Region Division for Large-scale Image Retrieval

Yunbo Rao*, Wei Liu

School of Information and Software Engineering, University of Electronic Science and Technology of China,
Chengdu, Sichuan, 610054, China
[e-mail: raoyb@uestc.edu.cn]
*Corresponding author: Yunbo Rao

Abstract

Large-scale retrieval algorithm is problem for visual analyses applications, along its research track. In this paper, we propose a high-efficiency region division-based image retrieve approaches, which fuse low-level local color histogram feature and texture feature. A novel image region division is proposed to roughly mimic the location distribution of image color and deal with the color histogram failing to describe spatial information. Furthermore, for optimizing our region division retrieval method, an image descriptor combining local color histogram and Gabor texture features with reduced feature dimensions are developed. Moreover, we propose an extended Canberra distance method for images similarity measure to increase the fault-tolerant ability of the whole large-scale image retrieval. Extensive experimental results on several benchmark image retrieval databases validate the superiority of the proposed approaches over many recently proposed color-histogram-based and texture-feature-based algorithms.

Keywords: Image retrieval, image color analysis, image texture, image similarity measurement

1. Introduction

$\mathbf{A}$s one of the main information retrieval services, the content-based image retrieval (CBIR) query has demonstrated exponential growth in the last decade. The limitation in metadata-based systems versus the various image retrieval query keeps the high efficient CBIR an active research area [1,2]. Traditional image retrieval approaches require humans to manually label each image in the target database [3]. Its has obvious disadvantages: based on manual method can not label images from internet network, such Jian's work[4].

Different from traditional content-based image retrieve approaches, text-based retrieve usually uses keywords, subject heading, captions, or natural language et al. CBIR aims at retrieving images with similar content as the query input through the analysis of image content, such as hierarchical-local-feature extraction approach [4], shape and spatial layout [5], Social-based anchor-unit graph regularized tensor completion [6]. Basically, indexing and similarity matching are two main steps for a CBIR system. Indexing is a technology to index images in the target database by their visual content descriptor and similarity is matching against the feature vector in the index by a feature descriptor extracted from the query image. If elements are non-negative, we can use Canberra distance metric. If we measure position number with corresponding symbols, hamming distance method should be used.

In CBIR systems, the precision & latency are two main evaluation criterions. The traditional image features (such as shape, color, texture, and edge orientation) can be use content-based image retrieval. However, the main challenge of image features is the 'semantic gap' problem [1]. Therefore, how to reduce this 'semantic gap' is become very import. Such as unsupervised hashing approach to facilitate large-scale semantic indexing of image data [12], global-based semantic information and the relative ordering correlation method [17]. The emphasis in different image features is different. Such as support-vector-machine-based classifier method for image retrieving [7].The color feature is mainly used to describe the color distribution of each pixel in the image, but would fail to describe the shape of objects and its spatial information [3]. The texture feature can achieve a good performance in image representation only when images have a low texture density. In addition, existing technologies of texture feature extraction can introduce unexpected noise during the extraction process which would reduce its performance [8]. The shape feature is only used to describe the image contours and boundaries. These methods have some limitations, such as: the proposed method need use more than one type features. Many of the existing CBIR studies adopt methods, in which many features are fused for image retrieval. Such as pattern distinctness and local contrast [9], color and texture feature fused method [10], proposed region-based method in [11].

In our work, local color histogram and texture feature are analyzed, and propose a high-efficiency image retrieve method, which use low-level feature to retrieve large-scale images. In summary, we highlight three contributions as follows:

- A category of image region division which divides images into five non-overlapping regions is designed for local color feature extraction. Meanwhile, those regions are utilized, a crude form of color localization distribution can be simulated.
- An image descriptor combining local color feature and Gabor texture is proposed. Moreover, the color & texture features quantization module are adopted to reduce the feature dimensions.
- An extended Canberra distance is introduced for image feature metrics, which increases fault-tolerant ability and stability of the CBIR system.

The paper is organized as follows: Section 2 describes the CBIR technology. Our proposed approach is descripted in Section 3. Section 4 shows the similarity measure of our method. In Section 5, the performances of different CBIR approaches are demonstrated. Experimental results and comparisons are presented in Section 6.

2. Related Work

Image feature extracting in CBIR is one of the main challenges for image retrieval well. The traditional low-level features extraction algorithms can be global color histogram [13,14] and local descriptor[15,16]. Global feature based algorithms extract the feature based on the whole frame. However, these approaches are not able to compare the regions division method. To resolve these issues, local feature based approaches are introduced to overcome this weakness by dividing the image into several sub-blocks to obtain the spatial distribution of pixels. These local feature based approaches can achieve a higher retrieval precision. Therefore, extracting local color features gain increasing interests in the studies of CBIR.

The color information is the most commonly used features in CBIR systems. The color feature can be described in multiple ways, such as color histogram [18,21], color structure descriptor [5,22], scalable color descriptor [19], color moment [20,23], and ROI-based color feature [24]. Color histogram-based algorithm have fast operation speed. However, using only the color histogram feature fails to model the color spatial structure and makes a high dimensional feature vector. Therefore, color histogram is often used in conjunction with color moment and other feature descriptors in the field of CBIR.

Extracting texture feature is another open challenge problem for CBIR system[3]. Many methods are proposed for texture features extraction, such as grey-level co-occurrence matrix (GCM) [25,26], Gabor filtering [27], local binary pattern (LBP) [8], markov random field model [28], co-occurrence of histogram of oriented gradients [29]. In order to improve retrieval precision, texture features and color features are fused, such as integrative co-occurrence matrix [30] , the color edge co-occurrence histogram [31], and local color contrast between the salient object and the image background [32]. Recently, many graph-based algorithms are applied, Jian et.al [33] proposed quaternionic distance method for detecting outliers in color image. This method takes into account the contrast of static target position and image color. Meanwhile, Jian et.al [34] also extended random walk algorithm to optimize the saliency map.

Additionally, similarity measure plays another important role for CIBR. In the process of CIBR, an image have large dimensionality. Usually, feature vector is sparse, i.e., some feature value may be zero. Such characteristics of image representation would be a severe challenge for similarity measure. Traditionally, some researchers used Euclidean geometry field based method for similarity metric. Similar to Euclidean distance metric, Manhattan distance [35], originated from the taxicab travel route metric, is another well-known similarity metric. However, both of the two metric is sensitive to the sample topology.

In recent years, lots of reseacher works have been done for similarity measure in image retrieval and clustering algorithm. Jarvis et al. [36] proposed that data points which are similar to the extent should share the same nearest neighbors and a similarity metric can be obtained with a series of similarity comparisons for these neighbors. Radhakrishna et al. [37] proposed a similarity measure method. In order to improve precision of similarity estimation, the proposed method use the standard score and normal probability. Lin et al. [38] proposed a

multiple case similarity measure method in which feature similarity measure are divided in three parts: the feature appears in both feature vectors, the feature appears in only one vector, and the feature appears in none of the vectors. Jian et al. [39] proposed new method, which used wavelet-based salient patch detector for image retrieval. Meanwhile, Radhakrishna et al.[40] also presented a new approach to retrieve temporal association patterns whose prevalence values are similar to those of the user specified reference.

3. Local Feature Extraction

In this section, a novel image retrieval method is proposed using local color histogram and texture feature. This method includes two stage: color histogram and Gabor texture features. In our work, color moment is used as a supplement to reduce the length of the color feature vector. In addition, an extended Canberra distance is applied as the metric to do similarity measure. As shown in **Fig. 1**, the general framework of region division-based image retrieval can be divided in two parts, building an index table and retrieving.

