points whose $R, G,$ or $B$ value was not within [0, 1] range were discarded. The remaining points defined the palette. The parameters defining a palette are the following:

- $N_L$ number of discrete luminance levels;
- $N_p$ number of generated pixels per slice;
- $\tau, \delta$ lattice parameters;
- $\alpha$ initial angle.

We experimented with four palettes of various sizes, created using the above procedure. All had the parameter $\tau$ set to $(\sqrt{5}-$

TABLE II
TRANSFORMATIONS APPLIED TO ORIGINAL IMAGES

| Image Transformations | Degree | Number of attacks |
|---|---|---|
| Scaling | 25%, 50%, 75%, 125%, 150%, 200% | 6 |
| Rotation | 10°, 20°, 30°, 90° | 4 |
| Cropping - both sides | 10%, 20%, 30% | 3 |
| Compression - JPEG | 25, 50, 75 (quality factor) | 3 |
| Blurring (median) | 3x3, 5x5, 7x7 | 3 |
| Combination attacks | Rotation 10°, cropping 10%, resizing to 25%, median filtering 5x5, compression with quality factor 50 | 1 |
| Additive noise-salt and pepper | 0.005, 0.01 (noise density) | 2 |
| Additive noise - speckle | 0.01, 0.02 (variance of uniformly distributed noise) | 2 |
| Additive noise - gaussian | 0.0025, 0.005 (noise variance) | 2 |
| Illumination change | 0.5, 1.5, 2.0 (gamma value) | 3 |
| Contrast adjustment | 0.02 (saturation value) | 1 |
| TOTAL | | 30 |

TABLE III
PARAMETERS USED FOR CREATING PALETTES OF VARIABLE NUMBER OF COLORS USING FIBONACCI LATTICES

| # of colors | 10 | 24 | 43 | 101 |
|---|---|---|---|---|
| $N_L$ | 3 | 4 | 6 | 9 |
| Luminance levels | 45, 75, 99 | 36, 51.6, 66.6, 81.3 | 0, 40, 60, 75, 90, 100 | 5, 20, 40, 50, 63, 75, 85, 94, 98 |
| $N_p$ | 10 | 15 | 30 | 56 |
| $\delta$ | 0.5 | 0.35 | 0.6 | 0.5 |
| $\alpha$ | 0.05 | 0.1 | 2 | 0.5 |

$1)/2$. Using the parameters given in Table III, four palettes were generated with sizes of 10, 24, 43, and 101 colors.

*2) Color Palette Based on Macbeth Color Checker Chart:* The second type of quantization involved the use of a palette based on a commercial color chart known as the Gretag Macbeth Color Checker. The Macbeth chart is a standard used to test color reproduction systems and it consists of 24 colors, scientifically prepared to represent a variety of different naturally occurring colors. The procedure for the generation of the color chart is reported in [12]. We created a color palette based on the Macbeth $xyY$ values found in [11, Table G.10]. The $xyY$ coordinates were transformed to $CIEXYZ$ using

$$X = x * \frac{Y}{y}, \quad Y = Y, \quad Z = (1 - x - y) * \frac{Y}{y}$$

The resulting $XYZ$ coordinates were transformed to $Lab$ values using the following equations:

Let $var\_X = X/95.047$, $var\_Y = Y/100.000$, $var\_Z = Z/108.883$, based on specifications for the D65 illuminant. Then the following pseudo-code below was used according to [11]:

$$\text{if } (\text{var\_X} > 0.008856) \text{ then var\_X} = \text{var\_X}\hat{}(1/3)$$
$$\text{else var\_X} = (7.787 * \text{var\_X}) + (16/116)$$
$$\text{if } (\text{var\_Y} > 0.008856) \text{ then var\_Y} = \text{var\_Y}\hat{}(1/3)$$
$$\text{else var\_Y} = (7.787 * \text{var\_Y}) + (16/116)$$
$$\text{if } (\text{var\_Z} > 0.008856) \text{ then var\_Z} = \text{var\_Z}\hat{}(1/3)$$
$$\text{else var\_Z} = (7.787 * \text{var\_Z}) + (16/116)$$
$$\text{L} = (116 * \text{var\_Y}) - 16$$
$$\text{a} = 500 * (\text{var\_X} - \text{var\_Y})$$
$$\text{b} = 200 * (\text{var\_Y} - \text{var\_Z})$$

The resulting palette of 24 $Lab$ coordinates, thereof defined as the Macbeth palette, is given in Table IV.

