

A Unified Fuzzy Feature Indexing Scheme for Region Based Online Image Querying

Ruofei Zhang, Zhongfei (Mark) Zhang, Jian Yao
Department of Computer Science
State University of New York at Binghamton, NY 13902, U.S.A.
Email: rzhang@binghamton.edu

Abstract

This paper describes a novel indexing and retrieval methodology integrating color, texture and shape information for content-based image retrieval in online image databases. This methodology, called PicSearcher, applies unsupervised image segmentation to partition an image into a set of regions, then fuzzy color histogram as well as fuzzy texture and shape properties of each region is calculated to be part of their signatures. The fuzzification procedures resolve the recognition uncertainty stemming from color quantization and human perception of colors. At the same time, this unified fuzzy scheme incorporates the segmentation-related uncertainties into the retrieval algorithm. Then an adaptive and effective measure for the overall similarity between images is developed by integrating properties of all the regions in the image. An implemented prototype system of PicSearcher has demonstrated a promising retrieval performance for an online test database containing 10,000 general-purpose color images, as compared with its peer systems in the literature.

1. Introduction

This work addresses the topic of general purpose content based image retrieval (CBIR). CBIR has received intensive attention in the literature since this area was started a few years ago, and consequently a broad range of techniques [1] is proposed. The ultimate goal of this project is to develop a semantics-sensitive, efficient image retrieval algorithm.

The majority of the early works focuses on global features of imagery, while more recently region-based approaches have been shown to be more effective [2]. The most fundamental and popularly used feature is the color histogram and its variants, which was used in the classic systems such as IBM QBIC [3] and Berkeley Chabot [4]. Since color histograms do not carry spatial information, which is directly related to the semantics of image content, efforts have been reported in the literature to change the color histograms in this regard. Pass and Zabih [5] proposed the Color Coherence Vector (CCV) to address this issue. Huang et al. [6] went further to use the Color Correlograms to integrate color and spatial information. Rao et al. [7] generalized the color spatial distribution as a Geometric Histogram.

A region-based retrieval system segments images into regions, and retrieves images based on the similarity between regions. Berkeley Blobworld [8] and UCSB Netra [9] compare images based on individual regions. To query an image, the user is required to select regions and the corresponding features for the similarity computation. Smith and Chang [10] partitioned an

image in regions using a sequential labeling algorithm. Wang et al. [11] recently proposed an integrated region matching scheme called IRM for CBIR, which allows for matching a region in one image to several regions of another image. As a result, the similarity between two images is defined as the weighed sum of distances, in the feature space, between all regions from different images.

In this paper we describe an efficient clustering-based fuzzy feature representation approach, called PicSearcher, to address the general-purpose CBIR problem. In this approach we integrate semantic-intensive clustering-based segmentation with unified fuzzy representation of color histogram, texture and shape to index image databases. A low computational as well as robust distance metric is developed to reduce the query time of the system. Below we describe the image segmentation and unified fuzzy feature indexing respectively, then region machine scheme of PicSearcher algorithm is presented. Finally the experimental evaluations are reported before we conclude this paper.

2. Image segmentation

In our system, the query image and all images in the database are first segmented into regions and the fuzzy feature of color, texture and shape are extracted to be the signature of the image. The image segmentation is based on color and spatial variation features using k-means algorithm [12]. To segment an image, the system first partitions the image into 4×4 blocks to compromise between texture effectiveness and computation time. We extract a feature vector consisting of six features from each block. Three of them are average color components in a 4×4 block. We use the LAB color space because of its desired property that the perceptual color difference is the numerical difference, these features are denoted as $\{C_1, C_2, C_3\}$. The other three features represent energy in the high frequency bands of the Haar wavelet transform [13], that is, the square root of the second order moment of wavelet coefficients in high frequency bands [14]. These three features of the block are denoted as $\{T_1, T_2, T_3\}$. They can be used to discern texture by showing variations in different directions.

After we get feature vectors for all blocks, we perform normalization on both color and texture feature to whiten them so the effects of different feature range are eliminated. Then the k-means algorithm is used to cluster the feature vectors into several classes with every class corresponding to one region in the segmented image. Because clustering is performed in the feature space, blocks in each cluster do not necessarily form a connected region in the image. This way, we preserve the natural clustering of objects in general-purpose images. The k-

means algorithm does not specify how many clusters to choose. We adaptively select the number of clusters C by gradually increasing C until a stop criterion is met.

