Content-Based Image Retrieval with Combined Features: Color and Gradient

Mahmut Kılıçaslan (✉ m.kilicaslan@ankara.edu.tr)

Ankara University: Ankara Universitesi <https://orcid.org/0000-0003-1117-7736>

Recep Demirci

Gazi University: Gazi Universitesi

Research Article

Keywords: colour quantization, colour histogram, content-based image retrieval

Posted Date: February 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-222235/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Content-Based Image Retrieval with Combined Features: Color and Gradient

Mahmut KILIÇASLAN^{a*}, Recep DEMİRCİ^b

^a Ankara University, Nallihan Vocational School, Computer Technologies Department

^b Gazi University, Technology Faculty, Computer Engineering Department

*corresponding author, m.kilicaslan@ankara.edu.tr, +90 507 5411125

ORCID:

Mahmut Kılıçaslan: <https://orcid.org/0000-0003-1117-7736>

Recep Demirci: <https://orcid.org/0000-0002-3278-0078>

Abstract: Feature extraction is fundamental stage of effective content based image retrieval (CBIR). However, it remains challenging issue to extract low-level features for retrieval systems. This paper puts forward an effective solution proposal for the aforementioned problem. Initially, images and their gradients are clustered with multi-level thresholding. A codebook is generated with threshold values. The size of the codebook generated depends on the number of thresholds. Consequently, every pixel in color image is included in a cluster by means of the codebook. Color reduction is performed by assigning the average values of pixels in the same cluster. A cluster-based one-dimensional histogram (CBH) is created with the numbers of pixels in every cluster represented with a single color. Then the cluster-based feature vectors with histogram are extracted from original image and gradient image. Accordingly, relevant features are combined. The developed feature vector is called as combined feature vector (CFV). The most important advantages of CFV are that it performs an effective color reduction technique and feature presentation by processing texture information with gradient operator. Therefore, the main contribution of the combined feature vector suggested is its high accuracy and stability for image retrieval. The proposed method has been tested with Corel-1K, Corel-5K Corel-10K and GHIM-10K datasets. In addition, performances of different image histogram similarity techniques such as cosine, histogram intersection and Euclidean distance have been verified with the developed algorithm. Experimental results have been analyzed in two categories. Initially, CBIR results produced with combined feature vectors which are generated by Otsu, Kapur and center of gravity of histogram (CGH) procedures have been evaluated. Then, the CBIR strategy based on CGH method has been compared with CBIR systems with local binary pattern (LBP) and gradient-structures histogram (GSH). It was observed that CBIR approach based on CGH technique has significantly outperformed.

Keywords: colour quantization; colour histogram; content-based image retrieval.

1. Introduction

The rapid increase in the number and usage areas of digital images have created large image databases. Retrieving any desired images from the related databases has become a critical issue. For this reason, image retrieval which is used in many fields such as medicine and health, aeronautics and military has been an important research area [1-3]. Text based systems were initially launched for the image retrieval problem. Related systems are based on indexing images with text data. However, manual indexing of all images in image databases requires costly labor [4,5]. In addition, indexing images with texts is a subjective approach. Content-based image retrieval (CBIR) systems that use low-level features of images such as color, texture, shape have been proposed to overcome the stated shortcomings of text-based retrieval systems [6].

A typical CBIR architecture consists of two basic stages: feature extraction and similarity measurement. The feature extraction process is the most vital stage of the relevant systems. Many researchers have focused on feature extraction procedures. Because the capability of the extracted features to represent images significantly affects the performance of CBIR systems. As a result, if the feature vectors attained for images are not defined properly, the achievements of the systems will be low. Feature extraction methods are generally evaluated globally and locally. Global methods work on the whole image and are computational methods with low cost. On the other hand, local approaches are the techniques that consider the local feature of regions in the image. Histogram [7], color moment [8], co-occurrence matrix [9] are generally used global features. LBP [10], GIST [11], SIFT [12] and SURF [13] are the typical used techniques for local features.

The histogram vector is a well-known color-based descriptor and can be used as a global feature. The easy calculation of related vector, its robustness to rotation and scaling made it frequently used for CBIR systems [14,15]. Liu and Yang adopted the edge orientation and the color difference between the two points into histogram [16]. Varish et al. devised a color and texture based retrieval system. In their suggested study, images were converted into HSV color space and used in histogram [17]. In another study, the edge histogram and wavelet transform were applied in the YbCbCr color space. It has been seen that the feature vector obtained by edge histogram and wavelet transformation increases the performance of the developed system [18]. A histogram, which is an effective method to represent images, is expressed in gray-level images with a one-dimensional vector. In color images, there are three different histograms. Therefore, the combination of histogram information created from each color channel is an issue that needs to be solved. In addition, using the color histogram in image retrieval systems

requires processing of a 3-dimensional array. However, since there are 2^{24} different colors in RGB color space, the computational cost will increase excessively. It has become imperative to use color reduction methods to overcome the aforementioned problems. Some of the popular color reduction approaches proposed so far are Median-Cut [19], Octree [20], K-means [21], C-means [22]. Although the global histogram is easy to use and frequently used representation tool, it does not calculate the spatial information of pixels. Also, it does not contain information about the texture properties of the image. The related inadequacies make it necessary to use a combined vector. It must include additional information such as texture, shape etc. apart from color features.

