

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Content-Based Image Feature Description and Retrieving

Nai-Chung Yang, Chung-Ming Kuo and
Wei-Han Chang

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/45841>

1. Introduction

With the growth in the number of color images, developing an efficient image retrieval system has received much attention in recent years. The first step to retrieve relevant information from image and video databases is the selection of appropriate feature representations (e.g. color, texture, shape) so that the feature attributes are both consistent in feature space and perceptually close to the user [1]. There are many CBIR systems, which adopt different low level features and similarity measure, have been proposed in the literature [2-5]. In general, perceptually similar images are not necessarily similar in terms of low-level features [6]. Hence, these content-based systems capture pre-attentive similarity rather than semantic similarity [7]. In order to achieve more efficient CBIR system, active researches are currently focused on the two complemented approaches: region-based approach [4, 8-10] and relevance feedback [6, 11-13].

Typically, the region-based approaches segment each image into several regions with homogenous visual prosperities, and enable users to rate the relevant regions for constructing a new query. In general, an incorrect segmentation may result in inaccurate representation. However, automatically extracting image objects is still a challenging issue, especially for a database containing a collection of heterogeneous images. For example, Jing et al. [8] integrate several effective relevance feedback algorithms into a region-based image retrieval system, which incorporates the properties of all the segmented regions to perform many-to-many relationships of regional similarity measure. However, some semantic information will be disregarded without considering similar regions in the same image. In another study [10], Vu et al. proposed a region-of-interest (ROI) technique which is a sampling-based ap-

proach called SamMatch for matching framework. This method can prevent incorrectly detecting the visual features.

On the other hand, the mechanism of relevance feedback is an online-learning technique that can capture the inherent subjectivity of user's perception during a retrieval session. In Power Tool [11], the user is allowed to give the relevance scores to the best matched images, and the system adjusts the weights by putting more emphasis on the specific features. Cox et al. [11] propose an alternative way to achieve CBIR that predicts the possible image targets by Bayes' rule rather than provides with segmented regions of the query image. However, the feedback information in [12] could be ignored if the most likely images and irrelevant images have similar features.

In this Chapter, a novel region-based relevance feedback system is proposed that incorporates several feature vectors. First, unsupervised texture segmentation for natural images is used to partition an image to several homogeneous regions. Then we propose an efficient dominant color descriptor (DCD) to represent the partitioned regions in image. Next, a regional similarity matrix model is introduced to rank the images. In order to attack the possible fails of segmentation and to simplify the user operations, we propose a foreground assumption to separate an image into two parts: foreground and background. The background could be regarded as the irrelevant region that confuses with the query semantics for retrieval. It should be noted that the main objectives of this approach could exclude irrelevant regions (background) from contributing to image-to-image similarity model. Furthermore, the global features extracted from entire image are used to compensate the inaccuracy due to imperfect segmentations. The details will be presented in the following Sections. Experimental results show that our framework improves the accuracy of relevance-feedback retrieval.

The Chapter is organized as follows. Section 2 describes the key observations which explain the basis of our algorithm. In Section 3, we first present a quantization scheme for extracting the representative colors from images, and then introduce a modified similarity measure for DCD. In Section 4, image segmentation and region representation based on our modified dominant color descriptor and local binary pattern are described. Then the image representation and the foreground assumption are explained in Section 5. Our integrated region-based relevance feedback strategies, which consider pseudo query image and relevant images as the relevance information, are introduced in Section 6. Experimental results and discussions of the framework are made in Section 7. Finally, a short conclusion is presented in Section 8.

2. Problem statement

The major goal in region-based relevance feedback for image retrieval is to search perceptually similar images with good accuracy in short response time. For nature image retrieval, conversational region-based relevance feedback systems use multiple features (e.g., color, shape, texture, size) and update weighting scheme. In this context, our algorithm is motivated by the following viewpoints.

1. Computational cost increases as the selected features increased. However, an algorithm with large number of features does not guarantee an improvement of retrieval performance. In theory, the retrieval performance can be enhanced by choosing more compact feature vectors.
2. The CBIR systems retrieve similar images according to the user-defined feature vectors [10]. To improve the accuracy, the region-based approaches [14, 15] segment each image into several regions, and then extract the image features, such as the dominant color, texture or shape. However, the correct detection of semantic objects involves many conditions [16] such as lighting conditions, occlusion and inaccurate segmentation. Since no automatic segmentation algorithm achieves satisfactory performance currently, segmented regions are commonly provided by the user to support the image retrieval. However, semantically correct segmentation is a strict challenge to the user, even some systems provide segmentation tools.
3. The CBIR technique helps the system to learn how to retrieve the results that users are looking for. Therefore, there is an urgent need to develop a convenient technique for region-of-interest analysis.

3. A modified dominant color descriptor

Color is one of the most widely used visual features for retrieving images from common semantic categories [12]. MPEG-7 specifies several color descriptors [17], such as dominant colors, scalable color histogram, color structure, color layout and GoF/GoP color. The human visual system captures dominant colors in images and eliminates the fine details in small areas [18]. In MPEG-7, DCD provides a compact color representation, and describes the color distribution in an image [16]. The dominant color descriptor in MPEG-7 is defined as

$$F = \{ \{c_i, p_i\}, i = 1, \dots, N \}, \quad (1)$$

where N is the total number of dominant colors in image, c_i is a 3-D dominant color vector, p_i is the percentage for each dominant color, and $\sum p_i = 1$.

In order to extract the dominant colors from an image, a color quantization algorithm has to be predetermined. A commonly used approach is the modified generalized Lloyd algorithm (GLA) [19], which is a color quantization algorithm with clusters merging. This method can simplify the large number of colors to a small number of representative colors. However, the GLA has several intrinsic problems associated with the existing algorithm as follows [20].

4. It may give different clustering results when the number of clusters is changed.
5. A correct initialization of the centroid of cluster is a crucial issue because some clusters may be empty if their initial centers lie far from the distribution of data.

6. The criterion of the GLA depends on the cluster “distance”; therefore, different initial parameters of an image may cause different clustering results.

In general, the conventional clustering algorithms are very time consuming [2, 21-24]. On the other hand, the quadratic-like measure [2, 17, 25] for dominant color descriptor in MPEG7 does not matching human perception very well, and it could cause incorrect ranks for images with similar color distribution [3, 20, 26]. In this Chapter, we adopt the linear block algorithm (LBA) [20] to extract the representative colors, and measure the perceptual similar dominant colors by the modified similarity measure.

Considering two dominant color features $F_1 = \{c_i, p_i\}, i=1, \dots, N_1\}$ and $F_2 = \{b_j, q_j\}, j=1, \dots, N_2\}$, the quadratic-like dissimilarity measure between two images F_1 and F_2 is calculated by:

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} p_i^2 + \sum_{j=1}^{N_2} q_j^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{i,j} p_i q_j \quad (2)$$

where $a_{i,j}$ is the similarity coefficient between color clusters c_i and b_j , and it is given by

$$a_{i,j} = \begin{cases} 1 - d_{i,j} / d_{\max} & d_{i,j} \leq T_d \\ 0 & d_{i,j} > T_d \end{cases} \quad (3)$$

The threshold T_d is the maximum distance used to judge whether two color clusters are similar, and $d_{i,j}$ is Euclidean distance between two color clusters c_i and b_j ; $d_{\max} = \alpha T_d$, notation α is a parameter that is set to 2.0 in this work.

The quadratic-like distance measure in Eq. (2) may incorrectly reflect the distance between two images. The improper results are mainly caused by two reasons. 1) If the number of dominant colors N_2 in target image increases, it might cause incorrect results. 2) If one dominant color can be found both in target images and query image, a high percentage q_j of the color in target image might cause improper results. In our earlier work [19], we proposed a modified distance measure that considers not only the similarity of dominant colors but also the difference of color percentages between images. The experimental results show that the measure in [20] provides better match to human perception in judging image similarity than the MPEG-7 DCD. The modified similarity measure between two images F_1 and F_2 is calculated by:

$$D^2(F_1, F_2) = 1 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{i,j} S_{i,j} \quad (4)$$

$$S_{i,j} = [1 - |p_q(i) - p_t(j)|] \times \min(p_q(i), p_t(j)), \quad (5)$$

where $p_q(i)$ and $p_t(j)$ are the percentages of the i th dominant color in query image and the j th dominant color in target image, respectively. The term in bracket, $1 - |p_q(i) - p_t(j)|$ is used to measure the difference between two colors in percentage, and the term $\min(p_q(i), p_t(j))$ is the intersection of $p_q(i)$ and $p_t(j)$ that represents the similarity between two colors in percentage. In Fig. 1, we use two real images selected from Corel as our example, where the color and percentage values are given for comparison.







Query image Q	Target image F1	Target image F2
{(33,31,33),0.794240}	{(66,41,29), 0.108795}	{(60,55,53), 0.378306}
{(184,179,180),0.20576}	{(203,47,71), 0.334035}	{(139,123,115),0.073598}
	{(207,193,59), 0.067861}	{(198,194,188),0.548096}
	{(228,98,161), 0.219045}	
	{(230,162,203),0.270264}	
		
		

Figure 1. Example images with the dominant colors and their percentage values. First row: 3-D dominant color vector c_i and the percentage p_i for each dominant color. Middle row: the original images. Bottom row: the corresponding quantized images.

In Fig. 1, we calculate this example by using the modified measure and quadratic-like measure for comparison. In order to properly reflect similarity coefficient between two color clusters, the parameter is set to 2 and $T_d = 25$ in Eq(3). Since the pair-wised distance between Q and F_1 in Fig. 1 is exceed T_d , the quadratic-like dissimilarity measure can be determined by

$$D^2(Q, F_1) = 0.6732 + 0.249 = 0.9222.$$

However, using the quadratic-like dissimilarity measure between the Q and F_2 is:

$$D^2(Q, F_2) = 0.6732 + 0.4489 - 2 \times (1 - 22/50) \times 0.20576 \times 0.548096 = 0.9958.$$

It can be seen that the comparison result of $D^2(Q, F_2) > D^2(Q, F_1)$ is not consistent with human perception. Whereas, using the dissimilarity measure in [19], we have

$D^2(Q, F_1)=1-0=1$

and

$D^2(Q, F_2)=1-\{(1-22/50)\times(1-|0.20576-0.548096|)\times0.20576\}=0.9242$



Figure 2. Example images with the dominant colors and their percentage values. First row: 3-D dominant color vector c_i and the percentage p_i for each dominant color. Middle row: the original images. Bottom row: the corresponding quantized images.

In DCD, the quadratic-like measure results incorrect matches due to the existence of high percentage of the same color in target image. For example, consider the quantized images in Fig. 2. We can see that the percentage of dominant colors of F_1 (rose) and F_2 (gorilla) are 82.21% and 92.72%, respectively. In human perception, Q is more similar to F_2 . However, the quadratic-like similarity measure is $D^2(Q, F_2)>D^2(Q, F_1)$. Obviously, the result causes a wrong rank. The robust similarity measure [19] is more accurate to capture human perception than that of MPEG-7 DCD. In our experiments, the modified DCD achieves 16.7% and 3% average retrieval rate (ARR) [27] improvements than Ma [28] and Mojsilovic [29], respectively. In this Chapter, the modified dominant color descriptor is chosen to support the proposed CBIR system.

4. Image segmentation and region representation

4.1. Image segmentation

It has been mentioned that segmentation is necessary for those region-based image retrieval systems. Nevertheless, automatic segmentation is still unpractical for the applications of region-based image retrieval (RBIR) systems [8, 30-32]. Although many systems provide segmentation tools, they usually need complicated user interaction to achieve image retrieval. Therefore, the processing is very inefficient and time consuming to the user. In the following, the new approach will propose to overcome this problem. In our algorithm, the user does not need to provide precisely segmented regions, instead, the boundary checking algorithm are used to support segmented regions.