### B. Color Descriptors

Color histograms have been used extensively in CBIR systems due to the simplicity of their calculation and their robustness to common image transformations. Many types of histograms exist in the literature falling mainly into two categories: those based only on the quantized colors and those incorporating information on the spatial color distribution. We examined the use of both kinds of color histograms. The first one was the *normalized color-only histogram*, providing the occurrence of a color in an image $I$ given by:

$$H_{I_i} = \frac{N_{I_i}}{N_I}, \qquad i = 1, \dots C_p,$$

TABLE IV
$Lab$ VALUES FOR MACBETH-BASED COLOR PALETTE

| Color Names | $L$ | $a$ | $b$ |
|---|---|---|---|
| Dark skin | 37.9338 | 14.6681 | 12.2931 |
| Light skin | 66.3849 | 17.1286 | 13.5070 |
| Blue sky | 51.0733 | 3.1987 | -25.8582 |
| Foliage | 43.1990 | -14.3516 | 19.7841 |
| Blue flower | 56.3595 | 16.0060 | -29.5825 |
| Bluish green | 71.5986 | -27.0852 | -2.9909 |
| Orange | 61.6999 | 31.3437 | 56.1121 |
| Purplish blue | 41.2180 | 20.3398 | -47.0477 |
| Moderate red | 51.5784 | 46.4966 | 11.9247 |
| Purple | 30.7850 | 28.2360 | -26.3364 |
| Yellow green | 72.4249 | -24.6892 | 56.8408 |
| Orange yellow | 71.5986 | 16.5210 | 64.5043 |
| Blue | 29.6898 | 29.8747 | -54.9689 |
| Green | 55.4742 | -38.4890 | 31.1622 |
| Red | 41.2180 | 53.9596 | 23.8431 |
| Yellow | 81.3498 | 0.2924 | 76.8184 |
| Magenta | 51.5784 | 52.5476 | -19.6328 |
| Cyan | 51.5784 | -18.9014 | -30.6055 |
| White | 96.0046 | 4.9936 | -5.2939 |
| Neutral 8 | 81.3498 | 4.3402 | -4.6013 |
| Neutral 6.5 | 66.6752 | 3.6860 | -3.9077 |
| Neutral 5 | 51.5784 | 3.0129 | -3.1941 |
| Neutral 3.5 | 35.9860 | 2.3177 | -2.4571 |
| Black | 20.5584 | 1.6299 | -1.7 |

where $N_{I_i}$ is the number of pixels with color $i$, $N_I$ is the total number of pixels in the image $I$ and $C_p$ is the number of colors in the palette.

The normalized color histogram depends only on the color properties of an image without providing any information on the spatial distribution of colors. In order to examine any advantages of using histograms incorporating color-spatial information for image fingerprinting, we also experimented with the *spatial chromatic histogram* proposed by Cinque *et al.* [13]. The spatial chromatic histogram descriptor gives information on color presence, and color spatial distribution. Technical details related to the calculation of the histogram can be found in [13]. Here we give a brief outline of this method.

Let $a_I(i)$ be the absolute number of the pixels in image $I$ of size nxm having $i$ colors and $h_I(i)$ be the normalized color histogram. Then the following features are defined:

Baricenter:$\mathbf{b}_I(i) = (\bar{x}_i, \bar{y}_i)$, where

$$\bar{x}_i = \frac{1}{n}\frac{1}{a_I(i)} \sum_{c(x,y)=i} x$$

and

$$\bar{y}_i = \frac{1}{m}\frac{1}{a_I(i)} \sum_{c(x,y)=i} y$$

for all palette colors found in the image. The baricenter gives a measure of the position of pixels having the same color. The standard deviation is defined as:

$$\sigma_I(i) = \sqrt{\frac{1}{a_I(i)} \sum_{c(\mathbf{p})=i} \mathrm{d}\left(\mathbf{p}, \mathbf{b}_I(i)\right)^2},$$

where $\mathbf{p}$ is a pixel in relative coordinates, $c(\mathbf{p}) = i$ denotes a pixel $\mathbf{p}$ having the color $i$, and $d(\mathbf{p}, \mathbf{b}_I(i))$ is the Euclidean distance between two pixels. The standard deviation gives a measure of the pixel spread around the baricenter. Then, the spatial chromatic histogram is a vector of the three above mentioned features

$$S_I(i) = \left(h_I(i), \mathbf{b}_I(i), \sigma_I(i)\right), \qquad i = 1, \ldots C_p,$$

using the notation mentioned above. For each image, we calculated the normalized color histogram and the spatial chromatic histogram using the palettes described in Section III-A. The resulting vectors were saved as indices for all the images in the database. The average size of the vectors per image, depending on the presence of zeros, was 0.18 KB, 0.36 KB, 0.56 KB, 0.95 KB, and 0.46 KB for the Fibonacci-10, Fibonacci-24, Fibonacci-43, Fibonacci-101 and Macbeth-24 palettes respectively.