Although various methods have been proposed as a features vector, the texture feature is one of the frequently used in image analysis [23-25]. Li et al. [26] proposed a medical image retrieval system using texture in their studies. Multilevel indexing has been used in their strategies which increase the performance of retrieval system. In the developed study, the relationship between symmetric pixels through a local window was analyzed and the texture feature was extracted. By using the obtained texture and color features together, the performance of the retrieval system has been improved [27]. It has been shown that performance increases with the use of multiple visual representation in their studies. The effect of similar approaches on CBIR systems has been confirmed in other studies [28-31]. For example, one of the frequently preferred techniques to extract texture is local binary pattern (LBP). LBP is texture feature methodology that takes into account the relationship between the center pixel and neighboring pixels [32]. Variants of LBP algorithm have been employed by many researchers [33-38]. Nevertheless, the most important drawbacks of LBP, which uses local spatial features overlooks macro texture information and is sensitive to noise.

One of the remarkable approaches for CBIR systems in recent years is deep learning based solutions [39,40]. However, these methodologies have higher computational complexity than traditional low-level feature-based approaches. Furthermore, the number of images in database must be enough learning process to be successful.

In this paper, a novel image retrieval system has been developed with the feature vectors obtained from the combination of gradient features and outcomes of cluster-based color histogram (CBH) algorithm. Consequently, the texture and color features have been unified to create feature vector. The multi-level thresholding process was calculated by the center of gravity of the histogram method. Then RGB color space is divided into sub- prisms with the thresholds calculated for each color channel. The averages of all pixels in the same sub-prism

are assigned for the same cluster. Thus, the color reduction process was implemented. Also color index based one dimensional histogram was created. The similar process was applied into original image and its gradients. Consequently, the global feature vector has created by combining the one-dimensional histogram information extracted from original image and its gradient.

2. Image Thresholding and Color Quantization

Image thresholding method are a process that incorporates similar pixels into the same group and provides an effective solution for image clustering. Kapur's entropy and Otsu techniques are the thresholding methods which give successful results for gray level images. Kapur's entropy aims to determine values that maximize the sum of local entropy at each gray level [41]. In the Otsu approach, the variance values that makes a maximum between classes are taken into account [42]. Both optimization methods are based on the image histogram. While the related methods work with a single histogram on gray-level images, they must be extended for three different histograms for color images. The histogram indicates the probability of pixels in the image and probability of i^{th} level pixel is calculated as

$$p_i = h_i / (M \times N) \quad (1)$$

where h_i number of pixels in i^{th} level and $M \times N$ is the size of the image. Otsu and Kapur methods determine threshold values through probability distribution given in Equation 1. When the multi-level thresholds are used for gray-level images such as $t_1, t_2, \dots, t_m$, the number of clusters becomes $m + 1$ and the clusters are defined as follows:

$$c_0 = \{0, \dots, t_1 - 1\}, c_1 = \{t_1, \dots, t_2 - 1\}, \dots, c_m = \{t_m, \dots, L - 1\} \quad (2)$$

Determination of threshold values is the main problem for image thresholding. The objective function of the Kapur to overcome the related problem is based on entropies. The related parameters are chosen as

$$J(t_1, t_2, \dots, t_m) = H_1 + \dots + H_m \quad (3)$$

where $H_1, \dots, H_m$ are partial entropies of each class and they are defined as

$$\begin{aligned}
H_1 &= -\sum_{i=0}^{t_1-1} \frac{p_i}{\omega_0} \ln \frac{p_i}{\omega_0}, \omega_0 = \sum_{i=0}^{t_1-1} p_i, \\
&\dots \\
H_m &= -\sum_{i=t_m}^{L-1} \frac{p_i}{\omega_m} \ln \frac{p_i}{\omega_m}, \omega_m = \sum_{i=t_m}^{L-1} p_i
\end{aligned} \tag{4}$$

where $\omega_0, \dots, \omega_m$ are partial probabilities of clusters. The values of variables, which make Equation 3 maximum are considered to be accurate thresholding for gray scale image clustering. Objective function of Otsu threshold approach is defined as

$$J(t_1, t_2, \dots, t_m) = \sigma_0 + \sigma_1 + \dots + \sigma_m \tag{5}$$

where $\sigma, \sigma_1, \dots, \sigma_m$ are defined as variance between clusters and are calculated as

$$\begin{aligned}
\sigma_0 &= \omega_0 (\mu_0 - \mu_T)^2, \\
\sigma_1 &= \omega_1 (\mu_1 - \mu_T)^2, \\
&\dots \\
\sigma_m &= \omega_m (\mu_m - \mu_T)^2
\end{aligned} \tag{6}$$

ω indicates partial probabilities, μ_i represents the class averages and $\mu_T = \sum_{i=0}^{L-1} ip_i$ gray level average. Related variables are calculated by Equation 7 as follows.

$$\begin{aligned}
\omega_0 &= \sum_{i=0}^{t_1-1} p_i, \quad \mu_0 = \sum_{i=0}^{t_1-1} \frac{ip_i}{\omega_0} \\
\omega_1 &= \sum_{i=t_1}^{t_2-1} p_i, \quad \mu_1 = \sum_{i=t_1}^{t_2-1} \frac{ip_i}{\omega_1} \\
&\vdots \\
\omega_m &= \sum_{i=t_m}^{L-1} p_i, \quad \mu_m = \sum_{i=t_m}^{L-1} \frac{ip_i}{\omega_m}
\end{aligned} \tag{7}$$