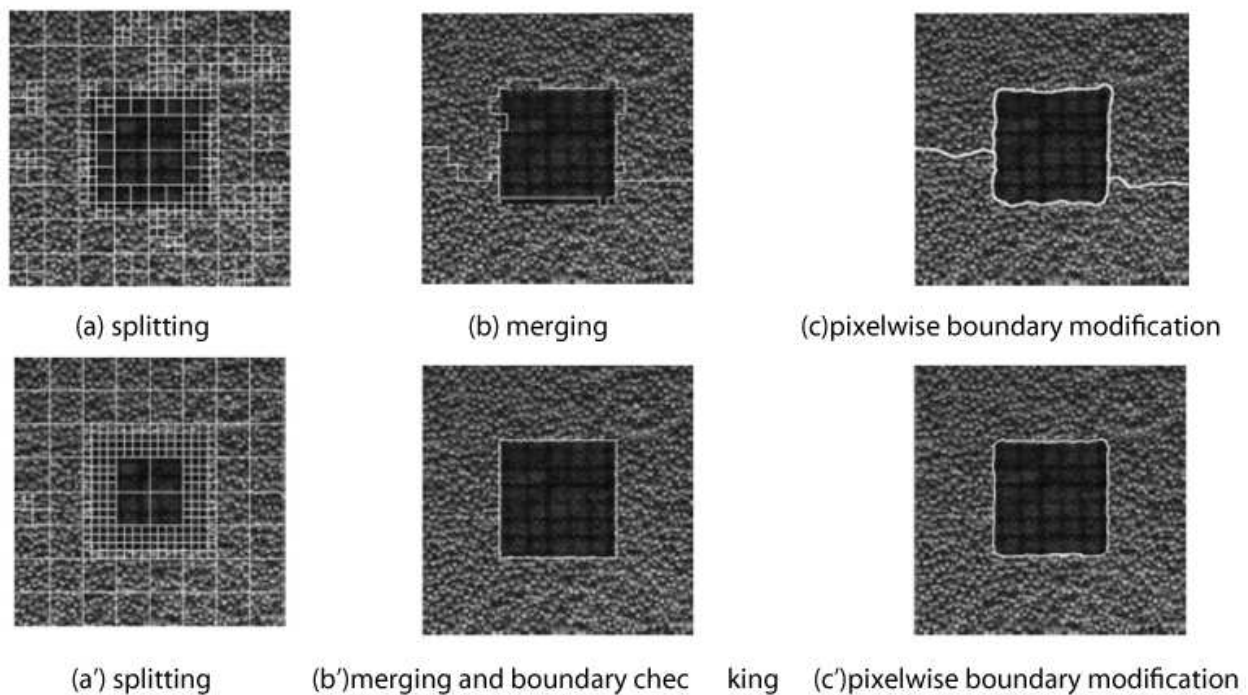


Figure 3. a), (b) and (c) are the results by using the method of T. Ojala et. al. (a'), (b') and (c') are the results by using our earlier segmentation method.

For region-based image retrieval, we adopt the unsupervised texture segmentation method [30, 33]. In [30], Ojala et al. use the nonparametric log-likelihood-ratio test and the G statistic to compare the similarity of feature distributions. The method is efficient for finding homogeneously textured image regions. Based on this method, a boundary checking algorithm [34] has been proposed to improve the segmentation accuracy and computational cost. For more details about our segmentation algorithm, we refer the reader to [33]. In this Chapter, the weighted distribution of global information CIH (color index histogram) and local information LBP (local binary pattern) are applied to measure the similarity of two adjacent regions.

An example is shown in Fig. 3. It can be seen that boundary checking algorithm segments the test image correctly, and it costs only about 1/20 processing time of the method in [30]. For color image segmentation, another example is shown in Fig. 4. In Fig. 4(c) Fig. 4(c'), we can see that the boundary checking algorithm achieves robustness segmentation for test image "Akiyo" and another nature image.

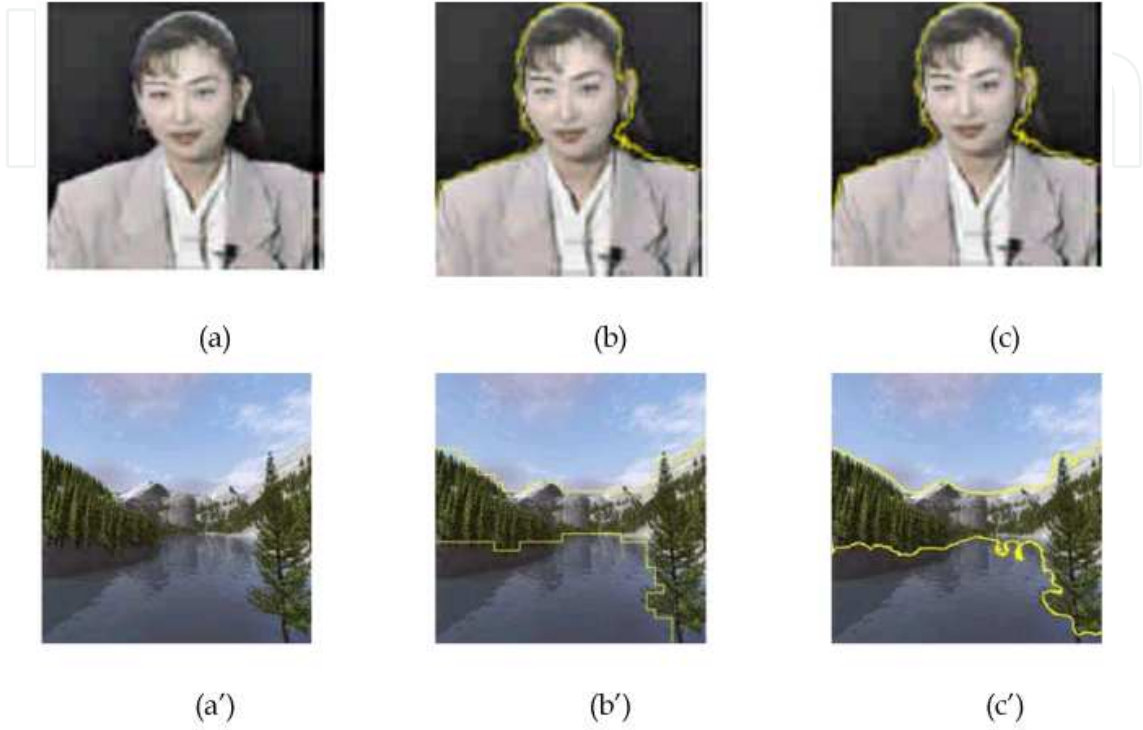


Figure 4. The segmentation processes for test image "Akiyo" and a nature image. (a), (a') Original image. (b), (b') Splitting and Merging. (c), (c') Boundary checking and modification.

4.2. Region representation

To achieve region-based image retrieval, we use two compact and intuitive visual features to describe a segmented region: dominate color descriptor (DCD) and texture. For the first one, we use our modified dominant color descriptor in [19, 26]. The feature representation of a segmented region R is defined as

$$R_{DCD} = \left\{ \left\{ R_{c_i}, R_{p_i} \right\}, 1 \leq i \leq 8 \right\}, \quad (6)$$

where R_{c_i} and R_{p_i} are the i th dominant color and its percentage in R , respectively.

For the second one, the texture feature of a region is characterized by the weighted distribution of local binary pattern (LBP) [6, 25, 32]. The advantages of LBP include its invariant property to illumination change and its low computational cost [32]. The value of k th bin in LBP histogram is given by:

$$R_{LBP_h_k} = \frac{n_K}{P}, \quad (7)$$

where n_K represents the frequency of LBP value at k th bin, and P is the number of pixels in a region. Therefore, the texture feature of region R is defined as

$$R_{texture} = \left\{ \left\{ R_{LBP_h_k} \right\}, 1 \leq k \leq 256 \right\}. \quad (8)$$

In addition, we define a feature R_{poa} to represent the *percentage of area* for region R in the image. Two regions are considered to be visual similar if both of their content (color and texture) and area are similar.

4.3. Image representation and definition of the foreground assumption

For image retrieval, each image in database is described by a set of its non-overlapping regions. For an image I that contains N non-overlapping regions, i.e., $I = \{I_{R^1}, I_{R^2}, \dots, I_{R^N}\}$, $\bigcup_{i=1}^N I_{R^i}$ and $I_{R^i} \cap I_{R^j} = 0$, where I_{R^i} represents the i th region in I . Although the region-based approaches perform well in [9, 11], their retrieval performances are strongly depends on success of image segmentation because segmentation techniques are still far from reliable for heterogeneous images database. In order to address the possible fails of segmentation, we propose a foreground assumption to “guess” the foreground and background regions in images. For instance, we can readily find a gorilla sitting on the grass as shown in Fig. 5. If Fig. 5 is the query image, the user could be interested in the main subject (gorilla) rather than grass-like features (color, texture, etc). In most case, user would pay more attention to the main subject.

The main goal of foreground assumption is to simply distinguish main objects and irrelevant regions in images. Assume that we can divide an image into two parts: foreground and background. In general, the foreground stands the central region of an image. To emphasize the importance of central region of an image, we define

$$\begin{aligned} R_{foreground} &= \{(x, y) : \frac{1}{8}h \leq x \leq \frac{7}{8}h, \frac{1}{8}w \leq y \leq \frac{7}{8}w\} \\ R_{background} &= \{(x, y) : x < \frac{1}{8}h \text{ or } x > \frac{7}{8}h, y < \frac{1}{8}w \text{ or } y > \frac{7}{8}w\}, \end{aligned} \quad (9)$$

where $R_{foreground}$ and $R_{background}$ are the occupied regions of foreground and background, respectively; h and w is height and width of the image.

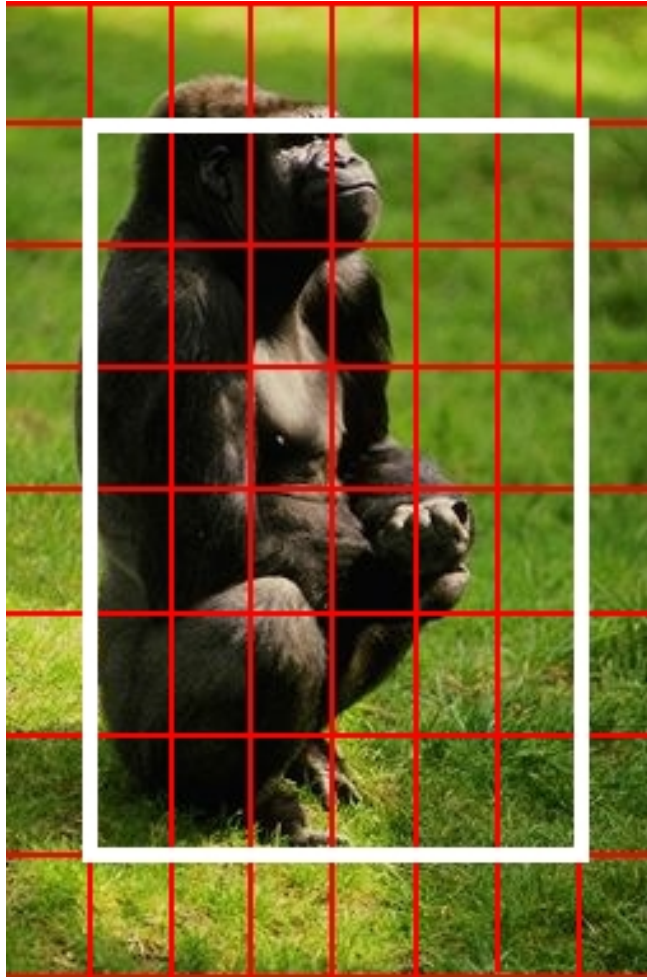


Figure 5. The definition of foreground and background based on foreground assumption.