To get a good performance and low latency, we improve the image feature extraction algorithm from three aspects: (1) dividing each image into five fixed non-overlapping regions to stimulate the color distribution of the images. (2) A CIELAB (10, 3, 3) quantization scheme is used to extract color features. (3) The real part of the Gabor filter is adopted to extract the image texture feature. In the following subsections, all these technologies are introduced in detail respectively.
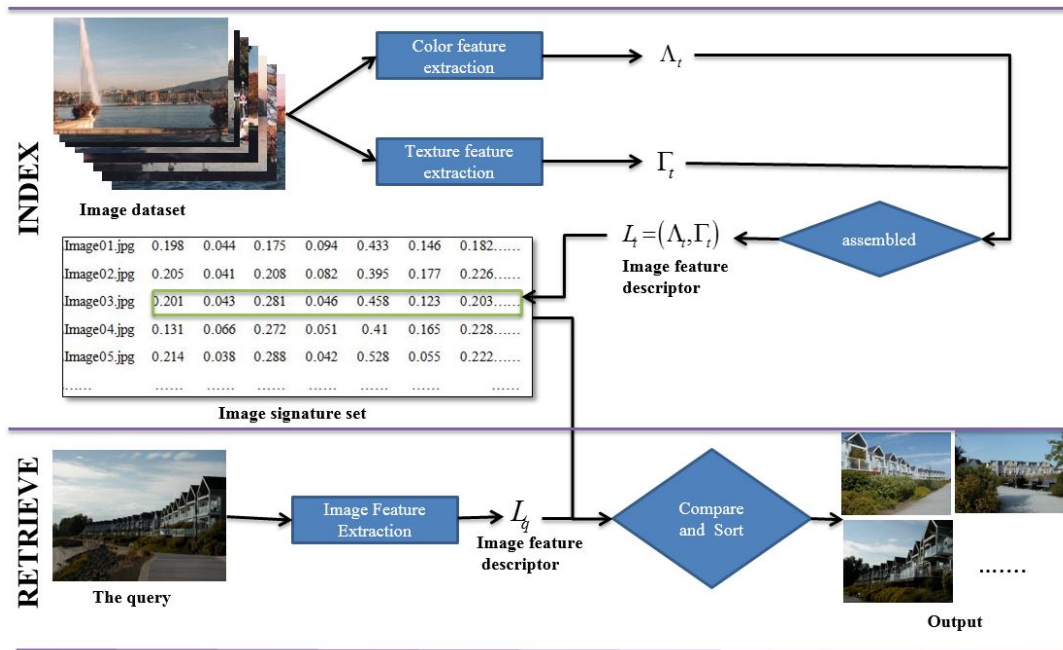

Fig. 1. General framework of content-based image retrieval

3.1 Image region division

Since the local features outperform the global ones with a better consideration of the spatial relationship of different regions, we propose to extract the local image descriptors for each region after fixed image region segmentation.

The aim of image region division is getting as much contextual related pixel together as possible. In our work, we can categorize the image region division algorithms into two major groups: image segmentation-based region division (ISBRD) and fixed region image division (FRD). Thoese two methods are different when it comes to image representation in CIBR. The ISBRD system automatically specifies the category of region division by relying image segmentation technology, such as, Shrivastava et al. [41] proposed a method to divide the image into a fixed number of blocks(including: color, texture, and shape features). However, image segmentation is not always reliable result in a reduction of retrieval accuracy. In this work, the system may not automatically designate a region correspond to the object that the users wishes to retrieve even though existing image segmentation techniques can accurately identify image regions. All these reason have limited its accuracy in CBIR systems. In contrast, the FRD approach facilitates the user to divide the image into several sub-blocks in advance, and extracts their features and matches them with the quer. Yue et al. [25] proposed a scheme method. All images will be divide into blocks. The fixed size is 3×3. However, this category of blocks division would lead to much more segmentation in ROIs (region of interest) and limit its capacity to reflect the user's intent in the process of retrieval.

Mangijaosingh et al. [10] proposed a method, which partitions an image into three equal horizontal regions. This method can approximate a regional distribution of color. **Fig. 2** show a example. From this case, we can get a proper segmentation result by adjusting the number of blocks and partition ratios. **Fig. 4** (a) show a proper segmentation result. However, from **Fig. 3**, the images have a relative complex content without an obvious color layer distribution, therefore the horizontal division makes no sense. To solve this problem, we assume that the most ROIs are concentrated in the middle, in fact it is true in most cases, and some strategy of region division can be used to get as much contextual related pixels together as possible.
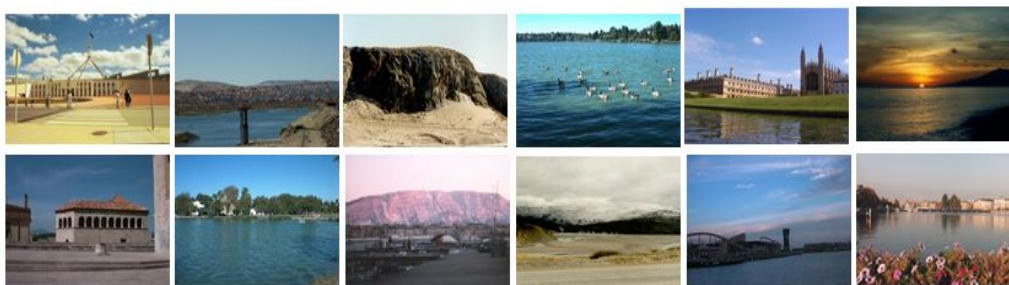

Fig. 2. Sample images of explicit color layer distribution

Fig. 3. Sample images without explicit color layer distribution

**Fig. 4.** Two categories of image region division

Applying this strategy, our approach partitions the image into five non-overlapping regions as shown in **Fig. 4** (b): (1) the top-left corner, (2) the bottom-left corner, (3) the bottom-right corner, (4) the top-right corner, and finally (5) the center of the image. By partitioning images into these regions, we are able to stimulate a coarse localization of the objects in the images, being able to represent image color distribution.

The basic idea of our image region segmentation is to help improve the accuracy of recognizing the major object that people intend to retrieval. In the most cases, the center region is believed to have more foreground information according to some basic rule of photography [42]. **Fig. 5** show an example of foreground extraction in several images with saliency object. several pictures with saliency object are used to verify our assumption by foreground segmentation [43]. The center region is able to represent the main visual saliency object by a reasonable adjustment of its size, at the same time, the rest of the four parts obtain a more accurate description of the background.

In our experiment, we use a concentric ellipse to represent the center region, the horizontal and the vertical central axis as the partition for the rest of four parts (shown in **Fig. 4** (b)). In order to let the ellipse, cover more of the major object information, its major radius and minor radius are set as following:

$$\begin{cases} x = \alpha width_{img}, \\ y = \beta height_{img}, \end{cases} \tag{1}$$

where x, y denote the major and the minor radius respectively, $width_{img}$ and $height_{img}$ represent the width and the height of the original image. α, β are percentages in scaling of the height and width of the image and set according to a specific condition. In our work, we set $\alpha = 70\%$ and $\beta = 80\%$ basing on vast experiment results.
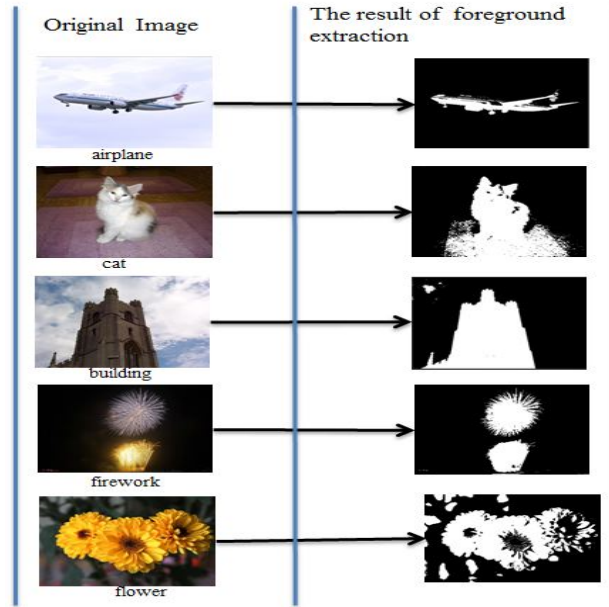
Fig. 5. Simple foreground extraction in several images with saliency object.