## IV. Color-Based Fingerprint Matching

In order to retrieve the transformed versions of an original image, the descriptors of that image were matched against all image descriptors from the Transformed-Image Set. Image matching between color histogram descriptors depends on the choice of similarity measures, so we investigated the use of four different measures given below to match the normalized color histograms. A fifth measure was used for the spatial chromatic histogram.

We used modified versions of some measures to scale them in the range 0 to 1, with 1 denoting a perfect match. Vectors $H_1$, $H_2$ denote the normalized color histograms extracted from images $I_1$, $I_2$ respectively, and $C_p$ denotes the number of colors in a palette. The matching measures used were the following.

i) *Scaled $L_1$-norm* distance, defined as

$$\begin{aligned}\mathrm{d}_{L_1}(H_1, H_2) &= 1 - 0.5 * \|H_1 - H_2\|_{L_1} \\ &= 1 - 0.5 * \sum_{i=1}^{C_p} |H_{1_i} - H_{2_i}|.\end{aligned}$$

ii) *Scaled $L_2$-norm* distance, defined as

$$\begin{aligned}\mathrm{d}_{L_2}(H_1, H_2) &= 1 - \frac{1}{\sqrt{2}} * \|H_1 - H_2\|_{L_2} \\ &= 1 - \frac{1}{\sqrt{2}} * \sqrt{\sum_{i=1}^{C_p} (H_{1_i} - H_{2_i})^2}.\end{aligned}$$

The above two measures are scaled versions of the $L_1$-, and $L_2$-norms which have been previously used for matching color histograms [14], [15].

iii) *Scaled Histogram Intersection* defined as

$$\mathrm{d}_{HI}(H_1, H_2) = \sum_{i=1}^{C_p} \min\left(H_{1_i}, H_{2_i}\right) * \left(1 - |H_{1_i} - H_{2_i}|\right).$$

This measure is a modified version of the Histogram Intersection measure [14]. Only colors present in the image contribute to this metric.

iv) *Histogram Quadratic Distance*, as defined in [16]

$$\mathrm{d}_{HQ}(H_1, H_2) = \sqrt{\sum_{i=1}^{C_p}\sum_{j=1}^{C_p} (H_{1_i} - H_{2_i}) * a_{ij} * (H_{1_i} - H_{2_i})}$$

where $a_{ij}$ is a similarity matrix between colors $i$ and $j$ and is expressed by

$$a_{ij} = 1 - \frac{\mathrm{d}_{ij}}{\max_{ij}(\mathrm{d}_{ij})},$$

and $\mathrm{d}_{ij}$ is the Euclidean distance between two colors $i$, $j$ in the $CIELab$ space given by

$$\mathrm{d}_{ij} = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}.$$

Finally, in order to compare the spatial chromatic histograms between images $I_1$, $I_2$, we used the *spatial chromatic distance* [13] defined as

$$\mathrm{d}_{Sch}(1,2) = \sum_{i=1}^{C_p} \min\left(h_1(i), h_2(i)\right) * \left(\frac{\sqrt{2} - \mathrm{d}\left(\mathbf{b}_1(i), \mathbf{b}_2(i)\right)}{\sqrt{2}} + \frac{\min\left(\sigma_1(i), \sigma_2(i)\right)}{\max\left(\sigma_1(i), \sigma_2(i)\right)}\right)$$

using the notation used in Section III-B.

## V. Results and Discussion

The use of color-based descriptors for the application of image fingerprinting was evaluated using Receiver Operator Characteristic (ROC) analysis [17]. Specifically, the evaluation consisted of taking the color descriptors from each original image and matching them against the descriptors from each of the 27 570 images in the Transformed-Images Set. Matches were determined by applying a threshold on the similarity measures and identifying those images with measures higher than the threshold. The goal was, for each image, to retrieve all of its 30 transformed versions without any false positive matches. In order to evaluate the performance of this experiment, the well-known measures *True Positive Fraction* (TPF or sensitivity) and *False Positive Fraction* (FPF) were used [17], defined as follows:

$$\mathrm{TPF_I} = \frac{\text{number of correctly retrieved versions of I}}{\text{total number of versions of I}},$$

