Content-Based Image Retrieval Using Multiresolution Color and Texture Features

Young Deok Chun, Nam Chul Kim, *Member, IEEE*, and Ick Hoon Jang, *Member, IEEE*

***Abstract*—In this paper, we propose a content-based image retrieval method based on an efficient combination of multiresolution color and texture features. As its color features, color autocorrelograms of the hue and saturation component images in HSV color space are used. As its texture features, BDIP and BVLC moments of the value component image are adopted. The color and texture features are extracted in multiresolution wavelet domain and combined. The dimension of the combined feature vector is determined at a point where the retrieval accuracy becomes saturated. Experimental results show that the proposed method yields higher retrieval accuracy than some conventional methods even though its feature vector dimension is not higher than those of the latter for six test DBs. Especially, it demonstrates more excellent retrieval accuracy for queries and target images of various resolutions. In addition, the proposed method almost always shows performance gain in precision versus recall and in ANMRR over the other methods.**

***Index Terms*—Content-based image retrieval, multiresolution representation, color and texture features.**

I. Introduction

SINCE the 1990s, content-based image retrieval (CBIR) has become an active and fast-advancing research area in image retrieval [1]–[6]. In a typical CBIR, features related to visual content such as shape, color, and texture are first extracted from a query image, the similarity between the set of features of the query image and that of each target image in a DB is then computed, and target images are next retrieved which are most similar to the query image. Extraction of good features which compactly represent a query image is one of the important tasks in CBIR. Shape is a visual feature that describes the contours of objects in an image, which are usually extracted from segmenting the image into meaningful regions or objects. However, since it is difficult to achieve such image segmentation for natural images, the use of shape features in image retrieval has been limited to special applications where the extraction of object contours is readily available such as in trademark images [1].

Color is one of the most widely used visual features and is invariant to image size and orientation [1], [2]. As conventional color features used in CBIR, there are color histogram [5], color correlogram [6], color structure descriptor (CSD), and scalable color descriptor (SCD). The latter two are MPEG-7 color descriptors [7]. Color histogram is the most commonly used color representation, but it does not include any spatial information. On the other hand, color correlogram describes the probability of finding color pairs at a fixed pixel distance and provides spatial information. Therefore color correlogram yields better retrieval accuracy in comparison to color histogram [6]. Color autocorrelogram is a subset of color correlogram, which captures the spatial correlation between identical colors only. Since it provides significant computational benefits over color correlogram [1], it is more suitable for image retrieval.

Texture is also a visual feature that refers to innate surface properties of an object and their relationship to the surrounding environment [8]. In conventional texture features used for CBIR, there are statistic texture features using gray-level co-occurrence matrix (GLCM) [8], edge histogram descriptor (EHD), which is one of the MPEG-7 texture descriptors [7], and wavelet moments [9]. Recently, BDIP (block difference of inverse probabilities) and BVLC (block variation of local correlation coefficients) features have been proposed which effectively measure local brightness variations and local texture smoothness, respectively [10]. These features are shown to yield better retrieval accuracy over the compared conventional features. They are extracted from 2×2 blocks into which a query image is partitioned to measure local image characteristics in great detail.

A feature extracted from an image is generally represented as a vector of finite dimension. The feature vector dimension is one of the most important factors that determine the amount of storage space for the vector, the retrieval accuracy, and the retrieval time (or computational complexity) [10]. As the dimension increases, not only the retrieval accuracy tends to improve generally but also the amount of storage space and the computational complexity increase. Thus, it is very important to select a proper feature vector dimension which results in good retrieval accuracy while not requiring a great amount of storage space and computational complexity in CBIR.

When an object is filmed in an image acquisition device, it is represented with different resolutions according to its distance from the device. As a result, there may be images of the same kind but of different resolutions in an image DB. Unlike color histograms invariant to image resolution [2], features using spatial relation between pixels in a fixed distance such as BDIP

Manuscript received May 09, 2007; revised March 16, 2008. Current version published October 24, 2008. This work was supported by the University Research Program of the Ministry of Information & Communication in Korea under Grant 04-FU-007 and by the Brain Korea 21 Project (BK21). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kiyoharu Aizawa.

Y. D. Chun is with the GPG1, S/W Laboratory, Mobile Communication Division, Samsung Electronics Co. Ltd., Gumi 730-350, Korea (e-mail: yd.chun@samsung.com).

N. C. Kim is with the Laboratory for Visual Communications, Department of Electronic Engineering, Kyungpook National University, Daegu 702-701, Korea (e-mail: nckim@ee.knu.ac.kr).

I. H. Jang is with the Department of Electronic Engineering, Kyungwoon University, Gumi 730-850, Korea (e-mail: ihjang@ikw.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2008.2001357

and BVLC may deteriorate retrieval performance for image DBs containing target images of various resolutions. One of the solutions for such a problem is to extract the features that give the spatial relation in a fixed distance as a form of multiresolution ones in the wavelet transform domain, as in [11], [12], and [13], which are used for image detection, texture analysis and classification, and image retrieval by using multiresolution analyses, respectively.

Most of the early studies on CBIR have used only a single feature among various color and texture features. However, it is hard to attain satisfactory retrieval results by using a single feature because, in general, an image contains various visual characteristics. Recently, active researches in image retrieval using a combination of color and texture features have been performed [13]–[15]. In [13], two-dimensional or one-dimensional histograms of the *CIELab* chromaticity coordinates are chosen as color features, and variances extracted by discrete wavelet frames analysis are chosen as texture features. In [14], SDHSV(r, r1) color histogram is used as a color feature and Haar or Daubechies' wavelet moment is used as a texture feature. The mean and covariance of RGB values and the energies of DCT coefficients are used as color and texture features in [15], respectively. In these methods, their feature vector dimension is not considered as an important factor in combining multiple features. It is shown that such a combination of features without increase of feature vector dimension does not always guarantee better retrieval accuracy [16]. Accordingly, for an advanced CBIR, it is necessary to choose efficient features that are complementary to each other so as to yield an improved retrieval performance and to combine chosen features effectively without increase of feature vector dimension.

In this paper, we propose a CBIR method which uses the combination of color autocorrelograms of hue and saturation component images and BDIP and BVLC moments of value component image in the wavelet transform domain. The dimension of the combined feature vector is determined at a saturation point above which the retrieval accuracy nearly increases as the dimension increases. The next section describes conventional features which will be adopted in the proposed method. Section III explains the proposed retrieval method, Section IV discusses experimental results, and Section V finally provides the conclusion.

II. Conventional Featrues

In this section, we explain the conventional features which are used in the proposed retrieval method: color autocorrelogram as a color feature and BDIP and BVLC as texture features.

A. Color Autocorrelogram

The color correlogram [6] was proposed to characterize not only the color distributions of pixels, but also the spatial correlation of pairs of colors. It describes the probability of finding a pixel of another special color at a fixed pixel distance for a given pixel of a color. If we consider all the possible combinations of color pairs, the size of the color correlogram will be very large. Therefore a simplified version of the feature called color autocorrelogram is often used instead [1]. The color autocorrelogram captures the spatial correlation between identical colors only.

The color autocorrelogram $\alpha^k(l)$ is given as the probability of finding a pixel p' of the identical color at a distance k from a given pixel p of the lth color, that is

$$\alpha^k(l) = \Pr[p' \in I \,\big|\, |p-p'| = k \text{ and } p' \in I(l) \text{ for } p \in I(l) \subset I], l \in \{0,1,\ldots,L-1\} \quad (1)$$

where $\Pr[\cdot]$ denotes the probability satisfying a given condition, I the set of all pixels in an image, $|p-p'|$ the distance between the two pixels, and $I(l)$ the set of pixels of the lth color which is requantized with L levels. Selecting the L_∞-norm to measure the distance between pixels, the $|p-p'|$ is expressed as

$$|p-p'| = \max\{|x-x'|, |y-y'|\} \quad (2)$$

where $p=(x,y)$ and $p'=(x',y')$.

As k in (1) varies, spatial correlation between identical colors in an image can be obtained in various resolutions. As a color space for autocorrelogram extraction, the HSV color space is known to provide better correspondence with human perception of color similarities than the RGB color space [17].

B. BDIP

In images, edges represent the regions which involve abrupt change of intensity, and valleys represent the regions which contain local intensity minima. They are very important features in human vision and, especially, valleys are fundamental in the visual perception of object shape [18], [19]. BDIP [10] is a texture feature that effectively extracts edges and valleys, which is defined as

$$\beta^k(l) = \frac{\frac{1}{|B_l^k|}\sum_{(x,y)\in B_l^k}\left(\max_{(x,y)\in B_l^k} I(x,y) - I(x,y)\right)}{\max_{(x,y)\in B_l^k} I(x,y)} \quad (3)$$

where $I(x,y)$ denotes the intensity of a pixel (x,y) in the block B_l^k of size $(k+1)\times(k+1)$, l the position index of the block in an image, and k the maximum distance of pairs of pixels in the block, and $|B_l^k|$ the block size, which means $|B_l^k| = (k+1)^2$. The numerator and denominator of (3) express the representative (maximum) of intensity variation in a block and the representative value in a block, respectively.