3.2 Color features extraction

How to extract color feature is very important for image retrieval. Usually, we extract image features based on RGB color space. However, RGB color space fails to stimulate humans' perception. In this work, we choose CIELAB color space model for color feature extraction. CIELAB color space is defined by one luminance and two chromatic channels to approximate human perception of color. The transformation between standard RGB and CIELAB can be referred to [44].

To reduce the dimensions of a color feature vector, we apply a color quantization scheme for each color vector. If a color RGB image of size of $P \times Q$, the (10, 3, 3) scheme is used to quantize the three channels. The L channel represent 10 bins, and the A and B channels represent 3 bins, respective. Therefore, $10 \times 3 \times 3 = 90$ color bins are obtained to represent the color histogram of each region.

To further reduce the dimension of a color feature vector and accelerate retrieval speed, color comments are used to describe the color histogram for the color information distribution concentrated in low order moment [13], viz. means and standard deviation. In our work, we use equation (2) to represent the mean and standard deviation (MSD) of the i-th channel in region R of an image:

$$\mu_{rj} = \frac{\sum_{j=1}^{N} I_{ij}}{N} \quad , \tag{2}$$

$$\delta_{ri} = \sqrt{\frac{\sum_{j=1}^{N} (I_{ij} - \mu_{ri})^2}{N}} \quad , \tag{3}$$

where I_{ij} represent the value of the i-th color channel in the j-th image pixel, r represent the number of divided regions, and N is the number of pixels in the region R. Therefore, the image color histogram feature vector in the r-th region is given as follows:

$$\alpha_r = \{\mu_{rl}, \sigma_{rl}, \mu_{ra}, \sigma_{ra}, \mu_{rb}, \sigma_{rb}\}, \qquad (4)$$

here, l represents the L channel, a and b represent the A and B channels, respectively. Since each image is divided into five regions and from each of the regions, two moments of color distribution are extracted from three-color channel; the color information of each image will be represented by a 30-dimension description vector. The whole image color histogram feature vector is defined by

$$\Lambda = \{\omega_1\alpha_1, \omega_2\alpha_2, \omega_3\alpha_3, \omega_4\alpha_4, \omega_5\alpha_5\}, \qquad (5)$$

here, ω is set as the weight of the color feature descriptor in each part.

Because some details are in the middle of the image, in our work, we also assigned a higher weight (1.2), as an emphasis since it plays a greater role in image similarity comparisons. Therefore, weights are allocated as $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 1$, $\omega_5 = 1.2$, empirically. The proposed color features extraction method can be concluded as algorithm 1:

Algorithm 1: proposed method of color features extraction

Input: RGB image, and convert RGB color space to CIELAB color space model for color feature extraction.

Output:

1: CIELAB color space

2: Using color quantization scheme to reduce the dimensions of each color feature vector.

3: Means and standard deviation are used to further reduce the dimension of a color feature vector and accelerate retrieval speed.

4: The whole image color histogram feature vector is defined to represent the color information of each image by a 30-dimension description vector.

3.3 Texture feature extraction

Texture is another commonly used visual feature, which can provide implicit pictorial information, like the smoothness, coarseness, etc. of objects in an image. Gabor filter is a commonly used method in texture exploration by means of signal processing in the frequency domain. In our work, we use Gabor filter to get more texture feature. 2D Gabor filter is a plane wave. The traditional method can use the sinusoidal with a Gaussian function. Its specific definition is as follows:

Complex:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda}\right) + \psi\right)\right), \qquad (6)$$

where:

$$\begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases},$$

Here, the frequency f of Gabor filter is represented as $f = 1/\lambda$.

2D Gabor filter would be cumbersome in practical computation. Therefore, most of researchers choose the real part as the filter [8][38][39].

Real:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \times \cos(2\pi \frac{x'}{\lambda} + \psi). \quad (7)$$

We find that frequency f and orientation θ are similar with Gabor filter expression [7]. We can choose different frequencies and orientations to obtain effective local texture features. In our work, repetitive pixels are as texture, where the number of patterns can be formed by a diverse pixel. In order to better represent texture features, the image is preprocessed by gray degree transformation, then we extract texture features in 4 direction, viz. 0, $\pi/4$, $\pi/2$, $3\pi/4$, and 3 frequencies, viz. 7, 11, 12, meaning 15 total filtered images are received, and each could get the required characteristics of different textures.

The objective of the texture-based method is to how to find similar texture features with high precision. To reduce the retrieval latency, we defined MSD as follows:

$$\mu_{mn} = \frac{\sum_x \sum_y | E_{mn}(x, y) |}{P \times Q}, \quad (8)$$

$$\sigma_{mn} = \sqrt{\frac{\sum_x \sum_y (| E_{mn}(x, y) | - \mu_{mn})^2}{P \times Q}}, \quad (9)$$

$$m = 0, 1, \cdots, M - 1; n = 0, 1, \cdots, N - 1 ,$$

where, $E_{mn}(x, y)$ is the texture feature of each point of the filtered image, $P \times Q$ is the original image size, M is the orientation numbers, and N represent the frequency number.

Fig. 6 gives flowchart of the process of texture feature extraction by Gabor filter. In this flowchart, 12 different Gabor filter kernels at different frequencies & orientations is used. Image texture representation vector is given as the following:

$$\Gamma = \left\{ \mu_{00}, \sigma_{00}, \mu_{10}, \sigma_{10}, \cdots, \mu_{(M-1)(N-1)}, \sigma_{(M-1)(N-1)} \right\}, \quad (10)$$
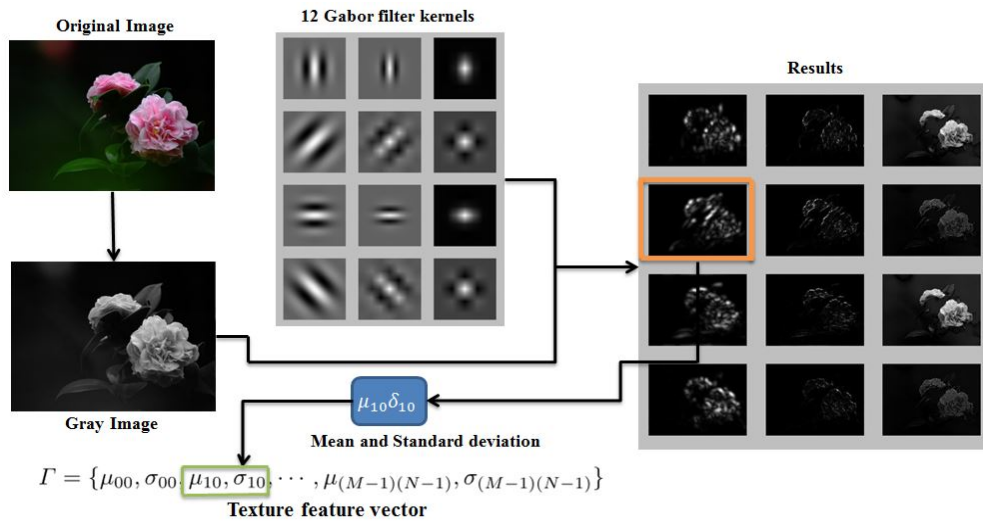
Fig. 6. Flowchart of the process of texture feature extraction by Gabor filter

