An Effective Visual Descriptor Based on Color and Shape Features for Image Retrieval

Atoany Fierro-Radilla[1], Karina Perez-Daniel[1], Mariko Nakano-Miyatakea[1], Hector Perez-Meana[1], and Jenny Benois-Pineau[2]

[1] Postgraduate Section of ESIME Culhuacan, Instituto Politecnico Nacional, Av. Santa Ana No. 1000, Col. San Francisco Culhuacan
afierror@hotmail.com, mnakano@ipn.mx
[2] University of Bordeaux I, Bordeaux, France
benois-p@labri.fr

Abstract. In this paper we present a Content-Based Image Retrieval (CBIR) system which extracts color features using Dominant Color Correlogram Descriptor (DCCD) and shape features using Pyramid Histogram of Oriented Gradients (PHOG). The DCCD is a descriptor which extracts global and local color features, whereas the PHOG descriptor extracts spatial information of shape in the image. In order to evaluate the image retrieval effectiveness of the proposed scheme, we used some metrics commonly used in the image retrieval task such as, the Average Retrieval Precision (ARP), the Average Retrieval Rate (ARR) and the Average Normalized Modified Retrieval Rank (ANMRR) and the Average Recall (R)-Average Precision (P) curve. The performance of the proposed algorithm is compared with some other methods which combine more than one visual feature (color, texture, shape). The results show a better performance of the proposed method compared with other methods previously reported in the literature.

Keywords: CBIR, color descriptor, shape descriptor, dominant color, color correlogram, PHOG.

1 Introduction

In the last years, due to the technological advances, a large amount of devices such as: digital cameras, smart phones and tablets, have been developed in order to capture images and video data and on the other hand, a technological advance on high-speed internet connection, as well as the increasing storage capacities, leads to a growing size of databases. As a result Internet has become the largest multimedia database; a huge amount of information becomes available for a large number of users [1]. With large databases, it is a challenge to browse and retrieve efficiently the desirable information. The traditional annotation heavily relies on manual labor to label images with keywords, which unfortunately can hardly describe the diversity and ambiguity for image contents [2].

The Content-Based Image Retrieval (CBIR) system is a useful tool to resolve the above mentioned problem. The typical CBIR system performs two major tasks, the

A. Gelbukh et al. (Eds.): MICAI 2014, Part I, LNAI 8856, pp. 336–348, 2014.

first one is the feature extraction, where a set of features is extracted to describe the content of each image in the database; and the second task is the similarity measurement between the query image and each image in the database, using the extracted features [3]. Generally the CBIR is performed using some low-level visual descriptors such as color-based, texture-based and shape-based descriptors, which extract feature vectors from the images.

There are many methods which combine more than one visual descriptor [2-6], improving the image retrieval effectiveness. In [2], authors combine Linear Block Algorithm (LBA) for the global color feature extraction, Steerable Filter for the texture features extraction and Pseudo-Zernike Moments for extraction of the shape features which are rotation invariant. Authors of [3] combine color and texture features, in which the color features are extracted using the Color Layout Descriptor (CLD) and the texture features are obtained using the Gabor Filters. Another method that combines more than one visual descriptors is proposed in [6], in which the image is divided into six blocks then, the color space of each block is converted from RGB to HSV and a cumulative histogram is computed in order to obtain the color features, whereas to obtain the texture features of each block, four statistic features, such as energy, contrast, entropy and inverse difference from the Gray-Level Concurrence Matrix (GLCM), are computed.

In this paper we propose a method that combines local and global color information using the Dominant Color Correlogram Descriptor (DCCD) [7] and local shape information using the Pyramid Histogram of Oriented Gradients (PHOG) [8]. The proposed scheme is performed through three stages: In the first stage, the algorithm obtains global color features using the Dominant Color Descriptor (DCD) proposed by MPEG-7 [9, 10] as well as the shape information using the PHOG [10]. In the second stage, using color correlogram the correlation between central pixel and its neighborhood is calculated from the image represented by only dominant colors, and in the third stage, color and shape features are combined in order to obtain a new visual descriptor which improves the image retrieval performance. In order to evaluate the proposed visual descriptor, we use three metrics commonly used in the CBIR systems, such as ARP (Average Retrieval Precision), ARR (Average Retrieval Rank) and ANMRR (Average Normalized Retrieval Rank), as well as RP-curves. The performance of our proposed scheme is compared with some methods reported in the literature and the results show that the proposed visual descriptor improves the image retrieval performance.