To understand the characteristics of BDIP more deeply, let us consider two operators which are defined as follows:

$$\text{operator 1} : |\overline{\nabla I}(x,y)| \quad (4)$$

$$\text{operator 2} : \frac{|\overline{\nabla I}(x,y)|}{\bar{I}(x,y)} \quad (5)$$

where $|\overline{\nabla I}(x,y)|$ denotes the magnitude of representative gradient in an $M\times N$ window whose center pixel is (x,y) and thus the first operator performs gradient operation. If we choose the

Fig. 1. Original image and the result images of Roberts gradient, Roberts sketch, and BDIP operators. (a) Original image, (b) Roberts gradient operator, (c) Roberts sketch operator, and (d) BDIP operator.

representative gradient as the averaged one within the window, it may correspond to Sobel or Roberts gradient operator [20]. The quantity $\bar{I}(x,y)$ in (5) means the representative pixel value in the window. Thus the second operator gives the result of gradient operator normalized by the representative pixel value, which yields a sketch-like image so as to be called a sketch operator. It becomes the BDIP operator if the numerator is selected as the magnitude of averaged nonlinear gradient, which is defined as the averaged difference between the maximum pixel value and each pixel value in the window, and the denominator as the maximum pixel value.

Fig. 1 shows an original image and the results of Roberts gradient operator, the sketch operator in which the result of Roberts gradient operator is normalized by the average pixel value in the window, and the BDIP operator. In Fig. 1(b)–(d), the negative images of the results of the operators are displayed. The distance k was chosen as $k = 1$ and the window or block size 2×2 because in our observation the smallest distance gave the most detailed local properties and the best performance for the image in Fig. 1(a). In Fig. 1(b), we can see that the boundaries of objects are extracted well. In Fig. 1(c), we can also see that the intensity variation in dark regions as well as object boundaries is extracted well. On the contrary, in Fig. 1(d), we can see that object boundaries are extracted well like in Fig. 1(c), but the intensity variation in dark regions is more emphasized.

Table I shows the average precision performance [3] in case that the texture features by the three operators are adopted in CBIR. The retrieval method used here is the same as in [10] and Corel DB and VisTex DB [10] are used as test DBs. In Table I, one can see that Roberts sketch operator yields the average performance improvement of 23.4% in both of the two DBs over Roberts gradient operator and the BDIP operator yields the average performance improvement of 1.4% over Roberts sketch operator. From these results, we see that the BDIP reflects the human visual system very well to be more sensitive to the intensity variation in dark regions than in bright regions, which is known as Weber's law. Also, as k in (3) varies, BDIP can show various properties of texture in an image.

TABLE I
AVERAGE PRECISION PERFORMANCE OF ROBERTS GRADIENT, ROBERTS SKETCH, AND BDIP OPERATORS

DB \ Operator	Roberts Gradient	Roberts Sketch	BDIP
Corel	59.8%	84.0%	85.2%
VisTex	61.9%	84.5%	86.1%

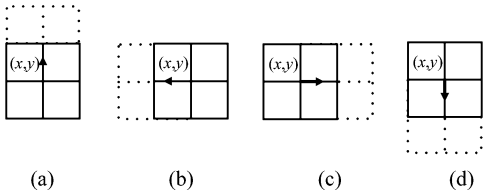


Fig. 2. Pixel configurations in 2×2 blocks and their corresponding blocks shifted in each of four orientations, which are required to compute (a) $(-k, 0)$, (b) $(0, -k)$, (c) $(0, k)$, and (d) $(k, 0)$.

C. BVLC

BVLC [10] represents the variation of block-based local correlation coefficients according to four orientations. It is known to measure texture smoothness well. Each local correlation coefficient is defined as local covariance normalized by local variance, as shown in (6), at the bottom of the page, where μ_l and σ_l the local mean and standard deviation of the block B_l^k, respectively. $\Delta(k) = (\Delta_x(k), \Delta_y(k))$ stands for shift in one of four orientations $\{(-k,0),(0,-k),(0,k),(k,0)\}$. Fig. 2 shows pixel configurations in 2×2 block $(k = 1)$ and their corresponding blocks shifted in each of four orientations. So $\mu_{l+\Delta(k)}$ and $\sigma_{l+\Delta(k)}$ represent the local mean and standard deviation of the lth block shifted by $\Delta(k)$, respectively. Then, the value of BVLC is determined as the difference between the maximum and minimum values of block-based local correlation coefficients according to four orientations. That is

$$\gamma^k(l) = \max_{\Delta(k)\in O_4}[\rho^k(l,\Delta(k))] - \min_{\Delta(k)\in O_4}[\rho^k(l,\Delta(k))],$$
$$O_4 = \{(-k,0),(0,-k),(0,k),(k,0)\}. \tag{7}$$

From (7), one can see that the higher the degree of roughness there is in a block, the larger the value of BVLC. It can also reveal various properties of texture in an image as k in (6) varies.

III. PROPOSED IMAGE RETRIEVAL METHOD

A. Overall Structure of the Proposed Method

Fig. 3 shows the block diagram of the proposed retrieval method. When an RGB query image whose components are I_R, I_G, and I_B enters the retrieval system, it is first transformed

$$\rho^k(l,\Delta(k)) = \frac{\frac{1}{|B_l^k|}\sum_{(x,y)\in B_l^k} I(x,y)I(x+\Delta_x(k),y+\Delta_y(k)) - \mu_l\mu_{l+\Delta(k)}}{\sigma_l\sigma_{l+\Delta(k)}} \tag{6}$$

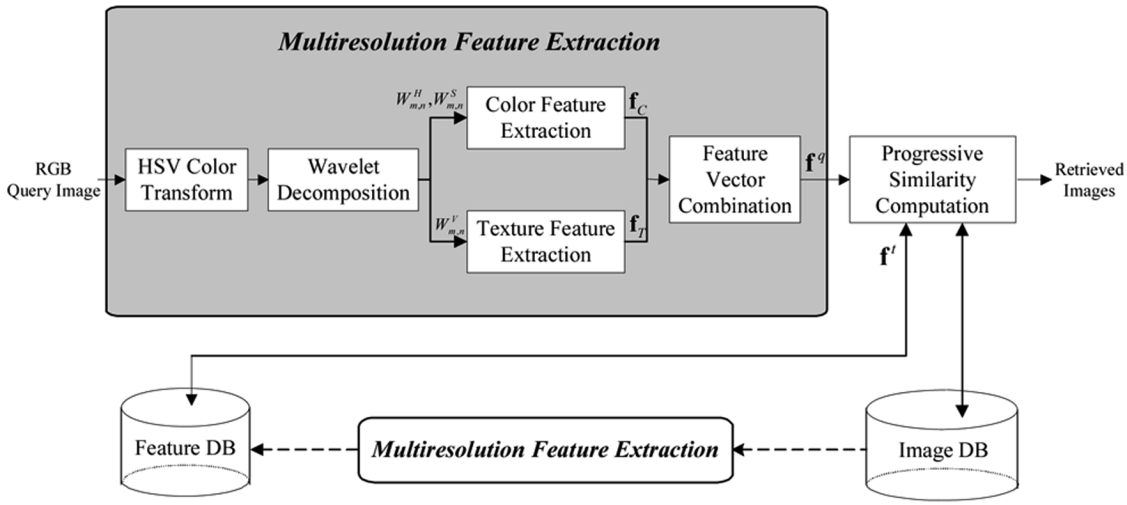


Fig. 3. Block diagram of the proposed method.