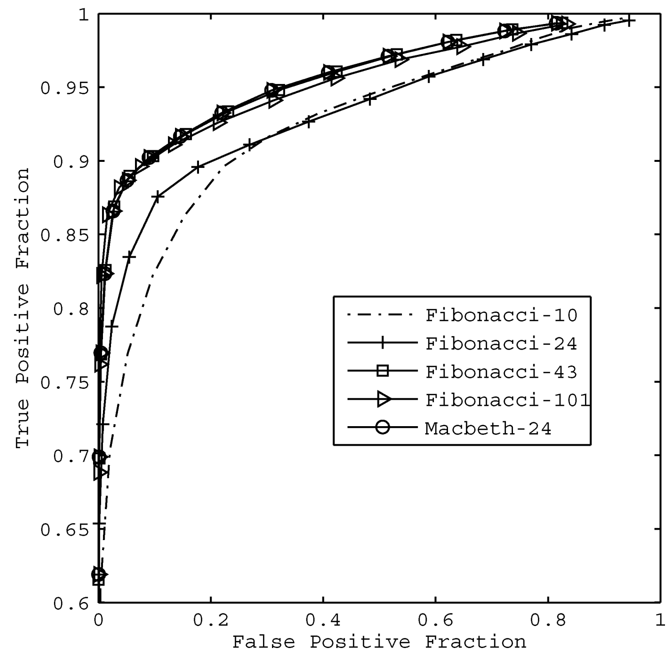



Fig. 1. ROC curves comparing different color palettes. The normalized color histogram with the scaled histogram intersection similarity measure was used to match images.

where *versions* refers to transformed images of I, and

$$\mathrm{FPF_I} = \frac{\text{number of falsely retrieved images}}{\text{total number of retrieved images}}.$$

By sweeping the threshold from 0 to 1 and averaging the measures of TPF and FPF over all images in the Original Image Set, pairs of (Average TPF, Average FPF) were collected and were used to construct ROC curves. An ideal ROC curve passes through the {TPF, FPF} pair point of 1.0,0.0.

The ROC curves in Fig. 1 compare the effect of the choice of palette on the performance of the system to match the 919 images in the All-Original set with their transformed versions. Each curve was constructed using one of the palettes and the *scaled histogram intersection* measure. It can be seen from the curves that for the four palettes created with the Mojsilovic method the performance increases as the number of colors increase from ten to 43 and is maintained at a similar level when the colors increase to 101. It is interesting to notice though, that almost equal performance with the palettes of 43 and 101 colors can be achieved with the Macbeth palette of only 24 colors. Moreover, comparing the two quantization methods for the equal number of 24 colors, the results with the Macbeth quantization outperform the ones using the Fibonacci palette. The small size of the descriptor is beneficial for the efficient storage and functionality of the fingerprinting system. A similar pattern was observed for the rest of the histogram measures.

Another parameter affecting the performance of the fingerprinting performance was the choice of similarity measures. The ROC curves for the five similarity measures described in Section IV are plotted in Fig. 2. The first four measures, *scaled $L_1$-norm*, *scaled $L_2$-norm*, *scaled histogram intersection* and *histogram quadratic distance*, apply to the normalized

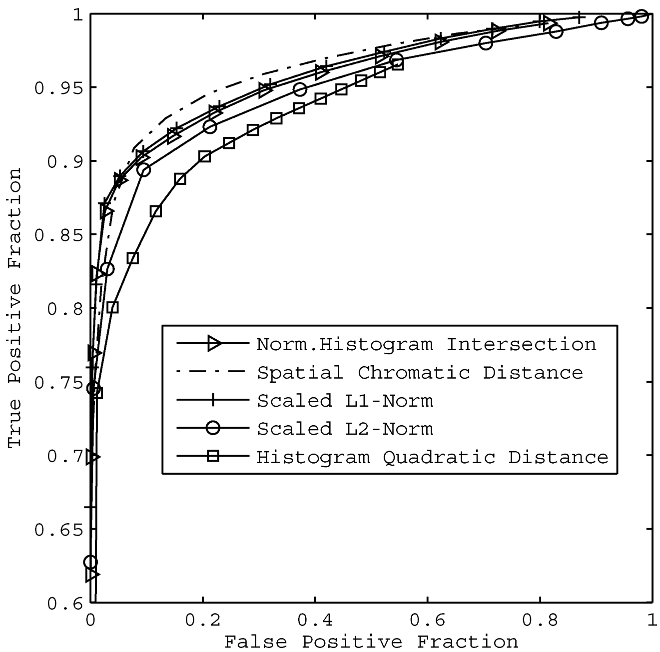



Fig. 2. ROC curves comparing the different similarity measures for matching of color histograms. The spatial chromatic distance was used to match spatial chromatic histograms whereas the rest of the measures were used to match normalized color histograms.

