

Image Similarity: A Genetic Algorithm Based Approach

R. C. Joshi, and Shashikala Tapaswi

Abstract—The paper proposes an approach using genetic algorithm for computing the region based image similarity. The image is denoted using a set of segmented regions reflecting color and texture properties of an image. An image is associated with a family of image features corresponding to the regions. The resemblance of two images is then defined as the overall similarity between two families of image features, and quantified by a similarity measure, which integrates properties of all the regions in the images. A genetic algorithm is applied to decide the most plausible matching. The performance of the proposed method is illustrated using examples from an image database of general-purpose images, and is shown to produce good results.

Keywords—Image Features, color descriptor, segmented classes, texture descriptors, genetic algorithm.

I. INTRODUCTION

MANY applications in digital library, entertainment industry, e-commerce domains require access and query of repositories of image data. Image retrieval has traditionally been based on manual caption insertion describing the scene, which can then be searched using keywords. Caption insertion is a very subjective method and becomes extremely tedious and time consuming, especially for large image databases which are becoming ever more common with the growing availability of digital cameras and scanners. The challenge is to bridge the gap between the physical characteristics of digital images (e.g. color, texture) that are used for comparison, and the semantic meaning of the images that are used by humans to query the database [8].

It is believed that the key to effective image retrieval lies in the ability to access the image at the *object* level. This is because users generally want to search for images containing particular object(s) of interest and thus the ability to represent, index and query images at the level of objects is critical [2]. Semantically precise image segmentation by an algorithm is very difficult.

The paper employs unsupervised segmentation method for segmentation of images into classes and computing the similarity between images using genetic algorithm based image

feature matching approach. As these segmented classes are homogeneous in terms of colour and texture in some sense, thus in the proposed approach colour and texture features are used, as they correlate well with the identity of objects. The images are decomposed as combinations of objects, querying becomes more significant and natural than it is with global image properties. This is obviously true for images with distinct foreground objects but the reasoning also holds for 'background' images where no interesting foreground objects are present. Images belonging to the latter category can be considered as consisting of combinations of classes with homogeneous colour and texture (such as, images of the seashore generally consist of the beach and the sea, images of landscapes scenes generally consist of green background), querying and similarity matching is made more effective by being based on these class combinations which characterize the scene.

The proposed implementation, first segment the images based on joint colour and texture features using unsupervised multiscale segmentation algorithm. The segmentation process is performed off-line for each image. Section 2, describes briefly the segmentation algorithm used. The representation of each image using effective and compact colour and texture descriptors of its classes is described in Section 3. In Section 4, we present the genetic algorithm based feature matching approach used to compute the similarity of images. In Section 5, the preliminary results from small image database test bed consisting of a range of natural images are presented. Section 6 gives the conclusion.

II. SEGMENTATION PROCEDURE

The following steps are used for segmentation algorithm:

- (i) Normalized colour and texture features (three for colour and two for texture) are mapped to a multidimensional feature space. Spatial information is incorporated into the process by including spatial features into the feature space. Generally, color spaces that are derived from human color perception models such as the CIE $L^*u^*v^*$ or $L^*a^*b^*$ are more useful for segmentation tasks. The reason is due to the fact that colour perception is by definition a human stimulus and in the absence of a specific application requirement, the segmentation of a colour image to agree with that performed by the human visual system is expected. In the proposed method, the colour space used is S-CIE $L^*a^*b^*$, the spatial extension of the perceptual uniform CIE $L^*a^*b^*$, originally developed by Zhang and Wandell [9]. This colour space takes into account the appearance of fine-patterned colours on the human visual system. Texture features are computed using the logarithm

Manuscript received March 15, 2007.

R. C. Joshi is a Professor in Electronics and Computer Engineering Department at Indian Institute of Technology, Roorkee, India.

Shashikala Tapaswi is an Associate Professor in Department of Information Technology at Atal Bihari Vajpayee Indian Institute of Information Technology and Management, Gwalior, India (corresponding author e-mail: stapaswi@hotmail.com).

- of the energies of the 2-D complex wavelet coefficients [5] and the top two *principal components* are considered.
- (ii) Significant features which correspond to clusters in the feature space are assumed to be representations of underlying regions, the recovery of which is achieved using the *mean shift* procedure [7], a robust kernel based decomposition method. The kernel size used is kept constant for all images, resulting in decomposition into an appropriate number of regions for each image.
 - (iii) By determining the number of classes and the properties of each class using step (ii), a Bayesian multiscale processing approach, which models the inherent uncertainty in the joint specification of class and position spaces using the Multiscale Random Field model [4], is used for the subsequent classification process.

