AN EFFICIENT LOCAL FUZZY COLOR AND GLOBAL COLOR-TEXTURE REPRESENTATION FOR IMAGE RETRIEVAL

Xiaojun Qi and Yutao Han
Computer Science Department, Utah State University, Logan, UT 84322-4205
xqi@cc.usu.edu and yhan@cc.usu.edu

ABSTRACT
An efficient local fuzzy color and global color-texture descriptor is proposed for content-based image retrieval. In our retrieval system, an image is represented by a set of color-clustering-based segmented regions and global color-texture descriptors. Each segmented region corresponds to an object or parts of an object and is represented by a fuzzified color feature. A fuzzy region matching scheme is then incorporated to address the issues associated with the color-related inaccuracies and segmentation-related uncertainties. The global color-texture features are also incorporated into our representation since they do not depend on the segmentation results. They are calculated by energy-distribution in a dimensionality reduced P_1P_2 color space. The overall similarity measure is defined as a weighted combination between the regional and global image level similarity measures incorporating all features. Our proposed retrieval approach demonstrates a promising retrieval performance for an image database of 1000 general-purpose images from COREL, as compared with some peer systems in the literature.

KEY WORDS
Image retrieval, image segmentation, and fuzzy matching.

1. Introduction

Content-based image retrieval (CBIR) has become an active research area since the early 1990s. It aims to develop an efficient visual-content-based-technique to search and browse large scale digital image collections. In general, the similarity comparison between a query and any target image is performed either globally based on visual content descriptors including a single or a combination of low-level features (e.g., color, texture, shape, and spatial layout of objects) or locally based on visual content descriptors derived from decomposed regions of the images.

Color is the most widely used visual feature in image retrieval. Three common color feature representations are color histogram, color moments, and color sets [1]. They are simple to calculate and can provide reasonable discriminating power in image retrieval. However, they do not include any spatial information and give many false positives. To retain the spatial information of the color features, many approaches including color coherence vector [1], color correlogram [1], spatial color histogram [2], and spatial chromatic histogram [3] have been proposed. Even though these approaches achieve better retrieval results, they are computationally intensive and fail to reflect the fuzzy nature of the color features inherent in the color histogram, moments, and sets.

Texture is another commonly used visual feature in image retrieval. It provides more information about the structural arrangement such as coarseness, contrast, directionality, linelikeness, regularity, and roughness. The common representations include Tamura features, Wold features, Gabor filter features, and wavelet transform features as summarized in [1]. These features characterize texture by the statistical distribution of the image intensity and are effective in image retrieval.

Since each of these two low-level features tends to capture only one aspect of image properties, many systems [4, 5] combine color, texture, and more features to get better retrieval results. However, such a combination does not have explicit semantic meanings to be applicable to all images. Consequently, region-based image retrieval methods have gained more attention 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 region-based retrieval systems are briefly reviewed here. Carson *et al.* [6] use expectation-maximization on the color and texture features to segment the image into coherent regions. The new segmentation-based color, texture, and spatial features are further utilized for retrieval. Ardizzoni *et al.* [7] use the color and texture features captured from the wavelet coefficients for both segmentation and retrieval. In [8], the 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. Wang *et al.* [9, 10] cluster the color

features computed in the Luv color space and the texture features computed from the wavelet coefficients for each 4×4 block by using the K-means algorithm. The cluster-based color, texture, and shape features for each region are utilized for retrieval. In [9], an integrated region matching (IRM) scheme, which allows a region in one image to match against regions from another image, is proposed to decrease the impact of inaccurate region segmentation. In [10], a unified feature matching (UFM) scheme is proposed, where cluster-based multiple fuzzy feature representations and fuzzy similarity measures are used in the above IRM to improve the retrieval accuracy.

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. The fuzzy representation [10] 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. Consequently, global color-texture features and global color moments are further utilized in our image representation. These global descriptors are extracted independent of segmentation and therefore provide more robustness against any issue associated with segmentation. The weighted similarity score is calculated in terms of both local fuzzy color and global color-texture features to retrieve images.

The remainder of the paper is organized as follows. Section 2 describes the general methodology of our proposed algorithm. Section 3 illustrates the experimental results. Section 4 draws conclusions.