The explained thresholding techniques above are applied in gray-level images. In addition, as the number of thresholds increases, the computational cost becomes high. Demirci and Okur [43] have recently developed an algorithm called as center of gravity of histogram (CGH) that provides multi-level thresholding. In their strategy histogram data and recursive means have been used for threshold values estimation. Initially, the global mean, μ_T was calculated. Subsequently, means determined from previous steps was re-used. Thus, the number threshold was increased with each step. The related algorithm based on the sum of partial probability distributions and partial means are calculated as follows:

$$\begin{aligned}
\omega_0 &= \sum_{i=0}^{\mu_T-1} p_i, & \mu_0 &= \sum_{i=0}^{\mu_T-1} \frac{ip_i}{\omega_0}, \\
\omega_1 &= \sum_{i=\mu_T}^{255} p_i, & \mu_1 &= \sum_{i=\mu_T}^{255} \frac{ip_i}{\omega_1}, \\
\omega_{00} &= \sum_{i=0}^{\mu_0-1} p_i, & \mu_{00} &= \sum_{i=0}^{\mu_0-1} \frac{ip_i}{\omega_{00}}, \\
\omega_{01} &= \sum_{i=\mu_0}^{\mu_T-1} p_i, & \mu_{01} &= \sum_{i=\mu_0}^{\mu_T-1} \frac{ip_i}{\omega_{01}}, \\
\omega_{10} &= \sum_{i=\mu_T}^{\mu_1-1} p_i, & \mu_{10} &= \sum_{i=\mu_T}^{\mu_1-1} \frac{ip_i}{\omega_{10}}, \\
\omega_{11} &= \sum_{i=\mu_1}^{255} p_i, & \mu_{11} &= \sum_{i=\mu_1}^{255} \frac{ip_i}{\omega_{11}}
\end{aligned} \tag{8}$$

As could be seen, each stage depends on the values obtained in previous stage. The threshold values to be determined with the proposed approach are presented in Table 1. While using the general average of the image for single threshold, recursive means are taken into account for more threshold.

Table 1. Number and values of thresholds in CGH

Number of thresholds	Values of thresholds
t_1	μ_T
$t_1 \ t_2$	$\mu_0 \ \mu_1$
$t_1 \ t_2 \ t_3$	$\mu_0 \ \mu_T \ \mu_1$
$t_1 \ t_2 \ t_3 \ t_4$	$\mu_{00} \ \mu_{01} \ \mu_{10} \ \mu_{11}$
$t_1 \ t_2 \ t_3 \ t_4 \ t_5$	$\mu_{00} \ \mu_{01} \ \mu_T \ \mu_{10} \ \mu_{11}$

The above-mentioned thresholding techniques are based on histogram, which is a feature vector that represents images globally. The low calculation cost, robustness to scaling and rotation makes it a useful feature for image retrieval systems. Single channel information is used in gray-level images for image retrieval. On the other hand, since the color image has three channels, three different histograms must be processed or combined for color image retrieval. The second solution for color image retrieval is to use the color histogram, which has a high calculation cost since the three-dimensional array needs to be processed. In addition, in color images, each color channel histogram consists of 256 levels. Therefore, color reduction methodologies are among the topics that researchers focus on.

Color reduction is a frequently used method for image compression and classification. It is also defined as the expression of the image with less colors. When the color quantization algorithm

is applied for the color image, the size of the image does not change, only the numbers of colors in image is reduced. In other words, the pixels in image are clustered as meaningful subsets.

In this study, a novel cluster-based one-dimensional histogram (CBH) which combines three different color channels with thresholding techniques has been proposed. Furthermore, color reduction structure has been devised for color image retrieval.

Firstly, RGB color space is divided into sub-cubes with threshold values determined for each color channel through thresholding techniques. Pixels located in the same sub-cubes are assigned to the same cluster. The average values of pixels in the same cluster have been used for color reduction. The number of sub cubes in RGB space depends on the number of thresholds. When the number of thresholds increases, the number of cubes will increase. Thus, the number of sub-cubes in the RGB is $k = (r + 1)^3$ where the number of threshold was chosen as r [44]. Subsequently, the labels of clusters are distributed as $\{c_0, c_1, \dots, c_{k-1}\}$. Figure 1 (a) and Figure 1 (b) show the partition of the color space with a single threshold and two thresholds, respectively. While 8 sub-prisms are formed with a single threshold, 27 sub-prisms are formed with two thresholds.

