Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition 38 (2005) 2449–2465
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

A novel fusion approach to content-based image retrieval

Xiaojun Qi*, Yutao Han

Computer Science Department, Utah State University, Logan, UT 84322-4205, USA

Received 29 April 2004; received in revised form 14 April 2005; accepted 14 April 2005

Abstract

This paper proposes a novel fusion approach to content-based image retrieval. In our retrieval system, an image is represented by a set of color-clustering-based segmented regions and global/semi-global edge histogram descriptors (EHDs). As a result, the resemblance of two images is measured by an overall similarity fusing both region-based and global/semi-global-based image level similarities. In our approach, each segmented region corresponds to an object or parts of an object and is represented by two sets of fuzzified color and texture features. A fuzzy region matching scheme, which allows one region to match several regions, is then incorporated to address the issues associated with the color/texture inaccuracies and segmentation uncertainties. The matched regions, together with the simple semantics for determining the relative importance of each region, are further used to calculate the region-based image level similarity. The global/semi-global EHDs are also incorporated into our retrieval system since they do not depend on the segmentation results. These EHDs not only decrease the impact of inaccurate segmentation and but also reduce the possible retrieval accuracy degradation after applying the fuzzy approach to the accurate segmentation for images with distinctive and relevant scenes. The Manhattan distance is used to measure the global/semi-global image level similarity. Finally, the overall similarity is computed as a weighted combination of regional and global/semi-global image level similarity measures incorporating all features. Our proposed retrieval approach demonstrates a promising performance for an image database of 5000 general-purpose images from COREL, as compared with some current peer systems in the literature.

Keywords: Content-based image retrieval; Similarity measure; Image segmentation; Edge histogram descriptors; Fuzzy features; Fuzzy region matching

1. Introduction

The evolution of the World Wide Web (WWW), the advances in computer technologies, and the recent information explosion in multimedia content have produced an enormous number of digital data archives in a variety of application domains such as entertainment, commerce, education, biomedicine, military, and web image classification and searching. Correspondingly, many techniques have been developed for fast indexing, retrieval, and manipulation of the digital images. However, traditional text-based (keyword-based) image retrieval methods do not work well as the image contents cannot be accurately described by human language and different persons may perceive the same image differently. In addition, it is time-consuming to manually annotate each image due to the enormous size of image databases. As a result, content-based image retrieval (CBIR) has been brought to the forefront since the early 1990s [1].

CBIR is an important alternative and complement to traditional text-based image searching and can greatly enhance the accuracy of the information being returned. It aims

* Corresponding author. Tel.: +1 435 797 8155; fax: +1 435 797 3265.
E-mail address: xqi@cc.usu.edu (X. Qi).

0031-3203/$30.00
doi:10.1016/j.patcog.2005.04.005

to develop an efficient visual-content-based technique to search, browse, and retrieve relevant images from large-scale digital image collections. Most proposed CBIR techniques automatically extract low-level features (e.g., color, texture, shapes, and spatial layout of objects) to measure the similarities among images by comparing the feature differences [1,2].

In this paper, we present a novel fusion framework for general-purpose image retrieval based on both regional color and texture features and the global and semi-global edge histogram descriptors (EHDs) expanded from the normative EHD for MPEG-7 [3,4]. We have developed a fast and automatic statistical-clustering-based segmentation method that provides reasonable segmentation results where each segmented region generally corresponds to an object or parts of an object. Each region-based feature is then individually fuzzified to incorporate the segmentation-related uncertainties into the retrieval algorithm. The global and semi-global EHDs, which do not depend on segmentation, have been further utilized to decrease the impact of inaccurate region segmentation and reduce the possible retrieval accuracy degradation for the accurate segmentation cases. The resemblance of two images is then defined as the overall similarity between two families of region-based fuzzy color and texture features, global and semi-global EHDs. This overall similarity is quantified by a computationally efficient distance metric which integrates properties of all fuzzy regions in the images, the normalized area percentage difference between matched regions, the normalized distance from the region center to the image center, different contributions from color and texture features, and different contributions from regional, global, and semi-global features. It is noteworthy that the objective of the proposed method is to match entire images, including backgrounds and main objects. It may not perform very well for situations where the goal is to find images containing a specific object where the background is not important.

The remainder of the paper is organized as follows. Section 2 briefly reviews the related work. Section 3 describes the general framework of our proposed retrieval system. Section 4 illustrates the experimental results. Section 5 concludes with a brief discussion of our approach and some proposed directions for future work.

2. Related work

In the CBIR system, the relevance between a query and any target image is ranked according to a similarity measure computed from the visual features. In general, the similarity comparison is performed either globally based on visual content descriptors including color, texture, and/or shape features or locally based on visual content descriptors derived from decomposed regions of the images.

Global visual content descriptors have been widely applied to many general-purpose image retrieval systems such as IBM QBIC [5], MIT Photobook [6], Virage System [7], Columbia VisualSEEK [8], Candid [9], Chabot [10], and MARS [11]. The commonly used effective global features include color histograms [12,13], Tamura features [14], Wold features [15,16], Gabor filter features [17], wavelet transform features [18–20], Fourier descriptors [21,22], and the like. However, several major drawbacks associated with these global features are:

1. They lack information about the spatial feature distribution.
2. They are sensitive to intensity variations and distortion.
3. They fail to narrow down the semantic gap (the difference between users' high-level query concepts for CBIR and the low-level features that are used for the querying) due to their limited description power based on objects.

Consequently, some approaches such as color coherence vector [23], color correlogram [24], spatial color histogram [25], and spatial chromatic histogram [26] have been proposed to overcome the spatial limitation by incorporating spatial information in the descriptor. However, most approaches have been proposed to focus on the local features to extend the capability of CBIR so users can retrieve images based on potential interest regions. These local-feature-based approaches are likely to provide a big step towards the semantics-based retrieval since human perceptions of certain visual contents could potentially be associated with interesting classes of objects/regions or semantic meanings of objects/regions in the image. Among various local-feature-based approaches, the region-based image retrieval methods have been widely studied since they have a strong correlation with real-world objects. In region-based image retrieval, each image is first segmented into homogenous regions and features for each region are extracted. The overall similarity between two images is calculated based on all the corresponding region-based features. Several important region-based retrieval systems are briefly reviewed here.

The UCSB NeTra system [27] uses an edge flow model to segment the image by using three user specified parameters (i.e., image features to be used, the preferred scale to localize the desired image boundaries, and the expected number of regions). The segmentation-based color, texture, shape, and spatial location features are further utilized to search and retrieve similar regions from the database. In Ref. [28], the same edge flow model is applied to segment the image. The dominant colors for each segmented region are obtained and a dominant-color-based similarity score is computed to measure the difference between two regions for retrieval. The Berkeley Blobworld system [29] applies the expectation-maximization algorithm on color, texture, and position features to segment the image into coherent regions. The new segmentation-based joint color–texture–position feature (blob representation) is then used for retrieval.

These three systems require significant user interaction in defining or selecting individual regions for image similarity

comparison. That is, to query an image, users must select interest regions for a similarity evaluation and images whose regional features are close to the features of the specified regions are retrieved. Such systems may be a burden to non-professional users especially when the entire scenes in the image are indistinctive and irrelevant or when more detailed feature information other than Boolean (yes/no) is required from users as mandatory in the NeTra system. To totally free the burden on the users' side, several systems have been proposed to develop similarity measures that combine information from all of the automatically segmented regions to compare the overall image similarity without any user control.

Ardizzoni et al. [30] use color and texture features captured from wavelet coefficients for both segmentation and retrieval. Similarity between images is assessed by combining the regional color-and-texture similarities computed from the Bhattacharyya distance at the image level. Li et al. [31] and Chen et al. [32] use color features computed in the LUV color space and texture features calculated from the wavelet coefficients for each 4×4 block to segment the image into regions. The region-based color, texture, and shape features are utilized for retrieval. In Ref. [31], an integrated region matching (IRM) scheme, which allows a region in one image to match against all regions from another image, is proposed to decrease the impact of inaccurate region segmentation. In Ref. [32], a unified feature matching (UFM) scheme is proposed, where region-based multiple fuzzy feature representations and fuzzy similarity measures are used to improve the retrieval accuracy. In Ref. [33], the region-growing approach is used to segment an image in the HSV color space. The color, shape, and position features of each region are extracted to represent the content of an image. A probabilistic framework is then utilized to simultaneously model both the first-order region properties and the second-order spatial relationships of all the regions in the image for similarity evaluation. Hsieh and Grimson [34] propose to represent each image by a set of templates and their spatial relations. Each template is characterized by a set of dominant segmented regions, which reflect different appearances of an object at different conditions. The visual similarity is computed based on the joint color–texture–geometric feature difference between each region and the template. It is further combined with the relation similarity to obtain the final similarity.