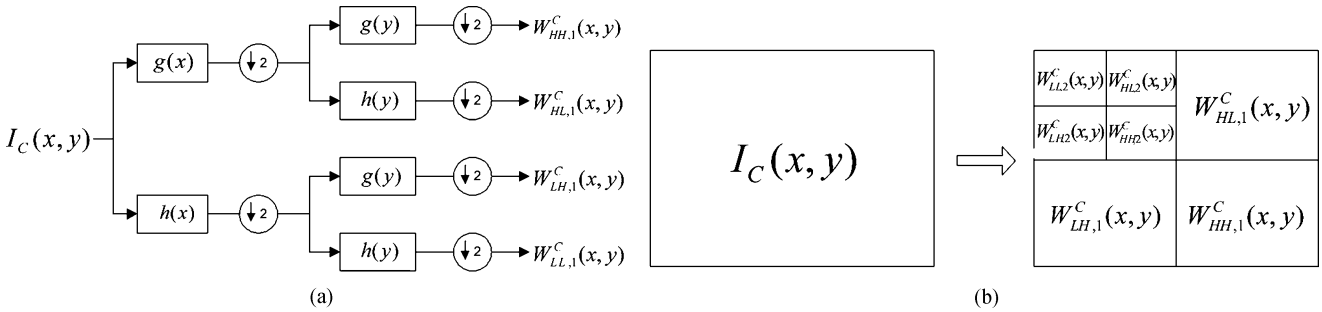


Fig. 4. Example of wavelet decomposition. (a) Block diagram for a 4-band wavelet decomposition and (b) configuration of 2-level wavelet decomposed images.

into an HSV color image whose components are I_H, I_S, and I_V. Then each component image $I_C, C \in \{H, S, V\}$ is wavelet decomposed into a wavelet image $W^C_{m,n}, C \in \{H, S, V\}$ as shown in Fig. 4, respectively, where $m \in \{LL, HL, LH, HH\}$ denotes the subband orientation and $n \in \{1, \ldots, Z\}$ the wavelet decomposition level. Since it seems that the V component image of wider bandwidth contains most of texture information and the H and S component images are closely related to chrominance information, a color feature vector $\mathbf{f}_C$ of dimension N_C is then formed with color autocorrelograms extracted from the $W^H_{m,n}$ and $W^S_{m,n}$ and a texture feature vector $\mathbf{f}_T$ of dimension N_T with BDIP and BVLC moments extracted from the $W^V_{m,n}$. The reason why we adopt the color autocorrelogram and the BDIP and BVLC moments is because it is known that the color autocorrelogram shows relatively excellent retrieval accuracy among conventional color features [21] and the BDIP and BVLC moments among conventional texture features [10]. Our previous work in [21] is an early version of this paper, where the color autocorrelogram and the BDIP and BVLC moments are redundantly extracted from V or Y luminance component and there is no consideration on efficient selection of feature vector dimension. The proposed retrieval system next combines the color feature vector $\mathbf{f}_C$ and the texture feature vector $\mathbf{f}_T$ to generate the query feature vector $\mathbf{f}^q$. It calculates the similarity between the query feature vector $\mathbf{f}^q$ and each target feature vector $\mathbf{f}^t$. According to the similarity ranks, it finally retrieves a given number of target images from the image DB.

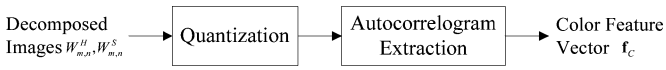


Fig. 5. Procedure of color feature vector extraction.

B. Color Feature Extraction

The procedure of extracting the color feature from the $W^H_{m,n}$ and $W^S_{m,n}$ is shown in Fig. 5. The $W^H_{m,n}$ and $W^S_{m,n}$ are first quantized into $Q^H_{m,n}$ and $Q^S_{m,n}$, respectively. Each LL band for $n \in \{1, \ldots, Z\}$ is uniformly quantized and the other subbands are non-uniformly quantized by the generalized Lloyd algorithm [22], considering the color distributions of subbands among which the ranges of LL band cover all the extent but those of the other subbands are condensed around zero.

The number of quantization levels of each subband in the $W^H_{m,n}$ and $W^S_{m,n}$ is decided as follows. Given the total number of quantization levels L of all subbands, L_i is allocated to the ith subband which results in the relation of

$$L = \sum_{i=0}^{K-1} L_i \tag{8}$$

where K denotes the number of subbands. The set of L_i's that minimizes the distortion generated by allocating the numbers of quantization levels to subbands is then given as [22]

$$\log_2 L_i = \frac{\log_2 L}{K} + \frac{1}{2}\log_2 \frac{\sigma_i^2}{\left[\prod_{j=0}^{K-1}\sigma_j^2\right]^{1/K}} \tag{9}$$

where σ_i^2 denotes the variance of the ith subband image.

After the numbers of quantization levels for all subbands are decided, the color autocorrelogram in (1) is extracted. To reduce computational complexity, we modify the color autocorrelogram as

$$\begin{aligned}&\alpha_{m,n}^k(C,l)\\ &\quad= \Pr\left[p' \in Q_{m,n}^C \middle| |p-p'| = k \text{ and } p' \in N_2(p)\right.\\ &\qquad\left. \text{and } p' \in Q_{m,n}^C(l) \text{ for } p \in Q_{m,n}^C(l) \subset Q_{m,n}^C\right],\\ &\quad C \in \{H,S\}, l \in \{0,1,\ldots,L_{m,n}-1\},\\ &\quad m \in \{LL,HL,LH,HH\}, n \in \{1,\ldots,Z\}\end{aligned} \tag{10}$$

where $N_2(p)$ denotes the set of two causal neighbors of a pixel p as shown in Fig. 6 and $Q_{m,n}^C(l)$ the set of pixels having the lth color in the $Q_{m,n}^C$ which is the result of quantizing $W_{m,n}^C$ with $L_{m,n}$ levels. As a result, $\alpha_{m,n}^k(C,l)$ in (10) represents the probability of finding a pixel p' of the lth color among the two causal neighboring pixels at a distance k from a given pixel p of the lth color in the $Q_{m,n}^C$. The color feature vector $\mathbf{f}_C$ is finally formed with the color autocorrelogram $\alpha_{m,n}^k(C,l)$ as

$$\begin{aligned}\mathbf{f}_C = &\left[\alpha_{m,n}^k(C,l)\right],\\ &C \in \{H,S\}, l \in \{0,1,\ldots,L_{m,n}-1\},\\ &m \in \{LL,HL,LH,HH\}, n \in \{1,\ldots,Z\}.\end{aligned} \tag{11}$$

C. Texture Feature Extraction

Fig. 7 shows the procedure of texture feature extraction from the $W_{m,n}^V(x,y)$. The $W_{m,n}^V(x,y)$ is first divided into nonoverlapping blocks of a given size, where the BDIP and BVLC are computed. To a block in each subband, the BDIP operator in (3) cannot be applied directly since it is defined for the processing in spatial domain where pixel values are nonnegative. In wavelet domain, there are pixels whose values are negative in each subband except the LL band. So the denominator in (3) may yield negative BDIP values in wavelet domain, which leads to invalid measurement of intensity variation. Therefore, for the processing in wavelet domain, we now modify the BDIP as shown in (12) at the bottom of the page, where $W_{LL,n}^V(x,y)$ denotes the LL band in the nth level. Comparing the denominator in (12) with that in (3), the HL, LH, or HH band is substituted with the LL band whose pixel positions correspond to those of HL, LH, or HH band. If the denominator in (12) is less than one, it is made to be one to avoid unstable behavior.

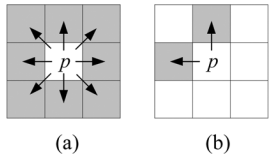


Fig. 6. Neighbors of a pixel p: (a) 8-neighbors and (b) two causal neighbors.

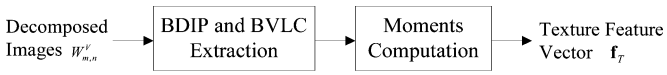


Fig. 7. Procedure of texture feature vector extraction.

To reduce computational complexity, we also modify the BVLC so that the local covariance in (6) is substituted with the mean absolute difference of pixels between two blocks and the local variance is substituted with the mean absolute difference of pixels in a block. The modified local correlation coefficient is given as shown in (13) at the bottom of the page, where $\Delta(k) = (\Delta_x(k), \Delta_y(k))$ stands for shift in one of two orientations $\{(-k,0),(0,-k)\}$ as shown in Fig. 2(a) and (b), respectively. ν_l denotes the mean absolute difference of four end pixels in the block B_l^k as shown in (14) at the bottom of the page. $\nu_{l+\Delta(k)}$ also represents the mean absolute difference in the lth block shifted by $\Delta(k)$. Therefore, the modified BVLC is expressed as the difference between the maximum

$$\beta_{m,n}^k(l) = \frac{\frac{1}{|B_l^k|}\sum_{(x,y)\in B_l^k}\left(\max_{(x,y)\in B_l^k} W_{m,n}^V(x,y) - W_{m,n}^V(x,y)\right)}{\max_{(x,y)\in B_l^k} W_{LL,n}^V(x,y)}, m \in \{LL,HL,LH,HH\}, n \in \{1,\ldots,Z\} \tag{12}$$

$$\rho_{m,n}^k(l,\Delta(k)) = \frac{\frac{1}{|B_l^k|}\sum_{(x,y)\in B_l^k}\left|W_{m,n}^V(x,y) - W_{m,n}^V(x+\Delta_x(k), y+\Delta_y(k))\right|}{(\nu_l + \nu_{l+\Delta(k)})/2} \tag{13}$$

$$\nu_l = \frac{1}{4}\left(\begin{array}{l}\left|W_{m,n}^V(x,y) - W_{m,n}^V(x,y+k)\right| + \left|W_{m,n}^V(x+k,y) - W_{m,n}^V(x+k,y+k)\right|\\ +\left|W_{m,n}^V(x,y) - W_{m,n}^V(x+k,y)\right| + \left|W_{m,n}^V(x,y+k) - W_{m,n}^V(x+k,y+k)\right|\end{array}\right) \tag{14}$$

and minimum values of the local correlation coefficients in (13) according to two orientations. That is

$$\gamma_{m,n}^{k}(l) = \max_{\Delta(k)\in O_2}\left[\rho_{m,n}^{k}(l,\Delta(k))\right] - \min_{\Delta(k)\in O_2}\left[\rho_{m,n}^{k}(l,\Delta(k))\right], \quad O_2 = \{(-k,0),(0,-k)\}. \tag{15}$$