In region-based retrieval procedure, segmented regions are required. It can be provided by the users or be generated by the system automatically. However, the criterion for similarity measure is based on the overall distances between feature vectors. If an image in database has background regions that is similar to the foreground object of the query image, this image will be considered as similar image based on the similarity measure. In this case, the accuracy of region-based retrieval system decreases. Therefore, we modify our region representation by adding a Boolean model $BV \in \{0, 1\}$ to determine whether the segmented region R belongs to the background of the query image or not.

$$BV = \begin{cases} 1 & R \in R_{background} \\ 0 & R \notin R_{background} \end{cases} \quad (10)$$

Note that the variable is designed to reduce the segmentation error.

On the other hand, we extract the global features for an image to compensate the inaccuracy of segmentation algorithms. The features F^I includes three feature sets: 1) dominant color $F_{R_{DCD}}^I$ for each region, 2) texture $F_{R_{texture}}^I$ for each region, and 3) dominant color F^I .

$$F_{R_{DCD}}^I = \left\{ \left\{ \left\{ R_{c_i}^j, R_{p_i}^j \right\}, 1 \leq i \leq 8 \right\}, R_{poa}^j, BV_j^j \right\}, 1 \leq j \leq N \quad (11)$$

$$F_{R_{texture}}^I = \left\{ \left\{ R_{LBP_{h_k}}^j \right\}, 1 \leq k \leq 256 \right\}, 1 \leq j \leq N \quad (12)$$

$$F^I = \left\{ F_{global}^I, F_{foreground}^I, F_{background}^I \right\} \quad (13)$$

where N is the number of partitioned regions in image I ; $F_{R_{DCD}}^I$ represents the dominant color vectors; $F_{R_{texture}}^I$ describes the texture distribution for each region; F_{global}^I , $F_{foreground}^I$ and $F_{background}^I$ represent the global, foreground and background color features, respectively. In brief, the images are first segmented using the fast color quantization scheme. Then, the dominant colors, texture distribution and the three color features are extracted in the image.

5. Integrated region-based relevance feedback framework

In region-based image retrieval, an image is considered as relevant if it contains some regions with satisfactory similarity to the query image. The retrieval system can reconstruct a new query that includes only the relevant regions according to user's feedback. In this way, the system can capture the user's query concept automatically. For example, Jing et al. [8] suggest that information in every region could be helpful in retrieval, and group all regions of positive examples by K-means algorithm iteratively to ensure the distance between all the clusters not exceeding a predefined threshold. Then, all regions within a cluster are merged into a new region. However, the computational cost for merging new regions is proportional to the number of positive examples. Moreover, users might be more interested in some specified regions or main objects rather than the positive examples.

To speed up the system, we introduce a similarity matrix model to infer the region-of-interest sets. Inspired by the query-point movement method [8, 31], the proposed system performs similarity comparisons by analyzing the salient region in pseudo query image and relevant images based on user's feedback information.

5.1. The formation of region-of-interest set

5.1.1. Region-based similarity measure

In order to perform region-of-interest (ROI) queries, the relevant regions are obtained by the measurement of region-based color similarity $R_S(R, R')$ and region based texture similarity $R_S_T(R, R')$ in Eq. (14) and (15), respectively. This similarity measure allows users to select their relevant regions accurately. Note that the conventional color histogram could not be applied on DCD directly because the images do not have exact numbers of dominant colors [12]. The region-based color similarity between two segmented regions R and R' can be calculated by

$$R_S(R, R') = R_S_c(R, R') \times R_S_{poa}(R, R') \quad (14)$$

$$R_S_c(R, R') = \sum_{i=1}^m \sum_{j=1}^n \min(R_{p_i}, R'_{p_j}) , \text{ if } d(R_{c_i}, R'_{c_j}) < T_d,$$

where m and n are the number of dominate colors in R and R' , respectively; $R_S_c(R, R')$ is the maximum similarity between two regions in similar color percentage. If the pair-wise Euclidean distance of two dominate color vector c_i and c_j is less than a predefined threshold T_d , it is set to 25 in our work. The notation $R_S_{poa}(R, R')$ is used to measure the similarity of the area percentage for region pair (R, R') . To measure the texture similarity between two regions, we define

$$R_S_T(R, R') = \frac{\sum_{k=1}^{256} \min(R_{LBP_hk}, R'_{LBP_hk})}{\min(R_{Pxl}, R'_{Pxl})}, \quad (15)$$

where R_{Pxl} and R'_{Pxl} represent the number of pixels in regions R and R' , respectively; $\min(R_{LBP_hk}, R'_{LBP_hk})$ is the intersection of LBA histogram for the k th bin.

Theoretically, visual similar is achieved when both color and texture are similar. For example, two regions should be considered as non-similar if they are similar in terms of color but not texture. This can be achieved by imposing

$$R_S > 0.8 \text{ and } R_S_T > 0.9. \quad (16)$$

5.1.2. Similarity matrix model

In the following, we introduce a region-based similarity matrix model. The regions of positive examples, which helps the system to find the intention of user's query, are able to exclude the irrelevant regions flexibly. The proposed similarity matrix model is described as follows.

The region similarity measure is performed for all regions. The relevant image set is denoted as $R_s = \{I^i; i=1, \dots, N\}$, where N represents the number of positive images from user's feedback, and each positive image I^i contains several segmented regions. See Fig. 6.

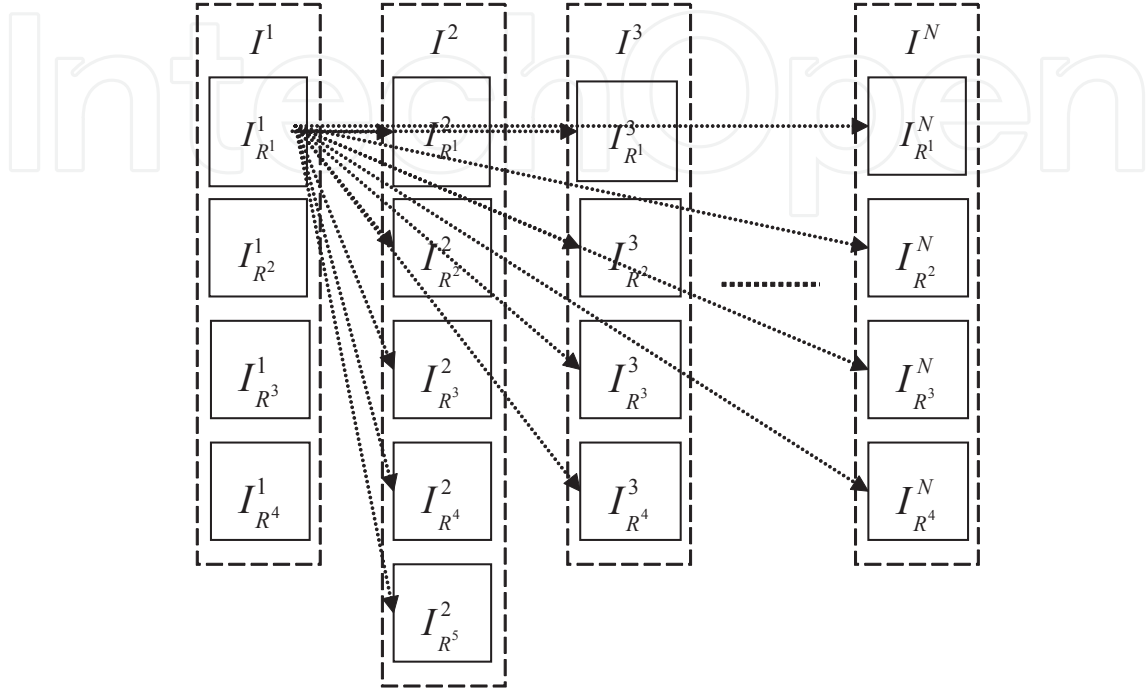


Figure 6. The similarity matching for region pairs.

As an example, let $R_s = \{I^1, I^2, I^3\}$ contains three relevant images, where $I^1 = \{I_{R^1}^1, I_{R^2}^1, I_{R^3}^1\}$, $I^2 = \{I_{R^1}^2, I_{R^2}^2, I_{R^3}^2\}$ and $I^3 = \{I_{R^1}^3, I_{R^2}^3\}$. Our similarity matrix model to infer the user's query concept is shown in Fig. 7, where the symbol "1" means that two regions are regarded as similar. On the contrary, the symbol "0" represents that two regions are non-similar in content.

To support ROI queries, we perform the one-to-many relationships to find a collection of similar region sets, e.g., $\{I_{R^1}^1, I_{R^1}^2, I_{R^1}^3\}$, $\{I_{R^2}^1, I_{R^2}^2\}$, $\{I_{R^3}^1, I_{R^2}^2, I_{R^3}^2\}$, $\{I_{R^1}^2, I_{R^1}^1, I_{R^2}^3\}$, $\{I_{R^2}^2, I_{R^2}^1, I_{R^3}^1\}$, $\{I_{R^3}^2, I_{R^3}^1\}$, $\{I_{R^3}^1\}$ and $\{I_{R^2}^3, I_{R^1}^1, I_{R^2}^1\}$, see Fig. 8. After this step, several region-of-interest sets can be obtained by merging all similar region sets. For example, the first set $\{I_{R^1}^1, I_{R^1}^2, I_{R^2}^3\}$ contains three similar regions. Each region will be merged together with the above eight similar region sets. In this example, three region-of-interest sets can be obtained by the merging operation, i.e., $\{I_{R^1}^1, I_{R^1}^2, I_{R^2}^3\}$, $\{I_{R^2}^1, I_{R^3}^1, I_{R^2}^2, I_{R^3}^2\}$ and $\{I_{R^3}^1\}$. Since user may be interested in some repeated similar regions, the single region set $\{I_{R^3}^1\}$ could be assumed to be irrelevant in our approach. Therefore, we have

$ROI^1 = \{I_{R^1}^1, I_{R^2}^2, I_{R^3}^3\}$ and $ROI^2 = \{I_{R^2}^1, I_{R^3}^1, I_{R^2}^2, I_{R^3}^2\}$ as shown in Fig. 8. The two sets are considered as region-of-interests that reflect user's query perception.

	$I_{R^1}^1$	$I_{R^2}^1$	$I_{R^3}^1$	$I_{R^1}^2$	$I_{R^2}^2$	$I_{R^3}^2$	$I_{R^1}^3$	$I_{R^2}^3$
$I_{R^1}^1$	×	×	×	1	0	0	0	1
$I_{R^2}^1$	×	×	×	0	1	0	0	0
$I_{R^3}^1$	×	×	×	0	1	1	0	0
$I_{R^1}^2$	1	0	0	×	×	×	0	1
$I_{R^2}^2$	0	1	1	×	×	×	0	0
$I_{R^3}^2$	0	0	1	×	×	×	0	0
$I_{R^1}^3$	0	0	0	0	0	0	×	×
$I_{R^2}^3$	1	0	0	1	0	0	×	×

Figure 7. Our proposed matrix structure comparison. x: no comparison for those regions in the same image, 1: similar regions and 0: non-similar regions.