The rest of this paper is organized as follows: In Section 2 we briefly describe some color-based descriptors commonly used in the literature. In Section 3, we briefly describe some shape-based descriptors reported in the literature. In Section 4, we present the proposed scheme. The results are shown in Section 5 and finally in Section 6 we present the conclusions of this work.

2 Color-Based Descriptors

Color is the basic element of image content and one of the main sensation features when a human distinguish images [9]. From the perspective of feature extraction, color-based image descriptor can be divided into two categories [12]: Global descriptor which takes into account the whole image in order to obtain the color features. In this group we have Histogram Intersection (HI) and Dominant Color Descriptor (DCD). On the other hand we have the local descriptors such as Color Correlogram (CC), Color Layout Descriptor (CLD) and Color Structure Descriptor (CSD), which obtain the color features by dividing the image into regions.

The Histogram Intersection (HI) was proposed by Swain Ballard [13]. This method is a global color descriptor and it is defined as: Given a pair of color histograms, I and M, with n bins each one, HI can be computed as:

$$HI(I,M) = \sum_{j=1}^{n} \min\left(I_j, M_j\right) \tag{1}$$

This method is robust to geometrical modification, such as rotation and scaling, as well as the variation of the image resolution [14]. The number of bins is an important factor, because the more bins are used, the image is better described but, the computational cost is increased. Another global color descriptor is the DCD, which was proposed in the MPEG-7 standard. This color descriptor replaces the whole image color information with a small number of representative colors [9]. The Dominant Color Descriptor can be defined as follows:

$$F = \{C_i, P_i\}, i = 1,2,\ldots,N, P \in [0,1], \tag{2}$$

where P_i is the percentage of the dominant color C_i.

The color correlogram (CC) is a local color descriptor which expresses how the spatial correlation of pairs of color changes with the distance [15] and it is defined as: For any pixel of color c_i in the image, the color correlogram ($\gamma_{c_i c_j}^{(k)}$) gives the probability that a pixel at distance k away from the given pixel c_i has a color c_j.The Color Layout Descriptor (CLD) is a compact descriptor which represents the color spatial distribution of visual data [16], where the color space used in this descriptor is the YCbCr. The extraction of this descriptor consists of four stages. In the first one the image is partitioned into 64 blocks where the size of each block is W/8 and H/8 with W and H denoting the width and height of the image. In the second stage, for each block, a single dominant color is selected, then in the third stage the three components of the color space are transformed into 8x8 DCT (Discrete Cosine Transform). And finally in the fourth stage, the DCT coefficients of Y, Cb and Cr color channel are quantized and their lower coefficients are extracted to form the CLD. Another local color descriptor is the Color Structure Descriptor, which was also proposed in the MPEG-7 standard. In which an image is represented by the color distribution, and the local spatial structure of color using structuring element. It is similar to color histogram but it is semantically different. The CSD is defined as $h(m), m = 1, \ldots, M$, where the bin value $h(m)$ is the number of structuring elements containing one or more pixels with color c_m. Denote I be the set of quantized index of an image and $S \in I$ be the set of quantized color index existing inside the sub-image region covered by the structuring element [17], the color histogram bins are accumulated according to

$$h(m) = h(m) + 1, m \in S \tag{3}$$

3 Shape-Based Descriptor

Shape feature is an important factor in order to identify objects as well as classification and indexing of the context semantically. In this section, three shape-based descriptors, which are Pseudo-Zernike Moments (PZM), Polar Harmonic Transform (PHT) and Pyramid Histogram of Oriented Gradients (PHOG), are described.