Typical segmentation outputs of some images from the database are shown in Fig. 1.

III. CLASS DESCRIPTION

A. Colour Descriptors

For representing the colour distribution of each class, the colour histograms of the pixels of the class are stored. This histogram is based on bins with width 10 in each dimension of the S-CIE L*a*b* colour space. This spacing yields 10 bins in the L* dimension and 40 bins in each of the a* and b* dimensions, for a total of 90 numbers as colour descriptors.

To evaluate the dissimilarity between the colour histograms of two classes/objects, the *Kolmogorov-Smirnov (K-S)* distance, as originally proposed in [6] is used. The K-S distance essentially measures the difference between two probability distribution functions. If $F_1(k)$ and $F_2(k)$ are two independent sample distribution functions (i.e. histograms) defined such that:

$$F_i(k) = \frac{1}{n} \#(i: y_i^t \leq k) \quad (1)$$

where n is the number of data samples, y_i^t so that $1 \leq i \leq n$, then the K-S distance is the maximum difference between the distributions over all k :

$$K-S(y^1, y^2) = \max |F_1(k) - F_2(k)| \quad (2)$$

The overall colour dissimilarity measure between two classes with colour histograms x^{COL} and y^{COL} is taken to be the root mean square of the K-S distances of each of the L*, a* and b* histograms:

$$d^{COL}(x^{COL}, y^{COL}) = \frac{1}{3} \left\{ [K-S(x_L^{COL} - y_L^{COL})]^2 + [K-S(x_a^{COL} - y_a^{COL})]^2 + [K-S(x_b^{COL} - y_b^{COL})]^2 \right\}^{1/2} \quad (3)$$

As the range of K-S distances lie between 0 and 1, the colour similarity measure, $s_{COL}(x^{COL}, y^{COL})$ is simply taken as:

$$S_{COL}(x, y) = 1 - d_{COL}(x, y) \quad (4)$$

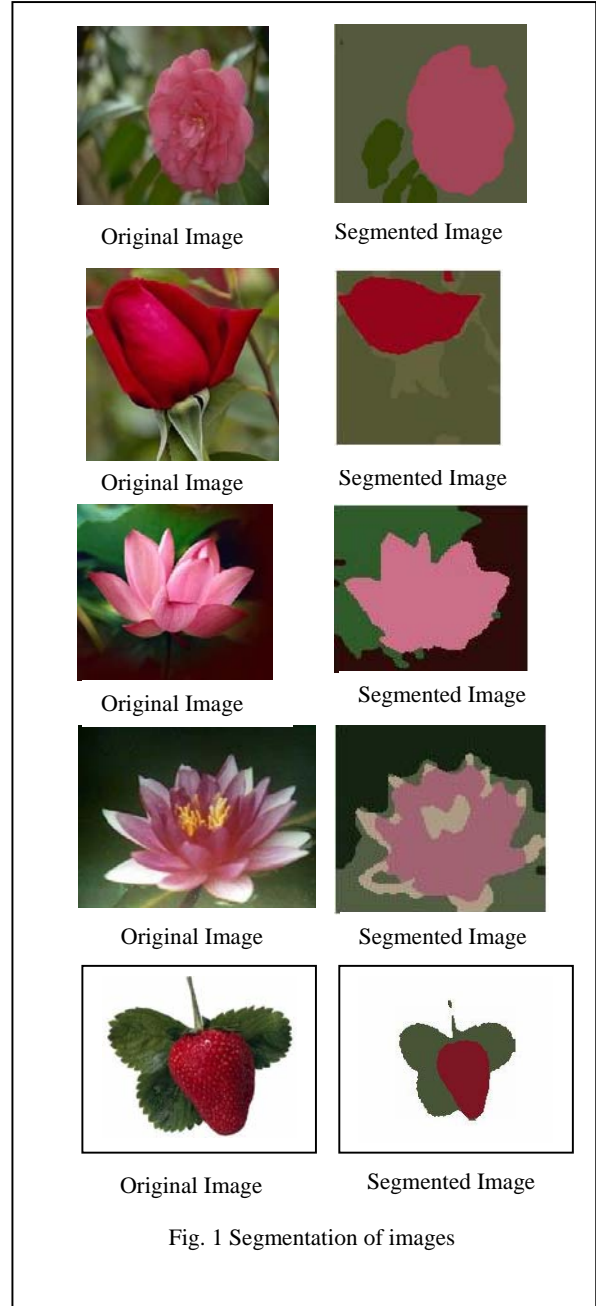


Fig. 1 Segmentation of images

B. Texture Descriptors

For each class, texture is described by the distribution of the magnitude of its complex wavelet coefficients, $f(x^{TEX})$. Using four levels of the 2-D complex wavelet transform (which yields six complex sub bands at every level) produces a total of 24 sub bands, the magnitude of each is converted into a histogram and modeled as a generalized Rayleigh distribution

