# Content Based Image Retrieval with Multiresolution Salient Points

Minakshi Banerjee
Machine Intelligence Unit, ISI
203, B.T. Road, Kolkata 700108, INDIA
minakshi_r@isical.ac.in

Malay. K. Kundu
Machine Intelligence Unit, ISI
203, B.T. Road, Kolkata 700108, INDIA
malay@isical.ac.in

## Abstract

*This paper presents a robust technique for Content Based Image Retrieval (CBIR) using salient points of an image. The salient points are extracted from different levels of the unsegmented image. Local contrast information at different resolution is embedded along with shape information. Fuzzy compactness vector is computed from the signature obtained at different thresholds. The resemblance of two images is defined as the similarity between the computed feature vectors.*

## 1. Introduction

Indexing image with suitable features has become an important research topic recently. Indexing means characterization of images based on one or more image properties. Such features are used extensively in image retrieval. Content Based Image Retrieval(CBIR) is the automated approach for retrieving images from a data base by measuring the similarity of visual contents between the query image and the images stored in the database. Natural images are mainly heterogeneous, where different parts of the image are having different characteristics. Recent research trend in this area is aimed to develop suitable low cost feature extraction mechanism, similarity measures in order to handle real life images. Global features (color, texture, shape) are inadequate to represent the some salient properties of an image. Local features need to be computed to obtain image index based on the local properties. However, such feature computations becomes too much time consuming, if the features are computed from the neighborhood of each pixel within the image. Local properties around visually significant points of an image can generate discriminating image signature for indexing an image. High curvature points mostly termed as corners characterize visually significant portions of an image and plays an important role in shape perception. These points are of interest in CBIR because local features of these regions can be used to index images without segmenting the images. Human beings are able to make decision to locate corners on a complicated shape quite easily. This indicates that there is a high level representation of corners in human perception system. Due to the inherent structural shape information, these points have become important for shape analysis and image retrieval. Classical corners have drawbacks when applied to various natural images for image retrieval, because visually significant features are not always at the corners, and corners may gather in small regions [7]. Also, methods using classical corner detectors lack the ability in handling the uncertainty arising due to the difference in human perception used in grouping the visually significant points. In an image, there may be a small number of points, which may be termed as salient points, though they may not be exactly corners. These salient points have significant variations in the local properties compared to other boundary points. Such points can be evaluated by taking into account its interaction with its neighbors which also varies at different scales of resolution [5]. These points gain importance for indexing, because feature computation is restricted to a limited number of points which appears to be the visually important to an observer. The algorithms used popularly as salient point detectors are mostly corner detectors based on the local properties. Of them, Harris detector[6], SUSAN detector are evaluated as the best benchmarking detector[7].

We have developed an algorithm to extract visually significant points, in which local contrast information at different resolution is embedded along with shape information. The edge points, where there is a change in curvature are the interesting points for capturing discriminating information from an image with limited number of pixels. These points are termed as salient points in our paper. The uncertainty arising in locating these points due to discretization noise and geometric distortion are represented by fuzzy rule base. Our salient point detector is able to extract points where certain variations occur in the local property and the interaction between these points are significant at different scales of resolution. Topological property (fuzzy compactness) [8],computed from the fuzzy salient point sets is used in indexing an image and retrieval performance. Retrieval performance is evaluated in terms of retrieving rotated, translated and scaled versions of the image.