3.1 Pseudo-Zernike Moment (PZM)

The PZM consists of a set of complex polynomials that form a complete orthogonal set over the interior of the unit circle, $x^2 + y^2 \leq 1$ [2]. These polynomials are denoted as

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{jm\theta}, \tag{4}$$

where $\rho = \sqrt{x^2 + y^2}$ is the distance from the origin to the pixel (x, y) and θ is an angle between vector ρ and the x-axis in the clockwise direction. The radial polynomial $R_{nm}(\rho)$ is defined as:

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} \frac{(-1)^s[(2n+1-s)!]\rho^{n-s}}{s!(n-|m|-s)!(n+|m|-s)!} \tag{5}$$

The PZM of order n with repetition m is defined as:

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y)V_{nm}^*(x, y)dxdy \tag{6}$$

For a digital image of size MxN, its PZM can be computed as:

$$\check{A}_{nm} = \frac{4(n+1)}{\pi MN} \sum_{i=1}^{M} \sum_{j=1}^{N} V_{nm}^*(x_i, y_j)f(x_i, y_j) \tag{7}$$

where $\Delta x = \frac{2}{M}, \Delta y = \frac{2}{N}$

The integer numbers n and m are defined in Table 1. A numerical instability, when high-order PZMs is required, is a serious problem of the PZM, due to the amount of factorial elements in the radial polynomial.

Table 1. Principal Pseudo-Zernike Moments

Order	Moments A_{nm}	No. Moments	Order	Moments A_{nm}	No. Moments
0	A_{00}	1	4	$A_{40,}, A_{42}, A_{44}$	3
1	A_{11}	1	5	A_{51}, A_{53}, A_{55}	3
2	A_{20}, A_{22}	2	6	$A_{60}, A_{62}, A_{64}, A_{66}$	4
3	A_{31}, A_{33}	2			

3.2 Polar Harmonic Transform (PHT)

The Polar Complex Exponential Transform (PCET) [18] is one of the PHT, which is defined as (8), when the order is n and the repetition is l, $|n| = |l| = 0,1, \dots, \infty$.

$$M_{nl} = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 [H_{nl}(r,\theta)]^* f(r,\theta) r dr d\theta \tag{8}$$

where

$$H_{nl}(r,\theta) = R_n(r) e^{il\theta} \tag{9}$$

The radial kernel is a complex exponential in the radial direction, that is

$$R_n(r) = e^{i2\pi n r^2} \tag{10}$$

For a digital image of size MxN, the PCET can be computed as:

$$M_{nl} = \frac{1}{\pi} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} [H'_{nl}(x_k, y_l)]^* f'(x_k, y_l) \Delta x \Delta y \tag{11}$$

where $\Delta x = \frac{2}{M}, \Delta y = \frac{2}{N}$, we finally obtain:

$$M_{nl} = \frac{4}{\pi M N} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} [H'_{nl}(x_k, y_l)]^* f'(x_k, y_l) \tag{12}$$

3.3 Pyramidal Histogram of Oriented Gradients (PHOG)

The Pyramidal Histogram of Oriented Gradients (PHOG) is a spatial shape descriptor which represents the spatial distribution of edges and it is formulated as a vector representation [8]. The operation of PHOG consists of the following four steps.

1. Edge contour extraction: The contour of input image can be extracted using the Canny edge detector.
2. Cell division: The edge detected binary image is divided into cells at several pyramid levels. For example, in the first pyramid level, the edge image is divided into 2×2 cells and in the second pyramid level, each cell furthermore is divided into 2×2 sub-cells. The cell division is repeated until desirable resolution levels of pyramid.
3. HOG calculation: The Histogram of Oriented Gradients (HOG) of each cell is calculated at each pyramid resolution level. The HOG of each cell in the same pyramid level is concatenated to form a vector.
4. PHOG extraction: The final PHOG is a concatenation of all HOG vectors generates in all pyramid levels.