2. Proposed Retrieval Algorithm

The first step of our region-based retrieval algorithm is to segment an image into different regions based on the color features by using an unsupervised K-means algorithm. Image indexing and retrieval is then taken based on the local fuzzy color feature in each region and the global color-texture feature and the RGB statistics (mean and variance) of the entire image.

2.1 Color-Based Image Segmentation

Most segmentation methods use either color and texture features [6, 7, 9, 10] or a complicated edge flow model [8] 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 nature images. This proposed method provides segmentations that are good enough for retrieval since the accurate segmentation is not required in image retrieval and our fuzzy matching scheme and our segmentation independent global features will decrease the impact of the inaccurate segmentation. The proposed segmentation 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, our region-based retrieval algorithm first partitions the image into non-overlapping small blocks. A color feature vector is then extracted for each block. The block size 2×2 is chosen since a small block preserves more details in an image. The color feature is constructed by the average color components in each 2×2 block. The Luv color space is used because the perceptual color difference of the human visual system is proportional to the numerical difference in the Luv color space. The texture feature is not utilized for segmentation based on the following two observations:

1. The texture feature of any block size smaller than 8×8 does not have enough discriminating power to yield effective segmentation results.
2. The texture feature requires more computation than the color feature since statistical distribution needs to be calculated.

An unsupervised K-means algorithm is used to cluster these color feature vectors into several regions, where each region in the color feature space corresponds to one spatial region in the image space. The unsupervised learning nature of this K-means algorithm enables us to adaptively update the number of regions C by gradually increasing C, which is initially set as 2. This iterative process will stop when a termination criterion is satisfied (i.e., the average distance between all pairs of cluster centers is less than a predetermined threshold value). Such an unsupervised iterative scheme also accommodates the fact that the number of regions in an image is unknown before the segmentation.

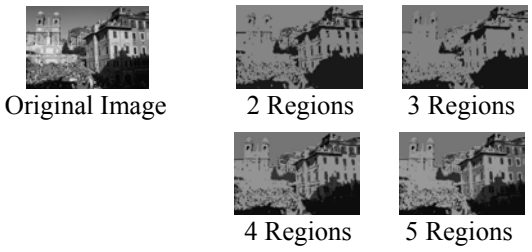


Fig. 1: Segmentation results by the unsupervised K-means clustering algorithm

Fig. 1 shows the intermediate segmentation results of one sample image from our test database by adaptively and gradually increasing the number of regions C. 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 black, gray, and brown portions in the image. 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.

2.2 Fuzzy Color Feature Representation and Fuzzy Region Matching

Unlike most region matching schemes [6-9], which directly compare the regional features between two images, we fuzzify each regional color feature and apply fuzzy matching on these fuzzified features to address the imperfect segmentation and inaccurate color issues.

In general, any membership function with a smooth transition between 0 and 1 can be selected for fuzzification. The Cauchy function [11] is utilized in our approach among some commonly used cone and exponential functions due to its good expressiveness and its high computational efficiency. To fuzzify each region j, the Cauchy function, Cau: $IR^k \rightarrow [0,1]$, defined as:

$$Cau(\vec{x}_{j,m}) = \frac{1}{1+\left(\frac{\|\vec{x}_{j,m} - \vec{f}_j\|}{d}\right)^{\alpha}} \qquad (1)$$

is used to calculate the membership of the color feature vector $\vec{x}_{j,m}$ for each 2×2 block m in the corresponding region j, where:

- α determines the smoothness of region j;
- $\vec{f}_j$ is the centroid of region j;
- d is the average distance between region centers. It determines the width of the region. It is calculated as:

$$d = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{k=i+1}^{C} \left\| \vec{f}_i - \vec{f}_k \right\| \qquad (2)$$

  where C is the number of regions obtained from the unsupervised K-mean algorithm.

This membership function indicates that the farther a feature vector $\vec{x}_{j,m}$ is away from the region center j, the lower its degree of membership to the fuzzy feature. This membership value illustrates the degree of wellness that $\vec{x}_{j,m}$ characterizes the region, and thus models the segmentation-related uncertainties.