Figure 1. Partitioned RGB color space (a) $r=1$, $k=8$ (b) $r=2$, $k=27$

The rules for partitioning of RGB color space are given in Table 2. Here t_r , t_g and t_b respectively show threshold values calculated from red, green and blue channels while n_i represents the number of pixels assigned to the i^{th} class. Since the number of all pixels in an image is equal to the sum of the number of pixels assigned to the classes, and number of pixels in image can be calculated as

$$M \times N = \sum_{i=0}^k n_i \quad (9)$$

Accordingly, the probability of any pixel being in i^{th} class could be written as follows,

$$p_i = \frac{n_i}{M \times N} \quad (10)$$

Subsequently, probability distribution function p_i demonstrates a cluster-based one-dimensional feature vector representing color images. It has been created by combination of three color channels. The rules in Table 2 refer to a codebook generation process. Color reduction is performed simultaneously by assigning the average values of pixels in each sub-cubes. Figure 2 (a) shows a randomly selected image from the Corel-1K dataset and Figure 2 (b) shows the distribution of its pixels in RGB color space. Figure 3 (a) indicates the reduced version of Figure 2 (a) by using CGH with single threshold. Figure 3 (b) shows the color distribution of the relevant image with 8 clusters. Figure 3 (c) represents the 27-cluster version of the same original image by using multilevel thresholds with the CGH method and Figure 3 (d) is its color distribution. Threshold values for red, green and blue channels with a single threshold are $t_{r,1}=115$, $t_{g,1}=81$ and $t_{b,1}=88$ respectively. The values with two thresholds are $t_{r,1}=62$ and $t_{r,2}=128$ for red channel, $t_{g,1}=26$ and $t_{g,2}=139$ for green channel and $t_{b,1}=35$ and $t_{b,2}=90$ for blue channel, respectively. Information loss is an expected case in color reduction. When Figure 3(a) is examined, it is seen that some details of road and bus in image have been lost. On the other hand, it is clear that the relevant details are preserved in Figure 3(c). In addition, consistency was not disturbed in both reduced images. When the color reduction was performed for 1, 2 and 3 thresholds with CGH, PSNR values are given in Table 3. As could be seen in Table 3, the similarity of reduced image has increased with the number of thresholds. Figure 4 shows the traditional histogram of bus image in Figure 2 (a) whereas Figure 5 (a) and Figure 5 (b) demonstrate the cluster-based histograms. Related histograms with 8 and 27 clusters are also feature vectors representing quantized images.

Table 2. The rules for partition of color space by single threshold

Class label (c_i)	Partition rules	Number of pixels (n_i)	Binary Code
c_0	$if(R \leq t_{r,1} \ \&\& \ G \leq t_{g,1} \ \&\& \ B \leq t_{b,1})$	n_0	000
c_1	$if(R \leq t_{r,1} \ \&\& \ G \leq t_{g,1} \ \&\& \ B \geq t_{b,1})$	n_1	001
c_2	$if(R \leq t_{r,1} \ \&\& \ G \geq t_{g,1} \ \&\& \ B \leq t_{b,1})$	n_2	010
c_3	$if(R \leq t_{r,1} \ \&\& \ G \geq t_{g,1} \ \&\& \ B \geq t_{b,1})$	n_3	011
c_4	$if(R \geq t_{r,1} \ \&\& \ G \leq t_{g,1} \ \&\& \ B \leq t_{b,1})$	n_4	100
c_5	$if(R \geq t_{r,1} \ \&\& \ G \leq t_{g,1} \ \&\& \ B \geq t_{b,1})$	n_5	101
c_6	$if(R \geq t_{r,1} \ \&\& \ G \geq t_{g,1} \ \&\& \ B \leq t_{b,1})$	n_6	110
c_7	$if(R \geq t_{r,1} \ \&\& \ G \geq t_{g,1} \ \&\& \ B \geq t_{b,1})$	n_7	111

(a)

(b)

Figure 2. Bus (a) original (b) color distribution

(a)

(b)

(c)

(d)

Figure 3. Bus quantized (a) $r=1, k=8$ (b) distribution of 8 clusters (c) $r=2, k=27$ (d) distribution of 27 clusters

Table 3. PSNR values of reduced Bus image

Number of Colors	R	G	B	Average
8	17,74	17,98	18,15	17,96
27	22,92	22,71	22,32	22,65
64	25,27	25,94	26,10	25,77

Figure 4. Conventional histogram of Bus

Figure 5. CBH of Bus a) 8 clusters b) 27 clusters

3. Image Retrieval based on Combined Feature Vector

The aim of CBIR system is to retrieve any desired image from databases. Feature databases are created from related image databases and CBIR utilize these feature vectors. So feature extraction is basic step of image retrieval. The histogram is an important feature vector to represent images. The corresponding vector is extracted by processing a single color channel in gray-level images. On the other hand, in color images three color channels must be processed separately. For this reason, computational complexity and storage costs of the process in color images are high. Also, combining three different color histograms is another problem. Kılıçaslan et al developed a new one-dimensional histogram feature vector for color images with the CGH thresholding technique[45]. However the produced vector does not contain the texture features of the image. They also used only one feature of images.

In this study, a novel combined feature vector (CVF), in which histograms of original and gradients of images are unified has been proposed. In the established approach, first of all texture features of image are extracted with the gradient operator. The original image and its gradient are reduced with threshold techniques. Accordingly, CBHs are made for original image and its gradient. CBHs have 8 colors for one threshold, 27 for two thresholds and 64 for three thresholds. CBHs obtained from original and gradient are feature vectors capable of representing of an image. Thus, texture and color features are integrated.

