

IMAGE RETRIEVAL WITH VISUALLY PROMINENT FEATURES USING FUZZY SET THEORETIC EVALUATION

M. Banerjee

Machine Intelligence Unit
I.S.I

203, B. T. Road, Kolkata - 700 108
minakshi_r@isical.ac.in

M. K. Kundu

Machine Intelligence Unit
I.S.I

203, B. T. Road, Kolkata - 700 108
malay@isical.ac.in

P. K. Das

Dept. of Computer Science
Jadavpur University

Jadavpur, Kolkata -700032
pkdas@ieee.org

Keywords: Content-based image retrieval, high curvature points, fuzzy feature evaluation index, color, invariant moments.

Abstract

This paper proposes a new image retrieval scheme using visually significant features. Clusters of points around significant curvature regions (high, medium, weak type) are extracted to obtain a representative image. Illumination, viewpoint invariant color features are computed from those points for evaluating similarity between images. Relative importance of the features are evaluated using a fuzzy entropy based measure computed from relevant and irrelevant set of the retrieved images marked by the users. The performance of the system is tested using different set of examples from general purpose image database. Robustness of the system has also been shown when the images have undergone different transformations.

1 Introduction

Effective image retrieval techniques from a large database is a difficult problem and still far being solved. Retrieval of relevant images, based on automatically derived imagery features (color, texture, shape etc.) is popularly known as content based image retrieval(CBIR). There are several popular CBIR techniques reported in the literature [19], [3], [5], [9]. The information in an image sometimes involve significant amount of reasoning about the meaning or the purpose of the objects or scenes depicted. As a result, it is still not possible to achieve desired accuracy from a fully automated CBIR system. The accuracy of a CBIR system may be improved by iterative process of refinement of queries and features, guided by users feedback [22] known as relevance feedback mechanism. Owing to these facts, derivation and selection of optimal set of features still remain a challenging issue in designing an efficient CBIR system.

Human visual system is highly efficient in sorting and selecting similar images from a very large collection. In this process of selection, a limited number of visually prominent features may be used to evaluate similarity between images. High curvature points play a significant role in characterizing an object with limited number of pixels compared to the total number of pixels. Designing a system using such features for retrieval

mechanism will be both fast and cost effective proposition. Image retrieval tasks based on visually significant points [10], [12] are reported in literature. In [11] local features (color, texture etc.) are computed on a window of regular geometrical shape surrounding the corner points. General purpose corner detectors [8] are also used for this purpose. However the curvature points may be of different types (sharp, medium, weak). The characteristics of sharp curvature points will be confined within a small region but for that of medium and weak type the region will be larger. These facts indicate that extracting the possible high curvature region of interest (roi) where shape and size of the extracted (roi) varies adaptively according to the nature of curvature type could be a better solution. It may be considered an alternative to segmenting an image and using it in overall scene matching applications. An efficient CBIR system should be able to handle imprecise image data, to some extent the difference arising due individual perception in evaluating similarity between images. A Fuzzy set theoretic based approach may be considered a good choice for handling uncertainties arising at different stages of processing and analysis of a CBIR system [4]. The proposed technique is based on the assumption that two visually similar images will have similar visual characteristics. Each feature has its individual significance where the importance of each feature may vary depending upon the query type and applications. Looking into these aspects, the basic contribution of the paper involves (i) A fuzzy set theoretic approach for extraction of clusters of different types of curvature points (sharp, medium, weak) whose centroid almost depicts a true corner. These points are considered as the candidate points for computation of features. (ii) The Invariant global moments of the extracted points set are source of similarity evaluation. Beside these features some global measurements have also been considered to improve the results further. (iii) A feature evaluation mechanism is provided to enhance the accuracy of the system further. The user marks the relevant images within the retrieved set. The individual feature weights are updated with a measure namely fuzzy feature evaluation index(FEI) [15] computed from 'intra set ambiguity' and the 'inter set ambiguity' of the relevant and irrelevant set of images. The results of the proposed methodology is compared with that of some well known techniques (a) integrated region based approach [21], [4] (b) color histogram method [17]. The organization of the paper is as follows, The proposed methodology and results are

described in section 2. The paper is concluded in section 3.