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

$$s(u,v) = \sup Cau(\vec{x}_{u\cap v,m}) \qquad (3)$$

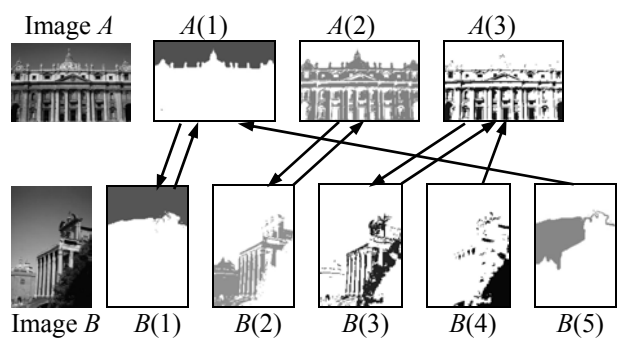
where $\vec{x}_{u\cap v,m}$ represents any block m in two regions u or v, and $Cau(\vec{x}_{u\cap v,m}) = \min[Cau_u(\vec{x}_m), Cau_v(\vec{x}_m)]$ with $Cau_u(\vec{x}_m)$ and $Cau_v(\vec{x}_m)$ being the membership values computed by (1). Because of the unimodal property of the Cauchy function, the fuzzy similarity (3) 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:

$$S(\vec{f}_u, \vec{f}_v) = \frac{(d_A + d_B)^{\alpha}}{(d_A + d_B)^{\alpha} + \left\| \vec{f}_u - \vec{f}_v \right\|^{\alpha}} \qquad (4)$$

where:

- $\vec{f}_u$ and $\vec{f}_v$ are the centroid of region u in image A and the centriod of region v in image B, respectively;
- d_A and d_B are the average distance between region centers in images A and B, respectively;
- α is the smoothness of the similarity function, which is empirically set to be 1.

This simplified fuzzy region similarity formula (4) indicates that the similarity between two fuzzified regions can be quickly and efficiently computed by using the non-fuzzified features and the fuzzy function shape parameter α. This fuzzy similarity measurement eliminates the need to compute the memberships of all blocks and therefore is another reason for the Cauchy function to be chosen in our fuzzification.



Matching Region Pairs		L
A→B	A.(1)→ B (1)	0.97414
	A (2)→ B (2)	0.94438
	A (3)→ B (3)	0.89681
B→A	B (1)→ A (1)	0.97414
	B (2)→ A (2)	0.94438
	B (3)→ A (3)	0.89681
	B (4)→ A (3)	0.81972
	B (5)→ A (1)	0.83261

Fig. 2: Fuzzy region matching scheme

The region-level similarities between all the regions in images A and B are utilized to generate a similarity vector

L. The length of ***L*** is the total regions in both images *A* and *B*. Each element in ***L*** is the maximum similarity score calculated by comparing each region in image *A* against all regions in image *B* or vice versa. It is easily observed that if two images are identical, they will have the same segmentation results and the elements in ***L*** will be all 1's. This fuzzy region matching scheme is illustrated in Fig. 2. Each of the three regions in image *A* is matched to the most similar region in image *B* by calculating the region similarity using (4). Similarly, each of the five regions in image *B* is matched to the most similar region in image *A*. The best similarities in terms of the color feature are stored in ***L***.

2.3 Global Color-Texture and Global Color Features

The global color-texture feature is derived from a family of reduced dimensionality color spaces [12]. Given the R, G, and B components of a color image, three reduced 2-D P_1P_2 color representations are computed:

$$P_1 = \frac{1}{\sqrt{2}}\frac{G-R}{R+G+B},\ P_2 = \frac{1}{\sqrt{6}}\frac{2B-R-G}{R+G+B} \quad (5)$$

$$P_1 = \frac{1}{\sqrt{2}}\frac{R-B}{R+G+B},\ P_2 = \frac{1}{\sqrt{6}}\frac{2G-R-B}{R+G+B} \quad (6)$$

$$P_1 = \frac{1}{\sqrt{2}}\frac{B-G}{R+G+B},\ P_2 = \frac{1}{\sqrt{6}}\frac{2R-B-G}{R+G+B} \quad (7)$$

The choice of a particular P_1P_2 color space is determined by the least significant color component, which is identified by the range to average ratios. For instance, suppose that blue (B) is the least significant color component (i.e., blue has a higher range to average ratio than red and green), P_1P_2 color space derived from (5) should be used.