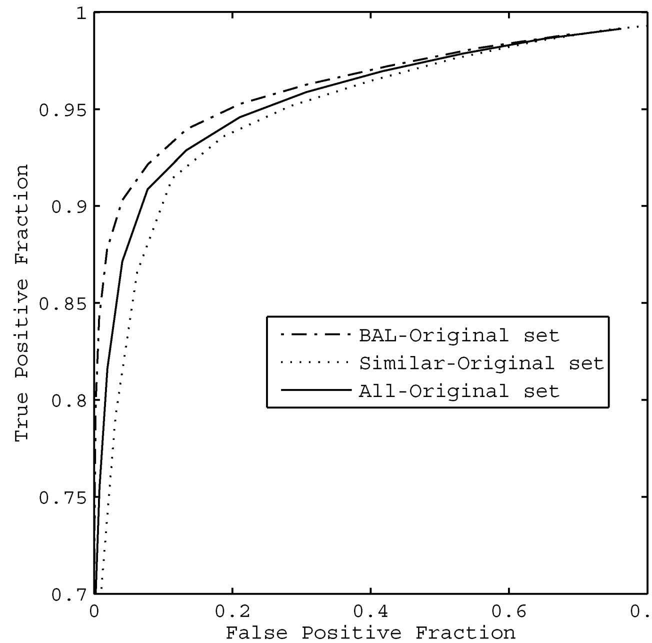



Fig. 3. ROC curves comparing the performance of the fingerprinting system for the different data sets. The plot shows that performance is degraded when similar images are included in the dataset.

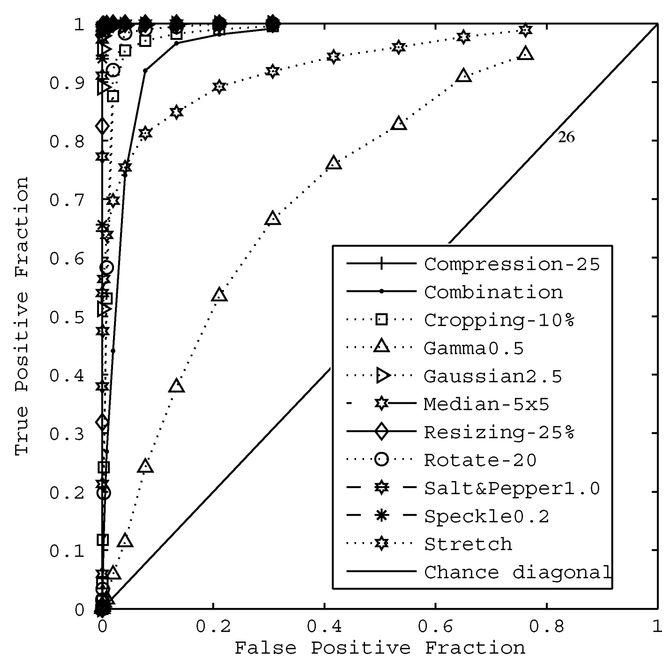



Fig. 4. ROC curves for different transformations. Matching was done using the spatial chromatic distance and the Macbeth color palette consisting of 24 colors.

color-only histogram, whereas the *spatial chromatic distance* was used to match spatial chromatic histograms. All similarity measures used the 24 colors of the Macbeth palette. It can be seen from the graph, that the spatial chromatic histogram shows the best performance for sensitivities above 90% whereas the quadratic histogram measure shows the worst performance. The color-spatial information was proven to be useful for this database that includes images that have the same colors but in different locations, resulting in more specific discrimination.

In order to examine the performance of the color descriptors when similar images to the query are present in the database, a comparison was made between results taken using the BAL-Original set and the Similar-Original set. As described in Section II, the BAL-Original set consisted of images randomly drawn from a generic art database whereas the Similar-Original set consisted of images belonging in specific subject categories. The results are demonstrated using the ROC curves of Fig. 3. The 24-color Macbeth palette and the spatial chromatic distance were used in this experiment. It can be seen from the plots that the overall system performance was degraded when the set of images included similar images since those images not only tend to share similar color vectors affecting the discrimination ability of the color descriptors but often have the same color layout.