In summary, all current CBIR systems use either global features or local features for retrieval, but not both. We observe that the segmentation-related uncertainties always exist due to inaccurate image segmentation especially for the images with indistinctive and irrelevant scenes. The fuzzy representation [32] of imprecise local features can somehow improve the retrieval accuracy and robustness against inaccurate image segmentation. However, it is mathematically difficult to find an effective and efficient fuzzy feature representation applicable to all kinds of images. Additional regional spatial relationships [33,34] may improve the retrieval accuracy. But humans can often correctly identify the object of interest at any possible location even when its boundary is blurry or when it is occluded by other objects. This fact indicates that the regional spatial relationships are not important factors for accurate retrieval. Consequently, our proposed novel retrieval system will fuse the global and semi-global EHDs and the region-based fuzzy color and texture features for better retrieval accuracy without using any spatial relationships. These global and semi-global EHDs are extracted independent of segmentation and therefore provide more robustness against any issue associated with segmentation.

3. Proposed CBIR system

The block diagram of our proposed novel fusion approach to CBIR is shown in Fig. 1. The first step of our algorithm is to segment an image into reasonable coherent regions based on exclusive color features. Image indexing and retrieval is then taken based on the semi-global and global EHDs and weighted independent region-based fuzzy color and texture features incorporating the normalized area percentage difference between matched regions and the normalized distance from the region center to the image center.

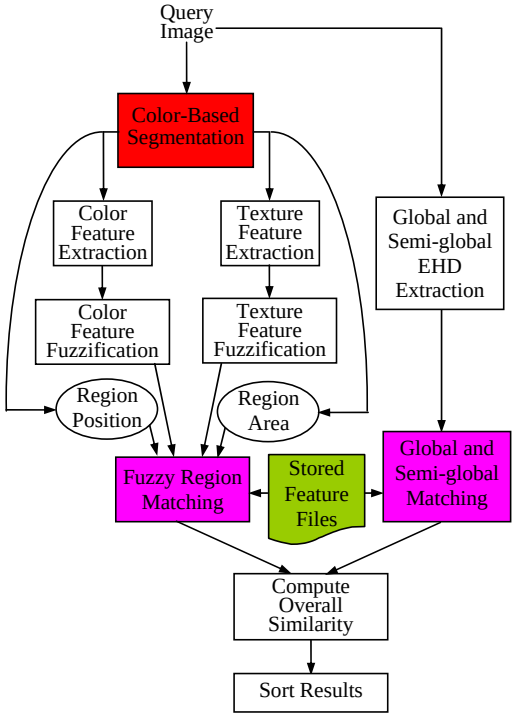


Fig. 1. Block diagram of our proposed CBIR system.

3.1. Color-clustering-based image segmentation

Most segmentation methods use either both color and texture features [29–32] or complicated edge flow models [27,28] or region-growing methods [33] to divide the image into several homogenous regions. Unlike these methods, we propose a fast and automatic color-clustering-based method to segment an image since color features are considered more important than other features in the domain of natural images [34]. This proposed method provides segmentations that are good enough for retrieval since the accurate segmentation is not required in image retrieval and our weighted fuzzy matching scheme and our segmentation-independent global and semi-global EHDs will decrease the impact of the inaccurate segmentation. The proposed segmentation method has the following advantages:

1. It dramatically reduces the computational cost since texture features are excluded, where the complicated statistical computation is usually involved.
2. It is fully automatic due to the adaptive learning nature of the clustering method.
3. It is robust in the sense that each segmented region generally corresponds to an object or parts of an object.

To segment an image into coherent regions, the image is first divided into non-overlapping square image-blocks and a color feature vector is extracted for each image-block. The size of the image-block is chosen to be 2×2 because fine details can be preserved. The LUV color space is used to extract color features for each image-block since the perceptual color abilities of the human visual system are proportional to the numerical difference in the LUV color space. The image-block-based color features are the means of each color component in the corresponding image-block.

After obtaining the color features for all 2×2 image-blocks, a statistical-clustering method, an unsupervised K-Means algorithm [35], is used to cluster these color features into several groups, where each group in the color feature space corresponds to one spatial region in the image space. The learning nature of this K-Means algorithm enables an automatic and iterative segmentation process, which accommodates the fact that the number of regions in an image is unknown before segmentation. Suppose there are M blocks $\{x_i = 1, \ldots, M\}$ for each image. The goal of the K-Means algorithm is to group each of M blocks into one of the K clusters, whose cluster centers are $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_K$, such that

$$D_K = \sum_{i=1}^{M} \min_{1 \leqslant j \leqslant K} (x_i - \hat{x}_j)^2 \quad (1)$$

is minimized. We start the segmentation with $K = 2$ and adaptively and gradually increase K until a reasonable segmentation is yielded. This iterative segmentation process will stop at the ith iteration (i.e., the total number of clusters K equals i) if D_i is less than a threshold or $|D_i - D_{i-1}|$ is less than another threshold. These two thresholds are empirically chosen so a reasonable segmentation can be achieved on all 5000 images from our test database.

Fig. 2 shows the intermediate segmentation results of two sample images from our test database by adaptively and gradually increasing the number of regions K. It clearly demonstrates that the segmentation results become better as the number of regions increases and the final segmentation results are reasonable for further retrieval. It also illustrates that each segmented region does not necessarily be connected, as shown in several separated red and black portions in the first sample image A and several separated black, gray, and brown portions in the second sample image B. However, this non-connected property preserves the natural clustering of an object or parts of an object in general-purpose images, which is required for accurate retrieval. That is, a uniform object will be grouped into a connected segmentation. A non-uniform object will be segmented into several non-connected clusters based on color and its distribution as shown in Fig. 2.

3.2. Fuzzy feature representation and fuzzy region matching

After segmentation, a set of regions will be obtained. In order to fully represent each segmented region, two sets of features including color and texture properties are used as underlying primitives to represent this region.

3.2.1. Region-based color and texture features

The color feature $\vec{f}_j^{\,c}$ for each region j is the K-Means' cluster center of this region, i.e., the average of color features of all the image-blocks in this region.

The texture feature $\vec{f}_j^{\,t}$ for each region j is computed as follows:

Step 1: Generate a "texture template" image by keeping all the pixels in region j intact and setting all the pixels outside region j as white.

Step 2: Convert this "texture template" image to the gray-scale image.

Step 3: Apply a two-level Haar wavelet transform to the gray-scaled "texture template" image obtained from Step 2.

Step 4: Calculate the average energy in each high frequency band (i.e., low–high, high–low, and high–high band) of level 1 and level 2 wavelet decompositions.

Step 5: Construct a texture feature vector $\vec{f}_j^{\,t}$ for region j by concatenating the six average energy values in an appropriately scaled measurement unit.

The computational cost of deriving the regional representative texture feature is minimal compared to most methods [29–32], where texture features of all the image-blocks are calculated and then averaged based on the segmentation results. Moreover, our proposed regional representative texture feature captures more accurate edge distribution along the region boundary and within the region than other methods which always generate small values due to the homogeneity of each merged image-block.

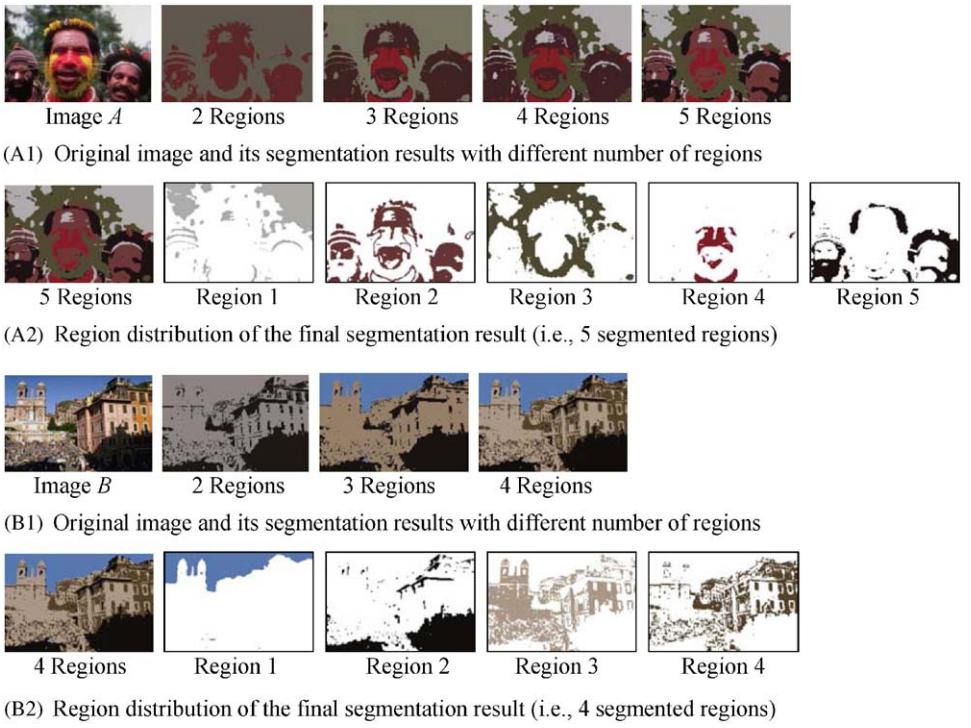


Fig. 2. Segmentation results by the unsupervised K-Means clustering algorithm.