Since the chosen P_1P_2 color space has the same significance, the complex Fourier transform can be applied to the complex image $C = P_1 + jP_2$ for the global color-texture feature extraction. This global color-texture feature consists of 40 elements: the first 8 elements are the normalized spectrum energy distribution of the 8 concentric, equal-distance disks; the rest is the normalized spectrum energy distribution of the 32 equally sized circular sectors, each spanning an angle of 11.25°.

The global color feature consists of 6 elements. They are the global color moments (i.e., mean and variance) of each R, G, and B component.

2.4 Similarity Measure

A weighted region-based fuzzy color feature similarity scheme is used based on the following observations:

- Important objects in an image tend to occupy larger areas near the image center;
- Regions adjacent to the image boundary provide more semantics information since semantically similar images always have similar background.

Consequently, more weight is assigned to the region that occupies more area or is near the image boundary or the image center.

The similarity score S_1 for the region-based fuzzy color features is computed as:

$$S_1 = \vec{w}^T L = ((1-\lambda)\vec{w}_a + \lambda\vec{w}_p)^T L \quad (8)$$

where $\vec{w}$ is the weight vector with $\vec{w}_a$ containing the normalized area percentages of the query and target images and $\vec{w}_p$ containing the normalized weights which favor regions near the image boundary or center, and λ adjusts the significance of $\vec{w}_a$ and $\vec{w}_p$ in the weight vector $\vec{w}$.

The similarity score S_2 for the global color-texture and color features is measured by the Euclidean distance between feature vectors of the query and target images.

The overall similarity is then measured by a weighted scheme in similarity measure between the regional and global matching. Since the region-based color feature captures more details in the image, more weight is assigned to this similarity measure. The overall similarity measure is computed as:

$$S = (1-\lambda)S_1 + \lambda(1-S_2) \quad (9)$$

where λ adjusts the significance of the regional and global similarity measure in the overall similarity and is set to be 0.2 in our implementation.

3. Experimental Results

To date, we have tested our region-based retrieval algorithm on a general-purpose image database with 1000 images from COREL. These images have 10 categories with 100 images in each category. The categories contain different semantics, namely architecture, beach, dinosaur, horse, snow-mountain, and the like. To evaluate the retrieval effectiveness of our algorithm, we randomly select five query images with different semantics (i.e., Africa, beach, horses, vehicles, and food). The top 11 results are shown in Fig. 3. A retrieved image is considered a correct match if and only if it is in the same category as the query image.