Another experiment examined the robustness of color-based descriptors for specific transformations. The ROC curves of Fig. 4, taken using the spatial chromatic distance, demonstrate the invariance of color-based descriptors to resizing, to JPEG compression, to additive noise, to median filtering, and to rotation of 90°, keeping in mind that for that particular angle no black frame was added to the image. As expected, the performance dropped for higher degrees of cropping and for rotations where a black frame is added, since some colors may get lost when a large piece of it is cut. It can be seen for the plots that color-based descriptors fail to match images when their illumination contrast is changed since the colors of the image are severely altered. The effect is maximal when nonlinear mapping was used and less severe when contrast stretching was used. The performance of the system for different transformations is shown in more detail in Fig. 5. Regarding additive noise, the system seems to be invariant to salt and pepper and speckle noise and is more affected by Gaussian noise. This can be explained by the fact that salt and pepper and speckle noise

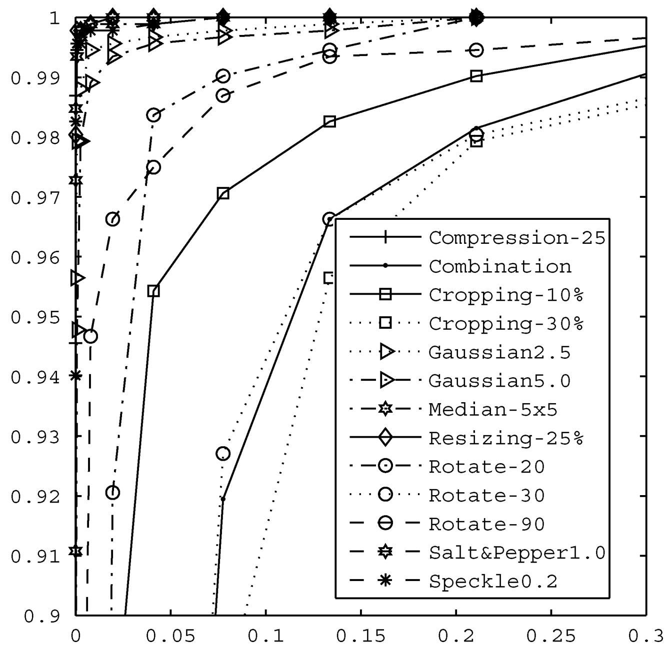



Fig. 5. Closeup of the ROC curves of Fig. 4 for different image transformations.

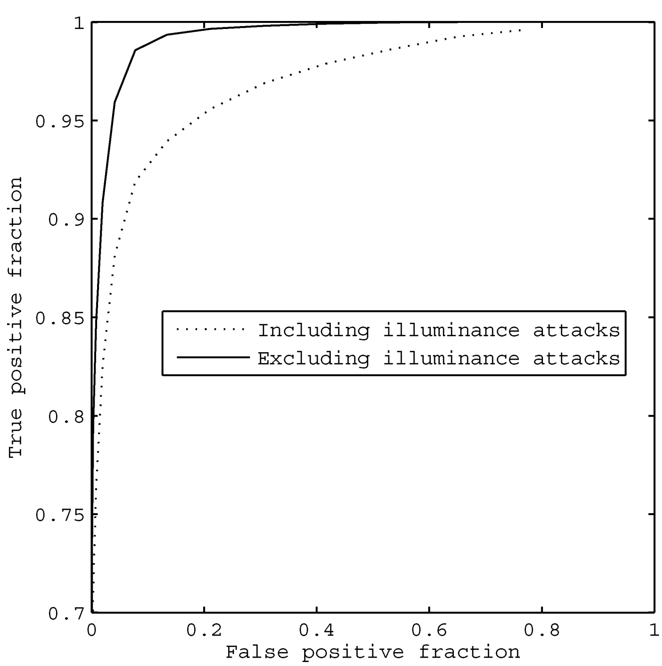



Fig. 6. Comparison of overall system performance with illumination contrast transformed images present or absent. At the false positive fraction of 4%, the system can achieve sensitivities of 96% and 88%, when illumination contrast transformations are absent and present respectively.

may alter the overall color distribution of an image but the majority of the pixels retain their color values.

Finally, the plot of Fig. 6 compares the overall performance of the system with all transformations included to the performance of the system with the illumination contrast transformations (four out of 30 transformations, or 13.3% of all images) removed. The spatial chromatic distance was used to match images. The overall performance of the system achieves a sensitivity of about 96% at a false positive fraction of 4% and a reduced sensitivity of 88% when illumination attacks are removed and are present respectively.