After the BDIP and BVLC are computed, the first and second moments of the BDIP and BVLC for each subband are extracted as components of a texture feature vector, which are written as

$$\mu(\beta_{m,n}^{k}) = <\beta_{m,n}^{k}(l)> \tag{16}$$

$$\sigma^2\left(\beta_{m,n}^{k}\right) = <\left(\beta_{m,n}^{k}(l) - \mu\left(\beta_{m,n}^{k}\right)\right)^2> \tag{17}$$

$$\mu\left(\gamma_{m,n}^{k}\right) = <\gamma_{m,n}^{k}(l)> \tag{18}$$

$$\sigma^2\left(\gamma_{m,n}^{k}\right) = <\left(\gamma_{m,n}^{k}(l) - \mu\left(\gamma_{m,n}^{k}\right)\right)^2> \tag{19}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the mean and standard deviation, respectively, and $<\cdot>$ the average of the quantity. As a result, the texture feature vector is given as

$$\mathbf{f}_T = \left[\left[\mu\left(\beta_{m,n}^{k}\right)\right], \left[\sigma\left(\beta_{m,n}^{k}\right)\right], \left[\mu\left(\gamma_{m,n}^{k}\right)\right], \left[\sigma\left(\gamma_{m,n}^{k}\right)\right]\right], \quad m \in \{LL,HL,LH,HH\}, n \in \{1,\ldots,Z\} \tag{20}$$

where $[\mu(\beta_{m,n}^{k})]$ and $[\sigma(\beta_{m,n}^{k})]$ denote the mean vector and standard deviation vector of BDIP, respectively, and $[\mu(\gamma_{m,n}^{k})]$ and $[\sigma(\gamma_{m,n}^{k})]$ the mean vector and standard deviation vector of BVLC, respectively.

D. Feature Vector Combination and Similarity Measurement

After the color and texture feature vectors are extracted, the retrieval system combines these feature vectors, calculates the similarity between the combined feature vector of the query image and that of each target image in an image DB, and retrieves a given number of the most similar target images. The color feature vector $\mathbf{f}_C$ in (11) and the texture feature vector $\mathbf{f}_T$ in (20) are combined as follows:

$$\begin{aligned}\mathbf{f} &= \left[\frac{\mathbf{f}_C}{N_C\cdot\boldsymbol{\sigma}_C}, \frac{\mathbf{f}_T}{N_T\cdot\boldsymbol{\sigma}_T}\right],\\ \mathbf{f}_C &= \left[\left[\alpha_{m,n}^{k}(H,l)\right], \left[\alpha_{m,n}^{k}(S,l)\right]\right]\\ \mathbf{f}_T &= \left[\left[\mu\left(\beta_{m,n}^{k}\right)\right], \left[\sigma\left(\beta_{m,n}^{k}\right)\right], \left[\mu\left(\gamma_{m,n}^{k}\right)\right], \left[\sigma\left(\gamma_{m,n}^{k}\right)\right]\right]\\ &\quad \text{for } m \in \{LL,HL,LH,HH\} \text{ and } n \in \{1,\ldots,Z\}\end{aligned} \tag{21}$$

where $\mathbf{f}$ denotes the combined feature vector, $\mathbf{f}_C$ and $\mathbf{f}_T$ the color and texture feature vectors of dimension N_C and N_T, and $\boldsymbol{\sigma}_C$ and $\boldsymbol{\sigma}_T$ their standard deviation vectors obtained from the DB, respectively. From (10) and (21), one can see that the N_C is determined as the total number of quantization levels L and the N_T is given as $16Z$. The vector division in (21) means the component-wise division. We can also see in (21) that each of the color and texture feature components is normalized by its dimension and standard deviation, which is for reducing the effect of different feature vector dimensions and component variances in the similarity computation. Using the combined feature vector $\mathbf{f}$ of dimension N, K_1 target images are retrieved from an image DB having K_0 target images. Here the vector dimension N is given as $N = N_C + N_T$.

As a similarity measure, we select the generalized Minkowski-form distance [23] of metric order one. So the distance used for computing the similarity between the query feature vector $\mathbf{f}^q$ and the target feature vector $\mathbf{f}^t$ is given as

$$D(\mathbf{f}^q,\mathbf{f}^t) = \sum_{i=1}^{N}|f^q(i) - f^t(i)| \tag{22}$$

where q and t represent the query and target images and $f^q(i)$ and $f^t(i)$ the ith components of the $\mathbf{f}^q$ and the $\mathbf{f}^t$, respectively. The number of additions for a query image in the similarity measurement of the retrieval is given as

$$K_0(2N-1). \tag{23}$$

The feature vector combination and the similarity measurement can also be progressively implemented to reduce the computational complexity in CBIR with a huge DB [16]. Previous studies related to progressive scheme can be found in [24], [25]. In [24], a multistep query processing is presented, where an index-based filter step produces a set of candidates and a subsequent refinement step performs the expensive exact evaluation of the candidates. In [25], a progressive retrieval is proposed, where multiresolution feature pyramids are formed, a set of candidate images is selected by feature matching at the lowest level, and a progressive refinement is performed as the level increases. Our progressive retrieval is composed of $Z+1$ steps. In the scheme, the color feature vector $\mathbf{f}_C$ and the texture feature vector $\mathbf{f}_T$ are combined as (21) for $m = LL$ and $n = 1$ at the first step, for $m \in \{LL,HL,LH,HH\}$ and $n = 1$ at the second step, and for $m \in \{LL,HL,LH,HH\}$ and $n \in \{1,2\}$ at the third step, ..., and for $m \in \{LL,HL,LH,HH\}$ and $n \in \{1,\ldots,Z\}$ at the $(Z+1)$th step. One can see that the texture vector dimension N_T is given as 4 at the first step and $16(j-1)$ at the jth step, $j = \{2,\ldots,Z+1\}$.

At each step, the combined feature vector $\mathbf{f}_j^q$ of dimension N_j for a query is matched with the combined feature vector $\mathbf{f}_j^t$ of the same dimension for each of K_{j-1} target images and so K_j target images with the best similarity are retrieved. As a result, the total number of additions for a query in the similarity measurement of the progressive retrieval is given as

$$\sum_{j=0}^{Z} K_j(2N_{j+1}-1) \tag{24}$$

where K_j is determined to decrease and the N_j increases as the retrieval step j increases so as to satisfy $K_j(2N_{j+1}-1) > K_{j+1}(2N_{j+2}-1), j = 0,\ldots,Z-1$.

One can see that the computational complexity per pixel of the feature vector extraction is constant in the case of a fixed k and the computational complexity of a retrieval method is given as the sum of that for the feature extraction and that for the similarity measurement. So, for an image of size $W\times H$, the number of additions per pixel of the nonprogressive retrieval is estimated as $\xi + K_0(2N-1)/WH$ and that of the progressive retrieval as $\xi + \sum_{j=0}^{Z} K_j(2N_{j+1}-1)/WH$, where ξ denotes the number of additions per pixel of the feature extraction. When K_0/WH approaches infinity, the ratio between the numbers of additions per pixel of the two schemes approaches $(2N-1)/(2N_1-1)$,

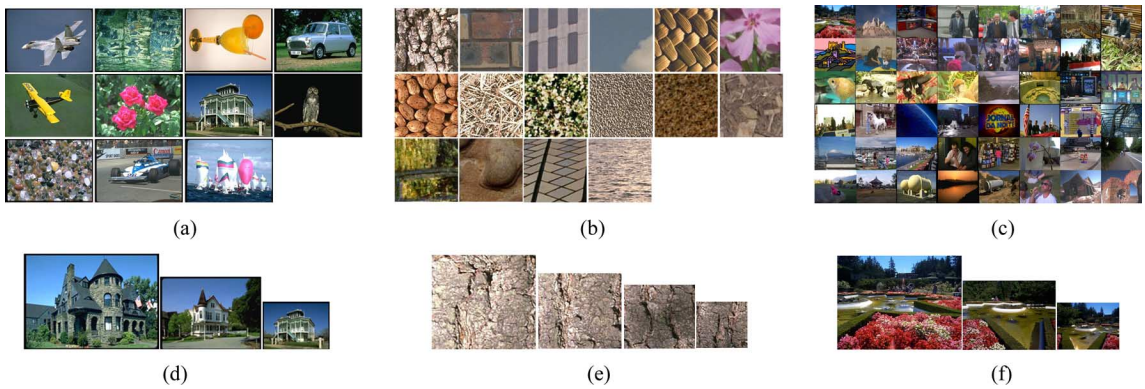


Fig. 8. RGB images sampled from (a) Corel DB, (b) VisTex DB, (c) MEGG-7 CCD, (d) Corel_MR DB, (e) VisTex_MR DB, and (f) MPEG-7 CCD_MR.

since the number of target images are usually determined as $K_0 \gg K_j, j = 1, \ldots, Z$. This means that the progressive retrieval can greatly reduce the computational complexity over the nonpregressive retrieval in the case of $K_0 \gg WH$.