(a) 9 matches out of 11, 18 matches out of 20

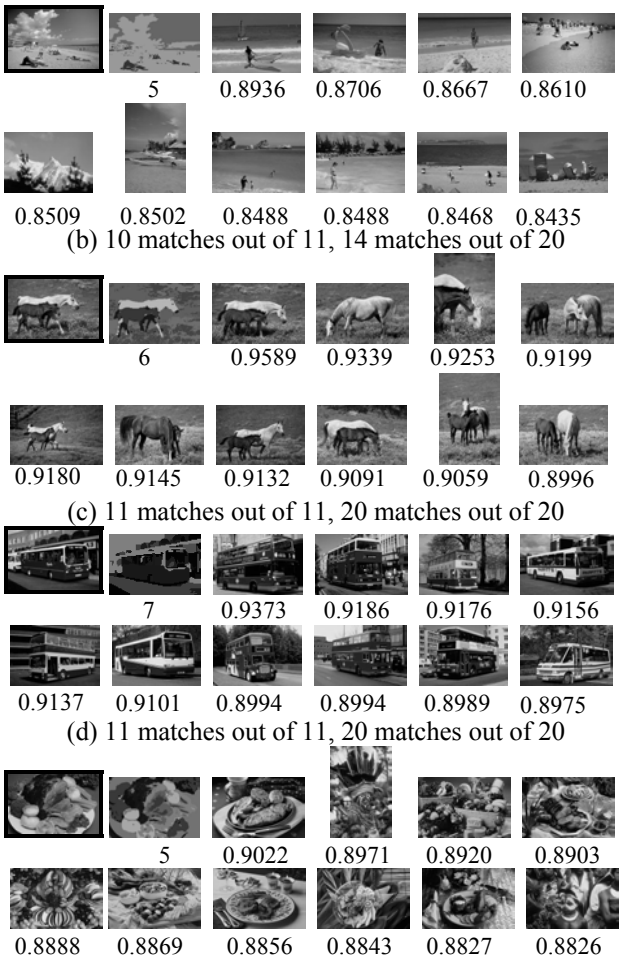


Fig. 3: Retrieval results of 5 queries. For each block of images, the query and the most matched image are the same and at the upper left corner. The segmentation of the query is shown at the right side of the query with the number of regions indicated below. The numbers below other images are the similarity scores.

To perform a more quantitative evaluation, we randomly choose 15 images from each category (i.e., 150 images in total) as query images. The precision, which is defined as the ratio of the relevant retrieved images (i.e., images belonging to the correct category) over the total retrieved images, is calculated by evaluating the top 50 returned results. Several peer retrieval methods are also used to compare the retrieval performance. These methods include our proposed method, our proposed method without global color-texture features (w/o GCT), UFM [10], MPEG-7 edge histogram descriptor (EHD) [13], global color histogram method with 32 color bins (HistC), and efficient color representation (ECR) [8] applied to our segmentation results. In order to ensure the fair comparison, we used the same 1000 images from COREL as test bed, the same 150 images as query, and the 50 images as returned retrieval images. Table 1 illustrates the average precision for each category by applying the above methods on the same query images.

The comparison results from Table 1 are:

- Our method vs. w/o GCT: The integration of global color and texture features improves the retrieval accuracy by 9.09%.
- Our method vs. ECR where no fuzzification is utilized: Fuzzy measurement improves the retrieval accuracy by 33.3%.
- Our method vs. EHD and HistC: Our method outperforms both EHD and HistC, two global features, by 50% and 71.43% respectively.
- Our method vs. UFM: Our method yields a comparable average retrieval precision as UFM. However, our method is more efficient than UFM since there are no local texture and shape features involved in the implementation. The local texture features require intensive computation since the wavelet decomposition is applied to each sub-block image to study the statistical distribution.

Table 1: Comparison of the average retrieval precision of different methods

Category	Our Method	w/o GCT	UFM	ECR	EHD	HistC
1	0.65	0.52	0.53	0.71	0.58	0.49
2	0.39	0.35	0.46	0.37	0.32	0.12
3	0.54	0.54	0.56	0.18	0.44	0.18
4	0.73	0.57	0.69	0.22	0.47	0.19
5	1.00	1.00	1.00	0.81	0.82	0.98
6	0.31	0.24	0.31	0.35	0.37	0.33
7	0.67	0.59	0.82	0.38	0.59	0.30
8	0.87	0.84	0.81	0.84	0.82	0.47
9	0.28	0.25	0.28	0.32	0.29	0.12
10	0.50	0.57	0.53	0.27	0.56	0.30
Average	**0.60**	**0.55**	**0.60**	**0.45**	**0.40**	**0.35**

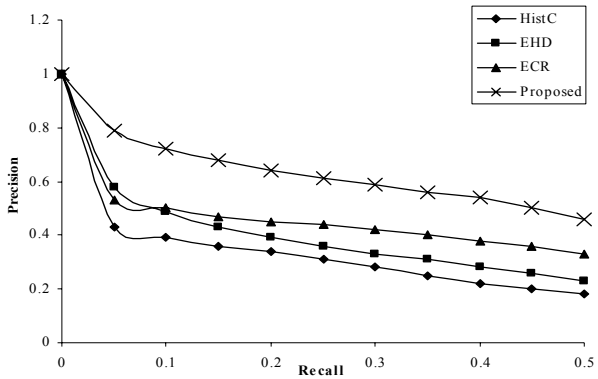


Fig. 4: Comparison of the average precision-recall curve of four different methods

Another quantitative evaluation is performed on our proposed retrieval method, ECR, EHD, and HistC by plotting a precision-recall curve as shown in Fig. 4. Recall (i.e., the ratio of the relevant retrieved images over the total relevant images in the database) increases as the retrieved images increase since the number of relevant images in the database for each query image is fixed.