Since these color and texture properties are constructed in different ways and have different measurement units, they are utilized as two sets of features for representing each region.

3.2.2. Region-based fuzzy features and similarity

Unlike most region matching schemes [27–30] which directly compare the regional features between two images, we independently fuzzify each regional color and texture feature and apply fuzzy matching on these fuzzified features to address the imperfect segmentation issue.

In general, any membership function with a smooth transition between 0 and 1 can be selected for fuzzification. The Cauchy function [36] is utilized in our approach among some commonly used cone and exponential functions due to its good expressiveness and its high computational efficiency. One example of the Cauchy function is illustrated in Fig. 3. The corresponding Cauchy function for this plot is defined as

$$C(\vec{x}) = \frac{1}{1+\left(\frac{\|\vec{x}-\vec{f}\|}{d}\right)^{\alpha}} = \frac{1}{1+\left(\frac{\|\vec{x}-\vec{0}\|}{30}\right)^{\alpha}} = \frac{1}{1+\left(\frac{\|\vec{x}\|}{30}\right)^{\alpha}}, \quad (2)$$

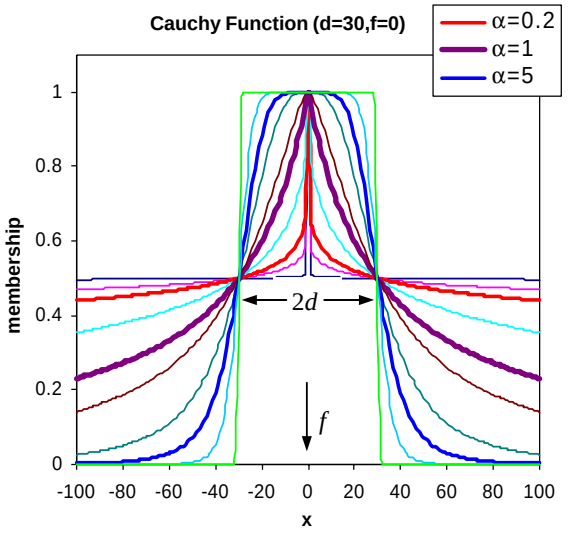


Fig. 3. One example of the Cauchy function with fixed d and $\vec{f}$ where $d = 30$ and $\vec{f} = \vec{0}$.

where d represents the width of the function; $\vec{f}$ represents the center location (point) of the fuzzy set; α represents the shape (or smoothness) of the function; $\vec{x}$ represents the feature to be fuzzified; and $C(\vec{x})$ is the degree of membership.

This Cauchy function plot clearly shows that the fuzziness increases as α decreases for a fixed d and the Cauchy function becomes sharper within its center region $[-d, d]$ and flatter outside. It also demonstrates that the farther a feature $\vec{x}$ is away from the region center (i.e., the representative feature which is shown as $\vec{f} = \vec{0}$ in Fig. 3), the lower its degree of membership. This membership value illustrates the degree of wellness that the $\vec{x}$ characterizes the region, and thus models the segmentation-related uncertainties.

To fuzzify the color feature in the region j, the Cauchy function defined as

$$C(\vec{x}^c_{j,m}) = \frac{1}{1 + \left(\frac{\|\vec{x}^c_{j,m} - \vec{f}^c_j\|}{d} \right)^{\alpha}} \tag{3}$$

is used to calculate the membership of the color feature $\vec{x}^c_{j,m}$ of each 2×2 image-block m in region j, where:

- $j \in \{1, \ldots, C\}$, C is the number of regions obtained from segmentation;
- $m \in \{1, \ldots, B_j\}$, B_j is the number of image-blocks in region j;
- $\vec{f}^c_j$ is the color feature of region j;
- d is the average distance between color features of all regions in the image and is calculated as

$$d = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{k=i+1}^{C} \|\vec{f}^c_i - \vec{f}^c_k\|; \tag{4}$$

- α represents the shape of the Cauchy function. It is chosen to be 1 in our implementation because it involves the least computation and its Cauchy function shape (shown as a bold purple line in Fig. 3) roughly corresponds to the gradual transition of region boundaries in an image with respect to the region center.

The fuzzy-color-based region similarity s^c between two regions u and v is defined as

$$s^c(u, v) = \sup[C(\vec{x}^c_{u \cap v, m})], \tag{5}$$

where $\vec{x}^c_{u \cap v, m}$ represents any image-block m in two regions u or v, and $C(\vec{x}^c_{u \cap v, m}) = \min[C_u(\vec{x}^c_m), C_v(\vec{x}^c_m)]$ with $C_u(\vec{x}^c_m)$ and $C_v(\vec{x}^c_m)$ being the membership values computed by Eq. (3). Because of the unimodal property of the Cauchy function, the fuzzy similarity (5) between the fuzzified color features of a region u in image A and a region v in image B can be computed in the reduced complexity form, whose proof is shown in the appendix:

$$S^c(u, v) = \frac{(d^c_A + d^c_B)^{\alpha}}{(d^c_A + d^c_B)^{\alpha} + \|\vec{f}^c_u - \vec{f}^c_v\|^{\alpha}}, \tag{6}$$

where α represents the shape of the Cauchy function, $\vec{f}^c_u$ is the color feature of region u in image A, $\vec{f}^c_v$ is the color feature of region v in image B and d^c_A and d^c_B are the average distances between color features of all regions in images A and B, respectively.

This simplified fuzzy region similarity formula (6) indicates that the similarity between two fuzzified regions can be quickly and efficiently computed by using the non-fuzzified features (i.e., the average distances d^c_A and d^c_B, and color features $\vec{f}^c_u$ and $\vec{f}^c_v$) and the fuzzy function shape parameter α. That is, using these average distances and color features eliminates the need to compute the memberships of all image-blocks. This is therefore another reason for the Cauchy function to be chosen in our fuzzification.

Similarly, the fuzzy similarity between the fuzzified texture features of region u in image A and region v in image B can be computed by

$$S^t(u, v) = \frac{(d^t_A + d^t_B)^{\alpha}}{(d^t_A + d^t_B)^{\alpha} + \|\vec{f}^t_u - \vec{f}^t_v\|^{\alpha}}, \tag{7}$$

where all the notations have the same meanings as the ones used in Eq. (6), except that the texture feature is used instead of the color feature.

The overall similarity between two fuzzy regions u and v is calculated by

$$S(u, v) = \lambda_1 S^c(u, v) + (1 - \lambda_1) S^t(u, v), \tag{8}$$

where $S^c(u, v)$ and $S^t(u, v)$ are calculated by using Eqs. (6) and (7), and λ_1 determines the contribution of color features in measuring the similarities. Since color is more important than texture in natural images [34], λ_1 is empirically set to be 0.85 in our system.

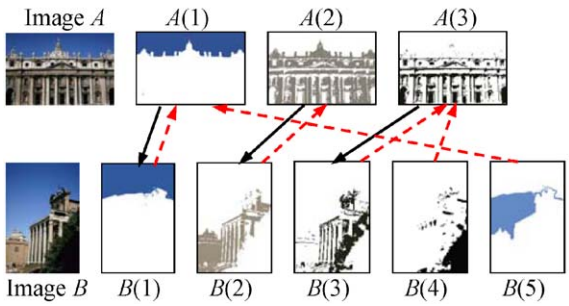
3.2.3. Fuzzy region matching

Since the resemblance of two images is conveyed through the similarities between regions from both images, the image-level similarity can be measured using region-level similarities. A fuzzy region matching scheme is used in our approach to measure all possible region similarities due to the imperfect segmentation which may cause a region to contain an object or parts of an object as illustrated in Fig. 2. That is, the fuzzy region matching scheme allows that a region in one image corresponds to several regions in another image and vice versa.

The best matched region in image B for an individual region u in image A is determined by the fuzzy region matching:

$$S(u, v_{best}) = \max\{S(u, v)\} \quad \forall v \in B, \tag{9}$$

where $S(u, v)$ is obtained by using Eq. (8). Based on Eq. (9), the region in image B, which yields the largest overall region similarity, is considered to be the most similar region for region u in image A. Its region–color-based similarity S^c and region–texture-based similarity S^t are saved in the similarity vectors $\boldsymbol{L}^c$ and $\boldsymbol{L}^t$ for calculating the image-level similarity. The same process will be applied to all the other regions in image A and all the regions in image B. As a



Matched Region Pairs		L^c	L^t
$A \rightarrow B$	$A(1) \rightarrow B(1)$	0.968	0.930
	$A(2) \rightarrow B(2)$	0.942	0.713
	$A(3) \rightarrow B(3)$	0.891	0.841
$B \rightarrow A$	$B(1) \rightarrow A(1)$	0.968	0.930
	$B(2) \rightarrow A(2)$	0.942	0.713
	$B(3) \rightarrow A(3)$	0.891	0.841
	$B(4) \rightarrow A(3)$	0.833	0.646
	$B(5) \rightarrow A(1)$	0.845	0.966

Fig. 4. Illustration of the fuzzy region matching scheme.

result, the lengths of $\boldsymbol{L}^c$ and $\boldsymbol{L}^t$ will be the total number of regions in images A and B. It is easily observed that two identical images will have the same segmentation results and all elements in $\boldsymbol{L}^c$ and $\boldsymbol{L}^t$ will be 1's.