4 Similarity Measure

4.1 Distance metrics

A good similarity measures is an important factor for retrieval accuracy, especially, in the case of multi-feature representation. The measurement of image content similarity remains a big challenge. Distance measurement is a commonly used method in image similarity measurement, which is defined by a certain distance between two images, such as Euclidean distance [46] and Canberra distance [45]. In this paper, we improve Canberra distance method. Canberra distance between vectors $X = [x_1, x_2, \ldots, x_p]$ and $Y = [y_1, y_2, \ldots, y_p]$ in a P-dimensional sreal vector space is defined as follows:

$$d(X, Y) = \sum_{i=1}^{P} \frac{|x_i - y_i|}{x_i + y_i} \qquad (x_i, y_i > 0), \qquad (11)$$

To avoid zero-division and achieve an equalization comparison result, a extend the Canberra distance is shown as follows:

$$d(X, Y) = \frac{1}{P} \sum_{i=1}^{P} \frac{|x_i - y_i|}{|x_i| + |y_i| + 10^{-10}}, \qquad (12)$$

Here, the image databases include two part. One is similarity measures of the query image. The other part is the target images databases. Color histogram similarity measure and texture feature similarity measure respectively. Let $d_1(\Lambda_q, \Lambda_t)$, $d_2(\Gamma_q, \Gamma_t)$ represent the distance metric of CIELAB color histogram vector and the distance metric of texture feature vector, respectively. We defined the global similarity distance metric $D(L_q, L_t)$ as follows:

$$D(I_q, I_t) = d_1(\Lambda_q, \Lambda_t) + d_2(\Gamma_q, \Gamma_t), \qquad (13)$$

Here, $I_q = (\Lambda_q, \Gamma_q)$ represent image feature descriptor of the query image, and $I_t = (\Lambda_t, \Gamma_t)$ represent the target image feature descriptor of an image in target database.

1. Cosine similarity distance

$$CX, Y = \frac{X.Y}{\|x\|.\|Y\|_{cos}}, \qquad (14)$$

where C_{cos} is used to measure the cosine.

2. Euclidean distance

$$D_{Euc}(X, Y) = [(X - Y)(X - Y)]^{\frac{1}{2}}, \qquad (15)$$

where X.Y is two vectors, and represent the inner product.

3. Manhattan distance

$$D_{man}(X, Y) = \sum_{i=1}^{P} |x_i - y_i|, \qquad (16)$$

4.2 Implement details

In our work, a ranking criterion [47] is used for evaluation. Given an input image, we first extract the image features vector as the image descriptor, which is denoted by $I_q = \{\Lambda_q, \Gamma_q\}$. Let $\Phi = \{I_1, I_2, \cdots, I_n\}$ denote the target database, which include n images. $I_i = \{\Lambda_i, \Gamma_i\}$, $I_i \in \Phi$ is image descriptor. We will assign rank for the target image by the similarity measure. We output the measure result by,

$$H(I_q, I_i) = \begin{cases} 1 & D(I_q, I_i) < sh \\ 0 & otherwise \end{cases}, \qquad (17)$$

where $H(I_q, I_i)$ denotes the classifier of similarity. If value is 1, it represents the query image. If value is 0, it represents the t-th image with a same label. Therefore, we can identify a pool of

k candidates, P={I^c_1,I^c_2,…,I^c_k}, if the similarity distance between I_q and I_i $(I_i \in \Phi)$ is lower than a threshold *sh* by which the output number of retrievals can be controlled. Plainly, the smaller Canberra distance equals to a higher similarity of two images.

4.3 Performance metric

In this paper, we evaluate the retrieved top k image with respect to a query image I_q by its precision and recall shown in the ranking-based criterion [44]. The precision and recall are defined as follows:

$$Precision @ k = \frac{\sum_{i=1}^{k} Re\, l(i)}{k} \quad, \tag{18}$$

$$Re\, c\, all@k = \frac{\sum_{i=1}^{k} Re\, l(i)}{Num} \quad, \tag{19}$$

Here, *Rel(i)* represent the ground truth relevance. Num is the number of related images in the target database. We consider the image label as the only category to measure the relevance of two images, where $Re\, l(i) \in \{0,1\}$.

5 Experiment Result and Analysis

In experiment part, we use three image data–sets, Ground Truth datasets [47], INRIA Holiday dataset [48], and CIFAR-10 dataset [49]. In order to get a more realistic experimental result, each category is selected 20 images as query images.

To justify how the distance metrics affect the retrieval result, several distance metrics are used in our experiments based on the same setup. Additionally, we have compared our proposed approach with several commonly used techniques, these methods are giving in [15, 21, 25], respectively.
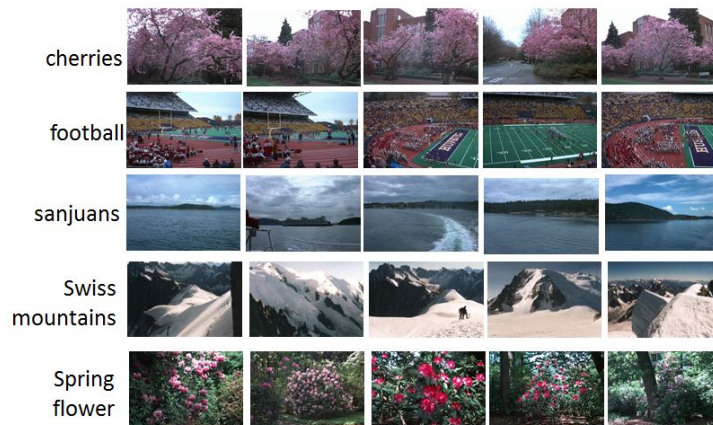
5.1 Image data-sets

**Fig. 7.** Sample images in Ground Truth data-sets

Ground Truth Dataset [47] is composed of more than 1300 images which are generally organized by a particular scene, such as animals, natural scenes, and people doing activities, and so on. In our experiment, 5 categories of classes are chosen as the query shown in **Fig. 7**,

there are Cherries, Football, Sanjuans, Swiss moyntans, and spring flower.

Fig. 8. Sample images in INRIA Holiday dataset

The INRIA Holidays data-sets [48] is one of a standard bench-marking data-sets for CBIR. This data-sets include 1491 images. **Fig. 8** show a example image of data-sets, such as:sea, building, fire effects, waterm and botany, etc.
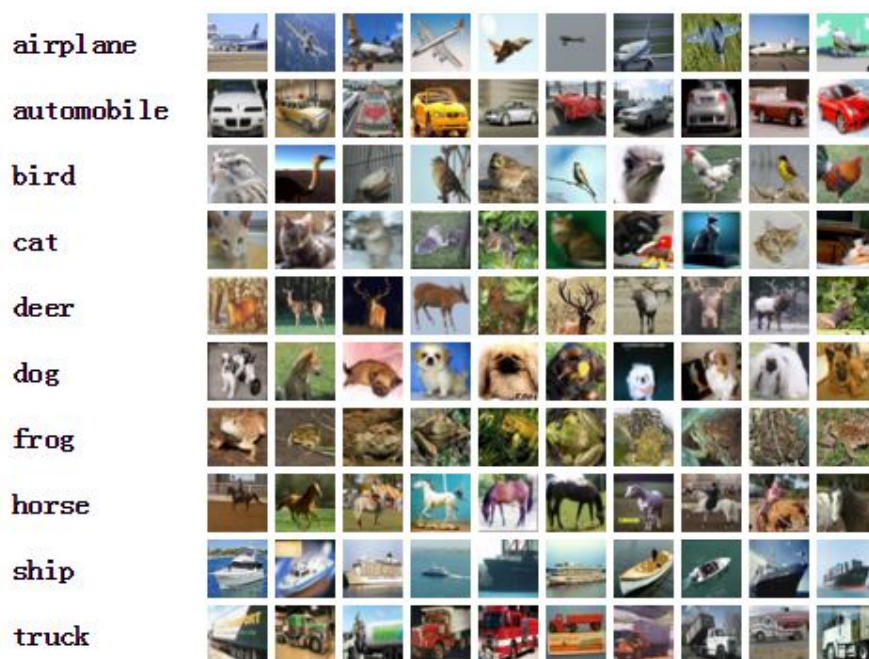