IV. Experimental Results

A. Image Database, Query Images, and Performance Measures

The Corel DB [10], VisTex DB [10], MPEG-7 common color dataset (CCD) [26], Corel_MR DB, VisTex_MR DB, and MPEG-7 CCD_MR were used to evaluate the performance of the proposed retrieval method. We derived the latter three from the former three so as to contain images of various resolutions. The Corel DB is composed of 990 RGB color images with 192×128 pixels and divided into 11 classes, each of 90 images. The VisTex DB is composed of 1200 RGB color images with 128×128 pixels and divided into 75 classes, each of 16 images. The MPEG-7 CCD is composed of 5420 color images that contain 48 common color queries (CCQ) and 332 ground truth sets (GTS), where the number of images in each GTS varies.

In the Corel_MR DB, a third comes directly from a third of all the images for each class in the Corel DB, another third is obtained by decimating another third for each class at the ratio of $(1.5 : 1)^2$, and the other by decimating the other for each class at the ratio of $(2 : 1)^2$. In the VisTex_MR DB, a fourth comes directly from all the images for each class in the VisTex DB, the others are obtained by decimating the others for each class equally at the ratios of $(1.5 : 1)^2$, $(1.75 : 1)^2$, and $(2 : 1)^2$, respectively. In the MPEG-7 CCD_MR, a third comes directly from a third of 48 images for CCQ and a third of 332 images GTS in the MPEG-7 CCD, and the others are obtained by decimating the others of the 48 and 332 original images equally at the ratios of $(1.5 : 1)^2$ and $(2 : 1)^2$, respectively. As a result, the sizes and numbers of classes of the three derived DBs are the same as those of the original DBs, but they contain images of various resolutions. Fig. 8 shows some RGB images sampled from the six DBs.

As measures of retrieval accuracy, we used precision versus recall [27] and ANMRR (average normalized modified retrieval rank) [28], [29]. For a query q, let us denote a set of retrieved images in a DB as $A(q)$ and a set of images relevant to the query q as $B(q)$. Then, the precision $P(q)$ and the recall $R(q)$ are given as [27]

$$P(q) = \frac{|A(q) \cap B(q)|}{|A(q)|} \tag{25}$$

$$R(q) = \frac{|A(q) \cap B(q)|}{|B(q)|} \tag{26}$$

where $|\cdot|$ returns the size of the set. So the precision $P(q)$ represents the ratio of the number of images relevant to the query q among retrieved images to the number of retrieved images. The recall $R(q)$ represents the ratio of the number of images relevant to the query q among retrieved images to the number of images relevant to the query in a DB. So the precision versus recall jointly represents what percent of the images relevant to the query are retrieved and what percent of the retrieved images are relevant. It gives different values according to the number of target images for a DB.

The ANMRR is a measure of retrieval accuracy used in almost all of the MPEG-7 color core experiments, which is described as follows [28], [29]. A rank $\mathrm{Rank}(k, q)$ for each image k in a set of images $B(q)$ relevant to the query q in a DB, that is $k \in B(q)$, is first found. The $\mathrm{Rank}(k, q)$ represents the rank of the similarity between k and q. A relevant rank $\mathrm{RR}(k, q)$ is next obtained using the $\mathrm{Rank}(k, q)$ and a number $K(q) \geq |B(q)|$ as

$$\mathrm{RR}(k, q) = \begin{cases} \mathrm{Rank}(k, q), & \text{if } \mathrm{Rank}(k, q) \leq K(q) \\ 1.25K(q), & \text{otherwise} \end{cases} \tag{27}$$

$$K(q) = \min[4 \cdot |B(q)|, 2 \cdot \max\{|B(q)|, \forall q\}]. \tag{28}$$

The $\mathrm{RR}(k, q)$ is next averaged so that an average rank $\mathrm{AVR}(q)$ for the query q is given as

$$\mathrm{AVR}(q) = \langle \mathrm{RR}(k, q) \rangle. \tag{29}$$

From the $\mathrm{AVR}(q)$, a normalized modified retrieval rank $\mathrm{NMRR}(q)$ for the query q is obtained as

$$\mathrm{NMRR}(q) = \frac{\mathrm{AVR}(q) - 0.5[1 + |B(q)|]}{1.25K(q) - 0.5[1 + |B(q)|]}. \tag{30}$$

The numerator of $\mathrm{NMRR}(q)$ in (30) denotes a modified retrieval rank $\mathrm{MRR}(q)$. The ANMRR of the DB is finally given by averaging the $\mathrm{NMRR}(q)$ over all the q's as

$$\mathrm{ANMRR} = \langle \mathrm{NMRR}(q) \rangle. \tag{31}$$

TABLE II
FEATURE VECTOR DIMENSION AND COLOR SPACE OF THE RETRIEVAL METHODS

Spec. / Method	Color space	Dimension	Remarks
Color histogram	RGB	128	number of quantization levels: 8 in R, 4 in G, 4 in B
SCD	HSV	128	-
CSD	HMMD	128	-
Color autocorrelogram	HSV	128	number of quantization levels: 8 in R, 4 in G, 4 in B
Wavelet moments	RGB	96	number of decomposition levels: 5
EHD	RGB	240	-
BDIP-BVLC	RGB, V	96	-
Color histogram + Wavelet moments	RGB, V	84 (64+20)	number of quantization levels: 4 in R, 4 in G, 4 in B number of decomposition levels: 3
SCD + BDIP-BVLC	HSV, V	96 (64+32)	-
CSD + EHD	HMMD, V	144 (64+80)	-
Proposed	HSV	92 (60+32)	number of decomposition levels: 2 $(N_1, N_2, N_3) = (20, 56, 92)$

The ANMRR gives just one value for a DB and a lower ANMRR value means more accurate retrieval performance.

In the Corel, VisTex, Corel_MR, and VisTex_MR DBs, each image in a DB was chosen as a query and the images in the same class were considered to be relevant. In the MPEG-7 CCD and MPEG-7 CCD_MR, 48 CCQs were chosen as queries and the images in the same GTS were considered to be relevant. The precision versus recall for a DB was obtained by averaging the precision versus recalls over all the queries. In the evaluation of the ANMRR, we ranked the retrieved images using the similarity adopted in each retrieval method.

B. Specifications of Retrieval Methods

The image retrieval methods chosen in our experiment for performance comparison are ones using color histogram, SCD, CSD, color autocorrelogram, wavelet moments, EHD, BDIP and BVLC moments (BDIP-BVLC), combination of color histogram and wavelet moments, combination of SCD and BDIP-BVLC, and combination of CSD and EHD. Since the ways of forming the feature vectors in the retrieval methods are somewhat different, it is not easy to fix all their feature vector dimensions to the same value for fair comparison. Therefore, except the EHD method of a fixed dimension, the dimensions for the other conventional methods were selected so as to be close to 92 which is a dimension chosen for the proposed method. The choice of the feature vector dimension of 92 for the proposed method was based on our experimental results that its performance improved very slightly over the dimension of 92.

Table II shows the color space and feature vector dimension of the retrieval methods selected in our experiment. In the color histogram, color autocorrelogram, and combination of color histogram and wavelet moments, the numbers of quantization levels allocated to color components were chosen as ones which yield the best performance for the test DBs. In the color autocorrelogram, BDIP-BVLC, and combination of SCD and BDIP-BVLC, the distance k in (1) and the maximum distance k in (3) and (7) were chosen as $k = 1$. This is also because in our observation the smallest distance gave the most detailed local properties and the best performance for the test image DBs. In the wavelet moments and combination of color histogram and wavelet moments, their decomposition levels were chosen as $Z = 5$ and $Z = 3$, respectively, and Haar filters [30] were chosen as the wavelet filters. In the case of combination methods, the combinational strategy was adopted to be the same as that of the proposed method in (21).