Figure 6 indicates the stages of proposed feature vector extracting algorithm. Figure 7 represents the structure of CFV for a single threshold. As seen in Figure 7, the corresponding vector has 16 elements that hold color and texture features together. The CFV vector has 54 and 128 elements for thresholds 2 and 3 respectively. Figure 8 is a block diagram of developed CBIR with CFV. The CFV is one-dimensional color histogram which contains color and texture features. Thus, there is no need to process three different histogram or three dimensional color histograms. Therefore, it is computational cost effective as well. Similarity measurement is another step of CBIR systems and any histogram-based similarity measurement technique can be used in the similarity measurement block of the proposed CBIR system. Euclidean, histogram intersection, and cosine similarity metrics are widely used approaches. So they have been used in our experiments as well. The euclidean similarity is calculated as follows

$$S_E = \sqrt{\sum_{i=0}^k (p_{1,i} - p_{2,i})^2} \quad (11)$$

where p_1 and p_2 are histograms of the queried and any image in database, respectively. k is the number of the color. The histogram intersection is an other technique which is wide preferred. The related similarity method can be defined as

$$S_I = \sum_{i=0}^k \min(p_{1,i}, p_{2,i}) \quad (12)$$

and cosine similarity formula is as follows

$$S_c = \frac{\sum_{i=0}^k p_{1,i} \cdot p_{2,i}}{\sqrt{\sum_{i=0}^k (p_{1,i})^2} \cdot \sqrt{\sum_{i=0}^k (p_{2,i})^2}} \quad (13)$$

Figure 6. Feature extraction step

Figure 8. Architecture of proposed CBIR based on CVF

4. Experimental Results and Discussion

The proposed algorithm was tested in different databases such as Corel-1K, Corel-5K, Corel-10K and GHIM-10K. Corel dataset is a frequently used data collection. Corel-1K has 10 different categories such as Africans, Beaches, Buildings, Buses etc. Relevant database has 100 images in each category. The resolutions of images are 384x256 or 256x384. Image sizes of Corel-5K are 187x126 or 126x187. It has images of 50 categories, including Lions, Bears, Drinks, Fighting etc., each category 100 images. The other image collection is the Corel-10K, which contains 10,000 images in 100 categories. Each category contains 100 images resolution 187x126 or 126x187 pixels. GHIM-10K is the last database where experiments are conducted. GHIM contains 10,000 images and 20 different categories. Each category consists of 500 images with sized 400x300 or 300x400.

The performances of CBIR system are measured by using precision (P), recall (R) and mean average precision (mAP). The P and R variables are calculated as follows:

$$P = \frac{\text{number of relevant images retrieved}}{\text{number of retrieved images}} \quad (14)$$

$$R = \frac{\text{number of relevant images retrieved}}{\text{total number of relevant images}}$$

The mAP value is defined as

$$mAP = \frac{1}{X} \sum_{j=1}^X \frac{1}{Q_j} \sum_{i=1}^{Q_j} P(I_i) \quad (15)$$

where Q_j is the number of relevant images for query j , X is total number of queries, $P(I_i)$ is precision of i^{th} relevant images. While obtaining the experimental results, each image in database was used as a query and the top 20 images were retrieved. Precision-Recall (P-R) curves were generated. The P-R curves are one of the most frequently used graphical methods to compare algorithms. Therefore, general performances in different databases are presented with the P-R curves.

Figure 9 and Figure 10 are P-R curves of experimental results achieved with Corel-1K and Corel-10K datasets, respectively. Figure 9(a) and Figure 10(a) are performances with cluster-based color histogram algorithm while Figure 9(b) and Figure 10(b) are performances of CFVs consisting of gradient and color information. When P values at the same R levels in Figure 9(a) and Figure 9(b) are examined, the outcomes of CFV have higher precision values. The similar performances are observed with Kapur, Otsu and CGH techniques. For example, in Figure 9(a), the P values at the level of $R = 0.3$ are 0.7, 0.76 and 0.79 for Kapur, Otsu and CGH, respectively. On the other hand, in Figure 9(b), the P values at the same R level are 0.8, 0.85, and 0.88, respectively. By combining the gradient feature with color, it is 10% superior with Kapur and 9% with Otsu and CGH. The success of integration of color and gradient for CBIR system is obvious with Corel-10K dataset. So it is apparent that the CBIR algorithm with CFV created by combining gradient histogram and color histogram has improved system performance. Therefore, it has retrieved more relevant images. CFV is more successful than traditional cluster-based color histogram [45] at distinguishing the searched images. Also, the corresponding vector uses low-level features. It does not require any parameters. Additionally, the related figures clearly demonstrate that the achievement of CGH approach in both databases are higher than Kapur and Otsu techniques. In other words, CGH technique has retrieved higher relevant images than Otsu and Kapur optimization approaches.

The Corel-10K dataset has more images than Corel-1K. Accordingly, it is unescapable that the increment in number of image in database lead to reduce in performance of CBIR system. For instance, the P values at the level of $R = 0.3$ are 0.48, 0.65 and 0.74 for Kapur, Otsu and CGH, respectively. When they are compared with results given in Figure9(b), the dropping in performance is clear. The performance of CGH has been reduced 14% while the Otsu and Kapur has been affected by 20% and 32%, respectively. Consequently, it could be concluded that the CGH is more robust than others.