Fig. 9. Sample images in CIFAR-10 dataset

CIFAR-10 data-sets [44] .
This data-set include about 60 thousand images, and total 10 classes. The image size is 32*32. **Fig. 9** shows sample images in CIFAR-10 dataset. In our experiment, we use five training batches about 1000 images. Meanwhile it include one test batch for ouw experiment. All sample images include dog, ship, truck, frog, deer, cat, bird, airplane, automobile et al.
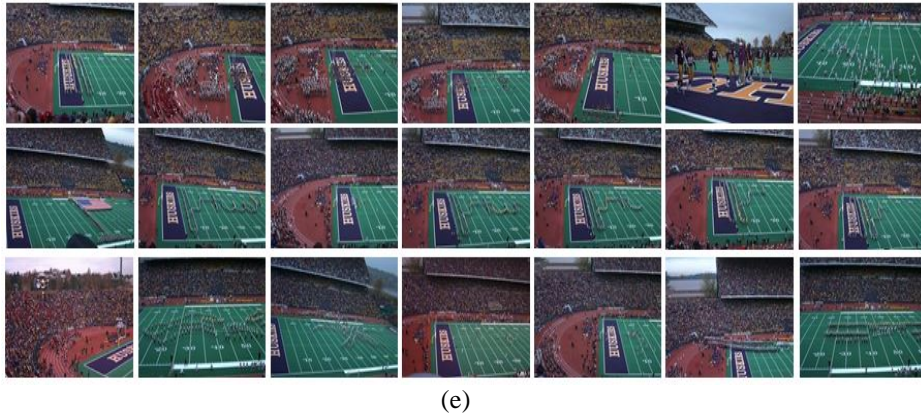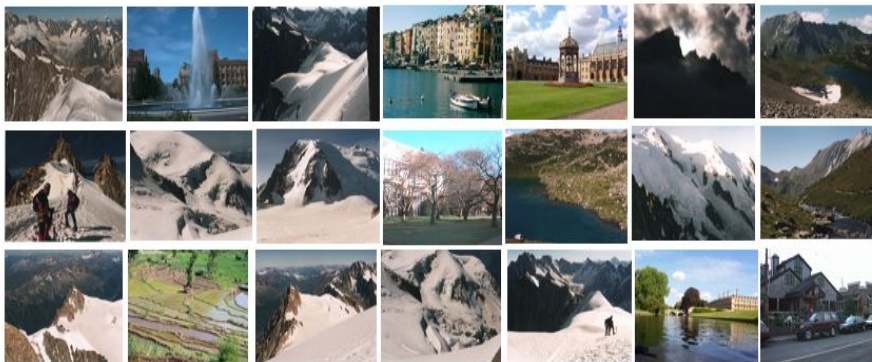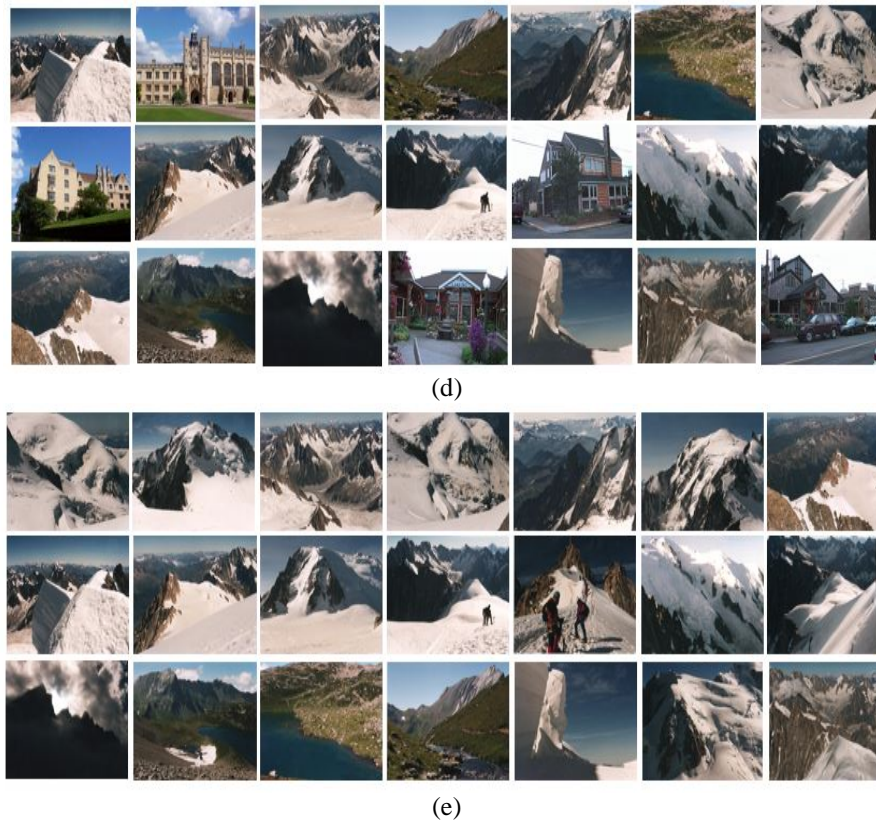
(a)

(b)

(c)

(d)

(e)
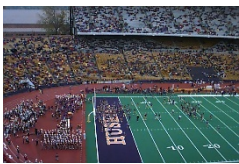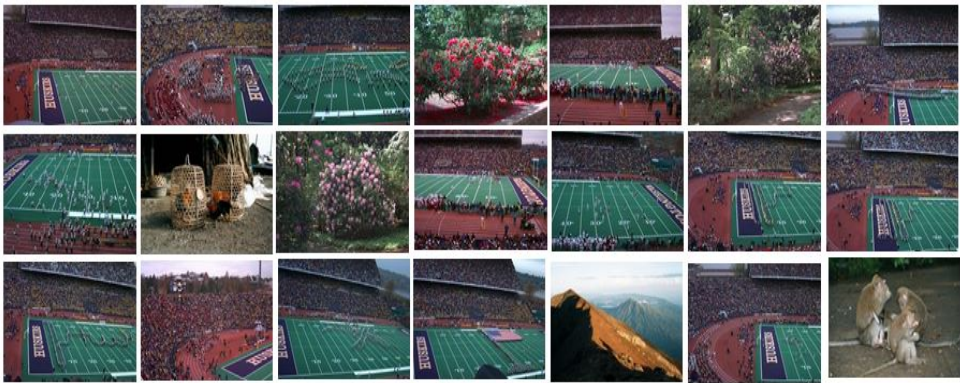
Fig. 10. Image retrieval results for football field image using different schemes. ((a) the query, (b) Huang Z.C. et al.'s method [15], (c) Singha M. et al.'s method [21], (d) J. Yue et al.'s method [25], and (e) our method.

(a)

(b)

(c)

(d)

(e)

Fig. 11. The image retrieval results for Snow berg using different schemes. (a) the query, (b) Huang Z.C. et al's method [15], (c) Singha M. et al's method [21], (d) Yue J. et al's method [25], and (e) our proposed method.