In the proposed method, the wavelet decomposition level was chosen as $Z = 2$. The total number of quantization levels L for each of the wavelet decomposed H and S component images was also chosen as $L = 30$. Based on (9), the number of quantization levels $L_{m,n}$ of each subband was determined as $\{L_{m,1}\} = \{8, 4, 4, 4\}$ and $\{L_{m,2}\} = \{4, 2, 2, 2\}$. As a result, the dimensions of the color and texture feature vectors were given as $N_C = 60$ and $N_T = 32$, respectively.

When the wavelet decomposition level is $Z = 2$, the proposed progressive retrieval has three steps. At the step of $j = 1, 2, 3$, the total number of quantization levels for each of the wavelet decomposed H and S component images is given as $L = 8, 20$, and 30, respectively. So the feature vector dimensions used in the three steps of the progressive retrieval were given as $(N_1, N_2, N_3) = (20, 56, 92)$. For the selected feature vector dimensions, as a way to satisfy the condition related to (24), the number of target images retrieved in the first and second steps was determined to decrease exponentially as the step j increases and depending on the number of target images K_3 retrieved in the third step as follows:

$$K_j = K_3 \cdot e^{(3-j)}, \quad j \in \{1, 2\}. \tag{32}$$

We also decided that each component of the feature vector at each step is uniformly quantized with 8 bits. Therefore, the number of bits required for storing a target feature vector of the proposed method is 736.

C. Results

1) Effect of Progressive Retrieval: To evaluate the effect of the progressive retrieval, we compare the computational complexity and the retrieval accuracy of the proposed method with progressive scheme and that with nonprogressive scheme. First, we compare the computational complexities of the two methods, the estimation of which is minutely described in [16]. Table III shows the computational complexity per pixel of various retrieval methods including the two methods for VisTex DB where the ratio K_0/WH is about 0.073. In Table III, we can see that

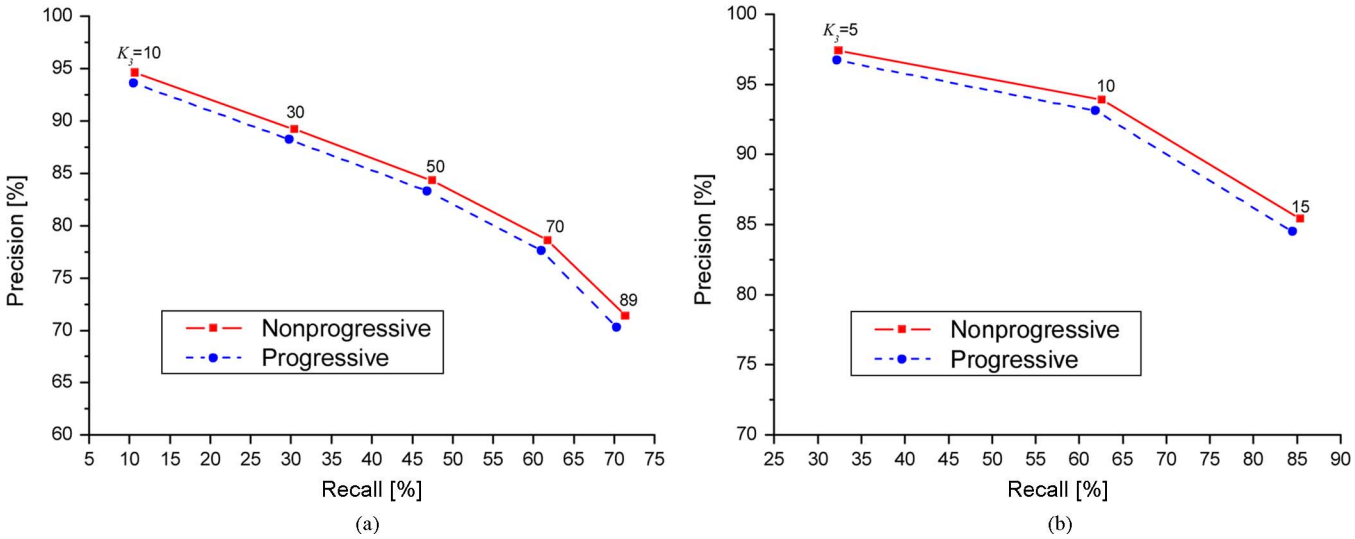


Fig. 9. Precision versus recall of the proposed method with nonprogressive scheme and that with progressive scheme. (a) Corel DB and (b) VisTex DB.

TABLE III
COMPUTATIONAL COMPLEXITY PER PIXEL OF VARIOUS RETRIEVAL METHODS FOR VISTEX DB WHERE $K_0/WH \approx 0.073$ [16]

Method	Computational complexity		
	Addition	Multiplication	Comparison
Color histogram	21.6	6	-
SCD	24.6	9	4
CSD	80.9	8	75
Color autocorrelogram	33.6	18.3	12
Wavelet moments	29.9	27	-
EHD	40.4	4.5	1.9
BDIP-BVLC	58.9	64	4.5
Color histogram + wavelet moments	20.5	12.6	-
SCD + BDIP-BVLC	34.9	28	5.5
CSD + EHD	90.7	9.5	75.6
Proposed with nonprogressive	47.9	18.6	15.1
Proposed with progressive	38.2	18.6	15.1

the proposed method with progressive scheme has the same numbers of the multiplications and comparisons as that with nonpressive scheme. However, the number of additions of the former is reduced by about 1.3 times over that of the latter.

We also evaluated the CPU times of the two methods, which were implemented on a Pentium IV PC with a 3.6 GHz processor and 1 GB of memory, running Microsoft Visual C++ 6.0 under the environment of Microsoft Windows XP Professional. Experimental results showed that the proposed retrieval method with progressive scheme gives 38 ms and that with nonprogressive scheme 44 ms for VisTex DB. We can see that the former is about 1.2 times faster than the latter. However, the CPU time of the progressive retrieval is expected to be much shorter than that of the nonprogressive retrieval for huge image DBs ($K_0 \gg WH$), as mentioned in Section III.

Fig. 9 shows the precision versus recall of the proposed retrieval method with nonprogressive scheme and that with progressive scheme for Corel DB and VisTex DB. In Fig. 9, the K_3 represents the number of target images finally retrieved. We can see from Fig. 9 that the former yields the average precision loss of 1.5% and that of 1.1% over the latter for Corel DB and for VisTex DB, respectively. Therefore, we can see that the performance of the proposed progressive retrieval is not so much degraded compared to that of the nonprogressive retrieval.

Fig. 10 shows the precision versus recall of the proposed method with progressive scheme according to each step. In Fig. 10, the retrieval in the second step is shown to yield the average precision gain of 6.9% and that of 5.7% over the retrieval in the first step for Corel DB and for VisTex DB, respectively. Also, the retrieval in the third step is shown to yield the average precision gain of 2.6% and that of 2.5% over the retrieval in the second step for Corel DB and VisTex DB, respectively. From the results, we can see that the retrieval accuracy is meaningfully improved as the step increases.

2) Performance Comparison of Retrieval Methods: First, we consider the computational complexity of the retrieval methods mentioned above, the estimation of which is also minutely described in [16]. In Table III, we can see that the proposed method with progressive scheme requires about 1.8 times the number of additions and about 3.1 times the number of multiplications of the color histogram method, which is the simplest one among the retrieval methods. We can see that the proposed method requires about 0.4 times the number of additions, about two times the number of multiplications, and about 0.2 times the number of comparisons of the retrieval method using a combination of CSD and EHD, which shows the highest computational load. We can also find that the proposed method with progressive scheme needs about 0.5 times the number of additions and about 2.3 times the number of multiplications and about 0.2 times the number of comparisons of the CSD method, which is known to produce the most excellent retrieval performance among the MPEG-7 color descriptors [29].