2 The proposed methodology

The feature extraction and feature evaluation methods are explained in the following subsections.

2.1 Extraction of corner signature

The proposed technique is based on a work reported in [2]. The potential fuzzy corner regions are extracted using the Topographic characteristics of intensity surfaces [18], [16]. The uncertainty arising in locating such points are handled using fuzzy set theoretic approach. The discontinuities in the intensity surfaces are the possible candidates for detecting curvature points. These points are characterized with gradient membership ($\mu_d(P)$) generated by a S type function shown in Fig. 1(a). The assignment of the membership value is based on the local gray level contrast [1].

$$\mu_d(P) = S(x; a, b, c) \quad (1)$$

, where x is determined from, ratio of contrast between two opposite pixels (X_{mr}) over a specified window shown in Fig. 1(b). $X_r = \left[\frac{1+|a_1-a_2|}{1+|c_1-c_2|}, \frac{1+|c_1-c_2|}{1+|a_1-a_2|}, \frac{1+|b_1-b_2|}{1+|d_1-d_2|}, \frac{1+|d_1-d_2|}{1+|b_1-b_2|} \right]$

$$X_{mr} = \min\{X_r\} \quad (2)$$

The parameters a and c are determined from the maximum and minimum value of X_{mr} , which maps the membership between 0.0 to 1.0. Two more memberships (μ_f, μ_b) are computed to estimate the strength of connectedness on both sides of the curvature junction. The memberships assigned are related to the value of the curvature subtended at the corresponding points. Different set of curvature points C_k are obtained, by selecting different membership values $\mu_d(P)$ as threshold and using fuzzy rules on the computed features (μ_f, μ_b). If the threshold is selected at ($\mu_d(P) \geq 0.5$) both higher and medium type of curvature points will be extracted together. Thresholding at values ($\mu_d(P) \leq 0.5$) are not considered because along with high and medium type of curvature points lot of spurious curvature points are simultaneously selected which may reduce the accuracy of the techniques. Experimentally it has been found that better results are with values $\mu_d(P) \geq 0.5$ typically (.6, .7, .8). The signature does not change significantly under varying imaging conditions as shown in Fig. 4(a), (b). The extracted signature for some images are shown in Figs. 2, 3. The cluster of points (roi) around the corners will not only carry shape information but also carry information about the spatial distribution of those points. The procedure is implemented for color images by converting RGB plane to HSI and considering only the intensity component for detecting the points.

2.2 Computation of global color moments at selective points

Among the different color models reported in literature, Normalized rgb representation, the illumination and viewing geometry invariant representations which mostly belong to HSI family of color model are popular. In addition to these traditional color spaces, new invariant color models (c_1, c_2, c_3) have been proposed in [6] which discounts the effects of shading and shadows also. We have chosen the (c_1, c_2, c_3) invariant feature model which are defined in the following (3). This model is able to denote the difference between two colors based on their perceptual difference. Invariant color representation although very popularly used in CBIR, these models have short comings under certain situations, due to some loss of discrimination power among images. The RGB plane is converted to (c_1, c_2, c_3)

$$\begin{aligned} c_1 &= \arctan(R/\max(G, B)) \\ c_2 &= \arctan(G/\max(R, B)) \\ c_3 &= \arctan(B/\max(R, G)) \end{aligned} \quad (3)$$

In the next step, the color property of the selected candidates of (roi) is extracted (using (3)) from each of the component planes (c_1, c_2, c_3) for computation of centralized second moments $\mu_{20}, \mu_{02}, \mu_{11}$. The normalized second central moments are computed using (4).

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (4)$$

where $\gamma = \frac{p+q}{2} + 1$

Based on these, the moment invariant to translation, rotation, scale is derived as shown in (5). These set of feature can also be considered as global descriptor of a shape with invariance properties and with an built in ability to discern and filter noise [14], [7].

$$\phi = \eta_{20} + \eta_{02} \quad (5)$$

The image is characterized in the following manner. The moment (ϕ) of the extracted significant spatial locations (roi) will help to identify color similarity of the identified regions. However for natural images (consisting of different objects) the representation obtained from shape signature, although important but may not be sufficient for discriminating them from other categories. Considering these facts three additional set of moments (considering all points from each of (c_1, c_2, c_3) plane) are computed. The components of the feature vector $F_k = [f_1, f_2, f_3, \dots, f_6]$ are as follows, f_1, f_2, f_3 represent the (ϕ) values computed, considering all points of each component plane obtained from (3). These values donot vary with the thresholding levels. f_4, f_5, f_6 contain the (ϕ) values as obtained from the representative locations (roi) of the generated signature C_k from each component plane.