**The proposed approach :** Human subjects determine corners from local properties along a curvature with support regions on both arms. Uncertainty arises in locating such points due to geometric distortion and also perceptual difference. Such imperfect situation can be handled with fuzzy rules. The uniform intensity surfaces of an image are termed as plateaus [9]. The border region between the uniform intensity surfaces of a blurred image [2] which constitute the edge pixels are extracted. These points are characterized with membership ($\mu_d(P)$) using $\pi$ type function, based on the local gray level contrast and represents the gradient strength. Each edge point is assigned gradient direction,at a step of 45 degree. The edge points are labeled as $\{0, 1, -1, inf\}$ depending upon its slopes. The image is thresholded above different membership values $\mu_d(P)$ using $\alpha$cuts to get different fuzzy edge signatures at different resolution planes of an multilevel image. Fuzzy rules, based on gradient directions of the edge pixels, are assigned to the selected candidates for which( $\delta\mu_d(P) > 0$) on the edge map to estimate the support on the arms, and group points from varying curvature. Due to the nonlinearity in the local contrast membership ($\mu_d(P)$),lower resolution edge pixels can be squeezed and rejected and the salient points resulting from the stronger edge pixels can be extracted at finer scale. At poor resolutions ambiguity arises in locating significant points and less important in generating the shape information. Significant portions mainly from the steeper gradient slope can be selected by using proper $\alpha$-cuts [8] to represent boundary pixels containing majority of shape information. Fuzzy compactness value [8] computed from the fuzzy salient point sets constitute the feature vector. This feature is invariant to rotation, translation and scaling by definition.

**Image as fuzzy sets :** An image X of $M \times N$ size and L levels can be considered as an fuzzy subset $A'$ in a space of points X=$\{x\}$. Each point in X can be characterized by a membership function $\mu_A(x_{mn})$. $A' = \{(\mu_A(x_{mn}), x_{mn}\}$ $m = 1, 2, ..M, n = 1, 2, ..N$ where $0 \leq \mu_A(x_{mn}) \leq 1.0$.
*fuzzy alpha cut :* $s_\alpha$ comprises all elements of X whose degree of membership in s is greater or equal to $\alpha$ where

$$s_\alpha = \{x \in X : \mu(x) \geq \alpha\} \quad (1)$$

where $0 \leq \alpha \leq 1.0$
*Fuzzy Compactness :*For a $M \times N$, $\mu_{mn}$ array the fuzzy compactness comp($\mu$) is defined as [8].

$$comp(\mu) = \frac{a(\mu)}{p^2\mu} \quad (2)$$

where $a(\mu)$ and $p(\mu)$ are the fuzzy area and perimeter of $(\mu)$.

## 2. Feature extraction

*Saliency estimation from local property :*
let X comprises of all elements of an image. The image consists of many uniform intensity surfaces termed as Plateaus [9]. If $S_{pti}$ represents Plateau Top and $S_{pbi}$ a Plateau Bottom, where (i=1.....n) over the entire image. The border region $S_{bi}$ between two uniform intensity plateaus ( $S_{pti}$ and $S_{pbi}$) represents the points which are eight neighbors of at least one element of $S_{pti}$,$S_{pbi}$ and constitute the edge subset. The edge subset $E_d$ can be represented as in (3).

$$E_d = \sum_{i=1}^{n} S_{bi} \quad (3)$$

and $E_d \subset X$ Each candidate P of $E_d$ is assigned membership value $\mu_d(P)$ representing the edge contrast magnitude [2].The direction $\phi(x)$ of an edge pixel is labeled shown in eqn.(5). The significant edge subset $E_{df}$ are those points for which $\mu_d(P) > 0$ as shown in (4).

$$E_{df} = \{P \in E_d : \mu_d(P) > 0\} \quad (4)$$

the corresponding directions are represented by (5)

$$A_{df} = \{P \in E_d : 0, 1, inf, -1\} \quad (5)$$

By thresholding the edge image above different membership value $\mu_d(P)$ using proper ($\alpha - cuts$ ) [8] where $0.0 \leq \alpha \leq 1.0$ we, are able to segregate the edge pixels of different resolution. Edge signature $E\_d\alpha$ represented by (6) is consisting of edge pixels above different resolution plane. Such thresholding is important to segregate edge pixels (strong, medium, weak) based on their contrast membership value. Edge points much less than $\mu_d(P) < 0.5$ are close to each other on the membership curve. Ambiguous points are generated in this region as there is no well defined curvature as seen in the bottom rectangle of Fig. 3(a). The points are widely separated where the membership well above the cross over points ($\mu_d(P) \geq 0.5$). A minimum level of threshold is chosen below which the variations are considered to be noise and no salient point designations are made. Only those edge pixels have corner like properties, for which the slope of $\mu_d(P) > 0$ i.e. ($\delta\mu_d(P) > 0.0$)in its local neighborhood. The value creates significant variations in the local properties in this region compared to other boundary points shown in Fig.1(b). A multi level plot of the local contrast measure $\mu_d(P)$, is shown in Fig. 1(b) where the values $0 \leq \mu_d(P) \leq 1.0$ is scaled with 0 to 255 gray levels.