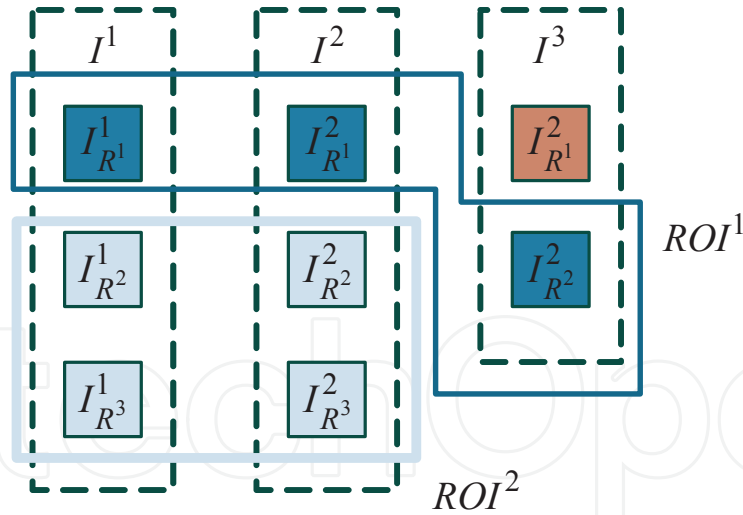


Figure 8. The region-of-interest sets based on the proposed matrix structure comparison.

If users are interested in many regions, the simple merging process can be used to capture the query concept. In Fig. 8, for example, $\{I_{R^2}^1, I_{R^3}^1\}$ and $\{I_{R^2}^2, I_{R^3}^2\}$ are the regions belong to the same relevant image I^1 and I^2 , respectively. It can be seen that the similar matrix approach is consistent with human perception and is efficient for region-based comparison.

5.1.3. Salient region model

To improve retrieval performance, all the region-of-interest sets from the relevant image set R_s will be integrated for the next step during relevance feedback. As described in previous subsection, each region-of-interest set could be regarded as a collection of regions, and extracted information can be used to identify the user's query concept. However, correctly capturing the semantic concept from the similar regions is still a difficult task. In this stage, we define salient region as all similar regions within each ROI set. The features of the new region are equal to the weighted average features of individual regions.

In order to emphasize the percentage of area feature, we modified the dominant color descriptor in Eq. (1). The feature representation of the salient region SR is described as

$$F_{SR} = \left\{ \left\{ \left\{ \bar{C}_i, \bar{P}_i \right\}, 1 \leq i \leq 8 \right\}, \bar{R}_{poa} \right\}, \quad (17)$$

where \bar{C}_i is the i^{th} average dominant color of similar region.

All similar regions in ROI can be determined from the eight uniformly divided partitions in RGB color space as shown in Fig. 9.

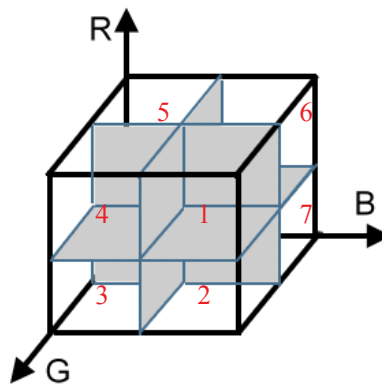


Figure 9. The division of RGB color space.

$$\bar{C}_i = \left(\frac{\sum_{j=1}^{N_{c_i}} R_{p_i}^j \times R_{c_i}^j(R)}{\sum_{j=1}^{N_{c_i}} R_{p_i}^j}, \frac{\sum_{j=1}^{N_{c_i}} R_{p_i}^j \times R_{c_i}^j(G)}{\sum_{j=1}^{N_{c_i}} R_{p_i}^j}, \frac{\sum_{j=1}^{N_{c_i}} R_{p_i}^j \times R_{c_i}^j(B)}{\sum_{j=1}^{N_{c_i}} R_{p_i}^j} \right), 1 \leq i \leq 8 \quad (18)$$

where N_{c_i} is the number of dominant colors in cluster i ; $R_{c_i}^j(R)$, $R_{c_i}^j(G)$ and $R_{c_i}^j(B)$ represent the dominant color components of R, G and B located within partition i for the region j , re-

spectively; $R_{p_i}^j$ represents the percentage of its corresponding 3-D dominant color vector in R^j ; \bar{P}_i is the average percentage of dominant color in the i th coarse partition, i.e.,

$$\bar{P}_i = \frac{\sum_{j=1}^{N_{c_i}} R_{p_i}^j}{N_{c_i}}; \bar{R}_{poa}$$

is the average percentage of area for all similar regions in ROI.

5.2. The pseudo query image and region weighting scheme

To capture the inherent subjectivity of user perception, we define a pseudo image I^+ as the set of salient regions, $I^+ = \{SR^1, SR^2, \dots, SR^n\}$. The feature representation of I^+ can be written as

$$F_{SR}^{I^+} = \left\{ \left\{ (\bar{C}_i^1, \bar{P}_i^1), 1 \leq i \leq 8 \right\}, \bar{R}_{poa}^1, \dots, \left\{ (\bar{C}_i^n, \bar{P}_i^n), 1 \leq i \leq 8 \right\}, \bar{R}_{poa}^n \right\}. \quad (19)$$

During retrieval, the user chooses the best matched regions what he/she is looking for. However, the retrieval system cannot precisely capture the user's query intention at the first or second steps of relevance feedback. With the increasing of the returned positive images, query vectors are then constructed to perform better results. Taking average [8] from all the feedback information could introduce redundant, i.e., information from irrelevant regions. Motivated by this observation, we suggest that each similar region in ROI should be properly weighted according to the amount of similar regions. For example, the ROI^2 in Fig. 8 is more important than in ROI^1 . The weights associated with the significance of SR in I^+ can be dynamically updated as

$$w_l = \frac{|ROI^l|}{\sum_{l=1}^n |ROI^l|}, \quad (20)$$

where $|ROI^l|$ represents the number of similar regions in region-of-interest set l , and n is the number of region-of-interest sets.

5.3. Region-based relevance feedback

In reality, inaccurate segmentation leads to poor matching result. However, it is difficult to ask for precise segmented regions from users. Based on the foreground assumption, we define three feature vectors, which are extracted from entire image (i.e., global dominant color), foreground and background, respectively. The advantage of this approach is that it provides an estimation that minimizes the influence of inaccurate segmentation. To integrate the two regional approaches, we summarize our relevance feedback as follows.

For the initial query, the similarity measure $S(F_{entireImage}^I, F_{entireImage}^{I'})$ for the initial query image I and target image I' in database are compared by using Eq. (4). Therefore, a coarse relevant-image set can be obtained. Then, all regions in the initial query image I and the positive images based on the user's feedback information are merged into relevant image set $R_s = \{I, I^1, I^2, \dots, I^N\}$. The proposed region-based similarity matrix model performs Eq. (14) and (15) to find the collection of the similar regions. The similar regions can be determined by Eq. (16), and then be merged into salient region SR . For the next iteration, the feature representation of I^+ in Eq. (19) could be regarded as an optimal pseudo query image that is characterized by salient regions.

It should be noted that I^+ and R_s defined above both contain the relevance information that reflects human semantics. The similarity measure for pseudo query image $F_{SR^l}^{I^+}$ and target image $F_{R_{DCD}^j}^{I'}$ is calculated by

$$S_{region_based}(I^+, I') = \sum_{l=1}^n \sum_{j=1}^m w_l \times \max R_S(F_{SR^l}^{I^+}, F_{R_{DCD}^j}^{I'}), \quad (21)$$

where n is the number of salient region sets in I^+ ; m is the number of color/texture segmented regions in target image I' ; w_l is the weight of salient region SR^l . In Eq. (21), the image-to-image similarity matching maximizes the value of region based color similarity by using Eq. (14). If the Boolean model $BV=1$ for a partitioned region in target image, then the background of the image will be excluded for matching in Eq. (21).

On the other hand, R_s is a collection of relevant images based on the user's feedback information. Since poor matches arise from inaccurate image segmentations, three global features $F_{entireImage}^I$, $F_{foreground}^I$ and $F_{background}^I$ in Eq. (13) are extracted to compensate the inaccuracy. The similarity between the relevant image set $R_s = \{I, I^1, I^2, \dots, I^N\}$ and target image I' in database is calculated by

$$\begin{aligned} S_{entireImage}(R_s, I') &= \sum_{i=1}^N \max S(F_{entireImage}^{R_s}, F_{entireImage}^{I'}) \\ S_{foreground}(R_s, I') &= \sum_{i=1}^N \max S(F_{foreground}^{R_s}, F_{foreground}^{I'}) \\ S_{background}(R_s, I') &= \sum_{i=1}^N \max S(F_{background}^{R_s}, F_{background}^{I'}) \end{aligned} \quad (22)$$

where $F_{entireImage}^{R_s}$, $F_{foreground}^{R_s}$ and $F_{background}^{R_s}$ are dominant colors, foreground and background for the i th relevant image in R_s , respectively. In Eq. (22), the similarity measure

maximizes the similarity score using Eq. (5). To reflect the difference between R_s and target image I' , the average similarity measure is given by

$$S_{\text{avg}}(R_s, I') = \frac{(S_{\text{entireImage}}(R_s, I') + S_{\text{foreground}}(R_s, I') + S_{\text{background}}(R_s, I'))}{3}. \quad (23)$$

It is worth to mention that our region-based relevance feedback approach defined above is able to reflect human semantics. In other words, user might aware some relevant image from the initial query, and then provides some positive image.

Considering the ability to capture the user's perceptions more precisely, the system determines the retrieved rank according to average of region-based image similarity measure in Eq. (21) and foreground-based similarity measure in Eq. (23).

$$S = \frac{S_{\text{region_based}}(I^+, I') + S_{\text{avg}}(R_s, I')}{2}. \quad (24)$$

6. Experimental results

We use an image database (31 categories about 3991 images) for general-purpose from Cor-el's photo to evaluate the performance of the proposed framework. The database has a variety of images including animal, plant, vehicle, architecture, scene, etc. It has the advantages of large size and wide coverage [11]. Table 1 lists the labels for 31 classes. The effectiveness of our proposed region-based relevance feedback approach is evaluated.

In order to make a comparison on the retrieval performance, both average retrieval rate (ARR) and average normalized modified retrieval rank (ANMRR) [26] are applied. An ideal performance will consist of ARR values equal to 1 for all values of recall. A high ARR value represents a good performance for retrieval rate, and a low ANMRR value indicates a good performance for retrieval rank. The brief definitions are given as follows. For a query q , the ARR and ANMRR are defined as:

$$\text{ARR}(q) = \frac{1}{NQ} \sum_{q=1}^{NQ} \frac{NF(\beta, q)}{NG(q)}, \quad (25)$$

$$\text{AVR}(q) = \sum_{k=1}^{NG(q)} \frac{\text{Rank}(k)}{NG(q)}, \quad (26)$$

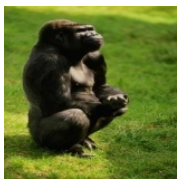
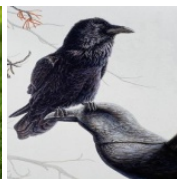
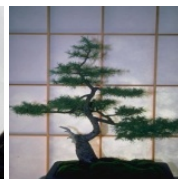
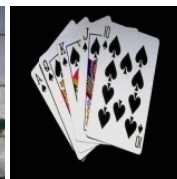
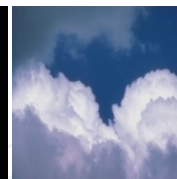
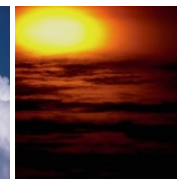
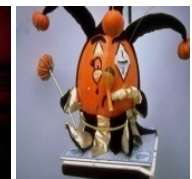


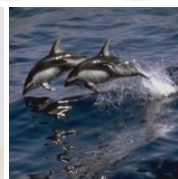
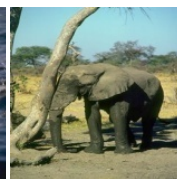