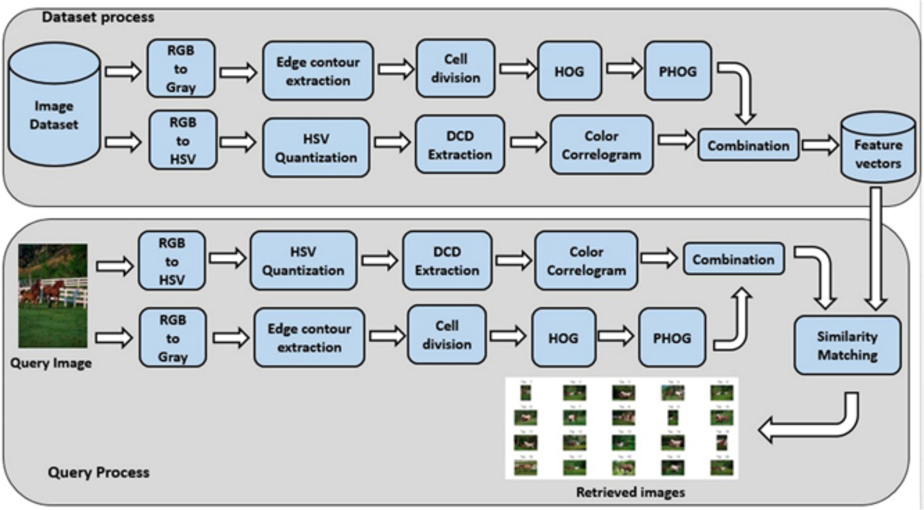

Fig. 1. Block diagram of the proposed scheme

4 Proposed Visual Descriptor

In this paper we propose a new visual descriptor which is a combination of the global and local color information based on the DCCD and the shape feature based on the PHOG. The block diagram of the proposed method is shown in fig. 1.

4.1 Global Color Information Extraction

In this paper we selected the HSV (Hue, Saturation, Value) color space because it is similar in the manner that human distinguish the colors [7]. In order to reduce computational cost, the quantization in the HSV color space is performed using a non-linear quantization, which is given by

$$H = \begin{cases} 0 \; if \; h & \in [316,20) \\ 1 \; if \; h & \in [20,40) \\ 2 \; if \; h & \in [40,75) \\ 3 \; if \; h & \in [75,155) \\ 4 \; if \; h & \in [155,190) \\ 5 \; if \; h & \in [190,270) \\ 6 \; if \; h & \in [270,295) \\ 7 \; if \; h & \in [295,316) \end{cases} \tag{13}$$

$$S = \begin{cases} 0\ if\ s\ \in [0,0.2] \\ 1\ if\ s\ \in (0.2,0.7] \\ 2\ if\ s\ \in (0.7,] \end{cases}, V = \begin{cases} 0\ if\ v\ \in [0,0.2] \\ 1\ if\ v\ \in (0.2,0.7] \\ 2\ if\ v\ \in (0.7,] \end{cases} \tag{14}$$

The three quantized components are then combined into one matrix:

$$C = 9 \times H + 3 \times S + V \tag{15}$$

As a result, it is obtained a matrix with only 8x3x3=72 colors. From the matrix calculated in (15) the dominant colors are extracted using the DCD operation [9,10].

4.2 Local Color Information and Shape Extraction

The color correlogram is computed using the structuring element of 3x3 pixels which scans the image, calculating the correlation between the central pixel and its neighborhood at a distance k=1. Thus, the proposed scheme gives spatial information. The color correlogram can be computed as:

$$\gamma_{c_i c_i}(I) \triangleq Pr_{p_1 \in I_{c_i}, p_2 \in I_{c_i}}[p_2 \in I_{c_i} \| p_1 - p_2 | = 1] \tag{16}$$

where c_i is *i-th* color and p_1 and p_2 are any two pixels in the input image *I*.

In the proposed scheme, we used auto-color correlogram, in which two pixels have the same the color. As a result we obtain a color-based descriptor called the Dominant Color Correlogram Descriptor (DCCD) and is defined as:

$$DCCD = \{C_i, CC_i\} \tag{17}$$

where CC_i is the color auto-correlogram of the *i-th* dominant color C_i.

The shape feature extraction is done using the PHOG descriptor mentioned in 3.3, in which we used 3 pyramid levels and 8 bins for computing the HOG in each level.