Figure 9. P-R Curves: Corel-1K with Euclidean Similarity: a) Color b) Color and Gradient

Figure 10. P-R Curve: Corel-10K with Euclidean Similarity a) Color b) Color and Gradient

The former results have been obtained with Euclidean similarity method. The proposed algorithm has been tested with cosine similarity metric and histogram intersection technique as well. Table 4 shows the performance of the cosine similarity metric on Corel-1K dataset. The related findings clearly indicate that three approaches produce better results with increasing of number of threshold. On the other hand, the CGH algorithm has retrieved the maximum number of relevant images. The mAP value of the CGH algorithm with $r = 3$ is 77.8% while they are 74.7% and 73.7% for Otsu and Kapur, respectively.

As the CGH algorithm proposed for CBIR system is more successful than Otsu and Kapur techniques in Corel-1K and Corel-10K dataset, its performance was compared with other image similarity methods. Furthermore, the number of threshold, r was set as 3. Accordingly, the performance of proposed system with different similarity approaches and various databases has been shown in Table 5. It can be seen that the Euclidean similarity metric is more efficient than other methods. As it is known, histogram intersection is an approach based on the minimum value of histogram bins. There are some clusters in CFV with a zero frequencies. Accordingly, intersection metric produces the worst results on databases.

Table 4 mAP performance of cosine similarity metric with different thresholds: CFV

Dataset	Number of threshold (r)	CFV with Kapur	CFV with Otsu	CFV with CGH
Corel-1K	1	0,697	0,702	0,709
	2	0,719	0,732	0,756
	3	0,737	0,747	0,778

Table 5 mAP performance of different similarity metrics (CFV): $r=3$.

Dataset	Similarity metric		
	Cosine	Euclidean	Intersection
Corel-1K	0,778	0,796	0,781
Corel-5K	0,682	0,698	0,677
Corel-10K	0,584	0,601	0,553

It was confirmed from previous experiments that the proposed CBIR system based on CFV and Euclidean similarity metric was successful within the same group of algorithms. To see outcomes of the developed algorithm, further experiments were carried out. Therefore, local binary pattern (LBP) and gradient structures histogram (GSH) based CBIR algorithms have been compared with the devised method. For comparison, two different databases: Corel-10K and GHIM-10K have been tested. The precision and Recall parameters of algorithms has been given in Table 6. The drawbacks of LBP are that it produces a long histogram and does not evaluate macro texture. The GSH strategy has a limitation that it needs the balanced color, density and edge components [46]. On the other hand, CFV not only consists of a combination of general color and texture information, but also does not require any parameters. Therefore, the CFV based system is more effective than others.

Table 6 Precision and recall parameters: Corel-10K and GHIM-10K

Dataset	Performance Parameter	Method		
		LBP	GSH	Proposed (CFV)
Corel-10K	Precision (%)	37,23	54,84	55,45
	Recall (%)	4,47	6,58	7,02
GHIM-10K	Precision (%)	46,89	63,38	65,84
	Recall (%)	1,13	1,52	1,87

5. Conclusion

Feature extraction stage is an important issue for an effective CBIR systems. In feature extraction, low-level features of images are converted into vectors. Images are represented by feature vectors. Histogram is a commonly used feature vector in CBIR systems. However, the histogram is a multidimensional matrix for color images. Therefore, computational complexity will be high if an operation is performed on the relevant matrix. In this study, a novel retrieval architecture using cluster-based one-dimensional histogram (CFV) which combines texture and color features has been suggested. Initially gradient operator was applied to color images. Subsequently, the original and gradient image were clustered using multi-level thresholding techniques. Color reduction is performed by assigning the average values of the related cluster in the reduced image. Cluster-based one-dimensional histograms (CBH) were produced for both original and gradient images. Accordingly, the CFVs were formed by combining them. Experiments were performed in different databases. It was observed that the CFV with CGH is significantly superior to CFV with Kapur and Otsu techniques. Consequently, a novel content based image retrieval system based on combined feature vector, which contains color and texture information has been developed. Its performance was confirmed by experiments.

Acknowledgments

The authors declare that they have no conflict of interest.

Funding

We confirm that we do not have a funding source.

Authors' contributions

MK provided drafted the manuscript, carried out implementations and simulations for this manuscript. RD provided core concepts proofread the manuscript and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