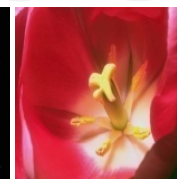


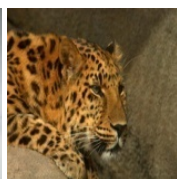
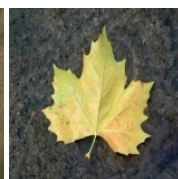
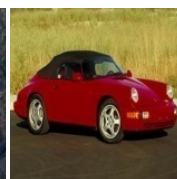


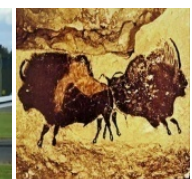
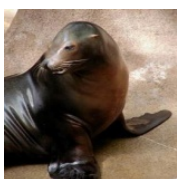
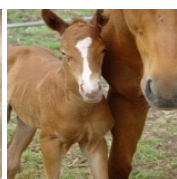
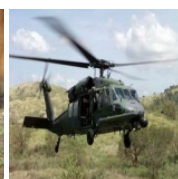

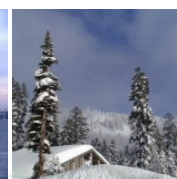
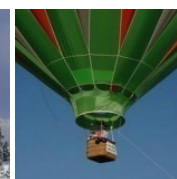
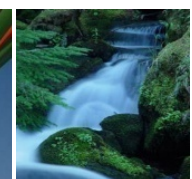
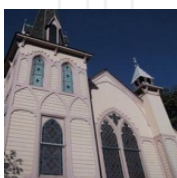
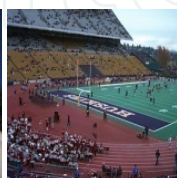
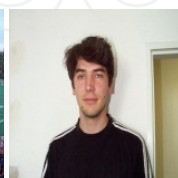
Class 1 (gorilla)	Class 2 (bird)	Class 3 (potted plant)	Class 4 (card)	Class 5 (cloud)	Class 6 (sunset)	Class 7 (pumpkin)
						
Class 8 (cake)	Class 9 (dinosaur)	Class 10 (dolphin)	Class 11 (elephant)	Class 12 (firework)	Class 13 (flower)	Class 14 (food)
						
Class 15 (duck)	Class 16 (leopard)	Class 17 (leaf)	Class 18 (car)	Class 19 (cactus)	Class 20 (airplane)	Class 21 (painting)
						
Class 22 (sea- elephant)	Class 23 (horse)	Class 24 (helicopter)	Class 25 (boat)	Class 26 (snow)	Class 27 (balloon)	Class 28 (waterfall)
						
Class 29 (building)	Class 30 (stadium)	Class 31 (people)				
						

Table 1. The labels and examples of the test database.

$$\text{MRR}(q) = \text{AVR}(q) - 0.5 - \frac{NG(q)}{2}, \tag{27}$$

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 \times NG(q)}, \quad (28)$$

$$ANMRR(q) = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q), \quad (29)$$

where NQ is total number of queries; $NG(q)$ is the number of the ground truth images for a query. The notation is a factor, and $NF(\beta, q)$ is number of ground truth images found within the first $\beta \cdot NG(q)$ retrievals. $Rank(k)$ is the rank of the retrieved signature image in the ground truth. In eq.(28), $K = \min(4 \cdot NG(q); 2 \cdot GTM)$, where $GTM = \max\{NG(q)\}$ for all queries. The NMRR and its average (ANMRR) are normalized to the range of [0 1].

To test the performance of our integrated approach for region-based relevance feedback, we first query an image with a gorilla sits on grass as shown Fig. 10(a).

As mentioned in Section 5.4, the dominant color between query image I and target image I' is used for similarity measure in the initial query. The retrieval results are shown in Fig. 10(b), the top 20 matching images are arranged from left to right and top to bottom in order of decreasing similarity score.

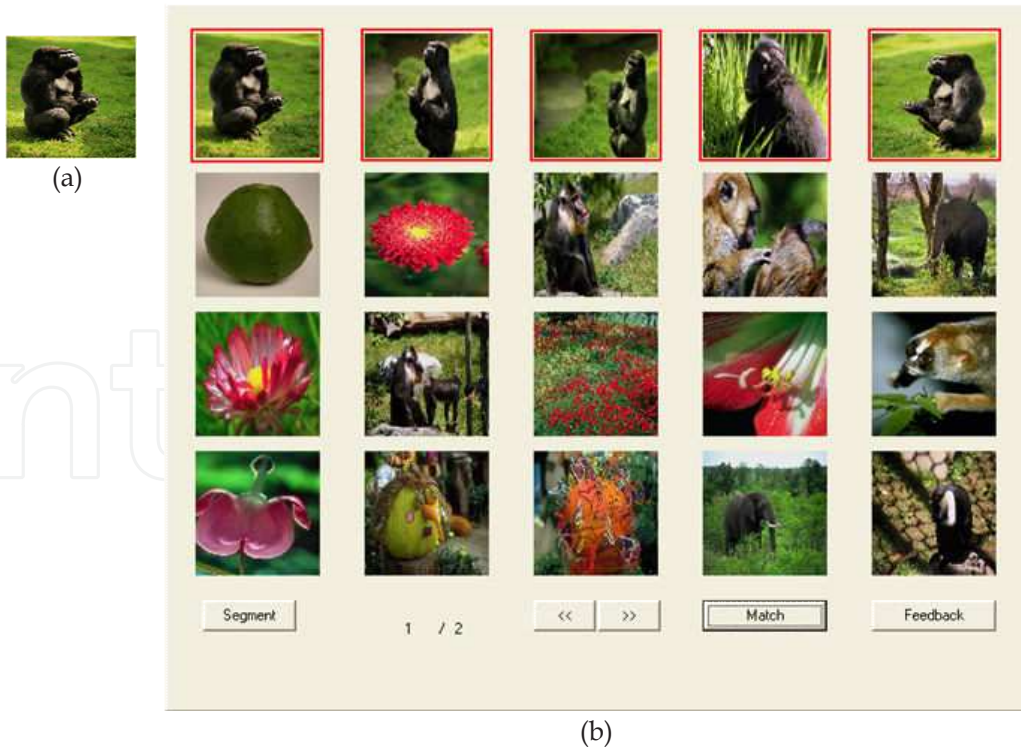


Figure 10. The initial query image and positive images. (a) Query image. (b) The 5 positive images in the first row are selected by user.

For better understanding of the retrieval results, the DCD vectors of the query image, rank 6th image and rank 8th image are listed, respectively. See Fig. 11. It can be seen that the query image and the image “lemon” are very similar in the first dominant color (marked by box). If we use the global DCD as the only feature for image retrieval, the system only returns eleven correct matches. Therefore, further investigation on extracting comprehensive image features is needed.

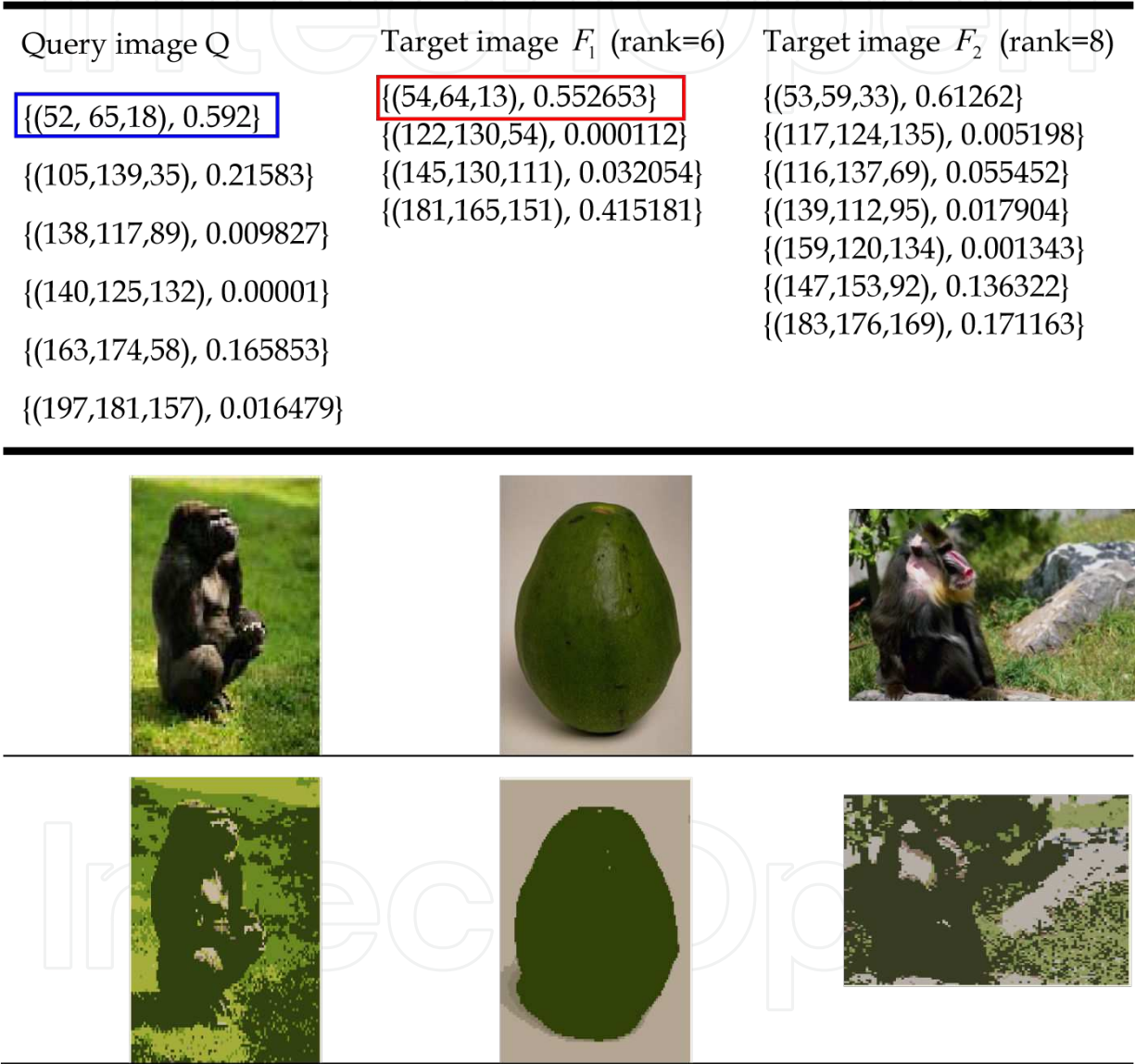


Figure 11. Example images with the dominant colors and their percentage values. First row: 3-D dominant color vector c_i and the percentage p_i for each dominant color. Middle row: the original images. Bottom row: the corresponding quantized images.

Assume that the user has selected five best matched images, marked by red box, as shown in Fig. 10(a). In conventional region-based relevance feedback approach, all regions in the initial query image I and the five positive images are merged into relevant image set

$R_s = \{I, I^1, I^2, ..., I^5\}$. The proposed similarity matrix model is able to find the region-of-interest region sets. For the next query, I^+ could be regarded as a new query image which is composed of some salient regions. The retrieval results based on the new query image I^+ are shown in Fig. 12. The following are discussions.

- 1. The pseudo query image I^+ is capable to reflect user's query perception. Without considering the Boolean model in Eq. (21), the similarity measure by Eq. (21) returns 16 correct matches as shown in Fig. 12.
- 2. Using the pseudo image I^+ as query image, the initial query image is not ranked first but fifth, as shown in Fig. 12.