$$E\_d\alpha = \{P \in E_{df} : \mu_d(P) \geq \alpha\} \quad (6)$$

where $0 \leq \alpha \leq 1.0$

$$E_c = \{P \in E\_d\alpha : \delta\mu_d(P) > 0.0\} \tag{7}$$

*Fuzzy corner ness rules :* The points $E_c$ represented by eqn.7 are points where there is a change in curvature and have corner-like properties. However dominant curvature points can be separated from spurious curvature points by estimating the arm support on both the sides. The region of support can be estimated from the fact that each boundary point have its own view of the point [10]. A dominant curvature point should have a view which constitutes a meaningful region of support and separates the view from non dominant points.

$E_c$ is corrupted with spurious corners ($E_{cn}$) which are mainly corrupted due to digitization error.

$$E_c = E_{ct} + E_{cn} \tag{8}$$

The directions of the edge pixels of the gradient map is plotted on Fig. 2.The labeled directions of eqn.5 is plotted with different gray values. Fuzzy rules are assigned to give a measure of arm support on the forward arm and the back arm of the curvature points of $E_c$.

If the direction of the test pixel is $\phi(x)$ then we label $\phi(x)$ + $\pi/4$ as relative forward and $\phi(x)$-$\pi/4$ as relative backward direction. That is if the labeled direction of the concerned pixel is (0 )then (+1) represents the relative forward direction and (-1) represents the relative backward direction. The change in direction can be observed from the different gray labels from the locality where there is a change in the curvature of Fig.2(a).

Two membership values ($\mu_f$, $\mu_b$) are assigned to estimate the connectivity of each points of $E_c$ with other edge points on the curvature in the local neighborhood.$\mu_f$ shown in eqn.(9) estimates the connectivity of the point with other same labeled edge points in the relative forward direction. Similarly $\mu_b$ shown in eqn.(10)is assigned to estimate the the connectivity of the points in the relative forward direction [2].

$$\mu_f = K * \exp(-x) \tag{9}$$

where $x = \frac{1}{f_c}$, $f_c$is the total nos of edge pixels possessing angle = $\phi(x)$ + $\pi/4$ i.e. on the relative forward direction in a fixed window in the local neighborhood of the concerned pixel.

Similarly $\mu_b$ is defined by

$$\mu_b = K * \exp(-x) \tag{10}$$

where $x = \frac{1}{b_c}$ ,$b_c$ is the total nos of pixels possessing angle = $\phi(x)$ - $\pi/4$ relative backward in a fixed window.

The value of K is estimated globally for an image by determining the maximum count of edge pixels possessing same direction in a fixed window in the local neighborhood of the concerned pixel.For this value of K the value of $\mu_f$ or $\mu_b$ should be maximum.

Each pixel is labeled (a) ('+', if $\mu_f > 0.0$ and $\mu_b \simeq 0.0$ ) (b) ('o', if $\mu_b > 0.0$ and $\mu_f \simeq 0.0$ ) ,(c)('$\diamond$', if $\mu_f > 0.0$ and $\mu_b > 0.0$ ) shown in Fig.2(b).

The typicality of these two features can be observed form Fig.2(b). It is seen from Fig.2(b) that such characteristics points ('+','o','$\diamond$') are found on the edge map where the is a change in curvature because the local features $\mu_f$ and $\mu_b$ varies significantly along these portions on the curve.

The saliency of the curvature region for specifying the arm support are assigned using the following fuzzy rules considering the relative changes in labels of neighboring pixels.