Fig. 12. Experimental results use our method on INRIA Holiday data-sets. The first image is as the query image. The top 11 image retrieval results are returned by our region division-based image retrieve method.

Fig. 13. Large-scale image retrieval by our proposed method on CIFAR-10 data-sets. The first image is as the query image. The top 29 image retrieval results are returned by our region division-based image retrieve method.

5.2 Influences of the Distance Metric

In this subsection, we investigate the influences of the distance metric. By following the experimental settings of our proposed method, we evaluated the influences of the five different distance metrics mentioned in Section 4.1 using retrieval precision comparison.

Table 1. Exeperimental results of our region division-based image retrieve method using several similarity metrics in the Ground Truth dataset.

| Category	Pricision(%) k=40				
	Cosine	Eudclidean	Manhattan	Canberra	our proposed
Green Land	85.3	84.7	78.6	86.0	89.1
Cherries	82.3	85.2	85.4	90.2	90.5
Football	85.5	87.5	73.7	88.4	94.5
Seaboard	84.2	86.1	75.2	84.7	90.5
Snow Mountains	85.2	88.3	81.3	90.5	89.6
Spring Flower	80.5	79.5	79.5	84.2	87.0
Iran	81.8	84.5	75.4	91.4	93.7
Average	83.5	85.1	78.4	87.9	90.7

Table 2. Exeperimental results of our region division-based image retrieve method using several similarity metrics in INRIA Holiday dataset

Category	Pricision(%) k=20				
	Cosine	Eudclidean	Manhattan	Canberra	our proposed
Building	85.0	85.5	88.8	88.7	90.0
Pyramid	88.7	93.2	84.5	90.7	93.4
Coral	77.1	73.1	64.2	81.7	85.1
Sailing	85.2	88.3	78.4	88.7	90.0
Lake	87.4	83.1	83.0	87.5	89.2
Venice	82.4	75.1	70.1	72.2	75.2
Average	84.3	83.1	78.1	84.9	87.2

Table 3. Exeperimental results of our region division-based image retrieve method using several similarity metrics in CIFAR-10 dataset

| Category	Pricision(%) k=80				
	Cosine	Eudclidean	Manhattan	Canberra	our proposed
Airplane	60.1	55.4	57.4	58.7	58.3
Automobile	78.4	83.2	84.5	80.4	83.4
Bird	57.1	63.1	54.2	61.7	70.1
Cat	80.8	82.3	78.4	85.6	87.2
Deer	84.9	88.1	82.0	88.1	88.0
Dog	85.4	78.3	74.1	83.2	85.3
Frog	87.0	90.5	82.1	93.5	94.7
Horse	82.1	84.3	80.4	86.0	90.7
Ship	75.3	77.6	64.2	78.4	77.6
Truck	70.2	76.5	62.5	71.5	73.4
Average	76.1	76.5	71.9	78.7	80.4

(a)

(b)

(c)
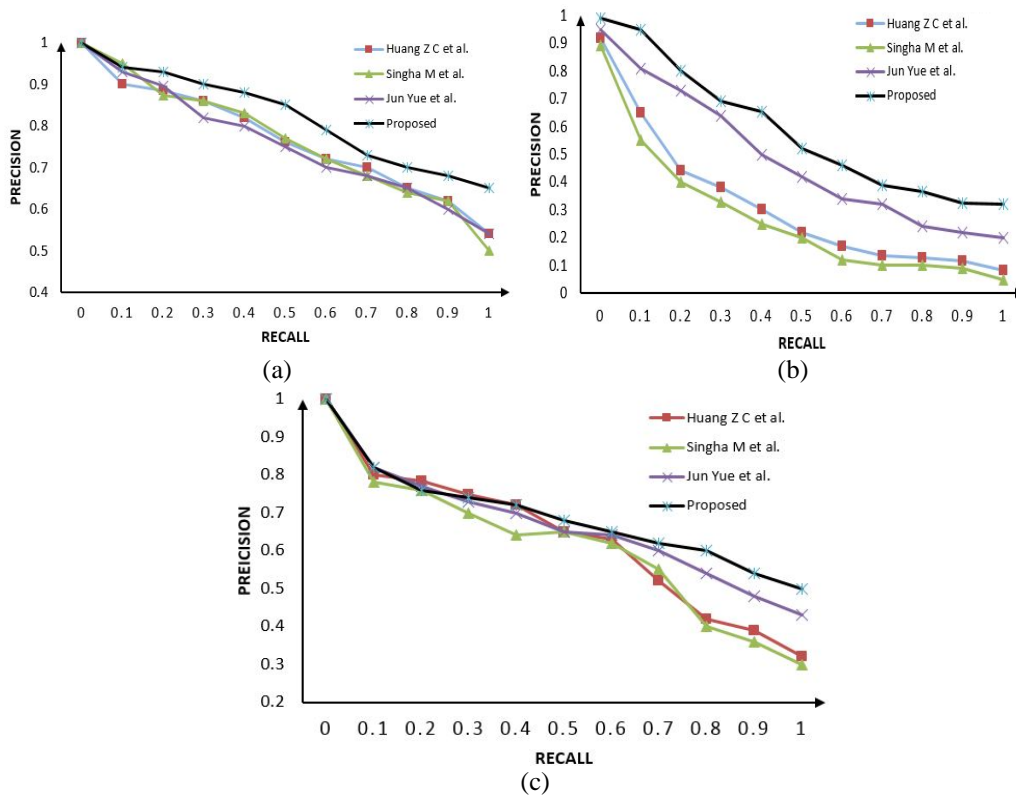
Fig. 14. Average Precision-Recall graph in the three image data-sets. (a) Ground Truth data-set, (b) INRIA Holiday data-set, and (c) CIFAR-10 data-set. We campare different method,including Huang Z.C. et al.'s method [15], Singha M. et al.'s method [21], Yue J. et al.'s method [25], and our region division-based image retrieve method.

Table 1, **Table 2** and **Table 3** illustrate the performance of our region division-based image retrieve method in the three experimental datasets(Ground Truth datasets, INRIA Holiday dataset, and CIFAR-10 dataset), respectively. From **Table 1-3**, the results of image retrieval will be influenced by distance metrics. The dimensionality of an image descriptor is large and its resulting feature vector is sparse, campare similarity metrics (Cosine, Eudclidean, Manhattan, Canberra), our proposed method achieves the best average performance.

Figs. 12 and 13 show two retrieval examples on the INRIA Holiday dataset and the CIFAR-10 dataset, respectively. In **Fig. 12**, the query is a pyramid image and top 11 matched images are returned by the system. In **Fig. 13**, the query is an air plane image and top 29 matched images are returned. Retrieval examples in the Ground Truth dataset are shown in **Fig. 10(e) and 11(e)**. Both of the two examples show good matches using the proposed color and texture features.