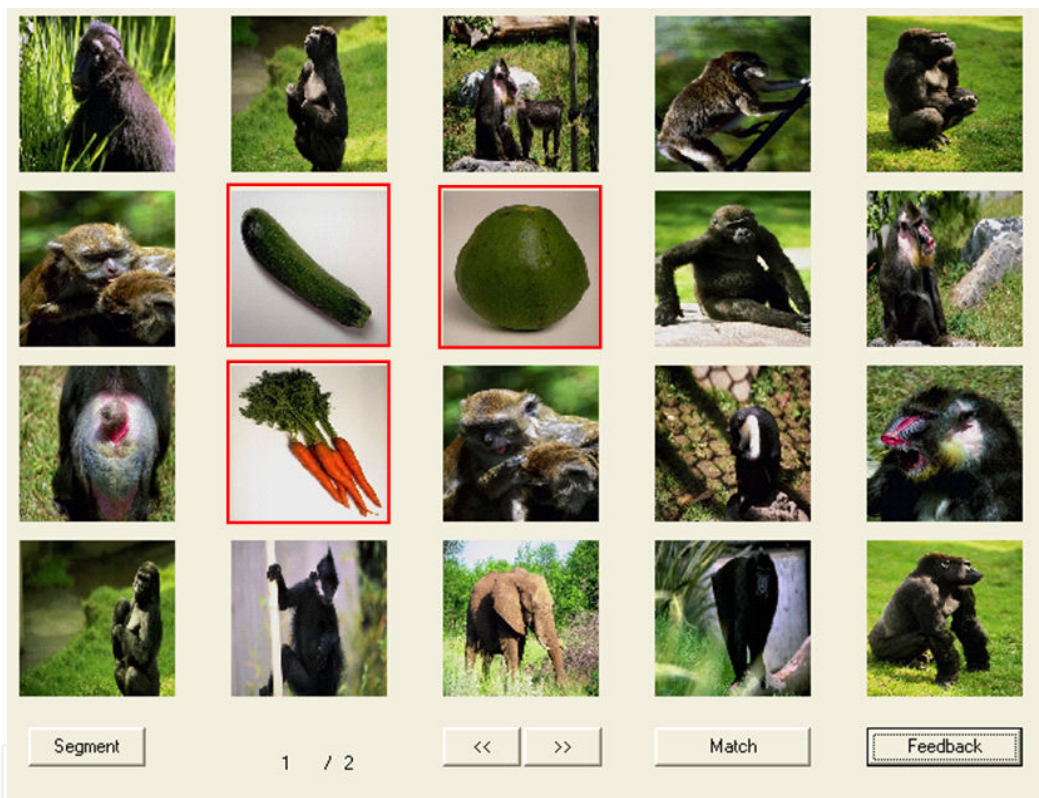


Figure 12. The retrieval results based on new pseudo query image I^+ for the first iteration.

- 3. The retrieval results return three dissimilar images (marked by red rectangle boxes), which ranks are 7th, 8th and 12th, respectively.
- 4. To analyze the improper result, the dominant color vectors and percentage of area of “cucumber” and “lemon” are listed. See Fig. 13. We can see that each of the images “gorilla”, “cucumber” and “lemon” contains three segmented regions. For each region, the number of the dominant colors, percentage of area and BV value are listed and colored red. For similarity matching, the dominant colors (i.e. region#1, region#2 and region#3) of initial image “gorilla” are similar to the dominant color (marked by red rectangle

box) of the image “cucumber”. In addition, the percentages of area (0.393911, 0.316813, 0.289276) of initial image “gorilla” are similar to the percentage of area (region#2, 0.264008) of the image “cucumber”. The other similarity comparisons between “gorilla” and “cucumber” image are not presented here because the maximum similarity between two regions in Eq. (14) is very small. In brief, without considering the exclusion of irrelevant regions, the region-based image-to-image similarity model in Eq. (21) could cause improper ranks in visualization.




The initial image “gorilla”	The image “cucumber”	The images “lemon”
Region#1 (5, 0.393911, 1) {(54,66,21), 0.358263} {(106,138,38), 0.459055} {(137,118,90), 0.002918} {(162,173,60), 0.175735} {(195,180,156), 0.004029} Region#2 (6, 0.316813, 0) {(54,66,21), 0.901682} {(106,138,38), 0.00366} {(137,118,90), 0.024017} {(142,127,130), 0.000032} {(162,173,60), 0.026554} {(195,180,156), 0.044053} Region#3 (5, 0.289276, 1) {(54,66,21), 0.571122} {(106,138,38), 0.116995} {(137,118,90), 0.003692} {(162,173,60), 0.304955} {(195,180,156), 0.003235}	Region#1 (4, 0.454976, 1) {(34,41,22), 0.003331} {(121,131,93), 0.000089} {(134,127,118), 0.000022} {(203,199,197), 0.996557} Region#2 (4, 0.264008, 0) <div>{(43,51,27), 0.985551}</div> {(121,131,93), 0.002004} {(134,127,118), 0.000501} {(203,199,197), 0.011945} Region#3 (3, 0.281016, 1) {(41,48,26), 0.049412} {(134,127,118), 0.012814} {(203,199,197), 0.937774}	Region#1 (3, 0.42462, 1) {(5,6,1), 0.035743} {(145,130,111), 0.006157} {(181,165,151), 0.9581} Region#2 (4, 0.25519, 0) {(52,62,12), 0.933309} {(122,130,54), 0.000438} {(145,130,111), 0.035119} {(181,165,151), 0.031133} Region#3 (3, 0.32019, 0) {(53,63,12), 0.998539} {(145,130,111), 0.000191} {(181,165,151), 0.001271}
		

Figure 13. The analysis of retrieval results using the conventional region-based relevance feedback approach. Top row: dominant color distributions and percentage of area P_{oa} for each region in initial query image, “cucumber” and “lemon” images. Bottom row: the corresponding segmented images.

The retrieval performance can be improved by automatically determining the user’s query perception. In the following, we would like to evaluate the advantages of our proposed relevance feedback approach. For the second query, the integrated region-based relevance feedback contains not only the salient-region information, but also the “specified-region” information based on relevant images set R_s . The retrieval results based on our integrated

region-based relevance feedback are shown in Fig. 14. Observations and discussions are described as follows.

1. The system returns 18 correct matches as shown in Fig. 14.
2. In Fig. 13, region#1 and region#3 in query image are two grass-like regions, which are labeled as inner region, i.e., $BV = 1$. On the other hand, the region#2 in image “cucumber” is a green region that is similar to the grass-like regions in query image. In our method, this problem can be solved by examining the BV value in Eq. (21). As we can see, none of the three incorrect images including “cucumber”, “lemon” and “carrot” in Fig. 12 appears in the top 20 images in Fig. 14.
3. In contrast, it is possible that the grass-like regions are parts of the user’s aspect. In this case, the three feature vectors including entire image, foreground and background can be used to compensate the loss of generality. In Fig. 14 retrieval results indicate that the high performance is achieved by using these features.
4. Our proposed relevance feedback approach can capture the query concept effectively. In Fig. 14, it can be seen that most of the retrieval results are considered to be highly correlated. In this example, 90% of top 20 images are correct images. In general, the features in all retrieval results look similar to gorilla or grass. The results reveal that the proposed method improves the performance of the region-based image retrieval.

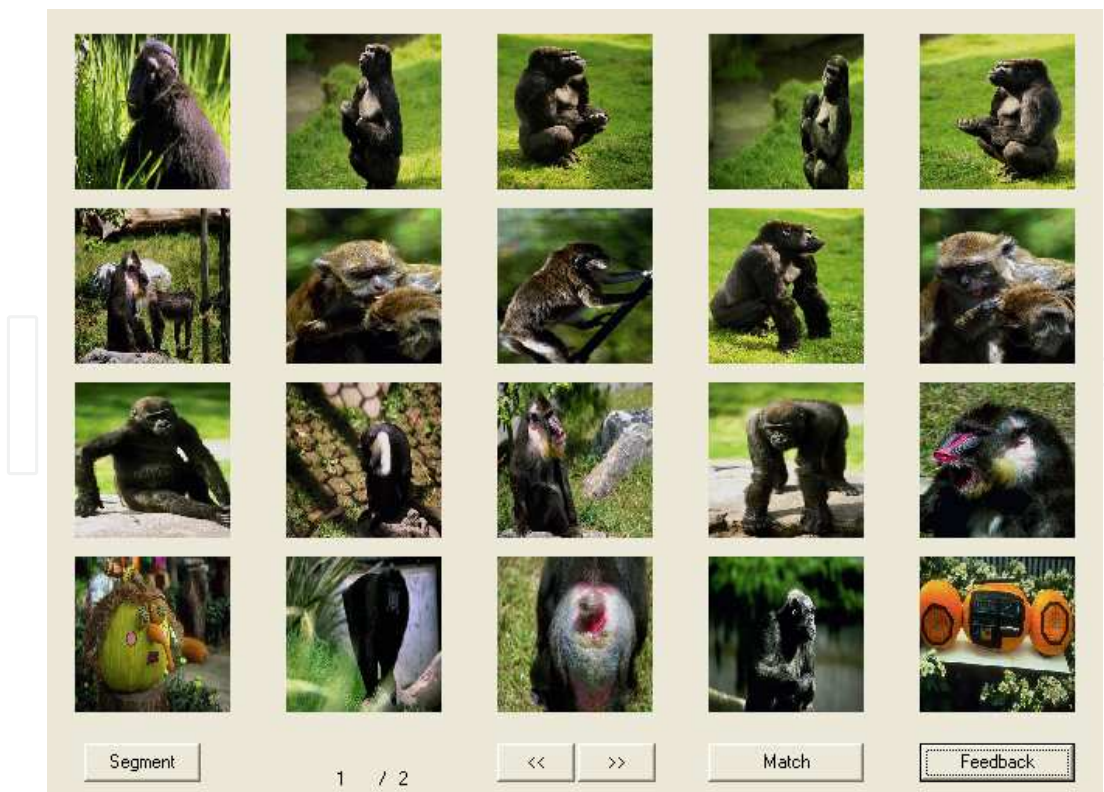


Figure 14. The retrieval results based on our integrated region-based relevance feedback.

In Fig. 15-17, further examples are tested to evaluate the performance of the integrated region-based relevance feedback for nature images. In Fig. 15, the contents of the query image include a red car on country road by the side of grasslands. If the user is only interested in the red car, four positive images marked by red boxes will be selected as shown in Fig. 15 (b). In this case, retrieval results ($RR=0.25$, $NMRR=0.7841$) are far from satisfactory performance for the initial query.