(i) If mark of change from a smooth part to a curve part is specified with points having ($\mu_f > 0.0$ and $\mu_b \simeq 0.0$). Such points ('+') estimates the connectivity support on the forward arm of a corner like point shown in Fig.2(b).

(ii) Then mark of change from curve part to a smooth part should have points ($\mu_b > 0.0$ and $\mu_f \simeq 0.0$). Such points ('o') estimates the connectivity support on the backward arm of a corner like point shown in Fig. 2(b).

(iii) Sharp turn of a curve with smooth parts on either sides is expected to have points with both $\mu_f > 0.0$ and $\mu_b > 0.0$ i.e. such points ('$\diamond$') have support on both the arms of a corner like point shown in Fig.2(b).

(iv) For straight line edge points, or change in curvature arising discretization effect is expected to have both $\mu_f \simeq 0.0$ and $\mu_b \simeq 0.0$ even if $\delta\mu_d(P) > 0.0$ in its local neighborhood.

The points (+,o,$\diamond$) signifies some change in curvature in their location shown in Fig.2(b).

Rules are summarized in Table.1.

*corner clusters :* Having obtained the two fuzzy sets $\mu_f$ and $\mu_b$ in the locality of each points of $E_c$,we can obtain a third fuzzy set $\mu$ which is separating $\mu_f$, from $\mu_b$.The subset $\mu$ which representing the cluster of salient points shown in Figs.4 are the points which have other points with ('+', $\mu_f > 0.0$ and $\mu_b \simeq 0.0$ ),('o', $\mu_b > 0.0$ and $\mu_f \simeq 0.0$ ) in it's local neighborhood. Having obtained an edge signature by thresholding at a value ($\mu_d(P) \geq \alpha$), by further thresholding on the values of $\mu_f$, $\mu_b$we are able to group the points from different curvature. The difference in the peak value of $\mu_f$ and $\mu_b$, th =abs ($max(\mu_f) - max(\mu_b)$) in the locality is used as threshold shown in Figs.8,9. The most representative point of each corner cluster is the resulting point $C_{ij}$ with coordinate equal to the average value of the co-ordinates of the cluster members.

$C_{ij}$ = $[\sum x_j/n, \sum y_j/n]$ (n number of points within individual corner cluster).

**Algorithm**

Step1 : Gaussian filtering is performed on the image.

Step2 : Edge candidates are identified and the image is converted to a fuzzy gradient map.
Step3 : Edge directions are labeled at a step of 45 degree.
Step4 : The edge candidates with $\delta\mu_d(P) > 0.0$ on the gradient slope are selected for different threshold value of $\mu_d(P)$.
Step5 : Each candidate is assigned membership $\mu_f$,$\mu_b$ based the relative change of direction of other edge pixels in $5 \times 5$ neighborhood of the concerned pixels. Step6 : The regions $\mu$ separating $\mu_b$ and $\mu_f$ constituting the salient points are extracted.
Step7 : The salient points are grouped from curvature of different sharpness by thresholding on the value of $\mathtt{th}$
Step8 :Most representative point is computed from each cluster.
Step7 : Topological features (fuzzy compactness vector with components $f_1$, $f_2$,.. $f_n$,) are computed from the salient point signature obtained at each threshold.
Step9 : Euclidean distance is computed between the feature vector elements of the query image and images in the database and ranked according to distance.

**Application as CBIR feature :** Feature extraction based on fuzzy reasoning can accommodate certain uncertainties and incompleteness in the data for (CBIR) [4], [1]. Local features are computed only on a limited number of pixels of $E_c$ without segmenting the images. It does not involve computation of complex geometrical operations along the curvature. The number of operations involved is less than the order $(MN)^2$ where the image is M by M and the neighborhood of gradient operator is $N^2$.