Fig. 11 shows the precision versus recall of the retrieval methods using single features and the proposed method with progressive scheme. In the figure, the proposed method yields

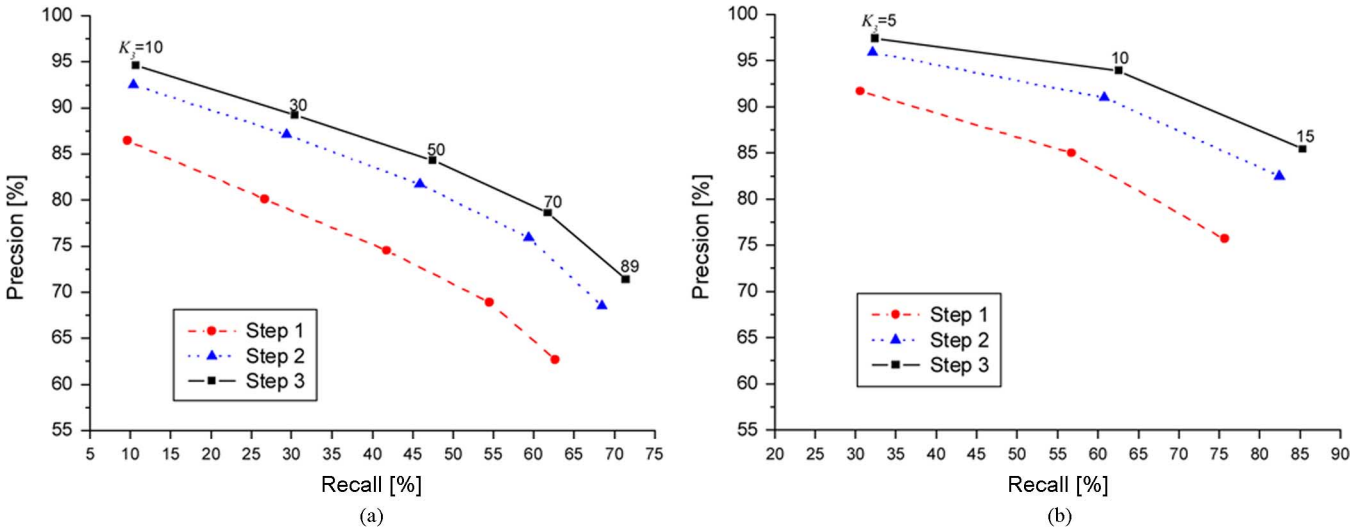


Fig. 10. Precision versus recall of the proposed progressive retrieval according to each step. (a) Corel DB and (b) VisTex DB.

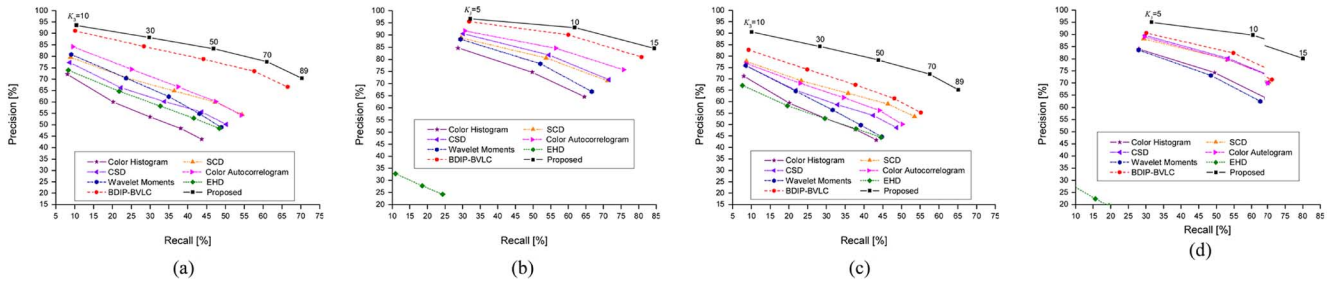


Fig. 11. Precision versus recall of the retrieval methods using single features and the proposed progressive retrieval method. (a) Corel DB, (b) VisTex DB, (c) Corel_MR DB, and (d) VisTex_MR DB.

3.8%–27.1% average precision gain for Corel DB [Fig. 11(a)], 3%–63.1% for VisTex DB [Fig. 11(b)], 9.8%–23.8% for Corel_MR DB [Fig. 11(c)], and 6.9%–65.1% for VisTex_MR DB [Fig. 11(d)] over the other methods, respectively. The proposed method also gives 3.8%, 14.7% average precision gain for Corel DB, 3%, 7.4% for VisTex DB, 9.8%, 15.4% for Corel_MR DB, and 6.9% and 8.8% for VisTex_MR DB over the method using BDIP-BVLC and that using color autocorrelogram, respectively. From these results, we can see that the effectiveness of combining the multiresolution color autocorrelogram and BDIP-BVLC is notable. Especially, these results show that the difference of retrieval accuracy between the proposed method and the two methods is greater for multiresolution DBs such as Corel_MR DB and VisTex_MR DB than for non-multiresolution DBs such as Corel DB and VisTex DB.

Fig. 12 shows a test query image and its relevant images, and Table IV shows the retrieval rank of the relevant images retrieved by the color autocorrelogram method, the BDIP-BVLC method, and the proposed method with progressive scheme for the query image in Fig. 12. In the case of Fig. 12(c), whose resolution is identical with that of the query image, retrieval ranks for the three methods appear similar. However, in the case of Fig. 12(b) and (d), whose resolutions are different from that of the query image, the retrieval ranks of the proposed method are much higher than those of the color autocorrelogram method and those of the BDIP-BVLC method. From the results in Fig. 12 and Table IV, we can see that the proposed method is more effective for multiresolution image DBs.

Fig. 12. Example of a query image (a) and its relevant images (b)–(d).

TABLE IV
RETRIEVAL RANKS OF THE RELEVANT IMAGES FOR THE QUERY IMAGE IN FIG. 12

Method / Image	Fig. 12(b)	Fig. 12(c)	Fig. 12(d)
Color autocorrelogram	54	2	15
BDIP-BVLC	17	1	43
Proposed with progressive	7	2	3

Fig. 13 shows the precision versus recall of the retrieval methods using combination of color and texture features and the proposed method with progressive scheme. In Fig. 13, the proposed method produces 4.6%–16.1% average precision gain for Corel DB [Fig. 13(a)], 3.9%–23.1% for VisTex DB

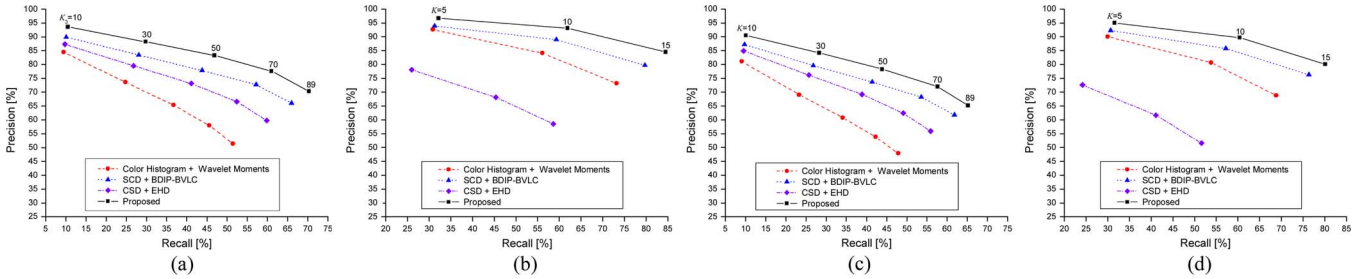


Fig. 13. Precision versus recall of the retrieval methods using combination of color and texture features and the proposed progressive retrieval method. (a) Corel DB, (b) VisTex DB, (c) Corel_MR DB, and (d) VisTex_MR DB.

TABLE V
ANMRR OF THE RETRIEVAL METHODS

Method \ DB	Corel	VisTex	MPEG-7 CCD	Corel_MR	VisTex_MR	MPEG-7 CCD_MR
Color histogram	0.4622	0.2637	0.2310	0.4657	0.2680	0.2576
SCD	0.3574	0.2047	0.1741	0.3649	0.2132	0.1959
CSD	0.4040	0.1949	0.0747	0.4310	0.2168	0.1237
Color autocorrelogram	0.3480	0.1705	0.1621	0.4054	0.2216	0.2288
Wavelet moments	0.3949	0.6831	0.4212	0.4272	0.7647	0.4983
EHD	0.2213	0.1193	0.2875	0.3220	0.1930	0.3812
BDIP-BVLC	0.3751	0.1853	0.2811	0.4109	0.2164	0.3443
Color histogram + Wavelet moments	0.3751	0.1853	0.2811	0.4109	0.2164	0.3443
SCD+BDIP-BVLC	0.2310	0.1291	0.1455	0.2675	0.1569	0.1916
CSD+EHD	0.2861	0.3133	0.1823	0.3238	0.3868	0.2488
Proposed with progressive	0.1837	0.0840	0.0864	0.2165	0.1131	0.1229

[Fig. 13(b)], 3.9%–15.5% for Corel_MR DB [Fig. 13(c)], and 3.5%–26.3% for VisTex_MR DB [Fig. 13(d)] over the other methods, respectively. From the results in Figs. 11 and 13, we can see that most of the combinational methods display higher retrieval accuracy than the method using single features which are used in the combinational methods. However, the method using CSD and EHD shows about 13% and 17.9% lower performance for the VisTex DB and VisTex_MR DB, respectively, than the method using only CSD.