$$f(x_i^{TEX}) = k_i x_i^{TEX} \exp\left(-\frac{x_i^{TEX}}{\sigma_i}\right)^{\beta_i} ; i = 1, 2, \dots, 24 \quad (5)$$

where, to achieve the same mean and variance as the input sample distribution:

* $k_i = \frac{\beta_i}{\sigma_i^2 \Gamma\left(\frac{2}{\beta_i}\right)}$ where $\Gamma(\cdot)$ denotes the gamma function

* $\sigma_i = \frac{m_1 \Gamma\left(\frac{2}{\beta_i}\right)}{\Gamma\left(\frac{3}{\beta_i}\right)}$ where $m_1 = E\left(x_i^{TEX}\right)$

* $\beta_i = F^{-1}\left(\frac{m_1^2}{m_2}\right)$ where $m_2 = E\left[\left(x_i^{TEX}\right)^2\right]$ and

$$F(x) = \frac{\left[\Gamma\left(\frac{3}{x}\right)\right]^2}{\Gamma\left(\frac{2}{x}\right)\Gamma\left(\frac{4}{x}\right)}$$

It is generally known that the distribution of (real) wavelet coefficients of texture takes the form of a generalized Gaussian density. This means that the real and imaginary parts of their complex wavelet coefficients will similarly have a generalized Gaussian distribution, with the resultant distribution of the magnitudes taking the form of a generalized Rayleigh density. This is the basic motivation for using generalized Rayleigh distribution.

Thus, for each class, the generalized Rayleigh model parameters, σ_i and β_i is calculated for each of the 24 histograms, for a total of 48 numbers as texture descriptors.

A segmented image can be viewed as a collection of regions, $\{R_1, \dots, R_C\}$. Equivalently, in the feature space, the image is characterized by a collection of feature sets, $\{F_1, \dots, F_C\}$, which form a partition of F . We could use the feature set F_j to describe the region R_j , and compute the similarity between two images based on F_j 's. Representing regions by feature sets incorporates all the information available in the form of feature vectors, but it has two drawbacks:

- It is sensitive to segmentation-related uncertainties. For any feature vector in F , under this region representation form, it belongs to exactly one feature set. But, in general, image segmentation cannot be perfect. As a result, for many feature vectors, a unique decision between in and not in the feature set is impossible.

- The computational cost for similarity calculation is very high. Usually, the similarity measure for two images is calculated based on the distances (Euclidean distance is the one that is commonly used in many applications) between feature vectors from different images. The amount of time is certainly too much for system users to tolerate.

In an improved region representation form [10], which mitigates the above drawbacks, each region (R_j) is represented by the center (\hat{f}_j) of the corresponding feature set (F_j) with \hat{f}_j defined as

$$\hat{f}_j = \frac{\sum_{\vec{f} \in F_j} \vec{f}}{V(F_j)} \tag{6}$$

which is essentially the mean of all elements of F_j , and in general may not be an element of F_j . While averaging over all features in a feature set decreases the impact of inaccurate segmentation, at the same time, lots of useful information is also submerged in the smoothing process because a set of feature vectors are mapped to a single feature vector. Moreover, the segmentation-related uncertainties are not explicitly expressed in this region representation form.

IV. IMPLEMENTATION OF GENETIC ALGORITHM

To measure the degree of match between two feature sets P and Q , fitness function is constructed by the partial bidirectional Hausdorff distance [11]. The output of GA is the feature set which has the highest value of fitness function.

A. Chromosome Code

The initial population is generated randomly. However, the generating scope of chromosomes is not arbitrary but limited within the image size [1]. Each chromosome's detailed construction method is seriating of the feature sets and coding them into binary codes.

Here binary coding is used because the image resolution is always limited. The feature sets can be represented by integers in the scope of the image resolution.