The experimental results demonstrate the robustness of color-based descriptors for the application of image fingerprinting, where the goal is to retrieve only transformed versions of the same image. It was shown that very high performance could be achieved with only 24 colors if the palette is carefully designed. The performance degrades significantly when the query space includes a large percentage of images whose illuminance was altered. However, there is a limit to the degree that an image can be changed without being transformed to a different image. The size of the descriptors is small enough for efficient storage and matching. It was also shown that a histogram incorporating information regarding the spatial distribution of colors did have an advantage over color-only measures since spatial information is useful in distinguishing between images with similar colors. The algorithm was evaluated against high degrees of the attacks: scaling, rotation, cropping, smoothing, additive noise, and compression, which can be performed easily with the availability of current image manipulation software. The results were very accurate keeping in mind that similar images where included and also that the Transformed Image Set included images that were badly deteriorated as those smoothed with a median filter of size $7 \times 7$, those cropped by 30% in both directions, those which had a black frame added around the rotated image, those with added significant noise, and the ones with combined attacks. Yet, the results are worse than the ones that can be obtained by watermarking, whose probability of errors are orders of magnitude smaller than the ones achieved using the proposed fingerprinting method. However, it must be noted that fingerprinting is a passive approach and can be installed very easily for broadcast monitoring. Also, image fingerprinting can be applied on images already distributed on the internet for which watermarking has not been applied. Another possible use of this approach is a validation step at the output of a CBIR method. Transformations of the retrieved images could be produced and matched to the source image in order to reduce false alarms. The combined use of watermarking and fingerprinting will be examined in future work, along with techniques to overcome attacks involving illumination contrast change.

## VI. Conclusion

In this manuscript, we presented an image fingerprinting system that was designed to retrieve transformed versions of a query image from a large database. We examined the use of two different quantization methods and found that very good results can be achieved with a palette of only 24 selected colors. Moreover, we examined the use of color-only and spatial chromatic histograms and the effect of different similarity measures. The system was evaluated on a database consisting of 919 original images and their corresponding 27 570 transformation images. Results showed that the spatial chromatic distance outperformed the other measures since it could differentiate between images with the same colors but different color layout. As expected, the results were worse for a data set that included similar images compared to a set of randomly-drawn images. It was also shown that the color-based descriptors used in this manuscript were robust to high degrees of attacks such as scaling, rotation, cropping, smoothing, compression, and additive noise and achieved a sensitivity of 96% at the small false positive fraction of 4%, despite the presence of similar images. Finally, the effect of illuminance changes was examined. It was found that at the same false positive fraction of 4%, the sensitivity was reduced to 88% when 13% of all transformations involved changing the illuminance of the images. The overall performance of the system is encouraging for the use of color, and particularly spatial chromatic descriptors for image fingerprinting.

## ACKNOWLEDGMENT

The authors would like to thank Dr. N. Nikolaidis for useful discussions regarding the contents of this manuscript, the Bridgeman Art Library for providing the art database for this study and the reviewers for their useful comments.

## REFERENCES


[1] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, 2000.

[2] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 530–535, 1997.

[3] A. Datta, N. Lobo, and J. Leeson, "Novel feature vector for image authentication," in *IEEE Int. Conf. Multimedia and Expo (ICME 2003)*, Baltimore, MD, 2003, vol. 2, pp. 221–224.

[4] N. Johnson, Z. Duric, and S. Jajodia, "On "fingerprinting" images for recognition," in *Fifth International Workshop on Multimedia Information Systems (MIS'99)*, Indian Wells, CA, 1999.

[5] J. Seo, J. Haitsma, T. Kalker, and C. Yoo, "Affine transform resilient image fingerprinting," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'03)*, 2003, vol. 3, pp. 61–64.

[6] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proc. 1996 Int. Conf. Image Processing (ICIP 1996)*, 1996, pp. 227–230.

[7] R. Venkatesan, S.-M. Koon, M. H. Jakubowski, and P. Moulin, "Robust image hashing," in *Proc. 2000 Int. Conf. Image Processing (ICIP 2000)*, 2000.

[8] M. Johnson and K. Ramchandran, "Dither-based secure image hashing using distributed coding," in *Proc. 2003 Int. Conf. Image Processing (ICIP 2003)*, 2003, pp. 495–498.

[9] R. Wolfgang and E. Delp, "Overview of image security techniques with applications in multimedia systems," in *Proc. SPIE Conf. Multimedia Networks, November 1997*, Dallas, TX, 1997, vol. 3228, pp. 297–308.

[10] A. Mojsilovic and E. Soljanin, "Color quantization and processing by fibonacci lattices," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1712–1725, Nov. 2001.

[11] A. Glassner, *Principles of Digital Image Synthesis*, ser. The Morgan Kaufman Series in Computer Graphics and Geometric Modeling. San Francisco, CA: Morgan-Kaufmann, 1995.