4.3 Combination of Features

In the literature there are many manners to combine more than one visual feature, in this paper we used the linear combination method used in [16], which is as follows.

$$nC = \frac{C}{\max(C)}, \ nS = \frac{S}{\max(S)} \tag{18}$$

$$Q = \omega nS + (1 - \omega)nC \tag{19}$$

where C and S are color and shape distances between the query and dataset image, respectively. Firstly the two distances are normalized by (18) and combined lineally by (19) to obtain the similarity Q, where ω is the weight of a particular visual feature. If this similarity is smaller, two images are considered as more similar.

Table 2. Performance of color-based descriptors using dataset 1

Method	ANMRR	ARR		ARP		
		α=2	α=1	α=1	α=0.5	α=0.25
CC [15]	0.3126	0.7577	0.5923	0.5923	0.7846	0.8923
HI [13]	0.2507	0.8115	0.6269	0.6269	0.7923	0.9077
DCD [9]	0.2576	0.8154	0.6154	0.6154	0.7846	0.8615
LBA [2,10]	0.3579	0.6808	0.5642	0.5642	0.7154	0.8154
CLD [3,16, 21]	0.3358	0.7385	0.5731	0.5731	0.7154	0.8000
CSD [17]	0.3145	0.7538	0.5846	0.5846	0.7538	0.8769
DCCD [7]	**0.2266**	**0.8231**	**0.6808**	**0.6808**	**0.8538**	**0.9231**

Table 3. Performance of color-based descriptors using dataset 2

Method	ANMRR	ARR		ARP		
		α=2	α=1	α=1	α=0.5	α=0.25
CC [15]	0.3228	0.7200	0.5870	0.5870	**0.7620**	**0.8880**
HI [13]	0.3174	0.7610	0.5760	0.5760	0.7380	0.8640
DCD [9]	0.3384	0.7420	0.5590	0.5590	0.6920	0.8480
LBA [2,10]	0.3478	0.7320	0.5500	0.5500	0.7040	0.8000
CLD [3,16, 21]	0.3194	**0.7620**	0.5740	0.5740	0.7280	0.8360
CSD [17]	0.4431	0.6190	0.4630	0.4630	0.6200	0.7680
DCCD [7]	**0.3086**	0.7590	**0.5960**	**0.5960**	0.7560	0.8840

5 Experimental Results

As we mentioned before, there are many visual descriptors, such as color-based, texture-based and shape-based, also, there are many methods which combine these descriptors in order to improve the image retrieval performance. We analyzed and evaluate some algorithms using three different datasets; the dataset 1 is composed by 500 images, divided into 25 categories with 20 ground truth images per category. The dataset 2 is composed by 1000 images, divided into 20 categories with 50 ground truth images per category, and the dataset 3 is called Corel Dataset 1k [19, 20] which is composed by 1000 images, divided into 10 categories with 100 ground truth images per category. The first two datasets (dataset 1 and dataset 2) are randomly selected by authors, considering different percentage of query images respect to the dataset size and the third one is commonly used in CBIR evaluation. The color-based descriptors were evaluated using dataset 1 and dataset 2 and the results are shown in table 2 and table 3.

In the table 2 and 3, we can observe that, the results obtained by the DCCD, proposed in [7], are better than the others. The RP-curves, shown in figures 2 and 3, can describe the performance of the descriptors in the image retrieval task:

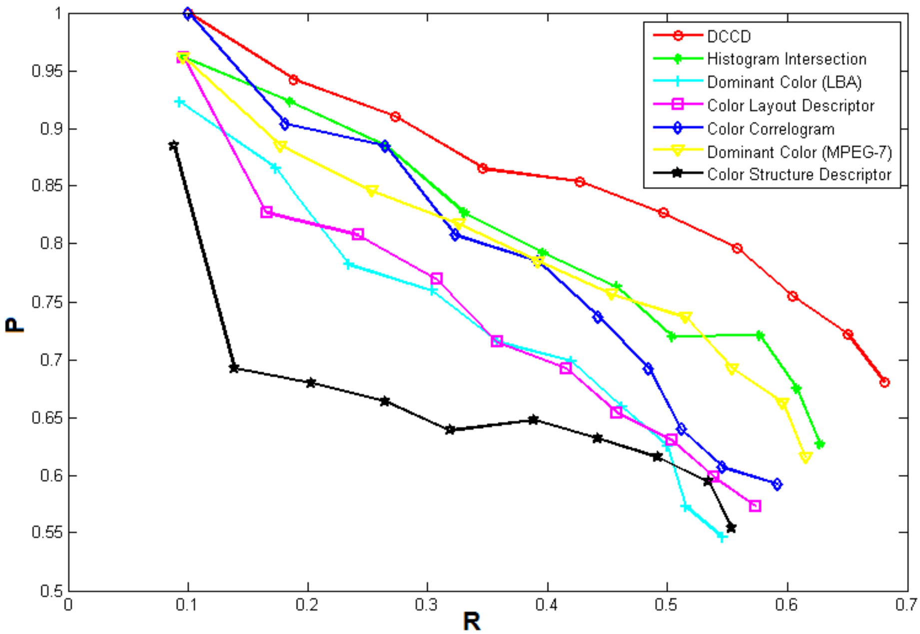

Fig. 2. RP-curve using dataset 1

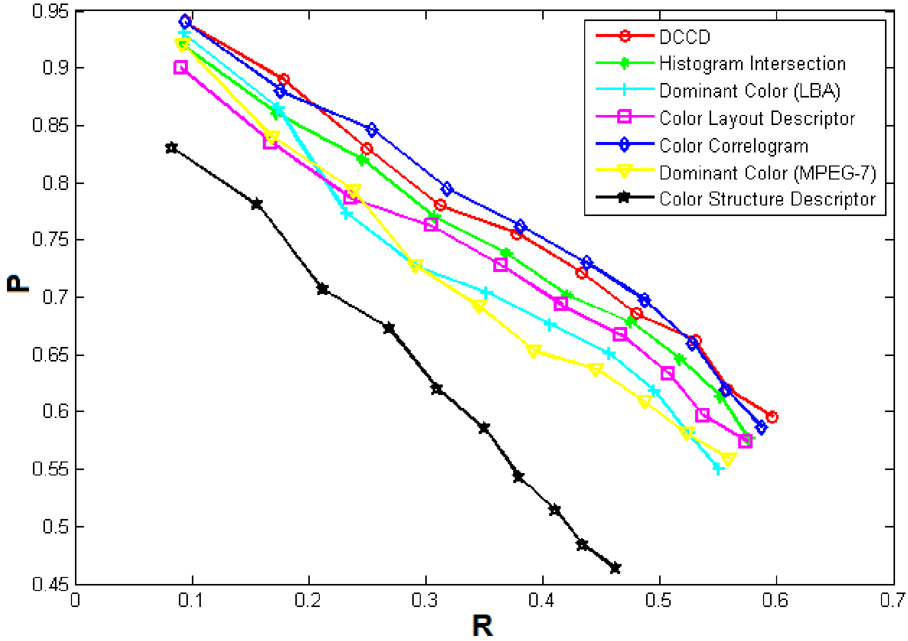

Fig. 3. RP-curve using dataset 2

The shape-based shape-based descriptors, PZM, PCET, PCT and PHOG, are evaluated. The evaluation was done using dataset 1, and the results are shown in the following table:

Table 4. Performance of shape-based descriptors using dataset 1

Method	ANMRR	ARR		ARP		
		α=2	α=1	α=1	α=0.5	α=0.25
PZM [2]	0.7177	0.3808	0.2000	0.2000	0.2308	0.3692
PCET [18]	0.7212	0.3642	0.2231	0.2231	0.2615	0.3692
PCT [18]	0.7066	0.3554	0.2423	0.2423	0.2538	0.3692
PHOG [8]	**0.5825**	**0.4462**	**0.3538**	**0.3538**	**0.4692**	**0.6308**

From tables 2, 3 and 4 and figures 2 and 3, we can observe that the DCCD and the PHOG perform better than the other descriptors, so we decided to combine the DCCD color-based descriptor and the PHOG shape-based descriptor in this paper. We used the combination of two feature vectors given by (19) with three different weight **ω** [16] and a simple concatenation of two feature vectors.