2.3 Estimation of relative importance of different features

For a given data base, different combination of features (color, texture, shape etc.) may be effective for handling different types of queries [13]. Within a selected feature set, the individual feature weights may be further updated for specifying its importance for improvement in precision. For a particular query, the relative importance of the features are estimated from a measure, fuzzy feature evaluation index (FEI) proposed by Pal et al., [15]. The fuzzy measure (FEI) defined from interclass and intraclass ambiguities are as follows. Let $C_1, C_2, \dots, C_j \dots C_m$ be the m pattern classes in an N dimensional $(x_1, x_2, x_q, \dots, x_N)$ feature space where class C_j contains, n_j number of samples. It is shown that the value of the proposed measures [16], ν_{qj} (index of fuzziness) or H_{qj} (fuzzy entropy) gives a measure of 'intra-set ambiguity' along the q th co-ordinate axis in C_j . For computing ν and H of C_j along q th component, an S-type function is considered and the parameters are set as, $b = (x_{qj})_{av}$, $c = b + \max\{|(x_{qj})_{av} - (x_{qj})_{\max}|, |(x_{qj})_{av} - (x_{qj})_{\min}|\}$, $a = 2b - c$, where $(x_{qj})_{av}$, $(x_{qj})_{\max}$, $(x_{qj})_{\min}$ denote the mean, maximum and minimum values respectively computed along the q th co-ordinate axis over all the n_j samples in C_j . Since $\mu(b) = \mu((x_{qj})_{av}) = 0.5$, the values of ν and H are 1.0 at $b = (x_{qj})_{av}$ and would tend to zero when moved away from b towards either c or a of the S function. The higher the value of ν or H the more would be the number of samples having $\mu(x)$ equal to .5 and hence greater would be the tendency of the samples to cluster around the mean value, resulting in less internal scatter within the class. After combining the classes C_j and C_k the mean, maximum and minimum values $(x_{qkj})_{av}$, $(x_{qkj})_{\max}$, $(x_{qkj})_{\min}$ respectively of q th dimension over the samples $(n_j + n_k)$ are computed.

The criteria of a good feature is that it should be invariant to within class variation while emphasizing differences between patterns of different types [15]. The value of ν or H would therefore decrease as the goodness of the q th feature in discriminating pattern classes C_j and C_k increases. The measure denoted as ν_{qjk} or H_{qjk} is called "Inter-set ambiguity" along q^{th} dimension between classes C_j and C_k . Considering the two types of ambiguities, the proposed Feature evaluation index (FEI) for the q th feature is,

$$(FEI)_q = \frac{d_{qjk}}{d_{qj} + d_{qk}} \quad (6)$$

where d stands for ν , H . The lower the value of $(FEI)_q$, the higher is the quality of importance of the q th feature in recognizing and discriminating different classes. The precision of retrieval can be improved by emphasizing the weight of the feature that helps in retrieving the relevant images while reducing the importance of the features that deter the process. Squire et.al [20] proposed a weight adjustment technique based on the variance of the feature values. Since each feature is weighted by its relative importance say w_i . The weighted feature vector is now represented as,

$$F = \sum_{i=1}^N w_i f_i \quad (7)$$

where w_i is the weight associated with the feature component f_i . The importance of each feature component (f_i) is evaluated as follows. Initially all components are considered to be equally important and i.e., $w_i = 1.0$ and the candidate images are retrieved using Euclidean distance metric as the similarity measure. Images are ranked according to this distance. From the first retrieved set of 20 images, the user marks the relevant set of images. The rest of the images from the set are considered irrelevant. The FEI for each feature component is evaluated from the two obtained classes relevant (intra-class), irrelevant (inter-class). Better results are obtained if weights are adjusted based on the FEI values as, $w_i = (FEI)^2$. A new retrieved set of the same query is then obtained. This process may be followed over a number of iterations.