References

- [1] Rodrigues, J., Cristo, M., and Colonna, J. G. (2020). Deep hashing for multi-label image retrieval: a survey,” *Artificial Intelligence Review*, 1-47.
<https://doi.org/10.1007/s10462-020-09820-x>
- [2] Bondur, V. G., and Murynin, A. B. (2016). Methods for retrieval of sea wave spectra from aerospace image spectra. *Izvestiya, Atmospheric and Oceanic Physics*, 52(9), 877-887.
<https://doi.org/10.1134/S0001433816090085>
- [3] Fu, H., Zhao, H., Kong, X. et al. BHoG: binary descriptor for sketch-based image retrieval. *Multimedia Systems* 22, 127–136 (2016). <https://doi.org/10.1007/s00530-014-0406-9>
- [4] Long, F., Zhang, H., and Feng, D. D. (2003). Fundamentals of content-based image retrieval. In *Multimedia information retrieval and management* (pp. 1-26). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-05300-3_1
- [5] Rui, Y., Huang, T. S., and Chang, S. F. (1999). Image retrieval: Current techniques, promising directions, and open issues. *Journal of visual communication and image representation*, 10(1), 39-62. <https://doi.org/10.1006/jvci.1999.0413>
- [6] Lin, C. H., Huang, D. C., Chan, Y. K., Chen, K. H., and Chang, Y. J. (2011). Fast color-spatial feature based image retrieval methods. *Expert Systems with Applications*, 38(9), 11412-11420. <https://doi.org/10.1016/j.eswa.2011.03.014>
- [7] Swain, M. J., and Ballard, D. H. (1991). Color indexing. *International journal of computer vision*, 7(1), 11-32. <https://doi.org/10.1007/BF00130487>
- [8] Stricker, M. A., and Orengo, M. (1995, March). Similarity of color images. In *Storage and retrieval for image and video databases III* (Vol. 2420, pp. 381-392). International Society for Optics and Photonics. <https://doi.org/10.1117/12.205308>
- [9] Tuceryan, M., and Jain, A. K. (1993). Texture analysis. In *Handbook of pattern recognition and computer vision* (pp. 235-276). https://doi.org/10.1142/9789814343138_0010

- [10] Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7), 971-987. <https://doi.org/10.1109/TPAMI.2002.1017623>
- [11] Oliva, A., and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3), 145-175. <https://doi.org/10.1023/A:1011139631724>
- [12] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [13] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359. <https://doi.org/10.1016/j.cviu.2007.09.014>
- [14] Raza, A., Dawood, H., Dawood, H., Shabbir, S., Mehboob, R., and Banjar, A. (2018). Correlated primary visual texton histogram features for content base image retrieval. *IEEE Access*, 6, 46595-46616. <https://doi.org/10.1109/ACCESS.2018.2866091>
- [15] Yuvaraj, D., Sivaram, M., Karthikeyan, B., and Abdulazeez, J. (2019). Shape, Color and Texture Based CBIR System Using Fuzzy Logic Classifier. *CMC-COMPUTERS MATERIALS and CONTINUA*, 59(3), 729-739. <https://doi.org/10.32604/cmc.2019.05945>
- [16] Liu, G. H., and Yang, J. Y. (2013). Content-based image retrieval using color difference histogram. *Pattern recognition*, 46(1), 188-198. <https://doi.org/10.1016/j.patcog.2012.06.001>
- [17] Varish, N., Pradhan, J., and Pal, A. K. (2017). Image retrieval based on non-uniform bins of color histogram and dual tree complex wavelet transform. *Multimedia Tools and Applications*, 76(14), 15885-15921. <https://doi.org/10.1007/s11042-016-3882-4>
- [18] Ashraf, R., Ahmed, M., Jabbar, S., Khalid, S., Ahmad, A., Din, S., and Jeon, G. (2018). Content based image retrieval by using color descriptor and discrete wavelet transform. *Journal of medical systems*, 42(3), 44. <https://doi.org/10.1007/s10916-017-0880-7>
- [19] Heckbert, P. (1982). Color image quantization for frame buffer display. *ACM Siggraph Computer Graphics*, 16(3), 297-307. <https://doi.org/10.1145/965145.801294>
- [20] Gervautz, M., and Purgathofer, W. (1988). A simple method for color quantization: Octree quantization. In *New trends in computer graphics* (pp. 219-231). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-83492-9_20
- [21] Celebi, M. E. (2011). Improving the performance of k-means for color quantization. *Image and Vision Computing*, 29(4), 260-271. <https://doi.org/10.1016/j.imavis.2010.10.002>
- [22] Schaefer, G., and Zhou, H. (2009). Fuzzy clustering for colour reduction in images. *Telecommunication Systems*, 40(1-2), 17. <https://doi.org/10.1007/s11235-008-9143-8>

- [23] Wang, J. Z., Li, J., and Wiederhold, G. (2001). SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on pattern analysis and machine intelligence*, 23(9), 947-963. <https://doi.org/10.1109/34.955109>
- [24] Huang, P. W., and Dai, S. K. (2003). Image retrieval by texture similarity. *Pattern recognition*, 36(3), 665-679. [https://doi.org/10.1016/S0031-3203\(02\)00083-3](https://doi.org/10.1016/S0031-3203(02)00083-3)
- [25] Li, C., Huang, Y., and Zhu, L. (2017). Color texture image retrieval based on Gaussian copula models of Gabor wavelets. *Pattern Recognition*, 64, 118-129. <https://doi.org/10.1016/j.patcog.2016.10.030>
- [26] Li, W., Pan, H., Li, P., Xie, X., and Zhang, Z. (2017). A medical image retrieval method based on texture block coding tree. *Signal Processing: Image Communication*, 59, 131-139. <https://doi.org/10.1016/j.image.2017.06.013>
- [27] Bhunia, A. K., Bhattacharyya, A., Banerjee, P., Roy, P. P., and Murala, S. (2019). A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern. *Pattern Analysis and Applications*, 1-21. <https://doi.org/10.1007/s10044-019-00827-x>
- [28] Liu, P., Guo, J. M., Chamnongthai, K., and Prasetyo, H. (2017). Fusion of color histogram and LBP-based features for texture image retrieval and classification. *Information Sciences*, 390, 95-111. <https://doi.org/10.1016/j.ins.2017.01.025>
- [29] Pham, M. T., Mercier, G., and Bombrun, L. (2017). Color texture image retrieval based on local extrema features and riemannian distance. *Journal of Imaging*, 3(4), 43. <https://doi.org/10.3390/jimaging3040043>
- [30] Liu, P., Guo, J. M., Wu, C. Y., and Cai, D. (2017). Fusion of deep learning and compressed domain features for content-based image retrieval. *IEEE Transactions on Image Processing*, 26(12), 5706-5717. <https://doi.org/10.1109/TIP.2017.2736343>
- [31] Pradhan, J., Pal, A. K., and Banka, H. (2019). Principal texture direction based block level image reordering and use of color edge features for application of object based image retrieval. *Multimedia Tools and Applications*, 78(2), 1685-1717. <https://doi.org/10.1007/s11042-018-6246-4>
- [32] Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1), 51-59. [https://doi.org/10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4)
- [33] Maenpaa, T., Pietikainen, M., and Viertola, J. (2002, August). Separating color and pattern information for color texture discrimination. In *Object recognition supported by user interaction for service robots* (Vol. 1, pp. 668-671). IEEE.
- [34] Murala, S., Maheshwari, R. P., and Balasubramanian, R. (2012). Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE transactions on image processing*, 21(5), 2874-2886. <https://doi.org/10.1109/TIP.2012.2188809>