B. Fitness Function

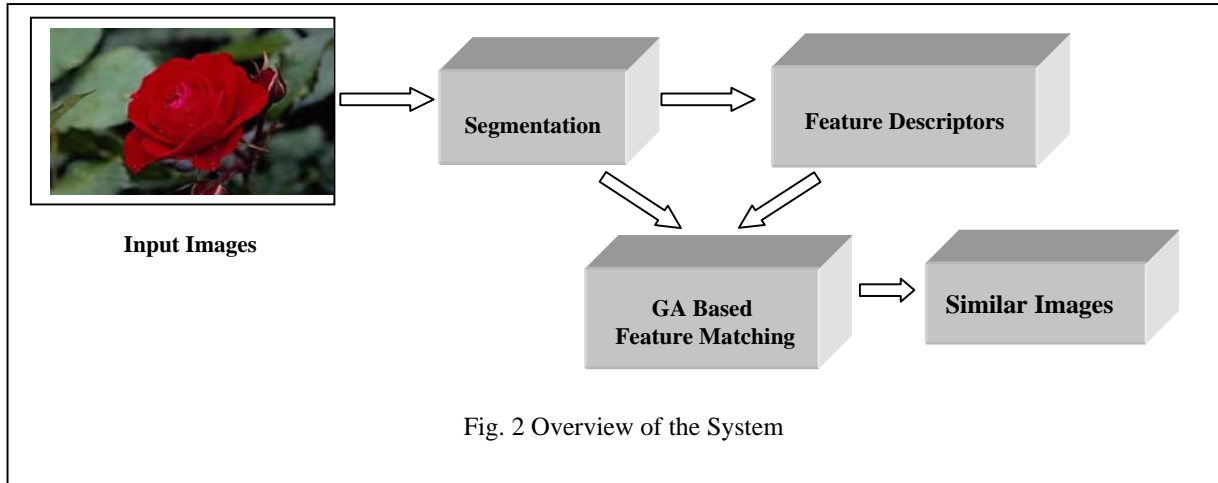
Clearly, if the match degree between P and Q can be measured, it is equivalent to evaluating the fitness of the chromosome [12]. Considering the partial bidirectional Hausdorff distance between feature sets P and Q , the smaller the distance is, the match degree between P and Q is larger. So, the fitness function can be selected as the inverse of the partial bidirectional Hausdorff distance

$$fitness = \frac{1}{a + H_{LK}(P, Q)} \tag{7}$$

where a is a positive constant. In order to avoid the denominator is a zero, the partial bidirectional Hausdorff distance is added a .

C. Genetic Operators

A pair of chromosome is randomly chosen from the population and is used as the parents to reproduce the offspring. The selection principle is the more chromosome number of its new offspring to the next generation, the bigger fitting function value F_i with the larger probability p_i .



The nature choice phenomenon of the biosphere is simulated by

$$p_s = \frac{F_i}{\sum_{j=1}^l F_j} \tag{8}$$

Crossing operator is obtained after resetting the parents facts which are got the selection processing. Crossing processing is performed according to a certain probability, which is called crossing probability p_c . The crossing effect is to produce better chromosome after the combination of the generating materials of the parents.

Here, single point crossover operator is adopted. Mutation operates on each binary bit of a chromosome in another predefined probability p_m the mutation operator is realized by reversing the value of the current binary bit, i.e. 0-1, 1-0.

D. Genetic Algorithm Based Feature Match

For two features sets P and Q , determine population size N , crossing probability p_c , mutation probability p_m , fractions f_L and f_K of the partial bidirectional Hausdorff distance and the maximum iterative steps G_{max}

Step 1: Randomly generate N feature sets in the test image and then convert them into chromosomes for initial generation.

Step 2: Evaluate the fitness of each chromosomes in current population and then create a new population by repeating following steps until the new population is complete.

- 1) Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected).
- 2) With a crossover probability cross over the parents to form new offspring (children). If no crossover was performed, offspring is the exact copy of parents.

- 3) With a mutation probability mutate new offspring at each locus (position in chromosome).
- 4) Place new offspring in the new population.

Step 3: Use new generated population for a further run of the algorithm.

Step 4: If the end condition is satisfied, **stop**, and return the best solution T_{best} in current population, where T_{best} be the feature sets determined by the best chromosome, and feature sets $T_{best}(P)$ and Q according to the simple nearest neighbor rule. If the maximum iterative step G_{max} is not reached, go to *Step 2*.

V. EXPERIMENTAL RESULTS

The experiments for computing the similarity have been performed on a small image database test bed, consisting of *jpg* images. The results of the image segmented regions and classes have been selected through an extensive series of experiments. The corresponding confusion matrix for a set of experiment is given in Table I. The implementation has been done using MATLAB and the preliminary results are encouraging. The image database is being expanded to include as many varied natural images as possible.