After segmentation, three extra features are calculated for each region to describe shape properties. They are normalized inertia [15] of order 1 to 3. For a region H in 2-dimensional Euclidean space \mathfrak{R}^2 (an image), its normalized inertia of order p is

$$l(H, p) = \frac{\sum_{(x,y) \in H} [(x-\hat{x})^2 + (y-\hat{y})^2]^{p/2}}{[V(H)]^{1/p/2}} \quad (1)$$

where $V(H)$ is the number of pixels in the region H and (\hat{x}, \hat{y}) is the centroid of H . The minimum normalized inertia is achieved by spheres. Denote the p^{th} order normalized inertia of spheres as L_p , we define following features to describe the shape of the region:

$$S_1 = l(H,1)/L_1, S_2 = l(H,2)/L_2, S_3 = l(H,3)/L_3 \quad (2)$$

3. Unified fuzzy feature indexing

In each segmented region, we apply fuzzy logic to define the fuzzy color, texture, and shape features, respectively. The motivation to incorporate fuzzy logic into the features is to address the uncertainty or inaccuracy issue to improve the robustness and effectiveness of the indexing scheme, when the features are extracted during indexing (e.g., color inaccuracy perceived by human vision, texture inaccuracy obtained by different texture measures or from varied regions, and shape uncertainty of a region from different region segmentations) such that a certain degree of variations of the feature values is allowed.

We use the color histogram as the color feature for a segmented region. Thus, the fuzzy color histogram is defined as follows. We assume that any color c is a fuzzy set. As a result, for any color c' of the color universe $\mu_c(c') : \mu \rightarrow [0,1]$ is the resemblance degree of the color c' to the color c . Note that a good fuzzy resemblance function should admit that the resemblance degree decreases as the inter-color distance increases. The most commonly used fuzzy resemblance functions are cone, trapezoidal, B-splines, exponential, Cauchy, and paired sigmoid functions [16]. Considering the computational complexity, we pick the Cauchy function to be the fuzzy resemblance function. The color resemblance in a region is defined as:

$$\mu_c(c') = \frac{1}{1 + \left(\frac{d(c,c')}{\sigma}\right)^\alpha} \quad (3)$$

where d is the Euclidean distance between color c and c' in LAB space, σ is the average distance between colors

$$\sigma = \frac{2}{B(B-1)} \sum_{i=1}^{B-1} \sum_{k=i+1}^B d(c, c') \quad (4)$$

where B is the number of bins in our color partition. Figure 1 illustrates Cauchy function in \mathfrak{R} with $v=0$, $d=36$, and α varying from 0.01 to 100. The experiments show that the performance changes insignificantly when α is in the interval [0.7, 1.5], but degrades rapidly outside the interval. So we set $\alpha = 1$ to simplify the computation.

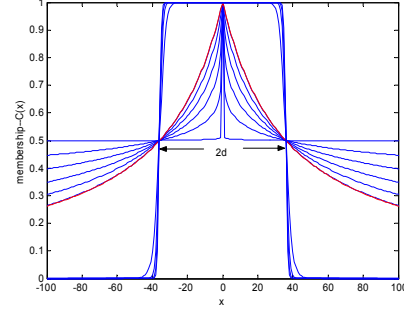


Figure 1. Cauchy functions in 1 dimension

This fuzzy color model enables to enlarge the influence of a given color to its neighboring colors, according to the uncertainty principle and the perceptual similarity. This means that each time a color C is found in the image, it will influence all the quantized colors according to their resemblance to the color C . Numerically, this could be expressed as:

$$h_2(c) = \sum_{c' \in \mu} h_1(c') \mu_c(c') \quad (5)$$

where μ is the color universe in the image and $h_1(c')$ is the usual normalized color histogram.

In our system, the LAB color space is quantized into 96 bins by using uniform quantization (L by 6, A by 4, B by 4). Then use above formula to calculate the fuzzy histogram for each region. $\mu_c(c')$ for each bin is pre-computed and implemented as a lookup table to reduce the online computation.