2.4 Experiment

The performance of image retrieval system is tested upon two databases (a) SIMPLiCity images (b) corel 10000 miscellaneous database down loaded from (http://bergman.stanford.edu/cgi-bin/www_wavesearch). The SIMPLiCity database consists of 1000 images from 10 different categories (Africa, Beach, Buildings, Buses, Dinosaurs, Elephants, Flowers, Mountains and Food). Each category is having around 100 images alongwith some images undergone changes due to (rotation, translation, scaling, noise injection, illumination) etc. Our main objective is to design a CBIR system using simple techniques involving low cost feature extraction mechanism. Moments generate a compact representation with fewer number of features compared to other sophisticated features. It may at the same time yield poor results if query complexity increases. The experiments are performed in the following manner. We started with the proposed feature set as explained in section 2.2, and obtained satisfactory results for almost all categories except for few cases. In such cases better results are obtained by computing invariant moments (ϕ) directly from the RGB component planes without using (3). We designate the feature set as computed from (c_1, c_2, c_3) model using (3) as set (A), and those computed directly from RGB plane as set (B). Such differences in performance as can be explained from the fact that, the RGB components are sensitive to varying imaging conditions but have better discriminating power among images. For an unknown database, feature set (B) becomes a good choice when there is less variation in the imaging conditions or perception. The retrieval score was further enhanced by combining both the set of features in a hierarchical fashion. This was tested on 10,000 images, where Illumination invariant set (A) is used first to get the short listed candidates around (100 images) and a second set of retrieval is performed on the short listed candidates using set (B). Each retrieved set can be further

subjected to feature updation scheme for generating still better results. The experimental results are shown from Figs. 2 to Figs. 11. The results are explained as follows : The shape signature of the image in Fig. 2(a) is shown in Fig.2(b) at $\mu_d(P) \geq 0.8$. Similarly for Fig. 3(a) the signature thresholded at $\mu_d(P) \geq 0.7, 0.8$ respectively is shown in Fig. 3(b), (c). The representative corners under varying imaging condition (blurred, illumination change) of Fig. 3(a) are shown in Fig. 4 (a), (b). The query results of SIMPLCcity data set are shown in (Figs. 5 - Figs. 10). The signature is generated at a threshold value of $\mu_d(P) \geq 0.8$. Images are displayed from left to right according to the Euclidean distance with the top left image as the query image.

The query as shown in Fig.5 is of a red flower, which is able to identify similar images undergone illumination change. The image in Fig. 6 is from horse category. The precision obtained is very high in this case. The retrieved images are less dependent to shadows. Fig. 7 shows the result when queried with dinosaur. The precision obtained is very high for such images having distinct objects. The feature is fairly invariant to linear transformations. The further improvement in precision can be seen from Fig. 8 from its ability to retrieve blurred, noisy images(at position 4th and 6th from left) after updating the weights calculated from the FEI values. The results when queried with a flower(yellow) is shown in Fig. 9. The images of this category have objects with some regularity in shape and background. The improvement is also observed in Fig. 10 from retrieving images with its literal color properties.

The results obtained from database (b) 10000 miscellaneous images are shown in Fig. 11 The results of Fig. 11(a) show the retrieved candidates combining the features in a hierarchical fashion. The results after iterative refinement is shown in Fig. 11(b). The result for a query (scene) is shown in Fig. 11(c). The results obtained proves to be satisfactory for retrieving scenes. Since proposed scheme can be compared and evaluated best when the results are tested and evaluated over the same data base. We benchmark our results with the well known standard image retrieval algorithms namely the [4], [21] and color histogram matching [17] using the same data set against the quantitative measure defined as weighted precision in (8). The weighted average of the precision values within k_1 retrieved images are computed as,

$$\bar{p}(i) = 1/(100) \sum_{k_1=1}^{100} n_k/100 \quad (8)$$

$k_1 = 1, \dots, 100$. n_k is the number of matches within the first k_1 retrieved images. The weighted precision as obtained from each category are shown in Table.I

SIMPLCcity method has reported better than color histogram method. It is difficult to obtain satisfactory results for retrieving from all categories using the same set of features. FEI can provide a good measure for improvement in precision. In