This precision-recall curve is generated by randomly selecting 150 query images over all categories. It clearly shows that the proposed method performs the best among the four methods.

4. Conclusions

An efficient local fuzzy color and global color-texture descriptor is proposed in this paper for content-based image retrieval. The local fuzzy color features are obtained from unsupervised segmentation and Cauchy fuzzification. The global color-texture feature is extracted from the P_1P_2 color space. The global color moments are computed from each RGB component. The proposed approach is efficient and effective because:

- The unsupervised adaptive K-means algorithm is exclusively performed on the 2×2 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 fuzzy region matching scheme allows a region in the query image to match against any region in the target image and vice versa. This scheme accommodates the imperfect segmentation and inaccurate color issues.
- The Cauchy function has been used for both fuzzification and fuzzy matching. This fuzzy membership function greatly reduces the computational cost as shown in (4).
- The global color-texture features are extracted from the reduced dimensionality color space. They have been utilized to decrease the impact of segmentation since they do not depend on segmentation.

The experimental results on 1000 images from COREL database demonstrate that the proposed algorithm achieves good retrieval accuracy.

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.

References:

[1] F. Long, H. Zhang, & D. Feng, *Multimedia Information Retrieval and Management:Technological Fundamentals and Applications* (Springer-Verlag, 2002).

[2] A. Rao, R. K. Srihari, & Z. Zhang, Spatial color histograms for content-based image retrieval, *IEEE Int. Conf. on Tools with Artificial Intelligence*, Chicago, Illinois, 1999, 183-186.

[3] L. Cinque, G. Ciocca, S. Levialdi, A. Pellicano, & R. Schettini, Color-based image retrieval using spatial-chromatic histogram, *Image and Vision Computing*, *19*, 2001, 979-986.

[4] T. Shin, J. Huang, C. Wang, J. Hung, & C. Kao, An intelligent content-based image retrieval system based on color, shape and spatial relations, *Proceedings of Natl. Sci. Counc. ROC(A),.25*(4), 2001, 232-242.

[5] K. Liang, and C.-C.J. Kuo, WaveGuide: a joint wavelet-based image representation and description system. *IEEE Trans on Image Processing, 8*(11), 1999, 1619-1629.

[6] C. Carson, S. Belongie, H. Greenspan, & J. Malik, Region-based image querying, *CVPR'97 Workshop on Content-Based Access of Image and Video Libraries*, San Juan, Puerto Rico, 1997, 42-49.

[7] S. Ardizzoni, I. Bartolini, & M. Patella, Windsurf: Region-based image retrieval using wavelets, *DEXA workshop,* Florence, Italy, 1999, 167–173.

[8] Y. Deng, B. S. Manjunath, & C. Kenney, An Efficient Color Representation for Image Retrieval, *IEEE Trans on Image Processing*, *10*(1), 2001, 140–147.

[9] J. Li, J. Wang, & G. Wiederhold, Simplicity: semantics-sensitive integrated matching for picture libraries, *IEEE Trans on PAMI*, *23*(9), 2001, 947–963.

[10] Y. Chen & J. Wang, A region-based fuzzy feature matching approach to content-based image retrieval, *IEEE Trans on PAMI*, *24*(9), 2002, 1252–1267.

[11] F. Hoppner, F. Klawonn, R. Kruse, & T. Runkler, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition* (John Wiley & Sons, 1999).

[12] C. Vertan, & N. Boujemaa, Color Texture Classification by Normalized Color Space Representation, *Int. Conf. on Pattern Recognition (ICPR'00)*, Barcelona, Spain, 2000, 3584-3587.

[13] B. S. Manjunath, P. Salembier, T. Sikora, *Introduction to MPEG-7 Multimedia Content Description Interface* (John Wiley & Sons, 2002).