Table V shows the ANMRR of the retrieval methods. The proposed method with progressive scheme yields 0.0376–0.2785 ANMRR gain for Corel DB, 0.0353–0.5991 for VisTex DB, 0.051–0.2492 for Corel_MR DB, and 0.0438–0.6516 for VisTex_MR DB over the other methods, respectively. For the MPEG-7 CCD, the proposed method yields a 0.0117 lower result than the CSD method and 0.0591–0.3348 gain over the other methods. For the MPEG-7 CCD_MR, the proposed method displays performance similar to that of the CSD method and 0.0687–0.3754 gain over the other methods. It is also shown that the performance degradation of the proposed method for the multiresolution DBs over its performance for the non-multiresolution DBs is much lower than that of the color autocorrelogram method and the BDIP-BVLC method. We can see from the above results that the proposed method almost always yields better performance in precision versus recall and in ANMRR over the other methods for the six test DBs.

In the proposed method, the feature vector is scalable according to the decomposition level Z in the wavelet transform domain. As the decomposition level Z increases, the feature vector has higher dimension. It was found in some experiments that the retrieval accuracies of $Z > 2$ are slightly better than those of $Z = 2$.

V. CONCLUSION

In this paper, a CBIR method has been proposed which uses the combination of H- and S-component color autocorrelograms and V-component BDIP-BVLC moments extracted in the wavelet transform domain. Experimental results for six test DBs showed that the proposed method yielded higher retrieval accuracy than the other conventional methods with no greater feature vector dimension. It was all the more so for multiresolution image DBs. In addition, the proposed method almost always showed performance gain in both of precision versus recall and ANMRR over the other methods for six test DBs. As further studies, the proposed retrieval method is to be evaluated for more various DBs and to be applied to video retrieval.

REFERENCES

[1] D. Feng, W. C. Siu, and H. J. Zhang, *Fundamentals of Content-Based Image Retrieval, in Multimedia Information Retrieval and Management—Technological Fundamentals and Applications*. New York: Springer, 2003.

[2] Y. Rui and T. S. Huang, "Image retrieval: Current techniques, promising, directions, and open issues," *J. Vis. Commun. Image Represent.*, vol. 10, pp. 39–62, Oct. 1999.

[3] 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 Anal. Machine Intell.*, vol. 22, pp. 1349–1380, Dec. 2000.

[4] R. Brunelli and O. Mich, "Image retrieval by examples," *IEEE Trans. Multimedia*, vol. 2, pp. 164–171, Sep. 2000.

[5] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, pp. 11–32, 1991.

[6] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, Jun. 1997, pp. 762–768.

[7] ISO/IEC 15938-3/FDIS Information Technology—Multimedia Content Description Interface—Part 3 Visual Jul. 2001, ISO/IEC/JTC1/SC29/WG11 Doc. N4358.

[8] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Trans. Syst. Man Cybern*, vol. SMC-8, pp. 610–621, Nov. 1973.

[9] J. R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases," in *Proc. IEEE Int. Conf. Image Processing*, Austin, TX, Nov. 1994, vol. 3, pp. 407–411.

[10] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 951–957, Sep. 2003.

[11] S. Liu, C. F. Babbs, and E. J. Delp, "Multiresolution detection of speculated lesions in digital mammorgrams," *IEEE Trans. Image Process.*, vol. 10, pp. 874–884, Jun. 2001.

[12] T. Chang and C. -C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Process.*, vol. 2, pp. 429–441, Oct. 1993.

[13] S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," *IEEE Trans. Multimedia*, vol. 6, pp. 676–686, Oct. 2004.

[14] A. Vadivel, A. K. Majumdar, and S. Sural, "Characteristics of weighted feature vector in content-based image retrieval applications," in *Proc. IEEE Int. Conf. Intelligent Sensing and Information Processing*, Chennai, India, Jan. 2004, pp. 127–132.

[15] H. Permuter, J. Francos, and I. H. Jermyn, "Gaussian mixture models of texture and colour for image database retrieval," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Hong Kong, Apr. 2003, vol. 3, pp. 569–572.

[16] Y. D. Chun, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features" Ph.D. thesis, Dept. Elect. Eng., Kyungpook National Univ., Daegu, Korea, Dec. 2005 [Online]. Available: http://vcl.knu.ac.kr/phd_thesis/ydchun.pdf

[17] T. Ojala, M. Rautiainen, E. Matinmikko, and M. Aittola, "Semantic image retrieval with HSV correlogram," in *Proc. 12th Scand. Conf. Image Analysis*, Bergen, Norway, Jun. 2001, pp. 621–627.

[18] D. E. Pearson and J. A. Robinson, "Visual communication at very low data rates," *Proc. IEEE*, vol. 73, pp. 795–812, Apr. 1985.

[19] Y. J. Ryoo and N. C. Kim, "Valley operator extracting sketch features: DIP," *Electron. Lett.*, vol. 248, pp. 461–463, Apr. 1988.

[20] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[21] Y. D. Chun, J. K. Sung, and N. C. Kim, "Image retrieval using combination of color and multiresolution texture features," in *Proc. SPIE Storage and Retrieval Methods and Applications for Multimedia*, San Jose, CA, Jan. 2005, pp. 195–203.

[22] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

[23] D. Androutsos, K. Plataniotiss, and A. Venetsanopoulos, "Distance measures for color image retrieval," in *Proc. IEEE Int. Conf. Image Processing*, Chicago, IL, Oct. 1998, vol. 2, pp. 770–774.

[24] M. Ankerst, H. P. Kriegel, and T. Seidl, "A multistep approach for shape similarity search in image databases," *IEEE Trans. Knowl. Data Eng.*, vol. 10, pp. 996–1004, Nov./Dec. 1998.

[25] B. C. Song, M. J. Kim, and J. B. Ra, "A fast multiresolution feature matching algorithm for exhaustive search in large image databases," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 673–678, May 2001.

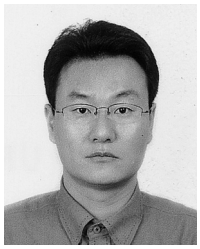
[26] Common Datasets and Queries in MPEG-7 Color Core Experiments Oct. 1999, ISO/IEC/JTC1/SC29/WG11/MPEG99, Doc. M5060.

[27] S. F. Chang, W. Chen, H. J. Meng, H. Sundaram, and D. Zhong, "A fully automated content-based video search engine supporting spatiotemporal queries," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 602–615, Sep. 1998.

[28] P. Ndjiki-Nya, J. Restat, T. Meiers, J. R. Ohm, A. Seyferth, and R. Sniehotta, "Subjective evaluation of the MPEG-7 retrieval accuracy measure (ANMRR)," in *ISO/WG11 MPEG Meeting*, Geneva, Switzerland, May 2000, Doc. M6029.

[29] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 703–715, Jun. 2001.

[30] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, *Wavelets for Computer Graphics: Theory and Applications*. San Mateo, CA: Morgan Kaufmann, 1996.

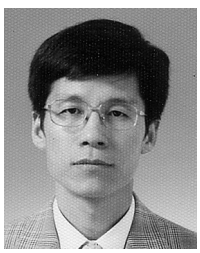


Young Deok Chun received the B.S. degree in electronic engineering from the Inje University, Gimhae, Korea, in 2000, and the M.S. and Ph.D. degrees in electronic engineering from the Kyungpook National University, Daegu, Korea, in 2002 and 2006, respectively.

Since March 2006, he has been with the GPG1, S/W Lab., Mobile Communication Division at Samsung Electronics Co. Ltd., Gumi, Korea, where he is currently a Senior Research Engineer. His research interests are image and video retrieval and mobile communication.

Nam Chul Kim (S'81–M'84) received the B.S. degree in electronic engineering from the Seoul National University, in 1978, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1980 and 1984, respectively.

Since March 1984, he has been with the Department of Electronic Engineering at Kyungpook National University, Daegu, Korea, where he is currently a Professor. During 1991–1992, he was on leave as a visiting scholar at the Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY. His research interests are image and video retrieval, image analysis and classification, biomedical image processing, and image and video coding.



Ick Hoon Jang (S'95–M'99) received the B.S., M.S., and Ph.D. degrees, all in electronic engineering from the Kyungpook National University, Daegu, Korea, in 1986, 1988, and 1998, respectively.

From 1988 to 1994, he was with the Agency for Defence Development, Daejon, Korea, as a Research Engineer. Since March 1998, he has been with the Department of Electronic Engineering at Kyungwoon University, Gumi, Korea, where he is currently an Associate Professor. His research interests are image and video retrieval, image analysis and classification, and image and video coding.