TABLE I
COMPARATIVE EVALUATION OF WEIGHTED AVERAGE PRECISION

| class | Our method | SIMPLCcity | Histogram based | FIRM |
|-----------|------------|------------|-----------------|------|
| Africa | .45 | .48 | .30 | .47 |
| Beach | .35 | .32 | .30 | .35 |
| Building | .35 | .35 | .25 | .35 |
| Bus | .60 | .36 | .26 | .60 |
| Dinosaur | .95 | .95 | .90 | .95 |
| Elephant | .60 | .38 | .36 | .25 |
| Flower | .65 | .42 | .40 | .65 |
| Horses | .70 | .72 | .38 | .65 |
| Mountains | .40 | .35 | .25 | .30 |
| Food | .40 | .38 | .20 | .48 |

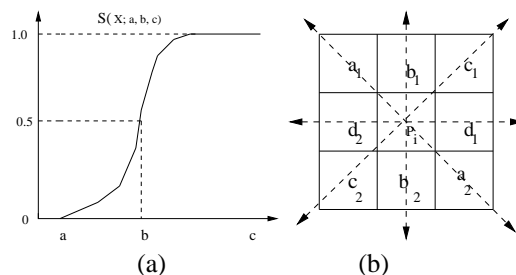


Fig. 1. (a) S type function (b) 3x3 neighborhood of a pixel

most of the cases the average precision for each category can be made better than SIMPLCcity and FIRM via the feature updation scheme. SIMPLCcity and FIRM methods are segmentation based. Better result is obtained with their algorithm when each object is having different textured property and the number of objects matches the number of classes. We have implemented the algorithm on a SUN Blade system with a 700 MHz Processor. The CPU time for computing the features is on the average 10 secs. The computation required for matching and sorting our results is of the order of $O(N \log N)$, N is the number of images in the database and does not depend upon the number of classes.

3 Conclusion

In the current work we have proposed an image retrieval scheme, where the the relative weights of the features can be updated adaptively to specify its importance. Although our CBIR system is not capable enough for handling very complex type of queries, it can identify relevant images which visually differ in some characteristics due to translation, rotation,

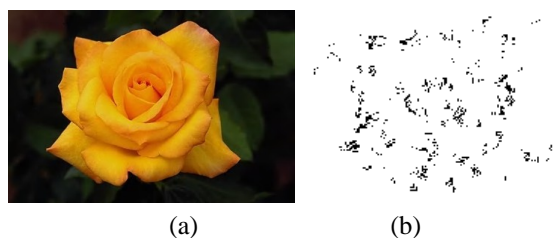


Fig. 2. (a) Flower image (b) Fuzzy corner signature at $\mu_d(P) \geq 0.8$

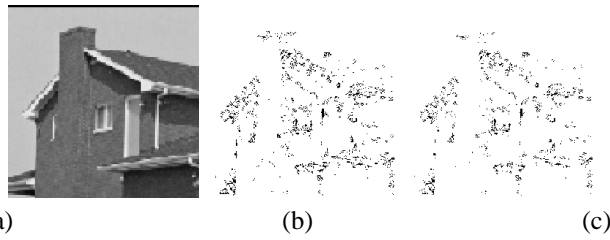


Fig. 3. (a) House image. (b) Fuzzy corner signature at $\mu_d(P) \geq 0.7$ (c) $\mu_d(P) \geq 0.8$

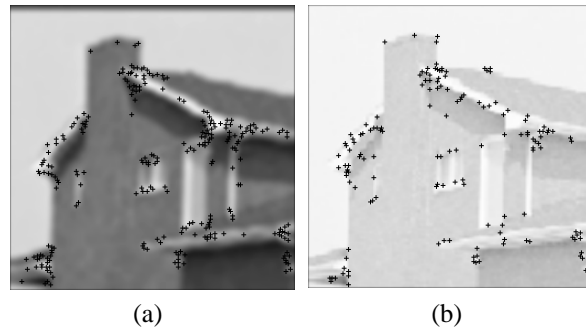


Fig. 4. (a) Representative points (centroid) of the clusters (a) blurred (b) illumination change