**feature vector, performance criteria, similarity metric :** The feature vector F = $[f_1, f_2, ..., f_n]$ is represented by the following feature components as $f_1$=fuzzy compactness value, $f_2$=fuzzy compactness value, signature thresholded at $\mu_d(P) \geq 0.6 and 0.9$ respectively. The variation of selection of thresholding level will not change the result drastically. The performance of retrieval are evaluated from the standard retrieval benchmarks **Recall rate** (R) is given by $\frac{n_1}{n_2} \times 100\%$ and **Precision rate** (P) is given by $\frac{n_1}{20} \times 100$ % [3]. Where $n_1$ is the number of images retrieved in top 20 positions that are close to the query. $n_2$ represent the number of images in the data base similar to the query. Euclidean distance metric is used to compute the dissimilarity value between two feature vectors.

## 3. Experimental Results

Experimental results from Figs.1 to Figs.10 have shown the effectiveness and robustness of the proposed algorithm. Fig.1(b), represents the multilevel gradient plot. Fig.2(a),represents the multilevel plot of the coresspond-

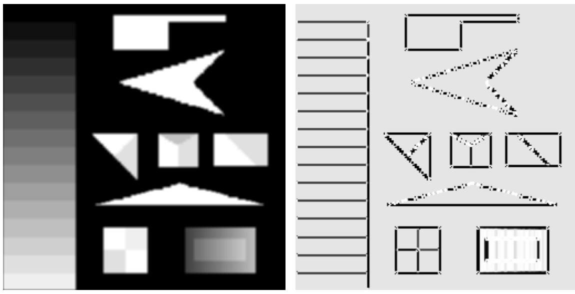
**Table 1. Fuzzy corner ness measure**

| $\mu_f$ | $\mu_b$ | cornerness | Straightness |
|---|---|---|---|
| high | low | medium | low |
| high | high | high | low |
| low | high | medium | low |
| low | low | low | high |

ing angles.The variation in the local property along the curvature can be observed from Figs.1(b),2(a) The gradient map of the image thresholded above different membership values using $\alpha$ -cuts are shown in Fig. 3(a),(b),(c) for ($\alpha \geq 0.0$), ($\alpha \geq 0.6$) ($\alpha \geq 0.9$). The points above the threshold values are plotted with dark edge pixels. The right dark region shown in Fig.3(a) represents poor resolution graded region, where lot of candidate edge pixels with lower membership values are selected. For higher thresholds significant edge points with higher membership values above threshold are selected as shown in Fig.3(c). The salient points(*)are represented on Fig.4(a),(b) points plotted on the edge map of Fig.3(a). The salient point signature at different thresholds are separately shown in Fig.5.The most representative point from the salient point cluster in shown in Fig.6. The distinguished points labeled as ('+', 'o', $\diamond$) are shown in In Fig. 2(b). A close observation in Fig.2(b) shows that the counts of (+),(o) along the curvature over the estimated window gives a measure of the arm support on both sides of the curvature points. This typicality in the local property is used in separating the curvature points based on the arm support. The salient points (*) are shown in Fig.8,9. Results obtained from the benchmarking corner detector of Harris and SUSAN is shown in Fig.7. By a comparative observation it is seen that the points extracted by our algorithm and may not represent the accurate location of corners but the salient point set shown in Fig.5 obtained at different threshold can serve suitable index for image retrieval and is expected to generate better result than point set matching. Retrieval performance is tested with SIMPLICITY database consisting of 1000 images of 10 distinct classes of images. Retrieval result of Fig. 10(c),(d) with the query image on the top left corner is evaluated in terms of precision and recall is shown in table 2. It is evident from the query result that, the feature capable of retrieving images invariant to rotation, translation and scaling with expected precision.

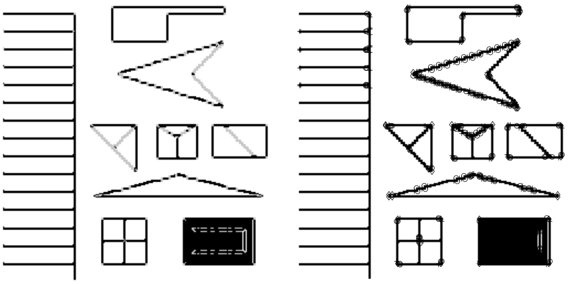
## 4. conclusion

The experimental results shows that the salient point features serve a good index for retrieving images from a heterogeneous database whose classes are not known apriori.The fact we have used a global measure (fuzzy compact-