This fuzzy region matching scheme is illustrated in Fig. 4. Each of the three regions in image A is matched to the most similar region in image B by Eq. (9). Similarly, each of the five regions in image B is matched to the most similar region in image A as shown in red dash arrows. The best similarity values in terms of both color and texture are stored in $\boldsymbol{L}^c$ and $\boldsymbol{L}^t$, respectively. It clearly shows that several regions in image B can be matched to the same region in image A, which is the important property of the proposed fuzzy matching scheme.

3.3. Global and semi-global edge histogram descriptors (EHDs) extraction

The fuzzy feature representation and fuzzy region matching have been developed in our retrieval system to address the imperfect segmentation issue. However, choosing a proper fuzzy membership function is an application-dependent problem. It is also mathematically difficult to find an effective and efficient fuzzy membership function with the appropriate parameters applicable to all kinds of images. Furthermore, possible degradation of the retrieval performance may occur when segmentation tends to be very accurate for images with distinctive and relevant scenes. Therefore, we add the global and semi-global EHDs in our retrieval system. The major advantages of using these EHDs are:

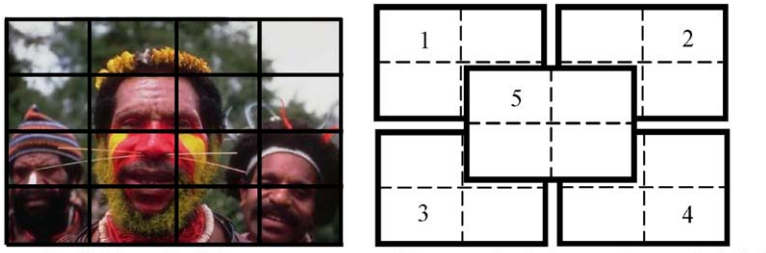
1. Edges are important features to represent the content of an image since human eyes are sensitive to edge features for image perception. They complement the region-based color and texture retrieval from a different perspective.
2. The global and semi-global EHDs do not depend on segmentation and therefore can provide more robustness against any issues associated with segmentation.
3. The global EHD represents the global features of the edge distribution and is invariant to translation, rotation, and scaling (RST).
4. The semi-global EHD represents the global features of the edge distribution in five predefined non-segmented regions and is robust against RST within each of these chosen regions.

The EHD is one of the three normative texture descriptors proposed for MPEG-7 [3,4,37]. It captures the spatial distribution of edges in an image and has been proven to be useful for image retrieval, especially for natural images with non-uniform textures and clip art images. Five types of edges, namely vertical, horizontal, 45° diagonal, 135° diagonal, and non-directional edges (i.e., the edges with no particular directionality), have been utilized to represent the edge orientation in 16 sub-images. These sub-images are obtained by dividing the entire image space into 16 non-overlapping sub-spaces as shown in Fig. 5(a). The EHD in MPEG-7 is therefore a total of 5×16 histogram bins, which represent local edge distribution for each sub-image in the image.

Based on the EHD, we construct global and semi-global edge histograms with various segments of the image to address any segmentation-related issues. The global edge histogram represents the edge distribution for the entire image space and has five bins since five edge types are considered in the calculation. For the semi-global edge histograms, we cluster four connected sub-images to generate five different clusters as shown in Fig. 5(b). The edge distributions of five different edge types are generated for each cluster to construct the semi-global EHD.

3.4. Similarity measure

Unlike other methods exclusively using global features [5–11] or local features [27–34] for image retrieval, we fuse local, global, and semi-global features into an integrated comprehensive feature for more accurate image retrieval. Correspondingly, three types of similarity measures, namely region-based, global, and semi-global similarity measures, will be computed and combined to automatically measure



(a) Definition of sub-images (b) Clusters of sub-images for semi-global histograms

Fig. 5. Global and semi-global edge histogram descriptors.

the overall resemblance between two images.

3.4.1. Local similarity measure at the image level

The image resemblance in terms of the segmented regions is measured by the overall region-based similarity. A weighted overall region-based similarity scheme is used based on the following three observations:

(1) Important objects in an image tend to occupy larger areas near the image center;
(2) Regions adjacent to the image boundary provide more semantic information because semantically similar images always have similar background;
(3) The matched regions between two images tend to have a similar area.

Consequently, the region position and the area difference between two matched regions are incorporated for computing the overall region-based similarity. That is, more weight is assigned to the region that is near the image center or is near the image boundary or has a similar area as the matched region.

The overall region-based image level similarity score S_l is computed based on the fuzzy region matching scheme in terms of both color and texture features:

$$S_l = \lambda_1 S_l^c + (1 - \lambda_1) S_l^t, \tag{10}$$

where λ_1 determines the contribution of color features in measuring the overall image level similarity and is set to be the same value (0.85) as used in Eq. (8); S_l^c and S_l^t are the region-based color and texture similarities, respectively. These two similarities are computed by using the similarity vectors $\boldsymbol{L}^c$ and $\boldsymbol{L}^t$ obtained from the fuzzy region matching, the region position, and the area difference between two matched regions:

$$\begin{aligned} S_l^c &= \vec{w}^T L^c = ((1 - \lambda_2)\vec{w}_a + \lambda_2 \vec{w}_p)^T \boldsymbol{L}^c, \\ S_l^t &= \vec{w}^T L^t = ((1 - \lambda_2)\vec{w}_a + \lambda_2 \vec{w}_p)^T \boldsymbol{L}^t. \end{aligned} \tag{11}$$

Here, $\vec{w}$ is the weight vector related to the region position and the area difference between two matched regions. In particular, $\vec{w}_a$ contains the normalized weights which favor the small normalized area percentage differences between two matched regions. $\vec{w}_p$ contains the normalized weights which favor regions near the image boundary or near the image center. λ_2 adjusts the significance of $\vec{w}_a$ and $\vec{w}_p$ and is experimentally set to be 0.1 since the size of the region is more important than the position. It is easily observed that the overall region-based image level similarity score S_l between two identical images will be 1.

3.4.2. Global similarity measure at the image level

The image resemblance in terms of the global edge distribution is measured by the global similarity S_g. The Manhattan distance is used to calculate the global similarity by pairwisely comparing each element in the global EHD of the query and target images. It is easily observed from this distance measure that the smaller the global similarity score S_g is, the similar the two images.

3.4.3. Semi-global similarity measure at the image level

The image resemblance in terms of the global edge distribution in five clusters of four connected sub-images is measured by the semi-global similarity $S_{semi\text{-}g}$. The Manhattan distance is used to measure the semi-global similarity by pairwise comparison of the semi-global EHD of the query and target images. This similarity measure ensures that the similar images will have a semi-global similarity score close to 0.

3.4.4. Overall image similarity

The overall image similarity is measured by a weighted scheme in similarity measures integrating the regional, global, and semi-global matching. The weights are determined as follows:

- The most weight is assigned to the region-based similarity measure since the region-based color and texture features capture more details in the image.

- The least weight is assigned to the semi-global similarity measure since the semi-global EHD has more elements and therefore a larger Manhattan distance than the global EHD.

The overall image similarity measure is computed as

$$S = \lambda_3 S_l + (1 - \lambda_3)(e^{-(\lambda_4 S_g + (1-\lambda_4) S_{semi\text{-}g})/a}), \quad (12)$$

where:

- λ_3 adjusts the contribution of the region-based (local) similarity measure in the overall similarity and is empirically set to be 0.7.
- λ_4 adjusts the contribution of the global similarity measure to scale the global-based and semi-global-based similarity scores. It is tested to be 0.8.
- $e^{-(\lambda_4 S_g + (1-\lambda_4) S_{semi\text{-}g})/a}$ is a non-linear transformation that converts the integrated global and semi-global similarity to a range of [0, 1] with 1 indicating the best match. Here, α controls the general shape of the non-linear transformation and is experimentally set to be 2.

4. Experimental results

To date, we have tested our retrieval algorithm on a general-purpose image database with 5000 images from COREL [38]. These images are stored in JPEG format with size 384 × 256 or 256 × 384. The entire database has 50 categories with 100 images in each category. Most categories contain distinct semantics including building, beach, dinosaur, horse, mountain, and the like. For each image, the region-based color and texture features, locations and areas of all its regions, the global EHD, and the semi-global EHD are stored in a feature file.