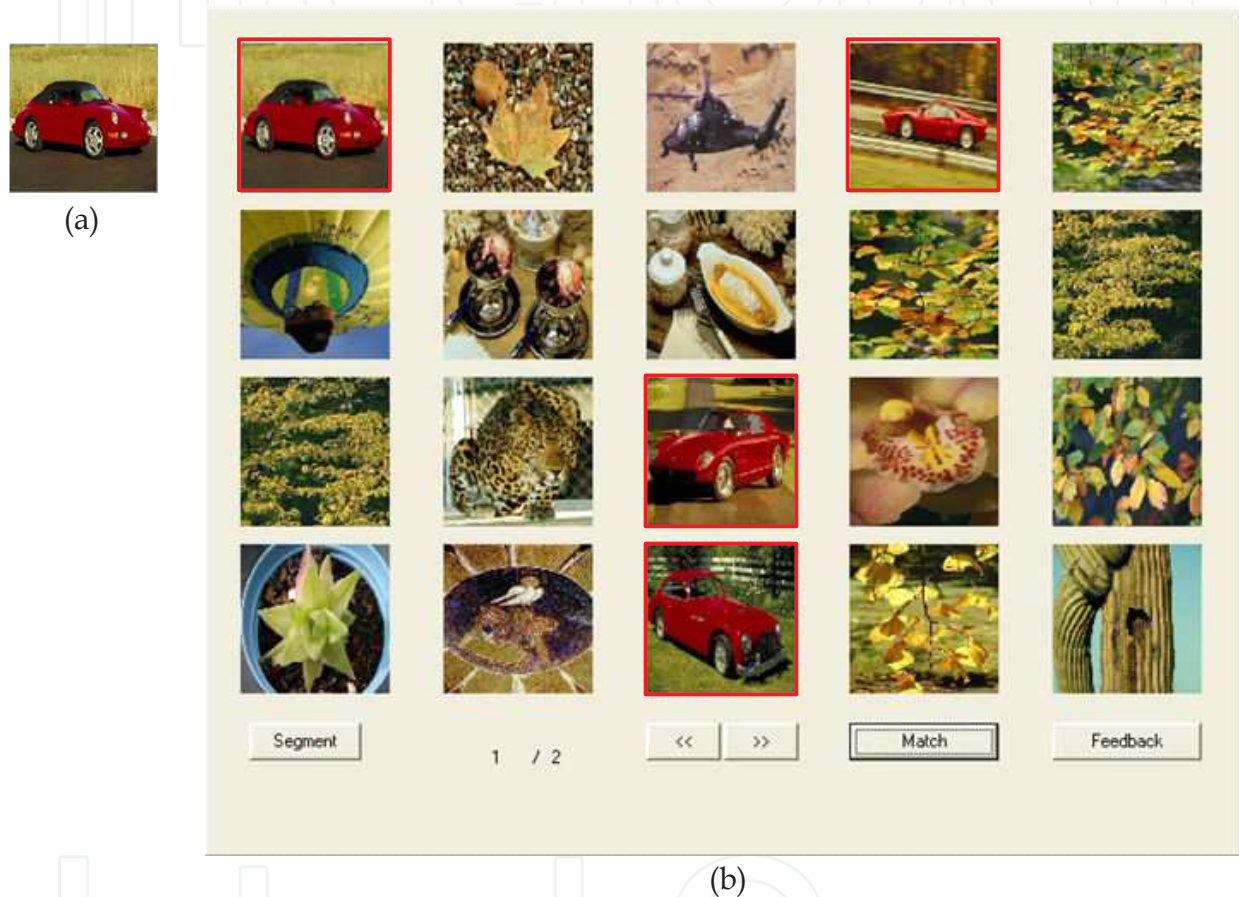


Figure 15. The initial query image and positive images. (a) Query image. (b) The 4 positive images marked by red boxes which are selected by user.

After the submission of pseudo query image I^+ and relevant images set R_s based on user's feedback information, the first feedback retrieval returns 10 images containing "red car" as shown in Fig. 16. For this example, the first feedback retrieval achieves an ARR improvement of 28.6%. More precise results can be achieved by increasing of the number of region-of-interest sets and relevant image set based for the second feedback retrieval as shown in Fig. 17. The retrieval results for the second feedback retrieval returns 11 images containing "red car", and achieve an NMRR improvement of 35% compared to the initial query. Furthermore, the rank order in Fig. 17 is more reasonable than that in Fig. 16.

To show the effectiveness of our proposed region-based relevance feedback approach, the quantitative results for individual class and average performance (ARR, ANMRR) are listed in Table 2 and 3, which show the comparison of the performance for each query. It can be seen that the performance of retrieving precision and rank are relatively poor for the initial query. Through the adding positive examples by user, feedback information could have more potential in finding the user’s query concept by means of optimal pseudo query image I^+ and relevant images set R_s as described in Section 5.4. In summary, the first feedback query improves 30.8% of ARR gain and 28% of ANMRR gain, and the second feedback query further improves 10.6% of ARR gain and 11% of ANMRR gain as compared with first feedback query. Although the improvement of retrieval efficiency is decreases progressively after two or three feedback queries, the proposed technique is able to provide satisfactory retrieval results in that few feedback queries.

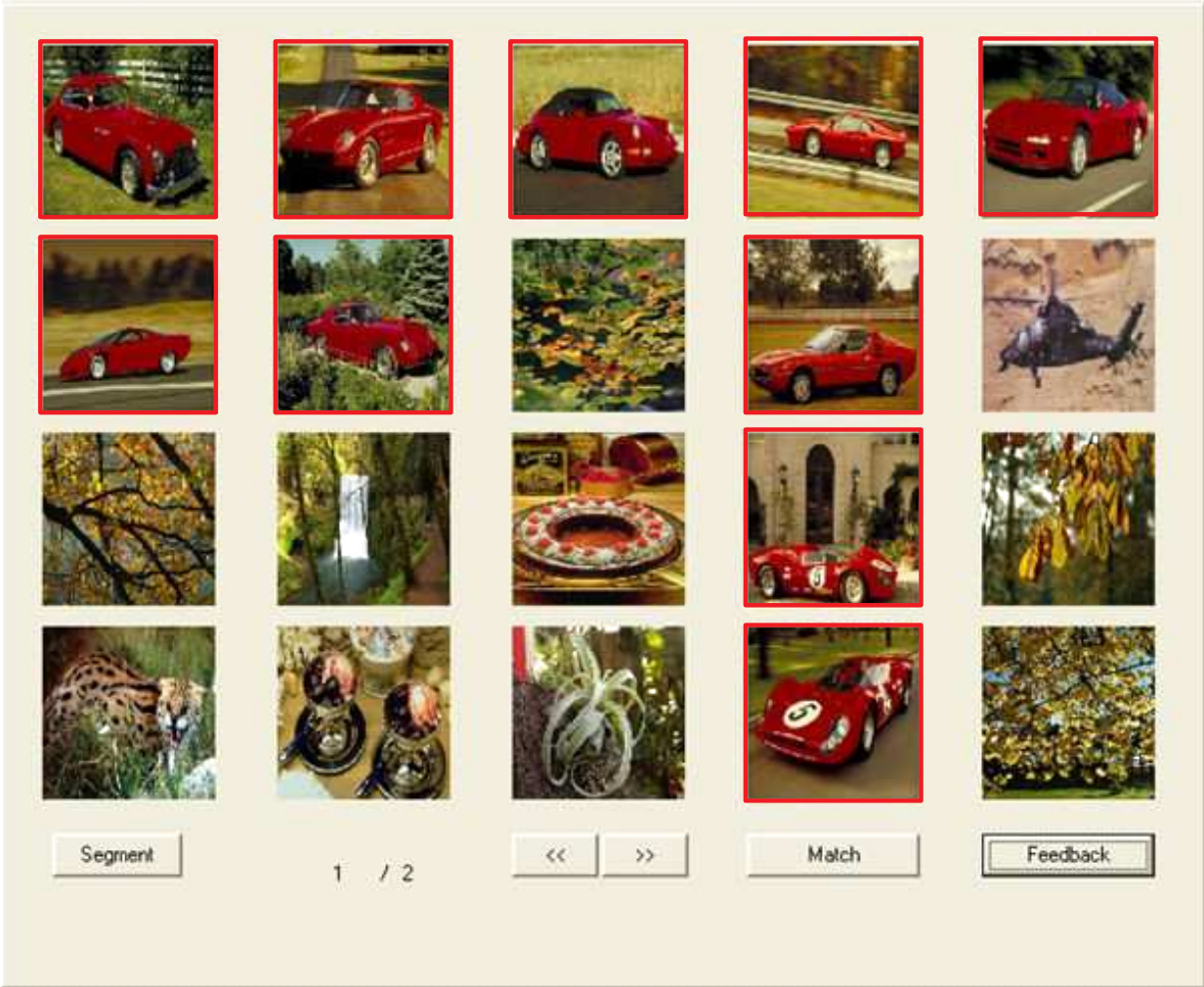


Figure 16. The retrieval results by our integrated region-based relevance feedback for the first iteration.

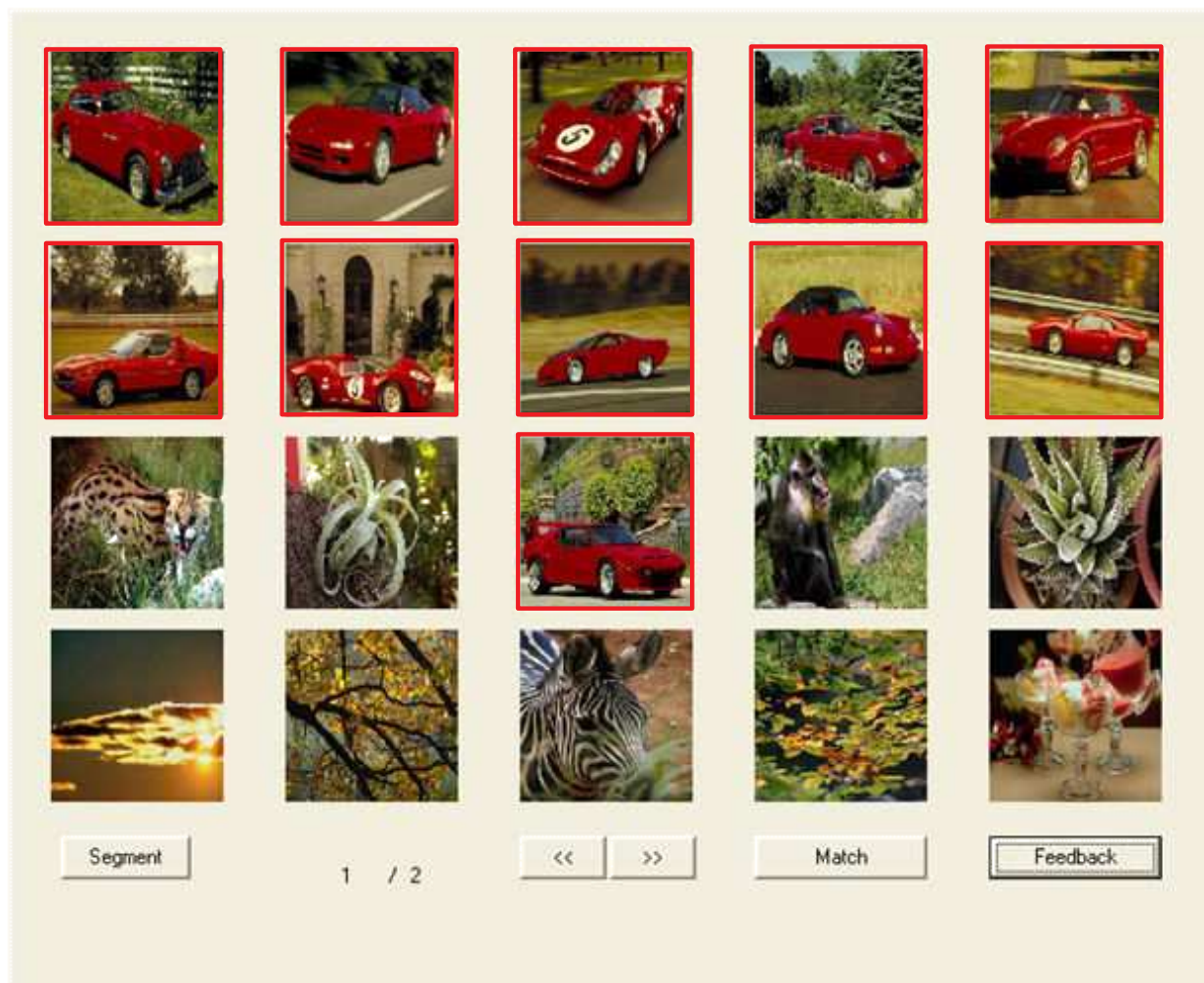


Figure 17. The retrieval results by our integrated region-based relevance feedback for the second iteration.