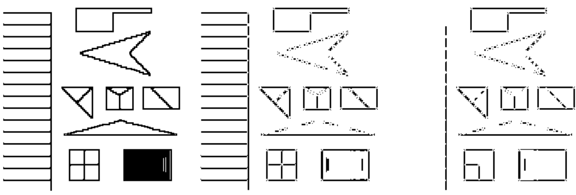
(a) (b)

**Figure 1. (a)Original image (b) Multilevel gradient map**



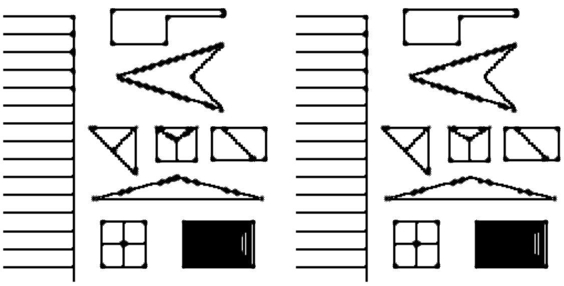
(a) (b)

**Figure 2. (a)Multilevel angle map (b) labeled points '+', 'o', '◇,' points plotted on the edge signature thresholded at** $\mu_d(P) > 0.0$



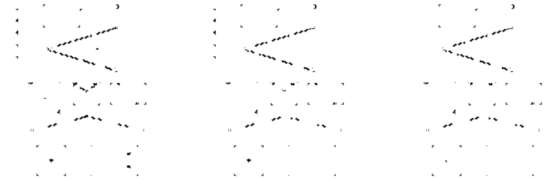
(a) (b) (c)

**Figure 3. (a)Edge signature for (**$\mu_d(P) > 0.0$**) (b) (**$\mu_d(P) \geq 0.6$**) (c) (**$\mu_d(p) \geq 0.9$**) points above threshold plotted as crisp edge points**



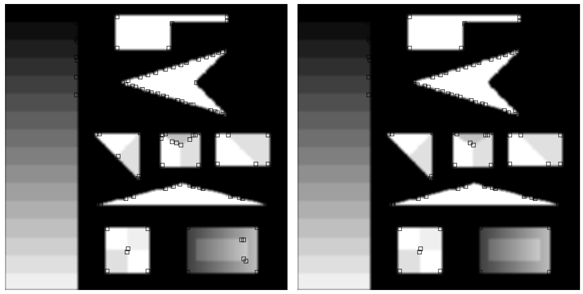
(a) (b)

**Figure 4. salient points ('*') (a)** $\mu_d(P) > 0.0$**(b)** $\mu_d(P) \geq 0.6$ **points plotted on the edge signature thresholded at** $\mu_d(P) > 0.0$



(a) (b)

**Figure 5. salient points (a)** $\mu_d(P) > 0.0$**(b)** $\mu_d(P) \geq 0.6$ **(c)** $\mu_d(P) \geq 0.9$

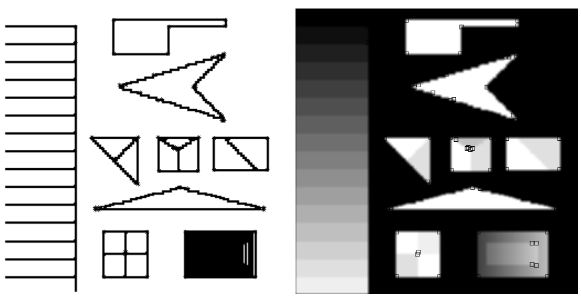


(a) (b)

**Figure 6. (a)Most Representative point from the cluster at(a)**$\mu_d(P) > 0.0$ **(b)**$\mu_d(P) \geq 0.6$


(a) (b)