For comparing the performance results were generated using a one-class default query. It is observed that the pre-segmentation stage performs better for all image categories tested. These results are consistent with the conviction that the key to effective image retrieval performance can be achieved if images are accessed at the level of objects.

TABLE I
CONFUSION MATRIX

	Confusion Matrix					
Roses	59.4	11.6	4.4	6.6	10.2	7.8
Lily	15.6	41.2	13.1	2.1	24.2	3.8
Sunflower	0.1	2.8	93.5	0.2	2.3	1.1
Bell flower	9.1	2.2	14.7	43.6	20.3	10.1
Bottle Brush	4.2	3.6	1.8	5.4	83.7	1.3
Lotus	0.0	0.0	0.0	0.0	0.0	100

VI. CONCLUSION

In the paper a general and powerful multiscale segmentation algorithm for segmentation of images based on pre-segmented classes is employed, which gives encouraging results. The new approach for region based similarity using genetic algorithm is discussed and demonstrated with its encouraging results. However, low level features like colour and texture are generally insufficient for effective retrieval of unconstrained imagery. The use of many individual populations and evolutionary operators overcomes the problems of greedy algorithms, but still we expect that genetic algorithm can be used for feature matching.

REFERENCES

- [1] H. Bandemer and W.Nather, "Fuzzy Data Analysis," Kluwer Academic Publishers, 1992.
- [2] C.Carson, M. Thomas, S. Belongie, J. Hellerstein, J. Malik, "Blobworld: A System for Region-based Image Indexing and Retrieval," Proceedings of the 3rd Intn'l Conference on Visual Information Systems, 1999.
- [3] C.Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, 2(2), 1998.
- [4] C.Bouman, M. Shapiro, "A Multiscale Random Field Model for Bayesian Image Segmentation," IEEE Transactions on Image Processing, 3(2): 162-177, 1994.
- [5] N.Kingsbury, "Shift Invariant Properties of the Dual-Tree Complex Wavelet Transform," Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing, Phoenix, Arizona, USA, 1999.
- [6] A.H.Kam,T.T.Ng, N.G.Kingsbury and W.J.Fitzgerald : "Content based image retrieval through object extraction and querying" ,Proc. IEEE Workshop on Content-based Access of Image and Video Libraries, Hilton Head Island, S Carolina, June 12, 2000.
- [7] K.Fukunaga and L. Hosteler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition," IEEE Transactions on Information Theory, 21:32-40, 1975.
- [8] J. Smith, "Integrated Spatial and Feature Image Systems: Retrieval, Compression and Analysis," Ph.D thesis, Columbia University, USA, February, 1997.
- [9] X. Zhang and B. Wandell, "Color Image Fidelity Metrics Evaluated Using Image Distortion Maps," Signal Processing, 70:201-214, 1998.
- [10] J.Li, J.Z.Wang and G.Wiederhold, "IRM: Integrated region matching for image retrieval," Proc 8th ACM Int'l Conference on Multimedia, pp.147-156, October 2000.
- [11] D.P.Huttenlocher, G.A.Klanderma, W.J.Rucklidge, "Comparing images using the Hausdroff distance", Transaction on Pattern Analysis and Machine Intelligence, v.15, pp.850-863, 1999.
- [12] J.T.Alander, "An Indexed Bibliography of Genetic Algorithms in Pattern Recognition", Report 94-1-PATTERN, University of Vaasa, 2001. URL :ftp://ftp.uvaasa.fi/cs/report94-1/.

Shashikala Tapaswi is an Associate professor in Information Technology Department, Atal Bihari Vajpayee Indian Institute of Information Technology and Management, Gwalior, India. She has obtained her Ph.D. (Computer Engineering) from Indian Institute of Technology, Roorkee, India, M.Tech (Computer Science) from University of Delhi, India and B.E.(Electronics Engineering) from Jiwaji University, Gwalior, India. Her primary research areas of interest are Artificial Intelligence, Neural Networks, Fuzzy Logic, Digital Image Processing, Computer Networks etc.

R. C. Joshi received his PhD Degree in Electronics and Computer Engineering from IIT, Roorkee. Since September 1987 he has been a Professor in the Department of Electronics and Computer Engineering, IIT, Roorkee. He has served as Head of the Department twice; first from January 1991-January 1994 and then from January 1997-December 1999. He has published several research papers in referred journals / conferences. His current research interests include database systems, data mining and knowledge discovery, mobile and distributed computing, sensor networks and security.