4.1. Retrieval effectiveness

To qualitatively evaluate the retrieval effectiveness of our algorithm over the 5000-image COREL database, we randomly select five query images with different semantics, namely, beach, building, horse, food, and sunset. For each query image, we examine the precision of the query results based on the relevance of the image semantics. The top 11 retrieved images by using our proposed retrieval method and the UFM method [32], which is the best region-based retrieval method to our knowledge, are shown side-by-side in Fig. 6. For each block of images, the query and the most matched images are the same and at the upper left corner. The segmentation of the query is shown at the right side of the query with the number of regions indicated below. The numbers below other images are the overall similarity scores and the number of segmented regions. We also provide the number of relevant images among the top 20 and 30 returned images for both methods.

Several important observations are:

- Both methods effectively separate the beach and building categories despite the similarity of the segmented regions for two query images. However, our retrieval results are obviously more relevant to the query semantics.
- Our method yields desirable retrieval results for the horse query whereas several irrelevant ones are obtained by the UFM method.
- Our method achieves much better retrieval results than the UFM method for both food and sunset queries. The chaotic content of the food query and the simple content of the sunset query do not severely degrade our retrieval performance as they do to the UFM method mainly due to the integration of the global and semi-global edge descriptions in our fusion approach.

4.2. Comparison of retrieval performance of 10 distinct image categories

A more quantitative evaluation is performed on 10 distinct image categories. We randomly choose 15 images from each category (i.e., 150 images in total) as query images and then calculate the average precision of each category by evaluating the top 20 returned results. Here, a retrieved image is considered as a correct match if and only if it is in the same category as the query image. Our proposed fusion method is compared with several peer retrieval methods, including the UFM method [32], the IRM method [31], the global HSV color histogram method with different bins (i.e., 32 and 64 bins) [37], the color indexing method [13], and the global EHD method [37]. These systems use a wide variety of global-based, region-based, and fuzzy-region-based CBIR approaches together with different features such as color, texture, shape, and their combinations. By comparing with these systems, the effectiveness of the proposed system can be validated. We implemented all these methods except for the UFM method whose executables were obtained directly from the authors.

In order to ensure fair comparisons, we use the same 5000 images from COREL as a test bed, the same 150 images as queries, and the 20 images as returned retrieval images. Table 1 lists the comparison details among different categories and methods in terms of the average retrieval precision. It is clear that our proposed method performs much better than all six approaches in almost all image categories.

- Our method outperforms the UFM method in all image categories and improves the overall average retrieval accuracy by 32.10%.
- Our method yields much better retrieval accuracy than the IRM method in all image categories except for the dinosaur and office queries (categories 7 and 10) which yield comparable retrieval accuracy. The overall average retrieval accuracy is improved by 39.96%.

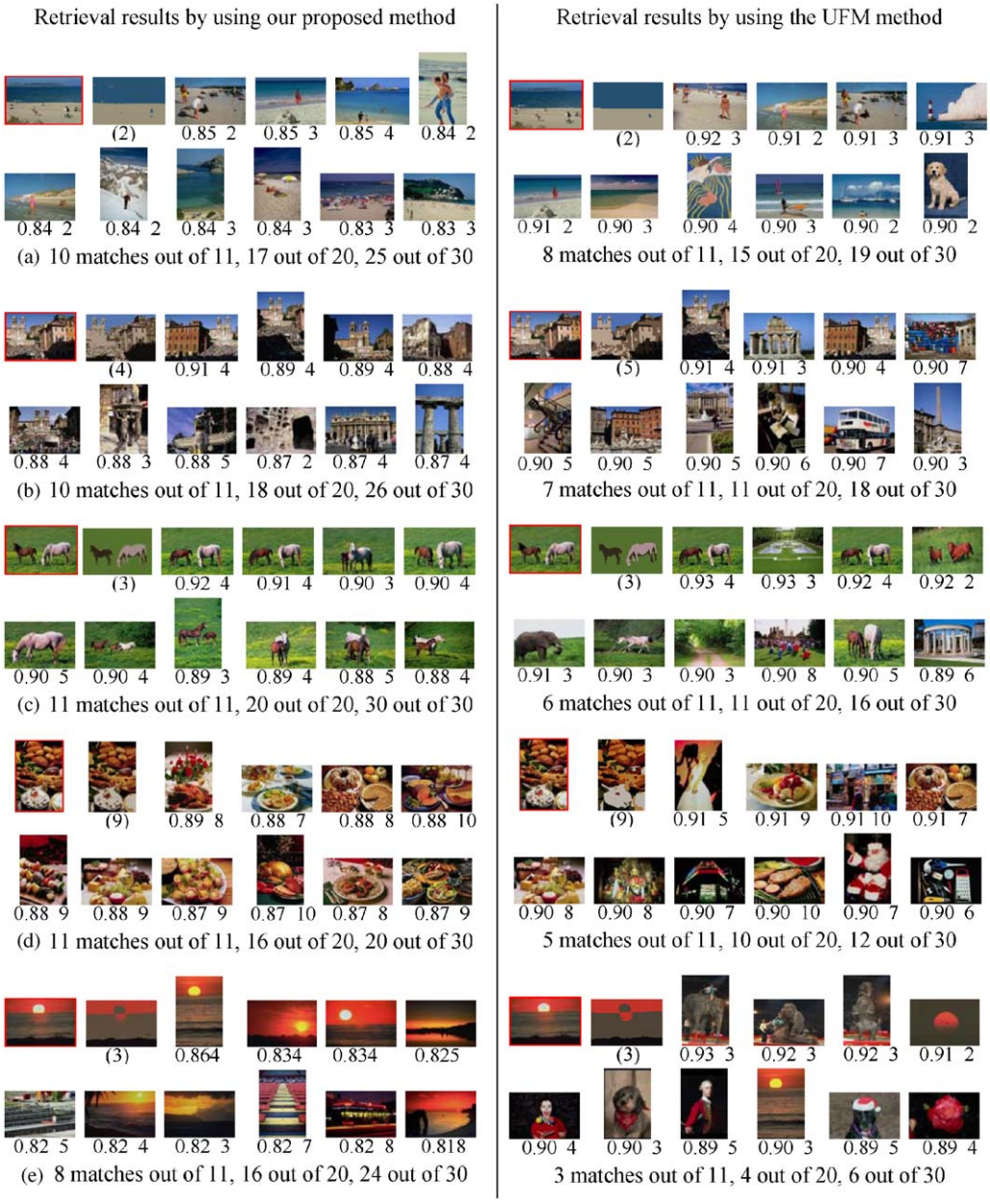


Fig. 6. Retrieval results of 5 queries by using our proposed method and the UFM method.

- Our method has better retrieval accuracy than the HSV 64-bin color histogram method in 9 categories and a little worse retrieval accuracy for the dinosaur queries. It improves the overall average retrieval accuracy by 39.51%.
- Our method has better retrieval accuracy than all the other three methods in all image categories. The overall average retrieval accuracy improvement over the HSV 32-bin color histogram method, the color indexing method, and the EHD method is 59.92%, 48.02%, and 91.72%, respectively.

4.3. Comparison of overall retrieval performance

More returned images are used to quantitatively measure the overall retrieval accuracy. Fig. 7 compares the overall average retrieval precision of the chosen 10 distinct image categories from top $20, 30, \ldots, 100$ returned images when seven methods are applied to the same 5000 images from COREL by using the same 150 query images. It clearly shows that our proposed method ranks the best with the highest overall retrieval accuracy

Table 1
Comparison of the average retrieval precision of each category by using seven different methods

	Proposed	UFM	IRM	HSV 32 bins	HSV 64 bins	Color Indexing	EHD
Beach	0.2800	0.2533	0.2800	0.1967	0.2567	0.2233	0.1067
Building	0.5467	0.4667	0.4133	0.2667	0.3400	0.4133	0.1367
Vehicle	0.5600	0.3167	0.1367	0.1933	0.2067	0.1633	0.4700
Flower	0.8233	0.6867	0.4933	0.5433	0.5967	0.3833	0.3100
Horse	0.8900	0.7633	0.8033	0.7400	0.8033	0.8367	0.4033
Food	0.5067	0.3000	0.2367	0.2033	0.2567	0.3133	0.1167
Dinosaur	0.7567	0.6100	0.7600	0.7267	0.8033	0.6567	0.6700
People	0.2433	0.2233	0.1433	0.1333	0.1133	0.1467	0.1400
Sunset	0.7100	0.3967	0.4333	0.3533	0.4100	0.3733	0.4500
Office	0.3333	0.2600	0.3367	0.1767	0.2633	0.3067	0.1433
Average	**0.5650**	**0.4277**	**0.4037**	**0.3533**	**0.4050**	**0.3817**	**0.2947**

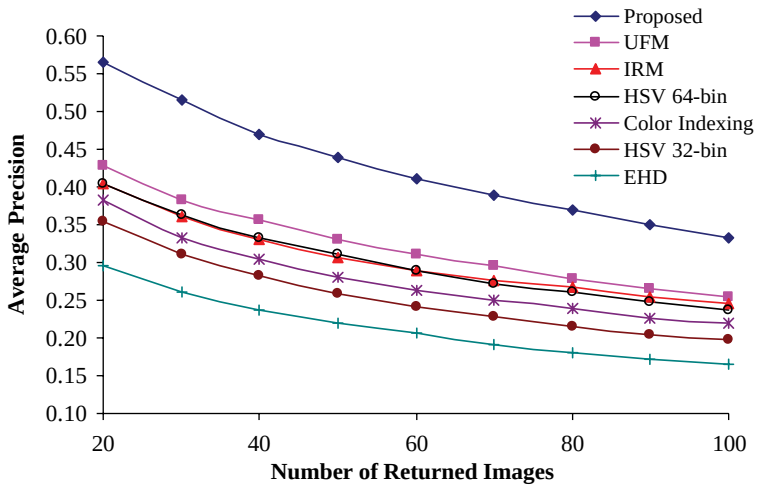


Fig. 7. Comparison of the overall retrieval precision of different returned image numbers by using seven different methods.

for different returned image numbers ranging from 20 to 100.