To accommodate the imprecise image segmentation and uncertainty characteristics for human perception, we propose to fuzzify each region generated from image segmentation by a fixed parameterized membership function. The parameter for the membership functions is calculated using the clustering results. The fuzzification of feature vectors brings in a crucial improvement on the region representation of an image: fuzzy features naturally characterize the gradual transition between regions within an image. A fuzzy feature set assigns weights, called degree of membership, to feature vectors of each block in the feature space. As a result, feature vector of a block usually belongs to multiple regions with different degrees of membership as opposed to the classical region representation, in which a feature vector belongs to exactly one region. This fuzzification technique has two major advantages: (i) it makes the retrieval system more accurate and robust to image alterations such as intensity variation, color distortion, and shape distortion, etc.; (ii) it's better in extracting useful information under the same uncertain conditions, i.e., it is more robust to imprecise segmentation.

To make our fuzzification scheme unified to be consistent to the fuzzy color histogram representation, we again use Cauchy operator to be our fuzzy membership function.

$$\mu_c(f) = \frac{1}{1 + \left(\frac{d(f, \hat{f}_i)}{\sigma}\right)^\alpha} \quad (6)$$

where $f \in \mathfrak{R}^k$ (in our approach, $k=3$) is the texture feature vector of each block, \hat{f}_i is the average texture feature vector of region i , d is the Euclidean distance between \hat{f}_i and f , σ represents

the average distance for texture features between cluster centers \hat{f}_j we get from the k-means algorithm, it's defined by:

$$\sigma = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{k=i+1}^C \|\hat{f}_i - \hat{f}_k\| \quad (7)$$

By this way, a region is described as a fuzzy set to which each block has a membership to avoid a hard segmentation so the inaccuracy stemming from image segmentation is addressed explicitly.

Accordingly by making use of this block membership functions, the fuzzy texture properties of region R_i is represented as

$$\hat{f}_i^T = \sum_{f \in U^T} f \mu_i(f) \quad (8)$$

where U^T is the feature space composed by texture features of all blocks.

Based on the same fuzzy membership function, we also fuzzify the shape property representation of region i by modifying Eq. (1) as:

$$l(i, p) = \frac{\sum_{f \in U^S} [(f_x - \hat{x})^2 + (f_y - \hat{y})^2]^{p/2} \mu_i(f)}{[N]^{1+p/2}} \quad (9)$$

where N is the number of blocks in an image and U^S is the block space in an image. According to Eq. (2) with Eq. (9), we get the fuzzy shape feature $\hat{f}_i^S = \{S1, S2, S3\}$ of each region.

4. Region matching and similarity calculation

Once we have the fuzzified color, texture, and shape feature vectors ready for a region of an image, we store the following information in the database for this region: (1) Fuzzy color histogram $h(c)$ (2) Fuzzy texture feature \hat{f}^T (3) Fuzzy shape feature \hat{f}^S (6) The relative size of the region to the whole image (7) The central coordinates of the region area (\hat{x}, \hat{y}) . Based on this stored information, we define the similarity for texture, shape, and color features between two regions p and q as the L2 distance functions:

$$d_{CT}^{pq} = \|\hat{f}_p^T - \hat{f}_q^T\| \quad (10)$$

$$d_{ST}^{pq} = \|\hat{f}_p^S - \hat{f}_q^S\| \quad (11)$$

$$d_C^{pq} = \sqrt{\frac{\sum_{i=0}^{B-1} [(h_p(i) - h_q(i))^2]}{B}} \quad (12)$$

where B is the number of bins of the color histograms, and $h_p(i)$ and $h_q(i)$ are fuzzy histogram of region p and q respectively.

Since typically color and texture features are more prominent and precisely measured than the shape feature in our system, we divide the overall distance between two regions into two parts: the color and texture component

$$d_{CT}^{pq} = \sqrt{d_C^{pq^2} + d_{ST}^{pq^2}} \quad (13)$$

and the shape component d_S^{pq} , the overall distance is the linear combination of the two components

$$DIST(p, q) = w d_{CT}^{pq} + (1-w) d_S^{pq} \quad (14)$$

where the weight w allows to favor the more reliable color and texture features.