Table 5. Combination of two feature vectors with different weght ω using dataset 1

Method	**ANMRR**	**ARR**		**ARP**		
		α=2	α=1	α=1	α=0.5	α=0.25
Q ω=0.8	0.3933	0.6385	0.5077	0.5077	0.7000	0.8000
Q ω=0.5	0.2364	0.8077	0.6769	0.6769	0.8154	0.9231
Q ω=0.3	**0.2150**	**0.8309**	**0.6846**	**0.6846**	**0.8615**	**0.9538**
Concatenation	0.2227	0.8231	0.6808	0.6808	0.8538	0.9385

The table 5 shows that using weight $\omega = 0.3$, the performance of image retrieval task is better, which means the contribution of color feature is more important than the shape feature in the CBIR. As we mentioned before, the proposed scheme combines color and shape features based on the DCCD and the PHOG using the combination method given by (19) to obtain the similarity Q. We compared the proposed scheme with the method proposed in [2], which combines color, texture and shape features, using the dataset 1 and dataset 2. In the evaluation, we employed three metrics commonly used in the CBIR, such as Average Normalized Retrieval Rank (ANMRR), Average Retrieval Rate (ARR) and Average Retrieval Precision (ARP) [7, 14]. The number of queries must be at least 1% of the dataset size [12]. For dataset 1, we used 13 queries for evaluation equivalent to the 2.6% and for dataset 2 we used 20 queries equivalent to the 2% of the dataset size. The comparison results between the proposed scheme and [2] are shown in table 6.

Table 6. Comparison with the previous method [2]

dataset 1						
Method	**ANMRR**	**ARR**		**ARP**		
		α=2	α=1	α=1	α=0.5	α=0.25
[2]	0.3425	0.7308	0.5346	0.5346	0.7308	0.8615
Proposed	**0.2150**	**0.8309**	**0.6846**	**0.6846**	**0.8615**	**0.9538**
dataset 2						
[2]	0.3672	0.6750	0.5420	0.5420	0.7020	0.8400
Proposed	**0.2698**	**0.7800**	**0.6550**	**0.6550**	**0.8120**	**0.9320**

Using Corel Dataset 1k, we compare the proposed algorithm with the algorithms proposed in [3-6]. The evaluation method used here is the same one used in [3], in which 80 queries are used corresponding to the 8% of the dataset size. The query images are the same used in [3] and the number of retrieved images to compute the

Average Precision (AP) is the same as well. The results are shown in table 7. From the table, we can conclude the proposed method globally outperforms four previously proposed methods [3-6].

Table 7. Comparison with several methods [3-6]

Class name	[3]	[4]	[5]	[6]	Proposed
Tribe	54 %	44.1%	32.3%	41%	**61%**
Beach	38%	30.6%	**61.2%**	32%	46.6%
Buildings	40%	38.2%	39.2%	37%	**41.09%**
Buses	64%	**67.6%**	39.5%	66%	62.93%
Dinosaurs	96%	97.2%	**99.6%**	43%	93.52%
Elephants	**62%**	33.8%	55.7%	39%	41.78%
Roses	68%	88.8%	**89.3%**	87%	77.51%
Horses	75%	63.2%	65.2%	35%	**77.35%**
Mountains	45%	31.3%	**56.8%**	34%	42.79%
Food	53%	34.9%	44.1%	31%	**68.59%**
Average	59.5%	52.97%	58.29%	44.5%	**61.26%**

6 Conclusions

In this paper we proposed a scheme which combines color and shape features for the content-based image retrieval (CBIR) task. The proposed scheme extracts both global and local color information using the DCCD and the shape feature based on boundary information of objects using the PHOG. Two features are combined by weighted linear combination. We set a greater color-weight because the color feature provides the most distinguishable information compared with the shape features. The comparison of the proposed algorithm with four previously proposed algorithms, which combine more than one visual descriptor, shows the better performance of the proposed algorithm.