The comparisons of the overall retrieval recall (i.e., the ratio between the number of correctly retrieved images and the total number of relevant images which are in the same category as the query image) for different returned image numbers are illustrated in Fig. 8. It clearly shows that our proposed method performs the best with the highest overall retrieval recall for any number of returned images.

4.4. Validity proof of the proposed method

To verify the effectiveness of the proposed approach, the overall average retrieval accuracy obtained by assigning different weights to the global-based and regional-based similarity scores is shown in Fig. 9, where G and R, respectively, represent global and regional weights. It also clearly shows the effectiveness of our fusion approach ($G: R = 3: 7$) as it achieves the best accuracy.

Additional experiments using the same test bed, the same 150 query images, and the 20 returned images are performed on several variants of our proposed method to experimentally illustrate the validity of our method. These variants include:

- Our global and semi-global method without using any local features (PrGl);
- Our fuzzy region-based method without using any global and semi-global features (PrRe1);
- Our fuzzy region-based method using only local color and no local texture and no global and semi-global edge features (PrRe2);
- Our fuzzy region-based method combined with the global HSV 32-bin color histogram [37] instead of the global and semi-global edge histograms (PrReC1);
- Our fuzzy region-based method combined with the global HSV 64-bin color histogram [37] (PrReC2);

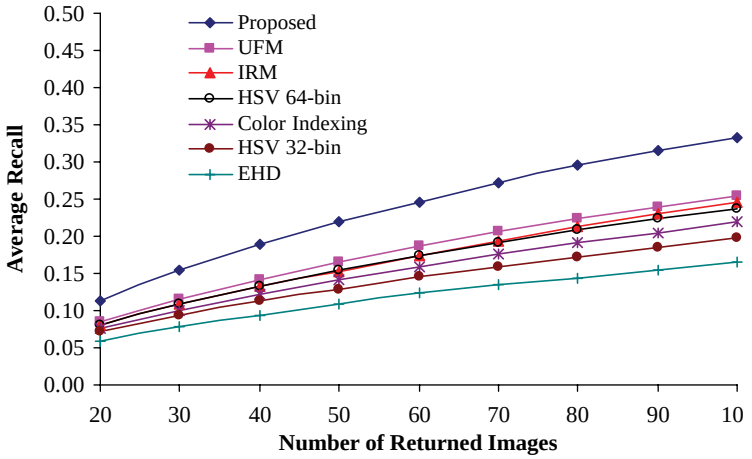


Fig. 8. Comparison of the overall retrieval recall of different returned image numbers by using seven different methods.

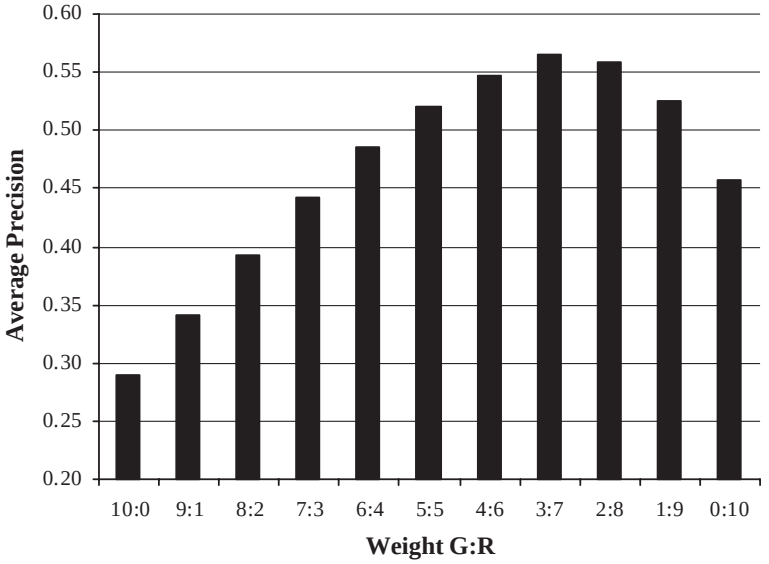


Fig. 9. Average retrieval accuracy for different weights.

- Our fuzzy region-based method combined with the color indexing [13] (PrReC3).
- Our fuzzy region-based method combined with the normative EHD for MPEG-7 [37] (PrReE).

Table 2 numerically lists the average retrieval precision of each category by applying our proposed method and its seven variants, namely, PrGl, PrRe1, PrRe2, PrReC1, PrReC2, PrReC3, and PrReE. The overall retrieval precision of the chosen 10 image categories for each retrieval method is also shown at the last row of Table 2. We also rank the overall retrieval performance of each method compared in Tables 1 and 2 in a descending order at the right-hand side of Table 2.

It is clear that our proposed method performs much better than all of its variant methods in almost all image categories except for a few ones with a comparable performance, as shown bold and italic in Table 2. In general, our proposed method and its five variants (i.e., PrReE, PrReC2, PrReC3, PrReC1, and PrRe1) outperform the six peer retrieval methods in terms of the overall retrieval precision. Several comparisons based on Table 2 are evaluated to illustrate the effectiveness of our fusion approach.

4.4.1. Effectiveness of fuzzy features and fuzzy matching

Our PrRe1 method is compared with the IRM method where only color and texture features are utilized for retrieval. Our PrRe1 method outperforms the IRM method

Table 2
Comparison of the average retrieval precision of each category by using our proposed method and its variants

	Proposed	PrGl	PrRe1	PrRe2	PrReC1	PrReC2	PrReC3	PrReE	Overall retrieval performance
Beach	0.2800	0.1133	0.2333	0.1900	0.2433	0.2633	0.2600	0.2433	Proposed: 56.50%
Building	0.5467	0.1600	***0.5500***	0.4733	0.5233	***0.5567***	***0.5533***	0.5300	PrReE: 53.97%
Vehicle	0.5600	0.4400	0.3067	0.2433	0.3033	0.3167	0.3033	0.5167	PrReC2: 48.33%
Flower	0.8233	0.4200	0.7200	0.5767	0.7167	0.7300	0.6533	0.7600	PrReC3: 47.87%
Horse	0.8900	0.2700	0.8500	0.7600	0.8867	***0.8933***	***0.9000***	***0.9167***	PrReC1: 45.70%
Food	0.5067	0.1400	0.3967	0.3200	0.3600	0.4167	0.4500	0.5033	PrRe1: 45.67%
Dinosaur	0.7567	0.4833	0.5333	0.3967	0.6033	0.6167	0.5800	0.7333	UFM: 42.77%
People	0.2433	0.1500	0.2000	0.1733	0.1833	0.2000	0.2300	0.2233	HSV 64 bins: 40.50%
Sunset	0.7100	0.5300	0.5133	0.4300	0.4767	0.4933	0.5100	***0.7167***	IRM: 40.37%
Office	0.3333	0.1867	0.2633	0.1800	0.2733	***0.3467***	***0.3467***	0.2533	Color Indexing: 38.17%
Average	**0.5650**	**0.2893**	**0.4567**	**0.3743**	**0.4570**	**0.4833**	**0.4787**	**0.5397**	PrRe2: 37.43%
									HSV 32 bins: 35.33%
									EHD: 29.47%
									PrG1: 28.93%

in almost all 10 image categories except for the beach, dinosaur, and office queries with a comparable retrieval performance. Our PrRe1 method improves the overall retrieval accuracy by 13.13% mainly due to the incorporation of the fuzzy features and fuzzy matching into the retrieval procedure.

4.4.2. Effectiveness of local texture features

We compare our PrRe1 and PrRe2 methods with respect to the average retrieval precision of each image category and the overall retrieval precision of all 10 image categories. The former has better retrieval performance in all 10 image categories and improves the overall retrieval accuracy by 22.01%. This result indicates that the additional local texture features do improve the retrieval accuracy.