5.3 Experimental comparison

In order to performance our region division-based image retrieve method well, we also comparison with related works of CBIR, such as [15][21][25]. Singha et al. [21] propsoed new color imge feature fused method with discrete wavelet transformation. This method also use color histogram, and get accuracy of image retrieval well. Huang et al. [15] proposed new image retrieval method by applied color histogram and Gabor filter for image feature descriptors. Both of those two methods are based on global feature extraction. The method proposed by Yue et al. [25] is a region based approach, which uses blocks of a fixed size (3×3) to quantize the color image and a co-occurrence matrix for texture feature representation.
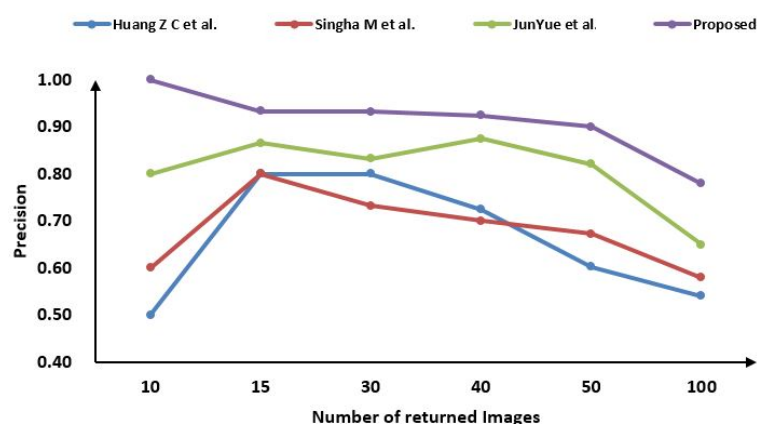
Fig. 15. Average retrieval precision based on the four schemes methods in CIFAR-10

Examples of the football field image and snow mountain image retrieval based on the above four schemes in the Ground Truth dataset is shown in **Fig. 10** and **Fig. 11,** respectively. **Figs. 10(a)-(c)** show the $Precision/Recall$ curve in the performance comparison of our region division-based image retrieve method on three techniques average image retrieval accuracy on the three experimental image data-sets. Clearly, a larger value of Recall will lead to a relative smaller Precision. **Fig. 14** shows a examples of average precision-recall graph. Our region division-based image retrieve method achieved the best performance.

In the process of CBIR, the number of results (k) returned from the system is an important influencing factoring in the retrieval precision. In order to investigate the influence of k on the retrieval accuracy, we varied the value of k and calculated the average retrieval precision in CIFAR-10 dataset. As shown in **Fig. 15**, if number of returned images is a small, it represents a relatively higher retrieval precision and our proposed method is competitive among the four schemes.

6 Conclusion

We propose a region division-based image retrieve method, which use local color histogram and Gabor texture feature. The proposed method is generalized without any model training and optimization by machine learning. Therefore, its can be considered as an improved low-level CBIR method. In the proposed algorithm, images in the target database were divided into five non-overlapping regions for color feature extraction. Weights were assigned to different regions according to the imbalanced information content distribution in each region to enhance the systems retrieval performance. In addition, an improved distance formula is present for image feature metrics. Experimental results on three image databases illustrate our proposed method having a relative higher retrieval performance than other commonly used methods using only color and texture features. The feedback technology can be introduced to improve retrieval efficiency in the future.

References

[1] P. Sandhaus, S. Boll, "Semantic analysis and retrieval in personal and social photo collections," *Multimedia Tools and Applications*, vol.51, no.1, pp.5-33, 2011. Article (CrossRef Link)

[2] Y, Yang, F. Shen, H. T. Shen, H. X. Li, X. L. Li, "Robust discrete code modeling for supervised . hashing," *Pattern Recognition*, vol.75, pp.128-135, 2018. Article (CrossRef Link)

[3] Y Liu, D Zhang, G Lu, W Y Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recogn*, vol.40, no.1, pp.262-282, 2007. Article (CrossRef Link)

[4] M.W Jian, Y.L Yin, J.Y Dong ,K.M Lam, "Content-based image retrieval via a hierarchical-local-feature extraction scheme," *Multimed Tools Appl.*, vol.77, pp.29099-29117, 2018. Article (CrossRef Link)

[5] H. Yu, W. Yang, G. S. Xia, G. Liu, "A color-texture-structure descriptor for high-resolution satellite image classification," *Remote Sensing*, vol.8, no.3, pp.2072-4292, 2016. Article (CrossRef Link)

[6] J.H Tang, X.B Shu, Z.C Li, Y.G Jiang, and Q.Tian, "Social anchor-unit graph regularized tensor completion for large-scale image retagging," *IEEE Transactions on Pattern Analysis and Macine Intelligence*, arXiv:1804.04397v2 [cs.CV] 3 Oct 2018. Article (CrossRef Link)

[7] Y.B Rao, W. Liu, B.J F an, J.L Song, Y.Yang, "A novel relevance feedback method for CBIR," *World Wide Web*, Jan. 2018. Article (CrossRef Link)

[8] F. Bianconi, E. Gonzalez, A. Fernadez, "Dominant local binary patterns for texture classification: Labelled or unlabelled," *Pattern Recogn Letters*, vol.65, no.1, pp.8-14, 2015. Article (CrossRef Link)

[9] M.W Jian, Q. Qi, J.Y Dong, Y.L Yin, K.M Lam, "Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection," *Journal of Visual Communication and Image Representation*, Vol.53, pp.31-41, 2018. Article (CrossRef Link)

[10] S. Mangijaosingh, K. Hemachandran, "Content based image retrieval based on the integration of color histogram, Color Moment and Gabor Texture," *International Journal of Computer Applications*, vol.59, no.17, pp.13-22, 2012. Article (CrossRef Link)

[11] G. Prasad, K. K. Biswas, S. K. Gupta, "Region-based image retrieval using integrated color, shape, and location index," *Computer Vision& Image Understanding*, vol.94, pp.193-233, 2004. Article (CrossRef Link)

[12] Y. Yang, F. Shen, H. T. Shen, H. X. Li, X. L. Li, "Robust discrete spectral hashing for large-scale image semantic indexing," *IEEE Transactions on Big Data*, vol.1, no.4, pp.162-171, 2015. Article (CrossRef Link)

[13] N. Fierro-Radilla, M. Nakano-Miyatake, H. Perez-Meana, M. Cedillo-Hernandez, F. Garcia-Ugalde, "An efficient color descriptor based on global and local color features for image retrieval," in *Proc. of 2013 10th Int. Conf. on Elec. Eng., Computing Science and Auto. Control (CCE) IEEE*, Mexico City, Mexico, pp.233-237, 2013. Article (CrossRef Link)

[14] M. Nor, J. M. Ogier, F. Manani, M. Z. M. JenuColor, "Color based properties query for CBIR: HSV global color histogram," in *Proc. of International Conf. on Graphic and Image Processing (ICGIP)*, Cairo, Egypt, p.82-85, 2011. Article (CrossRef Link)

[15] Z. C. Huang, P. K. Chan, W. W. Y. Ng, D. S. Yeung, "Content-based image retrieval using color moment and Gabor texture feature," in *Proc. of the ninth Inter. Conf. on Machine Le. and Cyb. (ICMLC)*, IEEE, Qingdao, Shandong, China, pp.1105-1109, 2010. Article (CrossRef Link)

[16] N. Shrivastava, V. Tyagi, "Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching," *Information Sciences*, vol.259, no.3, pp.212-224, 2014. Article (CrossRef Link)

[17] L Jin , X.B Shu , K.Li, Z.C Li , G.J Qi, and J.H Tang, "Deep ordinal hashing with spatial attention," *IEEE Transactions on Image Processing*, Vol.28, No.5, pp.2173-2186, MAY, 2019. Article (CrossRef Link)

[18] R. Brunelli, O. Mich, "Histograms analysis for image retrieval," *Pattern Recogn.*, vol.34 no.8, pp.1625-1637, 2001. Article (CrossRef Link)

[19] Y. G. Wen, S. Z. Peng, "Research on image retrieval based on scalable color descriptor of MPEG-7," *Advances in Control and Communication*, vol.137, pp.91-98, 2012. Article (CrossRef Link)

[20] T. Weng, Y. Yuan, L. Shen, Y. Zhao, "Clothing image retrieval using color moment," in *Proc. of International Conference on Computer Science & Network Technology*, pp.1016-1020, 2014. Article (CrossRef Link)