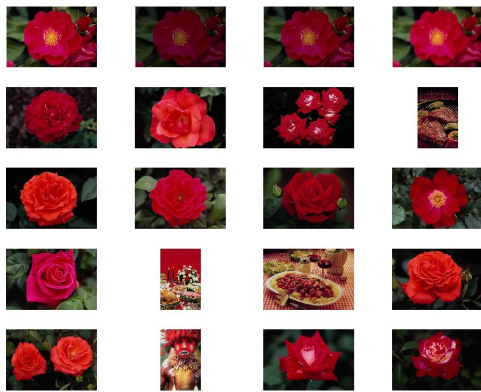


Fig. 5. Retrieved results, with the top left image as the query image. Test for illumination invariance with set (A).

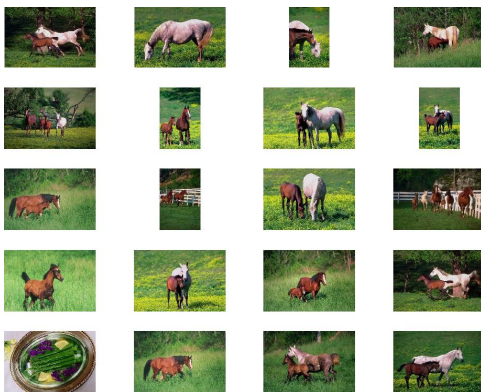


Fig. 6. Retrieved results using set (A), with the top left image as the query image.

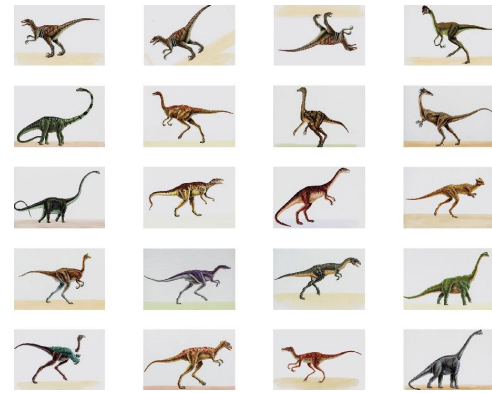


Fig. 7. Retrieved results, with the top left image as the query image. Test for (rot, trans, scale, noise invariance), using set(B)

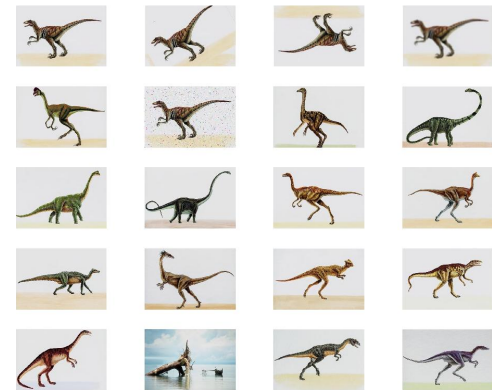


Fig. 8. Retrieved results with the same query of Fig.7. After feature updation with (FEI) the noisy and blurred images are retrieved. The top left image is the query image.

scaling, blurring, illumination change etc. We will try to incorporate some other mechanism like surrounding text, for web images and study the results with MPEG-7 to obtain better applications.

Acknowledgement

Minakshi Banerjee is grateful to the Department of Science and Technology, New Delhi, India, for providing her research fellowship under Women Scientist Scheme, vide, grant (no.SR/WOS- A/ET-111/2003) to carry out her research work.

REFERENCES

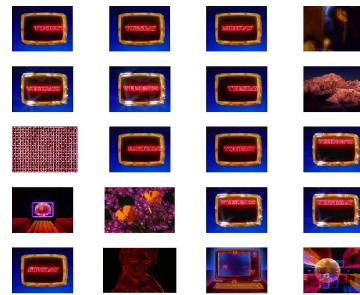
- [1] Minakshi Banerjee and Malay K. Kundu, "Edge based features for content based image retrieval," *Pattern Recognition*, vol. 36(11), pp. 2649–2661, (2003).
- [2] M. Banerjee and M. K. Kundu, "Content based image retrieval with multiresolution salient points," *Fourth Indian Conference Computer Vision, Graphics and Image Processing, ICVGIP 2004, India*, pp. 399–404, (2004).
- [3] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, and J. Malik, "Blobworld a system for region - based image indexing and retrieval," *Proceedings of Visual Information Systems*, pp. 509–516, (1999).