4.4.3. Effectiveness of our fuzzy region-based approach

Our PrRe1 method ranks the sixth in terms of the overall retrieval precision among all 14 methods. It has better retrieval performance (i.e., 6.78% improvement) than the UFM method, which is the best region-based retrieval method to our knowledge. Moreover, our PrRe1 method is more efficient than the UFM method in two perspectives:

- The exclusive color features are used for segmentation instead of the comprehensive features integrating both color and texture as proposed in the UFM method. Therefore, our PrRe1 method is faster than the UFM method in segmenting an image into several coherent regions.
- The color and texture are treated as two separate features to represent each region instead of one combined feature consisting of color, texture, and shape as proposed in the UFM method. Such a separation ensures that more accurate local texture features are derived and better retrieval precision is achieved without using any shape feature.

4.4.4. Effectiveness of our fusion approach

Our proposed fusion approach is compared with its several variants to experimentally illustrate the effectiveness of this fusion.

(1) Our method vs. our PrG1 method: Our method yields substantially better retrieval accuracy in all image categories. There is a 95.30% improvement in the overall retrieval accuracy, which indicates that the additional local features dramatically increase the retrieval accuracy.
(2) Our method vs. our PrRe1 method: Our method yields better retrieval accuracy in 9 image categories and a comparable performance for the building queries. Our method outperforms our PrRe1 method by 23.71% improvement in the overall retrieval accuracy. It is clear that the fusion of the global features does improve the retrieval performance.
(3) Our method vs. its variant fusion methods (PrReC1, PrReC2, PrReC3, and PrReE): Our method performs

much better than our variant fusion methods integrating different color histograms. Even though the global color histograms themselves yield better retrieval accuracy than our global and semi-global EHDs, the integration with the region-based color and texture features does not achieve better retrieval accuracy. This is mainly due to the fused edge descriptors provide more information which is complement to the region-based features. It is also interesting to notice that our method outperforms the PrReE method by 4.69% retrieval accuracy improvement despite that the EHD alone performs better than our global and semi-global EHDs. All these experimental results indicate that our global and semi-global edge features are more effective than both global color histograms and EHDs in the fusion procedure. We have further tested on several other possible combinations. The experiments show that our approach achieves comparable performance by using the smallest number of features.

4.5. Speed

The algorithm has been implemented using Matlab 6.5 on a Pentium IV 3.06 GHz PC running Windows XP operating system. Computing the feature vectors for 5000 color images of size 384×256 requires around 35 h, which can be easily reduced to at least one-twentieth if the entire algorithm is implemented by C language. In average, 25 s are needed to segment and compute all the features for an image. It should be faster than most region-based retrieval methods which use integrated color and texture features for segmentation, where complicated statistical approach is involved in computing texture features.

The time complexity for segmenting all 5000 images is $O(CN \times knd)$ where:

- $O(CN)$ measures the average total number of iterations required for the entire image database with C being the average number of regions of an image (i.e., the average number of iterations required for an image) and N being the number of images in the database;
- $O(knd)$ measures the time complexity for segmenting a single image with k being the number of clusters (i.e., number of regions), n being the total number of color features (i.e., the total number of image-blocks in the image), and d being the dimensionality of the color feature.

The time complexity for calculating local color features is $O(CN \times d)$. The time complexity for calculating local texture features is $O(CN \times Row \times Col)$ where *Row*'s and *Col*'s are the dimensionality of the image itself. The time complexity for calculating global and semi-global edge features is $O(N \times ne)$ where e is the total number of edge types. The time for matching images and sorting results in our fusion approach is $O(C^2N + N \log N)$.

For the COREL database, we have $N = 5000$, $C = 4.3$, $n = (384 \times 256)/(2 \times 2) = 24576$, $d = 3$, $e = 5$, and the average value of k is equal to C.

5. Conclusions and discussions

A novel fusion approach to CBIR is proposed in this paper. In this approach, an image is first segmented into regions by using a fast and automatic color-clustering-based segmentation method. Two sets of features including color and texture properties are then derived to represent each segmented region. In specific, the region-based color features are automatically derived along with the segmentation results. The region-based texture features are computed by the average energy values of the six high-frequency bands after applying a two-level wavelet transform to the gray-scaled "texture template" image. The region-based image level similarity is then measured by applying the fuzzy matching scheme to the fuzzified region-based color and texture features with respect to both the region position and the normalized matched region percentage difference. Such fuzzy feature representation and fuzzy region matching have been proven to be more effective than the UFM method for addressing the imperfect segmentation and the inaccurate color/texture issues, as shown in Section 4.4.3.

However, a major limitation of the proposed region-based scheme is the possible degradation of the retrieval performance when segmentation tends to be very accurate for images with distinctive and relevant scenes. Furthermore, it is extremely difficult to find an appropriate fuzzy membership function applicable to a variety of images. Therefore, we add segmentation-independent global and semi-global EHDs to resolve these limitations. These EHDs are expanded from the normative EHD for MPEG-7 and are computed to measure the global image level similarity by using the Manhattan distance.

These two image level similarities are finally fused to one integrated overall similarity measure to compare the resemblance between the query and target images. The effectiveness, efficiency, and uniqueness of the proposed approach are summarized as follows:

- The unsupervised adaptive K-Means algorithm is exclusively performed on the 2×2 image-block-based color features to automatically, quickly, and efficiently segment an image into coherent regions, where each segmented region generally corresponds to an object or parts of an object.
- The region-based texture features are derived after segmentation instead of during the segmentation procedure. This derivation ensures that more accurate texture information is captured along the region boundary and within the region.
- The region-based color and texture are treated as two separate features to represent each region. Such a separa-

tion achieves better retrieval performance than the other schemes combining the color and texture into one comprehensive feature (e.g., IRM and UFM methods) as illustrated in Sections 4.4.1 and 4.4.3.

- Each independent color and texture feature is fuzzified to incorporate the segmentation-related uncertainties into the retrieval algorithm.
- The fuzzy region matching scheme is also implemented to allow one region in image A to match several regions in image B and vice versa. This scheme accommodates the imperfect segmentation and inaccurate color/texture issues.
- The normalized area percentage difference between matched regions and the normalized distance from the region center to the image center are incorporated into regional features. These two region-related properties are derived from the semantics observation and are used to determine the importance of a specific region when computing the region-based image level similarity.
- The Cauchy function has been used for both fuzzification and fuzzy matching. This fuzzy membership function greatly reduces the computational cost as proven in the appendix.
- The global and semi-global EHDs have been utilized to decrease the impact of segmentation since they do not depend on segmentation.
- The resemblance of two images is measured by the overall similarity fusing together two families of region-based fuzzy color and texture features, global and semi-global EHDs with different contribution parameters.

The experimental results on 5000 images from the COREL database demonstrate that the proposed algorithm achieves good retrieval accuracy with fast speed due to the small feature vector size (i.e., 3 elements for color, 6 elements for texture, 5 elements for global EHDs, and 25 elements for semi-global EHDs).

Shape or spatial information is not considered in our implementation for the efficiency consideration. It may be further integrated into the retrieval system to improve the accuracy with a compromised efficiency. Other fuzzy membership functions and other global feature representations may be further studied to improve the retrieval accuracy. Instead of searching the image database sequentially as currently implemented in our system, some tree-structure-based indexing methods may be explored to further speedup the retrieval.

Appendix A

Proof of Eq. (6) is depicted as follows. Consider two fuzzy sets $C_u(\vec{x})$ and $C_v(\vec{x})$ for different regions u and v in images A and B, let $\vec{f}_u$ and $\vec{f}_v$ be the corresponding fuzzy centers with the assumption of $\vec{f}_u \leqslant \vec{f}_v$. The illustration of such two fuzzy sets with respect to the fuzzy centers is shown in Fig. 10.

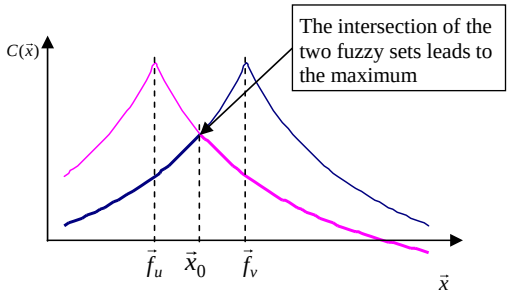


Fig. 10. Two fuzzy sets and their fuzzy centers.