[12] C. McCamy, H. Marcus, and J. Davidson, "A color-rendition chart," *J. Appl. Photographic Eng.*, vol. 2, no. 3, pp. 95–99, 1976.

[13] L. Cinque, G. Ciocca, S. Levialdi, A. Pellicano, and R. Schettini, "Color-based image retrieval using spatial-chromatic histograms," *Image Vis. Comput.*, vol. 19, pp. 979–986, 2001.

[14] M. Swain and D. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11–32, 1991.

[15] M. Stricker and M. Swain, "The capacity of color histogram indexing," in *Proc. Computer Vision and Pattern Recognition (CVPR'94)*, 1994, pp. 704–708.

[16] J. Haffner, H. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 17, no. 7, pp. 729–736, 1995.

[17] J. Swets and R. Pickett, *Evaluation of Diagnostic Systems: Methods From Signal Detection Theory*. New York: Academic, 1982.

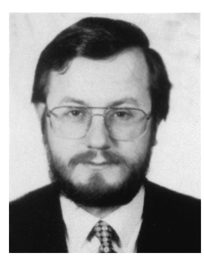

**Marios A. Gavrielides** received the B.S. and M.S. degrees in electrical engineering in 1994 and 1996, respectively, from the University of South Florida, Tampa, and the Ph.D. degree in biomedical engineering from Duke University, Durham, NC, in 2001.

From 2002 to 2004, he was a Postdoctoral Researcher at the Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece. Currently, he is a Research Fellow at the Joint Laboratory for the Assessment of Medical Imaging Systems between the Center for Devices and Radiological Health (CDRH) of the Food and Drug Administration (FDA) and the National Institute of Biomedical Imaging and Bioengineering (NIBIB), Rockville, MD. His research interests include medical imaging processing and the development of computer-aided detection and diagnosis methods in the fields of digital mammography, computed tomography, and microscopy imaging.

Dr. Gavrielides was a recipient of a Susan G. Komen Breast Cancer foundation Dissertation Award in 2000 for his Ph.D. thesis work on automatic diagnosis of breast cancer.

**Elena Šikudova** received the M.S. degree in mathematics in 1994 from the Comenius University, Bratislava, Slovakia, where she is currently pursuing the Ph.D. degree in the Department of Applied Informatics, Faculty of Mathematics, Physics and Informatics.

From 2002 to 2004 she was a Researcher with the Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece. Currently, she is a Lecture Assistant at Comenius University. Her research interests include color image processing and image retrieval.



**Ioannis Pitas** (SM'94) received the Diploma of electrical engineering in 1980 and the Ph.D. degree in electrical engineering in 1985, both from the Aristotle University of Thessaloniki, Thessaloniki, Greece.

Since 1994, he has been a Professor with the Department of Informatics, Aristotle University of Thessaloniki. From 1980 to 1993, he served as Scientific Assistant, Lecturer, Assistant Professor, and Associate Professor with the Department of Electrical and Computer Engineering at the same University. He served as a Visiting Research Associate with the University of Toronto, Canada, University of Erlangen-Nuernberg, Germany, Tampere University of Technology, Finland, as Visiting Assistant Professor with the University of Toronto, and as a Visiting Professor with the University of British Columbia, Vancouver, Canada. He was a lecturer in short courses for continuing education. He has published more than 145 journal papers, 380 conference papers, and contributed in 18 books in his areas of interest. He is the coauthor of the books *Nonlinear Digital Filters: Principles and Applications* (Boston, MA: Kluwer, 1990), and *3-D Image Processing Algorithms* (New York: Wiley, 2000) *Nonlinear Model-Based Image/Video Processing and Analysis* (New York: Wiley, 2001); and author of *Digital Image Processing Algorithms and Applications* (New York: Wiley, 2000). He is the editor of the book *Parallel Algorithms and Architectures for Digital Image Processing, Computer Vision and Neural Networks* (New York: Wiley, 1993). His current interests are in the areas of digital image and video processing and analysis, multidimensional signal processing, watermarking, and computer vision.

Dr. Pitas has also been an invited speaker and/or member of the program committee of several scientific conferences and workshops. In the past, he served as Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, IEEE TRANSACTIONS ON NEURAL NETWORKS, IEEE TRANSACTIONS ON IMAGE PROCESSING, *EURASIP Journal on Applied Signal Processing* IJIG, IEICE, *Circuits Systems and Signal Processing* (CSSP), co-editor of *Multidimensional Systems and Signal Processing*, been member of the editorial board of six journals and guest editor of six special journal issues. He is member of the National Research Council of Greece.