Fig. 9. Retrieved results using set (B), with the top left image as the query image.



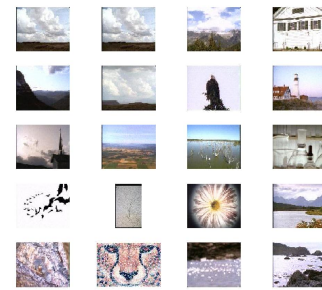
Fig. 10. Retrieved results, After feature updation with (FEI), for the query of Fig. 9.



(a)



(b)



(c)

Fig. 11. Retrieved results, with top left image as the query image. (a) First set of retrieved candidates (b) After feature updation with (FEI) (c) Retrieved results for scene query.

[4] Y. Chen and J. Z.Wang, "A region-based fuzzy feature approach to content-based image retrieval.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(9), pp. 1–16, (2002).

[5] T. Gevers and A. W. M. Smeulders., "Combining color and shape invariant features for image retrieval.," *Image and Vision computing*, vol. 17(7), pp. 475–488, (1999).

[6] T. Gevers and A. W. M. Smeulders, "Color - based object recognition.," *Pattern Recognition*, vol. 32, pp. 453–464, (1999).

[7] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Wiley, New York, (1985).

[8] C. Harris and M. Stephens, "A combined corner and edge detector.," *4th Alvey Vision Conference*, pp. 147–151, (1988).

[9] A. K. Jain and Vailaya, "Image retrieval using color and shape.," *Pattern Recognition*, vol. 29, pp. 1233–1244, (1996).

[10] D. Lowe, "Distinctive image features from scale invariant keypoints.," *International Journal of Computer vision*, vol. 2(60), pp. 91–110, (2004).

[11] Etinne Loupias and Nicu sebe, "Wavelet-based salient points: Applications to image retrieval using color and texture features.," in *Advances in visual Information Systems, Proceedings of the 4th International Conference, VISUAL 2000*, pp. 223–232, (2000).

[12] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors.," *International Journal of Computer vision*, vol. 1(60), pp. 63–86, (2004).

[13] H. Muller, W. Muller, D. Squire, S. M. Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: overview and proposals.," *Pattern Recognition Letters*, vol. 22, pp. 593–601, (2001).

[14] M. S. Nixon and A. S. Aguado, *Feature extraction and Image Processing*, Reed Educational and Professional Publishing Ltd., Oxford, England, (2002).

[15] S. K. Pal and B. Chakraborty, "Intraclass and interclass ambiguities (fuzziness) in feature evaluation.," *Pattern Recognition Letters*, vol. 2, pp. 275–279, (1984).

[16] S. K. Pal and D. Dutta Majumder, *Fuzzy mathematical Approach to Pattern Recognition*, Willey Eastern Limited, New York, (1985).

[17] Y. Rubner, L. J. Guibas, and C. Tomasi, "The earth mover's distance,

multi- dimensional scaling, and color- based image retrieval.," *Proceedings of DARPA Image understanding Workshop*, pp. 661–668, (1997).

[18] A. Rosenfeld, "Fuzzy digital topology.," in *Fuzzy Models For pattern Recognition*, James C. Bezdek and Sankar K. Pal, Eds., pp. 331–339. IEEE press, (1991).

[19] J. R. Smith and S. F. Chang, "VisualSEEK a fully automated content-based image query system.," *Proc. ACM Multimedia*, pp. 87–98, (1996).

[20] D. M. Squire, W. Muller, H. Muller, and T.Pun, "Content based query of image databases: inspirations from text retrieval.," *Pattern Recognition Letters*, vol. 21, pp. 1993–1998, (2000).

[21] J. Z. Wang, J. Li, and G. Wiederhold, "Simplicity: Semantics-sensitive integrated matching for picture libraries.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(9), pp. 947–963, (2001).

[22] X. S. Zhou and T. S. Huang, "Relevance feedback in content based image retrieval: some recent advances.," *Information Sciences*, vol. 148, pp. 129–137, (2002).