Acknowledgements. We thanks to the National Council of Science and Technology of Mexico (CONACyT), to ANR of France and the National Polytechnic Institute for the financial support during the realization of this research.

References

1. Penatti, O., Valle, E., Torres, S.: Comparative study of global color and texture descriptors for web image retrieval. J. Vis. Comm. Image Representation 23, 359–380 (2012)
2. Wang, X.Y., Yu, Y.J., Yang, H.Y.: An effective image retrieval scheme using color, texture and shape features. Computer Standards & Interfaces 33, 59–68 (2011)
3. Jalab, H.A.: Image retrieval system based on color layout descriptor and global filters. In: IEEE Conf. on Open System, pp. 32–36 (2011)
4. Pujari, J., Hiremath, P.: Content-based image retrieval based on color, texture and shape features. Signal and Image Processing, 239–242 (2010)
5. Hafiane, A., Zavidovique, B.: Local relational string and mutual matching doe image retrieval. Information Processing &Manegement 44, 1201–1212 (2008)
6. Kavitha, C., Prabhakara, B., Govardhan, A.: Image retrieval based on color and texture features of the image subblocks. Int. J. Comput. Appl. 15, 33–37 (2011)
7. Fierro, A., Perez, K., Nakano, M., Benois, J.: Dominant color correlogram descriptor for content-based image retrieval. In: 3rd Int. Conf. on Image, Vision and Computing (2014)
8. Yang, B., Guo, L., Jin, L., Huang, Q.: A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition. In: 16th IEEE Int. Conf. on Image Processing, pp. 3305–3308 (2009)
9. Shao, H., Wu, Y., Cui, W., Zhang, J.: Image retrieval based on MPEG-7 dominant color descriptor. In: 9th Int. Conf. for Young Scientist, pp. 753–757 (2008)
10. Yang, N., Chang, W., Kuo, C., Li, T.: A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval. J. of Vis. Commum. & Image Representation 19, 92–105 (2008)
11. Johnson, G., Song, X., Montag, E., Fairchild, M.: Derivation of a color space for image color difference measurement. Color Research and Appl. 5, 387–400 (2010)
12. Talib, A., Mahmuddin, M., Husni, H., Loay, E.G.: A weighted dominant color descriptor for content-based image retrieval. J. of Vis. Commun. & Image Representation 24, 345–360 (2013)
13. Swain, M., Ballard, D.: Color indexing. Int. J. of Computer Vision 7, 11–32 (1991)
14. Fierro, A., Nakano, M., Perez, H., Cedillo, M., Garcia, F.: An efficient color descriptor based on global and local color features for image retrieval. In: Int. Conf. on Elect. Eng. Comput. Science and Automatic Control, pp. 233–238 (2013)
15. Huang, J., Kumar, S., Mitra, M., Zhu, W., Zabih, R.: Image indexing using color correlogram. In: Conf. on Computer Vision Pattern Recognition, pp. 762–768 (1997)
16. Bleschke, M., Madonski, R., Rudnicki, R.: Image retrieval system based on combined MPEG-7 texture and color descriptors. In: Int. Conf. Mixed Design of Integrated Circuit and Systems, pp. 635–639 (2009)
17. Wong, K., Po, L., Cheung, K.: Dominant color structure descriptor for image retrieval. IEEE Trans. on PAMI 32, 1259–1270 (2010)
18. Yap, P., Jiang, X., Kot, C.: Two-dimensional polar harmonic transform for invariant image representation. IEEE Trans. on PAMI 32, 1259–1270 (2010)
19. Li, J., Wang, J.: Automatic linguistic indexing of pictures by a statistical modeling approach. IEEE Trans on PAMI 25, 1075–1088 (2003)
20. Wang, J., Li, J., Wiederhold, G.: SIMPLIcity semantic-sensitive integrated matching for picture libraries. IEEE Trans. on PAMI 23, 947–964 (2001)
21. Kasutani, E., Yamada, A.: The MPEG-7 color layout descriptor: A compact image feature description for high-speed image/video segment retrieval. In: Int. Conf. on Image Processing, pp. 674–667 (2001)