According to (5), the similarity between two regions u and v is defined as

$$s(u,v) = \sup[C(\vec{x}_{u\cap v,m})] = \sup\{\min[C_u(\vec{x}_m), C_v(\vec{x}_m)]\}. \tag{A.1}$$

The minimum of the two fuzzy sets for any given $\vec{x}$ is illustrated as bold in Fig. 10. It is clearly shown that the maximum occurs at the intersection $\vec{x}_0$ between the two fuzzy sets $C_u(\vec{x})$ and $C_v(\vec{x})$. This intersection $\vec{x}_0$ can be derived by:

$$\begin{aligned} C_u(\vec{x}_0) = C_v(\vec{x}_0) &\Leftrightarrow \frac{1}{1+\left(\dfrac{\|\vec{x}_0-\vec{f}_u\|}{d_A}\right)^{\alpha}} \\ &= \frac{1}{1+\left(\dfrac{\|\vec{x}_0-\vec{f}_v\|}{d_B}\right)^{\alpha}} \Leftrightarrow \frac{\|\vec{x}_0-\vec{f}_u\|}{d_A} \\ &= \frac{\|\vec{x}_0-\vec{f}_v\|}{d_B} \Leftrightarrow \frac{\vec{x}_0-\vec{f}_u}{d_A} \\ &= \frac{\vec{f}_v-\vec{x}_0}{d_B} \Leftrightarrow d_B(\vec{x}_0-\vec{f}_u) \\ &= d_A(\vec{f}_v-\vec{x}_0) \Leftrightarrow \vec{x}_0 = \frac{d_A\vec{f}_v + d_B\vec{f}_u}{d_A+d_B}. \end{aligned} \tag{A.2}$$

Substitute this $\vec{x}_0$ into (A.1) to obtain the similarity between two regions u and v represented by two fuzzy sets.

$$\begin{aligned} s = C_u(\vec{x}_0) &= \frac{1}{1+\left(\dfrac{\|\vec{x}_0-\vec{f}_u\|}{d_A}\right)^{\alpha}} \\ &= \frac{1}{1+\left(\dfrac{\left\|\dfrac{d_A\vec{f}_v+d_B\vec{f}_u}{d_A+d_B}-\vec{f}_u\right\|}{d_A}\right)^{\alpha}} \\ &= \frac{1}{1+\left(\dfrac{\|d_A\vec{f}_v+d_B\vec{f}_u-\vec{f}_u d_A-\vec{f}_u d_B\|}{d_A(d_A+d_B)}\right)^{\alpha}} \\ &= \frac{(d_A+d_B)^{\alpha}}{(d_A+d_B)^{\alpha}+\|\vec{f}_u-\vec{f}_v\|^{\alpha}}. \end{aligned}$$

References

[1] Y. Rui, T.S. Huang, S.F. Chang, Image retrieval: current techniques, promising directions and open issues, J. Visual Commun. Image Represent. 10 (1) (1999) 1–23.

[2] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, J. Ramesh, Content-based image retrieval at the end of the early years, IEEE Trans. Pattern Anal. Mach. Intell. 22 (12) (2000) 1349–1380.

[3] S.J. Park, D.K. Park, C.S. Won, Core experiments on MPEG-7 edge histogram descriptor, MPEG Document M5984, Geneva, May, 2000.

[4] ISO/IEC/JTC1/SC29/WG11, Core experiment results for edge histogram descriptor (CT4), MPEG Document M6174, Beijing, July, 2000.

[5] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by image and video content: the QBIC system, IEEE Comput. 28 (9) (1995) 23–32.

[6] A. Pentland, R.W. Picard, S. Sclaroff, Photobook: content based manipulation for image databases, Int. J. Comput. Vision 18 (1996) 233–254.

[7] A. Gupta, R. Jain, Visual information retrieval, Commun. ACM 40 (5) (1997) 70–79.

[8] J.R. Smith, S.F. Chang, VisualSEEK: a fully automated content-based query system, in: Proceedings of the ACM Multimedia, 1996, pp. 87–98.

[9] P. Kelly, M. Cannon, D. Hush, Query by image example: the CANDID approach, in: SPIE Proceedings on Storage and Retrieval for Image and Video Databases, 1995, pp. 238–248.

[10] V. Ogle, M. Stonebraker, Chabot: retrieval from a relational database of images, IEEE Comput. 28 (9) (1995) 40–48.

[11] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, T. Huang, Supporting similarity queries in MARS, in: Proceedings of the ACM Conference on Multimedia, 1997, pp. 403–413.

[12] M. Stricker, M. Swain, The capacity and the sensitivity of color histogram indexing, Technical Report 94-05, University of Chicago, March, 1994.

[13] M. Swain, D. Ballard, Color indexing, Int. J. Comput. Vision 7 (1) (1991) 11–32.

[14] H. Tamura, S. Mori, T. Yamawaki, Texture features corresponding to visual perception, IEEE Trans. Systems Man Cybernet. Smc-8 (6) (1978) 464–473.

[15] J.M. Francos, Orthogonal decompositions of 2D random fields and their applications in 2D spectral estimation, in: N.K. Bose, C.R. Rao (Eds.), Signal Processing and Its Application, 1993, pp. 20–27.

[16] F. Liu, R.W. Picard, Periodicity directionality and randomness: Wold features for image modeling and retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 18 (7) (1996) 722–733.

[17] A.K. Jain, F. Farroknia, Unsupervised texture segmentation using Gabor filters, Pattern Recogn. 24 (12) (1991) 1167–1186.

[18] M.H. Gross, R. Koch, L. Lippert, A. Dreger, Multiscale image texture analysis in wavelet spaces, in: Proceedings of the IEEE International Conference on Image Processing, 1994, pp. 412–416.

[19] A. Laine, J. Fan, Texture classification by wavelet packet signatures, IEEE Trans. Pattern Anal. Mach. Intell. 15 (11) (1993) 1186–1191.

[20] T. Chang, C.C.J. Kuo, Texture analysis and classification with tree-structured wavelet transform, IEEE Trans. Image Process. 2 (4) (1993) 429–441.

[21] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, IEEE Trans. Comput. C-21 (3) (1972) 269–281.

[22] E. Persoon, K.S. Fu, Shape discrimination using Fourier descriptors, IEEE Trans. Systems Man Cybernet. Smc-7 (2) (1977) 170–179.

[23] G. Pass, R. Zabith, Histogram refinement for content-based image retrieval, in: IEEE Workshop on Applications of Computer Vision, 1996, pp. 96–102.

[24] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu, R. Zabih, Image indexing using color correlogram, in: IEEE International Conference on Computer Vision and Pattern Recognition, Puerto Rico, June 1997, pp. 762–768.

[25] A. Rao, R.K. Srihari, Z. Zhang, Spatial color histograms for content-based image retrieval, in: IEEE International Conference on Tools with Artificial Intelligence, 1999, pp. 183–186.

[26] L. Cinque, G. Ciocca, et al., Color-based image retrieval using spatial-chromatic histogram, Image Vision Comput. 19 (2001) 979–986.

[27] W.Y. Ma, B.S. Manjunath, NeTra: a toolbox for navigating large image databases, Multimedia Systems 7 (1999) 184–198.

[28] Y. Deng, B.S. Manjunath, C. Kenney, An efficient color representation for image retrieval, IEEE Trans. Image Process. 10 (1) (2001) 140–147.

[29] C. Carson, S. Belongie, H. Greenspan, J. Malik, Blobworld: image segmentation using expectation-maximization and its application to image query, IEEE Trans. Pattern Anal. Mach. Intell. 24 (8) (2002) 1026–1038.

[30] S. Ardizzoni, I. Bartolini, M. Patella, Windsurf: region-based image retrieval using wavelets, DEXA workshop, 1999, pp. 167–173.

[31] J. Li, J. Wang, G. Wiederhold, SIMPLIcity: semantics-sensitive integrated matching for picture libraries, IEEE Trans. Pattern Anal. Mach. Intell. 23 (9) (2001) 947–963.

[32] Y. Chen, J. Wang, A region-based fuzzy feature matching approach to content-based image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 24 (9) (2002) 1252–1267.

[33] T. Wang, Y. Rui, J.G. Sun, Constraint based region matching for image retrieval, Int. J. Comput. Vision 25 (1–2) (2004) 37–45.

[34] J.W. Hsieh, W.E.L. Grimson, Spatial template extraction for image retrieval by region matching, IEEE Trans. Image Process. 12 (11) (2003) 1404–1415.

[35] J.A. Hartigan, M.A. Wong, Algorithm AS136: a K-means clustering algorithm, Appl. Stat. 28 (1979) 100–108.

[36] F. Hoppner, F. Klawonn, R. Kruse, T. Runkler, Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition, Wiley, New York, 1999.

[37] B.S. Manjunath, P. Salembier, T. Sikora, Introduction to MPEG-7 Multimedia Content Description Interface, Wiley, New York, 2002.

[38] http://www.corel.com/

About the author—XIAOJUN QI received her Ph.D. degree in Computer Science in 2001 from Louisiana State University. She is an assistant professor in Computer Science Department at Utah State University since 2002. Her research interests include image processing, pattern recognition, computer vision, and machine learning.

About the author—YUTAO HAN received his M.S. degree in Electrical Engineering from Utah State University in 2003. He currently works on his M.S. degree in Computer Science at Utah State University. His research interests include signal and image processing, pattern recognition, machine learning, and networking.