- [35] Zhu, C., Bichot, C. E., and Chen, L. (2013). Image region description using orthogonal combination of local binary patterns enhanced with color information. *Pattern Recognition*, 46(7), 1949-1963.
<https://doi.org/10.1016/j.patcog.2013.01.003>
- [36] Dubey, S. R., Singh, S. K., and Singh, R. K. (2015). Local diagonal extrema pattern: a new and efficient feature descriptor for CT image retrieval. *IEEE Signal Processing Letters*, 22(9), 1215-1219.
<https://doi.org/10.1109/LSP.2015.2392623>
- [37] Verma, M., and Raman, B. (2016). Local tri-directional patterns: A new texture feature descriptor for image retrieval. *Digital Signal Processing*, 51, 62-72.
<https://doi.org/10.1016/j.dsp.2016.02.002>
- [38] Chakraborty, S., Singh, S. K., and Chakraborty, P. (2017). Local directional gradient pattern: a local descriptor for face recognition. *Multimedia Tools and Applications*, 76(1), 1201-1216.
<https://doi.org/10.1007/s11042-015-3111-6>
- [39] Alzu'bi, A., Amira, A., and Ramzan, N. (2017). Content-based image retrieval with compact deep convolutional features. *Neurocomputing*, 249, 95-105.
<https://doi.org/10.1016/j.neucom.2017.03.072>
- [40] Xu, H., Huang, C., and Wang, D. (2019). Enhancing semantic image retrieval with limited labeled examples via deep learning. *Knowledge-Based Systems*, 163, 252-266.
<https://doi.org/10.1016/j.knosys.2018.08.032>
- [41] Kapur, J. N., Sahoo, P. K., and Wong, A. K. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision, graphics, and image processing*, 29(3), 273-285.
[https://doi.org/10.1016/0734-189X\(85\)90125-2](https://doi.org/10.1016/0734-189X(85)90125-2)
- [42] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.
<https://doi.org/10.1109/TSMC.1979.4310076>
- [43] Demirci, R., and Okur, Ü. (2019). Renkli Görüntülerin Ortalama Tabanlı Çok Seviyeli Eşiklenmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 7(1), 664-676.
<https://doi.org/10.29130/dubited.471040>
- [44] Rahkar Farshi, T., Demirci, R., and Feizi-Derakhshi, M. R. (2018). Image clustering with optimization algorithms and color space. *Entropy*, 20(4), 296.
<https://doi.org/10.3390/e20040296>
- [45] Kılıçaslan, M., Tanyeri, U., and Demirci, R. (2020). Tekrarlı Ortalama Yardımıyla Renk İndirgeme ve Görüntü Erişimi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 8(1), 1042-1057. <https://doi.org/10.29130/dubited.643351>
- [46] Yuan, B. H., and Liu, G. H. (2020). Image retrieval based on gradient-structures histogram. *Neural Computing and Applications*, 32(15), 11717-11727.
<https://doi.org/10.1007/s00521-019-04657-0>

Figures

Figure 1

Partitioned RGB color space (a) $r=1, k=8$ (b) $r=2, k=27$

Figure 2

Bus (a) original (b) color distribution

Figure 3

Bus quantized (a) $r=1,k=8$ (b) distribution of 8 clusters (c) $r=2,k=27$ (d) distribution of 27 clusters

Figure 4

Conventional histogram of Bus

Figure 5

CBH of Bus a) 8 clusters b) 27 clusters

Figure 6

Feature extraction step

Figure 7

Structure of CFV, $r=1$

Figure 8. Architecture of proposed CBIR based on CVF

Figure 8

Architecture of proposed CBIR based on CVF

Figure 9

P-R Curves: Corel-1K with Euclidean Similarity: a) Color b) Color and Gradient

Figure 10

P-R Curve: Corel-10K with Euclidean Similarity a) Color b) Color and Gradient