Given the definition of the distance between two regions, we are ready to compute the global similarity between two images. Suppose we have M regions in image 1 and N regions in image 2, the following algorithm computes the global similarity between image 1 and image 2:

Step 1: compute the distance between one region in image 1 and all regions in image 2. For each region R_i in image 1, the distance between this region and image 2 is defined as:

$$R_{i \text{ Image2}} = \text{Min}\{DIST(i, j)\} \quad (15)$$

where j enumerates each regions in image 2. This definition captures the minimum distance between a region and all the regions in an image, which maximizes the potential similarity between the region and the image.

Step 2: similarly, the distance between a region R_j in image 2 and image 1 is defined as

$$R_{j \text{ Image1}} = \text{Min}\{DIST(j, i)\} \quad (16)$$

where i enumerates each regions in image 1.

Step 3: now we have $M+N$ distances. We define the distance between two images (1 and 2) as follows:

$$\text{Dist Image}(1,2) = \frac{\sum_{i=1}^M w_{1i} R_{i \text{ Image2}} + \sum_{j=1}^N w_{2j} R_{j \text{ Image1}}}{2} \quad (17)$$

where w_{1i} is the weight for region i in image 1, and w_{2j} is the weight for region j in image 2. Since we think a region with a larger area plays a more significant role in contributing to the overall similarity value between two images than a region with a smaller area, we define $w_{1i} = \frac{N_{1i}}{N_1}$, where N_{1i} is the number of

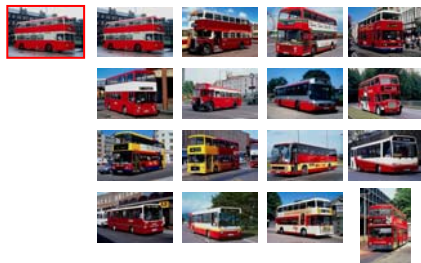
blocks in region i and N_1 is the total number of blocks in image 1; define w_{2j} similarly for image 2. Then the distance is normalized to $[0, 1]$ for comparison.

This definition of the overall similarity between two images captured by the overall distance between the images is a balanced scheme in similarity measure between regional and global matching. As compared with many existing similarity measures in the literature, this definition strives to incorporate as much semantic information as possible and at the same time also achieves a computational efficiency.

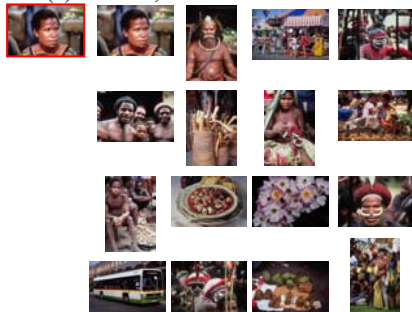
5. Experiments and results

We have tested our approach in an online general-purpose image database containing 10,000 images from COREL. These images are manually divided into 10 well-defined groups such as people, nature scene, building, vehicles, etc. No pre-restriction such as different camera models, lighting conditions, etc. are specified in the image database for the testing. These images are all in JPEG format

We implemented the system as a web application so users can retrieve the online database conveniently. To evaluate the accuracy of the approach we randomly pick query images from each group and the semantic relevance evaluation of the retrieval results is based on the group membership of the query image, which is done by human being's subjective observation. Figure 2 shows two examples of the retrieval results.



(a) Vehicle; 16 matches out of 16.



(b) African people; 12 matches out of 16.

Figure 2. The retrieval results of two query images; the top-left corner image is the query and the ranking goes rightward and downward

To evaluate our approach more quantitatively, we compare the performance of our system with the IRM [11] system and geometric histogram based system [7]. For each group we randomly select 1000 images as queries submitting to the system and calculate the average precision for each group based on the returned top 100 images of these two approaches. Since the number of relevant images in the database for each query image is same, we don't calculate the recall explicitly since it's proportional to the precision in this case. The precision comparison is shown in Figure 3:

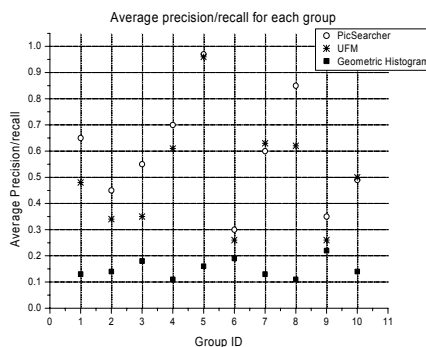


Figure 3. Average precision/recall comparison between PicSearcher, IRM, and Geometric Histogram systems

We can see from these figures that the performance of our approach is much better than the spatial histogram because the histogram approach is one of low level comparing method and is not semantically related. Experiment results also show that our approach outperforms the IRM approach in most groups. The performance gaining stems in part from the explicit addressing to the image segmentation uncertainties as well as more detail color representation in our system.