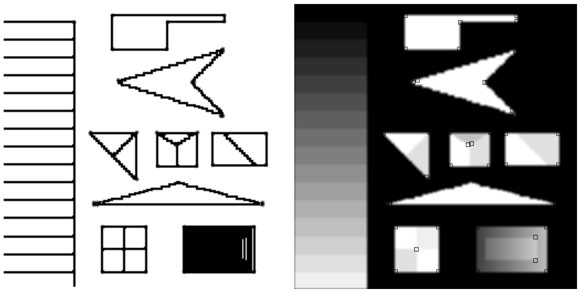
**Figure 7. (a) Points from SUSAN detector (b) Harris detector**



(a) (b)

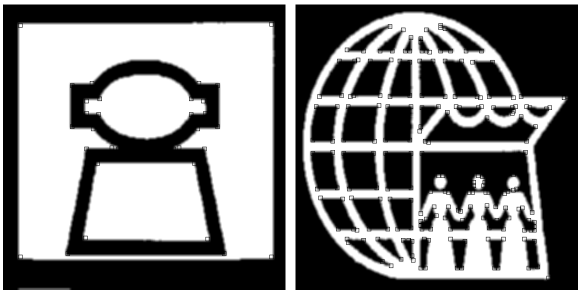
**Figure 8. salient points (*), (a) from threshold th=.2, plot on the edge signature thresholded at** $\mu_d(P) > 0.0$ **(b) Most representative points**

ness) for computing similarity,generates best results when queried with single objects and having non textured background.The method could be further improved if we use some other features like texture in association with the present work.

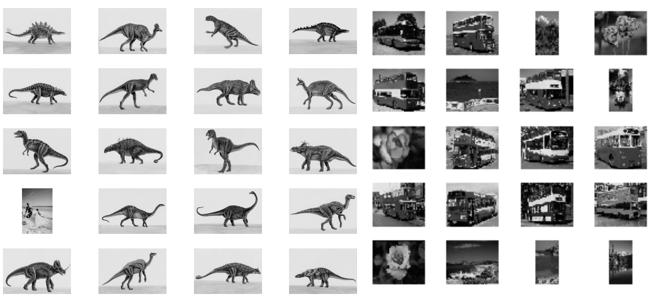


(a) (b)

**Figure 9. salient points(*), (a)from threshold th=.3,plot on the edge signature thresholded at $\mu_d(P) > 0.0$ (b) Most representative points**



(a) (b)



(c) (d)

**Figure 10. (a),(b) Most representative point plot, signature obtained at $\mu_d(P) > 0.0$ (c),(d) Retrieved result, with top left image as the query image.**

**Table 2. Performance comparison %**

| Indexing property | Edge based [2] | Salient point |
|---|---|---|
| Precision(Fig.10(c) ) | 70 | 95 |
| Recall | 14 | 19 |
| Precision (Fig.10(d)) | 55 | 60 |
| Recall | 11 | 12 |

## References


[1] M. Banerjee and M. K. Kundu. Content based image retrieval with fuzzy geometrical features. *IEEE International conference in fuzzy systems*, pages 932–937, 2003.

[2] M. Banerjee and M. K. Kundu. Edge based features for content based image retrieval. *Pattern Recognition*, 36(11):2649–2661, 2003.

[3] A. D. Bimbo. *Visual Information Retrieval*. Morgan Kaufmann Publishers,Inc., San francisco, USA, 2001.

[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*, 24(9), 2002.

[5] M. Fischler and H. C.Wolf. Locating perceptually salient points on planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):113–129, 1994.

[6] C. Harris and M. Stephens. A combined corner and edge detector. *4 th Alvey Vision Conference*, pages 147–151, 1988.

[7] E. Loupias and N. sebe. Wavelet-based salient points: Applications to image retrieval using color and texture features,Tech report,Leiden institute of advanced computer science, Netherlands,http://www.nicu@wileidenuniv.nl.

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

[9] A. Rosenfeld. Fuzzy digital topology. In J. C. Bezdek and S. K. Pal, editors, *Fuzzy Models For pattern Recognition*, pages 331–339. IEEE press, 1991.

[10] C. Teh and R.T.Chin. On the detection of dominant points on digital curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):859–872, 1989.