7. Conclusion

The conventional existing region-based relevance feedback approaches work well in some specified applications; however, their performances depend on the accuracy of segmentation techniques. To solve this problem, we have introduced a novel region-based relevance feedback for image retrieval with the modified dominant color descriptor. The term “specified area”, which combines main objects and irrelevant regions in image, has been defined for compensating the inaccuracy of segmentation algorithm. In order to manipulate the optimal query, we have proposed the similarity matrix model to form the salient region sets. Our integrated region-based relevance feedback approach contains relevance information including pseudo query image I^+ and relevant images set R_s , which are capable to reflect the user's query perception. Experimental results indicate that the proposed technique achieves precise results in general-purpose image database.

Class	Initial query	The 1st feedback query	The 2nd feedback query
1	0.28	0.465	0.635
2	0.56	0.785	0.845
3	0.31	0.53	0.535
4	0.8375	0.85	0.9
5	0.19	0.275	0.32
6	0.255	0.355	0.385
7	0.2	0.29	0.3
8	0.165	0.235	0.245
9	0.73	0.985	1
10	0.345	0.525	0.625
11	0.23	0.345	0.4
12	0.835	1	1
13	0.33	0.52	0.63
14	0.235	0.38	0.4
15	0.655	0.885	0.98
16	0.435	0.625	0.705
17	0.365	0.465	0.515
18	0.235	0.275	0.275
19	0.32	0.505	0.59
20	0.34	0.59	0.635
21	0.37	0.76	0.865
22	0.22	0.355	0.495
23	0.15	0.21	0.225
24	0.31	0.46	0.565
25	0.25	0.43	0.465
26	0.38	0.515	0.61
27	0.245	0.34	0.395
28	0.385	0.415	0.46
29	0.195	0.325	0.41
30	0.4125	0.8	0.8875
31	0.3	0.51	0.61
Avg.	0.357097	0.51629	0.577661

Table 2. Comparisons of ARR performance with different iterations by our proposed integrated region-based relevance feedback approach.

Class	Initial query	The 1st feedback query	The 2nd feedback query
1	0.735	0.399	0.306
2	0.624	0.395	0.326
3	0.741	0.519	0.503
4	0.246	0.135	0.118
5	0.745	0.694	0.643
6	0.744	0.643	0.581
7	0.783	0.721	0.633
8	0.762	0.578	0.537
9	0.215	0.155	0.132
10	0.745	0.571	0.553
11	0.794	0.619	0.557
12	0.331	0.156	0.144
13	0.683	0.591	0.517
14	0.807	0.728	0.709
15	0.514	0.256	0.161
16	0.687	0.559	0.416
17	0.712	0.579	0.554
18	0.836	0.81	0.798
19	0.763	0.512	0.438
20	0.699	0.548	0.488
21	0.716	0.311	0.293
22	0.805	0.664	0.581
23	0.851	0.809	0.797
24	0.725	0.691	0.556
25	0.782	0.645	0.623
26	0.699	0.587	0.503
27	0.791	0.688	0.628
28	0.642	0.613	0.561
29	0.851	0.687	0.649
30	0.662	0.321	0.287
31	0.779	0.587	0.514
Avg.	0.692548	0.541	0.48729

Table 3. Comparisons of ANMRR performance with different iterations by our proposed integrated region-based relevance feedback approach.

Acknowledgement

This work was supported by the National Science Counsel of Republic of China Granted NSC. 97-2221-E-214-053-.

Appendix

BV : Boolean model, which is used to determine whether the segmented region R belongs to the background or foreground.

F : dominant color descriptor

D^2 : similarity measure (dominant color descriptor)

I_{R^i} : the i th non-overlapping region in I

R_{DCD} : dominate color descriptor (DCD) of a segmented region R

$R_{LBP_h_k}$: the value of k th bin in LBP histogram

R_S : region-based color similarity

R_S_c : the maximum similarity between two regions in similar color percentage

R_S_{poa} : similarity of the area percentage

R_S_T : region based texture similarity

$R_{background}$: defined background based on foreground assumption

$R_{foreground}$: defined foreground based on foreground assumption

R_{poa} : the percentage of area for region R in the image

R_s : relevant image set

$R_{texture}$: texture feature of region R

$a_{i,j}$: similarity coefficient between two color clusters (dominant color descriptor)

c_i : dominant color vector (dominant color descriptor)

$d_{i,j}$: Euclidean distance between two color clusters (dominant color descriptor)

p_i : percentage of each dominant color (dominant color descriptor)

Author details

Nai-Chung Yang¹, Chung-Ming Kuo¹ and Wei-Han Chang²

1 Department of Information Engineering, I-Shou University Tahsu, Kaohsiung, Taiwan R.O.C.

2 Department of Information Management, Fortune Institute of Technology, Daliao, Kaohsiung, Taiwan R.O.C.

References

- [1] A. Gaurav, T. V. Ashwin and G. Sugata, An image retrieval system with automatic query modification, *IEEE Trans. Multimedia*, 4(2) (2002) 201-214.
- [2] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore, and H. Shin, An efficient color representation for image retrieval, *IEEE Trans. Image Process.*, 10(1) (2001) 140–147.
- [3] Y. Yang, F. Nie, D. Xu, J. Luo, Y. Zhuang and Y. Pan, A Multimedia Retrieval Framework Based on Semi-Supervised Ranking and Relevance Feedback, *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(4) (2012) 723-742.
- [4] M.Y. Fang, Y.H. Kuan, C.M. Kuo, C.H. Hsieh, "Effective image retrieval techniques based on novel salient region segmentation and relevance feedback," *Multimedia Tools and Applications*, 57(3), (2012) 501-525.
- [5] S. Murala, R. P. Maheshwari and R. Balasubramanian, Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval, *IEEE Trans. Image Process.*, 21(5) (2012) 2874-2886.
- [6] G. Ciocca and R. Schettini, Content-based similarity retrieval of trademarks using relevance feedback, *Pattern Recognit.*, 34(8) (2001) 1639-1655.
- [7] X. He, O. King, W. Y. Ma, M. Li, and H. J. Zhang, Learning a Semantic Space from User's Relevance Feedback for Image Retrieval, *IEEE Trans. Circ. Syst. Vid. technol.*, 13(1) (2003) 39-48.
- [8] F. Jing, M. J. Li, H. J. Zhang and B. Zhang, Relevance Feedback in Region-Based Image Retrieval, *IEEE Trans. Circ. Syst. Vid. technol.*, 14(5) (2004) 672-681.
- [9] T. P. Minka and R. W. Picard, Interactive learning using a society of models, *Pattern Recognit.*, 30(4) (1997) 565–581.
- [10] K. Vu, K. A. Hua and W. Tavanapong, Image Retrieval Based on Regions of Interest, *IEEE Trans. Knowl. Data Eng.*, 15(4) 2003 1045-1049.

- [11] R. Yong, T. S. Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval, *IEEE Trans. Circ. Syst. Vid. technol.*, 8(5) (1998) 644 – 655.
- [12] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas and P. N. Yianilos, The Bayesian Image Retrieval System, *PicHunter: Theory, Implementation, and Psychophysical Experiments*, *IEEE Trans. Image Process.*, 9(1) (2000) 20–37.
- [13] Y. H. Kuo, W. H. Cheng, H. T. Lin and W. H. Hsu, Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement, *IEEE Trans. Multimedia*, 14(9) (2012) 1079-1090.
- [14] C. Gao, X. Zhang and H. Wang, A Combined Method for Multi-class Image Semantic Segmentation, *IEEE Transactions on Consumer Electronics*, 58(2) (2012) 596-604.
- [15] J. J. Chen, C. R. Su, W. L. Grimson, J. L. Liu and D. H. Shiue, Object Segmentation of Database Images by Dual Multiscale Morphological Reconstructions and Retrieval Applications, *IEEE Trans. Image Process.*, 21(2) (2012) 828-843.
- [16] A. Pardo, Extraction of semantic objects from still images, *IEEE International Conference on Image Processing (ICIP '02)*, vol. 3, 2002, pp. 305 -308.
- [17] A. Yamada, M. Pickering, S. Jeannin and L. C. Jens, MPEG-7 Visual Part of Experimentation Model Version 9.0-Part 3 Dominant Color, *ISO/IEC JTC1/SC29/WG11/N3914*, Pisa, Jan. 2001.
- [18] A. Mojsilovic, J. Hu and E. Soljanin, Extraction of Perceptually Important Colors and Similarity Measurement for Image Matching, Retrieval, and Analysis, *IEEE Trans. Image Process.*, 11 (11) (2002) 1238-1248.
- [19] S. P. Lloyd, Least Squares Quantization in PCM, *IEEE Trans. Inform. Theory*, 28(2) (1982) 129-137.
- [20] N. C. Yang, W. H. Chang, C. M. Kuo and T. H. Li, A Fast MPEG-7 Dominant Color Extraction with New Similarity Measure for Image Retrieval, *Journal of Visual Communication and Image Representation*, 19(2) (2008) 92-105.
- [21] Y. W. Lim and S. U. Lee, On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques, *Pattern Recognit.*, 23(9) (1990) 935-952.
- [22] S. Kiranyaz, M. Birinci and M. Gabbouj, Perceptual color descriptor based on spatial distribution: A top-down approach, *Image and Vision Computing* 28(8) (2010) 1309-1326.
- [23] P. Scheunders, A genetic approach towards optimal color image quantization, *IEEE International Conference on Image Processing (ICIP'96)*, vol. 3, 1996, pp. 1031-1034.
- [24] W. Chen, W. C. Liu and M. S. Chen, Adaptive Color Feature Extraction Based on Image Color Distributions, *IEEE Trans. Image Process.*, 19(8) (2010) 2005-2016.

- [25] Text of ISO/IEC 15 938-3, "Multimedia Content Description Interface—Part 3: Visual. Final Committee Draft," ISO/IEC/JTC1/SC29/WG11, Doc. N4062, Mar. 2001.
- [26] N. C. Yang, C. M. Kuo, W. H. Chang and T. H. Lee, A Fast Method for Dominant Color Descriptor with New Similarity Measure, 2005 International Symposium on Communication (ISCOM2005), Paper ID: 89, Nov. 20-22, 2005.
- [27] W. Y. Ma, Y. Deng and B. S. Manjunath, Tools for texture/color based search of images, SPIE Int. Conf. on Human Vision and Electronic Imaging II, 1997, pp. 496- 507.
- [28] A. Mojsilovic, J. Kovacevic, J. Hu, R. J. Safranek and S. K. Ganapathy, Matching and Retrieval Based on the Vocabulary and Grammar of Color Patterns, IEEE Trans. Image Process., 9 (1) (2000) 38-54.
- [29] T. Ojala and M. Pietikainen, Unsupervised texture segmentation using feature distributions, Pattern Recognit., 32(9) (1999) 447-486.
- [30] N. Abbadeni, Computational Perceptual Features for Texture Representation and Retrieval, IEEE Trans. Image Process., 20(1) (2011) 236-246.
- [31] M. Broilo, and F. G. B. De Natale, A Stochastic Approach to Image Retrieval Using Relevance Feedback and Particle Swarm Optimization, IEEE Trans. Multimedia, 12(4) (2010) 267-277.
- [32] W. C. Kang and C. M. Kuo, Unsupervised Texture Segmentation Using Color Quantization And Color Feature Distributions, IEEE International Conference on Image Processing (ICIP '05), vol. 3, 2005, pp. 1136 - 1139.
- [33] S. K. Weng, C. M. Kuo and W. C. Kang, Color Texture Segmentation Using Color Transform and Feature Distributions, IEICE TRANS. INF. & SYST., E90-D(4) (2007) 787-790.
- [34] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan and A. Yamada, Color and Texture Descriptors, IEEE Trans. Circ. Syst. Vid. technol., 11(6) (2001) 703-714.