6. Conclusions

A novel image indexing and retrieval methodology, called PicSearcher, is presented for the goal of semantics-sensitive, efficient CBIR. As compared with the existing techniques and systems, our approach exhibits the following distinctive advantages: (1) PicSearcher proposes a new indexing method based on fuzzy logic to incorporate color, texture, and shape information into a region based approach to robust effectiveness for image indexing (ii) it develops a balanced scheme in similarity measure between regional and global matching in order to capture as much semantic information as possible without sacrificing the efficiency

7. References

- [1] M. De Marsicoi, L. Cinque, S. Levialdi. Indexing pictorial documents by their content: a survey of current techniques. *Image and Vision Computing*, 15: 119-141, 1997
- [2] Arnlod W. M. Smeulders, Marcel Worring et al., "Content-Based Image Retrieval at the End of the Early Years", *IEEE Trans. On PAMI*, Vol. 22, No. 12, Dec. 2000
- [3] M. Flickner et al., "Query by Image and Video content: The QBIC system", *IEEE Computer*, 28(9):23-32, Sep. 1995
- [4] Virginia Ogle, Michael Stonebraker. Chabot: Retrieval from a relational database of images. *IEEE Computer*, 28(9), Sep. 1995, pp. 40-48
- [5] Greg Pass, Ramin Zabih. "Histogram Refinement for Content-Based Image", *Proc. of the IEEE Workshop on Applications of Computer Vision*, Sarasota, Florida, Dec. 1996
- [6] Jing Huang, S Ravi Kumar et al., "Image Indexing Using Color Correlograms", *Proc. of the IEEE Conf. Computer Vision and Pattern Recognition*, Puerto Rico, Jun. 1997.
- [7] Aibing Rao, Rohini K. Srihari, and Zhongfei Zhang, "Geometric Histogram: A Distribution of Geometric Configuration of Color Subsets", *Internet Imaging, Proc. of SPIE, Vol. 3964*, San Jose, Jan. 2000, pp. 91-101
- [8] C. Carson, M. Thomas, S. Belongie, et al., "Blobworld: A System for Region-Based Image Indexing and Retrieval", *Proc. of the third international conf. on Visual Information System*, Amsterdam, Netherlands, Jun. 1999, pp. 509-516.
- [9] W.Y. Ma, B. Manjunath, "NaTra: A Toolbox for Navigating Large Image Databases", *Proc. IEEE Int'l Conf. Image Processing*, Santa Barbara, Oct. 1997, pp. 568-571
- [10] John R. Smith, and Shih-Fu Chang, "VisualSEEK: a fully automated content-based image query system", *ACM Multimedia 96*, Nov. 1996
- [11] James Z. Wang, Jia Li, Gio Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries", *IEEE Trans. On PAMI*, Vol. 23, No.9, Sep. 2001
- [12] J.A. Hartigan, M.A. Wong. Algorithm AS136: "A k-means Clustering Algorithm", *Applied Statistics*, Vol. 28, 1979, pp. 100-108
- [13] I. Daubechies. Ten Lectures on Wavelets, Capital City Press, 1992
- [14] M. Unser, "Texture Classification and Segmentation Using Wavelet Frames", *IEEE Trans. Image Processing*, Vol. 4, No.11, Nov. 1995, pp. 1549-1560
- [15] A.Gersho, "Asymptotically Optimum Block Quantization", *IEEE Trans. On Information Theory*, Vol. 25, No. 4, 1979, pp. 373-380
- [16] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, "Fuzzy cluster Analysis: Methods for Classification, Data Analysis and Image Recognition", *John Wiley & Sons, LTD*, 1999