[21] M. Singha, K. Hemachandran, "Content based image retrieval using color and texture," *Signal & Image Processing*, vol.3, no.1, pp.271-273, 2012. Article (CrossRef Link)

[22] M. Q. Hu, Y, Yang, F. Shen, L. M. Zhang, H. T. Shen, X. L. Li, "Robust web image annotation via exploring multi-facet and structural knowledge," *IEEE Transactions on Image Processing*, vol.26, no.10, pp.4871-4884, 2017. Article (CrossRef Link)

[23] S. R. Dubey, S. K. Singh, R. K. Singh, "A multi-channel based illumination compensation mechanism for brightness invariant image retrieval," *Multimedia Tools and Applications*, vol.74, no.24, pp.11223-11253, 2015. Article (CrossRef Link)

[24] Y. K. Chan, Y. A. Ho, Y. T. Liu, R. C. Chen, "A ROI image retrieval method based on CVAAO," *Image & Vision Computing*, vol.26,11, pp.1540-1549, 2008. Article (CrossRef Link)

[25] J. Yue, Z. Li, L. Liu, Z. T. Fu, "Content-based image retrieval using color and texture fused features," *Mathematical & Computer Modeling*, vol.54, no.34, pp.1121-1127, 2011. Article (CrossRef Link)

[26] N. Severoglu, "Mammogram images classification using gray level co-occurrence matrices," in *Proc. of 2016 IEEE conference on Signal Processing & Communication Application Conf.* (SIU), Hong Kong, China, 2016. Article (CrossRef Link)

[27] B.S. Manjunath, W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.18, no.8, pp.837-842, 1996. Article (CrossRef Link)

[28] P Ahmadvand Kabiri, "Multispectral MRI image segmentation using markov random field model," *Signal Image and Video Processing*, vol.10, no.2, pp.251-258, 2016. Article (CrossRef Link)

[29] M. Q. Hu, Y. Yang, F. M. Shen, N. Xie, H. T. Shen, "Hashing with angular reconstructive embeddings," *IEEE Transactions on Image Processing*, vol.27, no.2, pp.545-555, 2018. Article (CrossRef Link)

[30] S. Beura, B. Majhi, R. Dash, "Mammogram classification using two-dimensional discrete wavelet transforms and gray-level co-occurrence matrix for detection of breast cancer," *Neurocomputing*, vol.154, no.1, pp.1-14, 2014. Article (CrossRef Link)

[31] S. X. Tian, U. Bhattacharya, S. J. Lu, C. L. Tan, "Multilingual scene character recognition with co-occurrence of histogram of oriented gradients," *Pattern Recogn.*, vol.51,no.1, pp.125-134, 2016. Article (CrossRef Link)

[32] M.W Jian, W.Y Zhang, H Yu, C.R Cui, X.S Nie, H.X Zhang, Y.L Yin, "Saliency detection based on directional patches extraction and principal local color contrast," Journal of Visual Communication and Image Representation, Vol.57, pp.1-11, 2018. Article (CrossRef Link)

[33] M.W Jian, Q. Qi, J.Y Dong, X.S Sun, Y.J Sun, K.M Lam, Saliency detection using quaternionic distance-based weber local descriptor and level priors, Multimed Tools Appl, vol.77, pp.14343-14360, 2018. Article (CrossRef Link)

[34] M.W Jian, R.X Zhao, X. Sun, H.J Luo, W.Y Zhang, H.X Zhang, J.Y Dong, Y.L Yin, K.M Lam, "Saliency detection based on background seeds by object proposals and extended random walk," *Journal of Visual Communication and Image Representation*, vol.57, pp.202-211, 2018. Article (CrossRef Link)

[35] G.D. Guo, A.K. Jain, W. Y. Ma, H. J. Zhang, "Learning similarity measure for natural image retrieval with relevance feedback," *IEEE Transactions on Neural Networks*, vol.13, no.4, pp.811-820, 2002. Article (CrossRef Link)

[36] R. A. Jarvis, E. A. Patrick, "Clustering using a similarity measure based on shared near neighbors," IEEE Transactions on Computers, vol.22, no.11, pp.1025-1034, 2006. Article (CrossRef Link)

[37] V. Radhakrishna, P. V. Kumar, V. Janaki, "SRIHASS-a similarity measure for discovery of hidden time profiled temporal associations," *Multimed Tools Appl.*, vol.6, no.1, pp.1-50, 2017.

[38] Y. S. Lin, J. Y. Jiang, S. J. Lee, "A similarity measure for text classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol.26, no.7, pp.1575-1590, 2014. Article (CrossRef Link)

[39] M.W Jian, K.M Lam, J.Y Dong, and L.L Shen, "Visual-patch-attention-aware saliency detection," *IEEE Transactions on Cybernetics*, vol. 45, no. 8, pp.1575-1586, August 2015. Article (CrossRef Link)

[40] V. Radhakrishna, S. A. Aljawarneh, P. V. Kumar, K. R. Choo, "A novel fuzzy gaussian-based dissimilarity measure for discovering similarity temporal association patterns," *Soft Computing*, pp.1-17, 2016. Article (CrossRef Link)

[41] N. Shrivastava, V. Tyagi, "An integrated approach for image retrieval using local binary pattern, Multimedia Tools and Applications," vol.75, no.11, pp.6569-6583, 2016. Article (CrossRef Link)

[42] S. Dhar, V. Ordonez, T.L. Berg, "High level describable attributes for predicting aesthetics and interestingness," in *Proc. of 2011 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Colorado Springs, Colorado, USA, pp.1657-1664, 2011. Article (CrossRef Link)

[43] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man.*, Cyber, Vol.9, No.1, pp.62-66, 1979. Article (CrossRef Link)

[44] G. H. Liu, J. Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recogn.*, Vol.46, No.1, pp.188-198, 2012. Article (CrossRef Link)

[45] M. Tsai, C. P. Lin, K. T. Huang, "Defect detection in colored texture surfaces using Gabor filters," *The Imaging Science Journal*, vol.53, No.1, pp.27-37, 2015. Article (CrossRef Link)

[46] Q. Gao, F. Gao, H. Zhang, X. Wang, "Two-dimensional maximum local variation based on image Euclidean distance for face recognition," *IEEE Trans. on Image Processing Biological Cybernetics*, vol.22, no.10, pp.807-3817, 2013. Article (CrossRef Link)

[47] The Ground Truth image database, http://imagedatabase.cs.washington.edu/groundtruth

[48] J. Li, J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.25, no.9, pp.1075-1088, 2003. Article (CrossRef Link)

[49] Krizhevsky, "Learning multiple layers of features from tiny images," *Computer Science Department, University of Toronto, Tech. Report*, vol. 61, no.2, pp.103-113, April, 1989.

Yunbo Rao received his B.S. degree and M.E. degree from Sichuan Normal University and University of Electronic Science and Technology of China in 2003 and 2006, respectively, and Ph.D. degree from University of Electronic Science and Technology of China, Chengdu in 2012, both in School of Computer Science and Engineering (SCSE). He has been as a visiting scholar of Electrical Engineering of the University of Washington from Oct 2009 to Oct 2011, Seattle,USA. Since May. 2012, he joined School of Information and Software Engineering, University of Electronic Science and Technology of China. Currently, he is an associate professor at University of Electronic Science and Technology of China. His research interests include video enhancement, computer vision, and crowd animation.

Wei Liu has received the B.S. degree from the School of Computer Science and Engineering, at Chongqing Three Gorges University, China, in 2015. Now, he is a graduate student at School of Information and Software Engineering, University of Electronic and Science Technology of China (UESTC). His research interests include video enhancement, computer vision, and crowd animation.