A Novel Method For Image Retrieval Based On Visually Significant Feature Point Maps

Mangipudi Partha Sarathi
*Department of Electronics & Communications Engineering, Amity School of Engineering & Technology,
Amity University,Sector-125, Noida, UP-201301, India
Email:psmangipudi@amity.edu*

M. A. Ansari
*Islamic University, Madinah, KSA
(on EOL from Gautam Buddha University, NCR, UP-201308, India)
Email: ma.ansari@ieee.org*

Abstract

In this paper, we present a new algorithm for visually significant feature points detection for efficient image indexing and retrieval. We propose a hybrid feature point detector based on edge maps obtained via spatial and wavelet methods. Proposed method is compared with other existing interest point detectors. Robustness of feature points to geometric transformations like rotation and scaling is analyzed. Hausdorff distance measure is used for indexing and ranking of results.

Keywords: Wavelets, Edge maps, feature points, Hausdorff distance, Retrieval

1. Introduction

The numerous applications of Image Retrieval based on image content representation has drawn much attention in the past decade. Representation of image content in a low dimensional feature space is still a challenging task. In a recent paper [1] , Datta et al. presented an excellent survey of state-of-art approaches and techniques for Image Retrieval. Various CBIR techniques exist in literature today. It is of great interest to note different kinds of approach to the same class of image indexing and retrieval problem. In Traditional CBIR approach, image information represented by a feature vector, is a point in a n-dimensional feature space. Feature vector usually consist of low level visual features like color , gray level, texture etc. It is important to select the weighted combination of such independent visual features to suitably represent an image. The goal is to try and bridge the semantic gap that exists between the image content and its representation.

In recent years, vast number of speedy algorithms have been devised for image retrieval based on interest point detectors. The basic idea is to represent large image information by using only a few points. The interest points thus represented only the significant information needed for retrieval. This will enable fast and speedy retrieval. Newer task is then the trade-off between speed and accuracy. Corners are among the first low-level features used for image analysis [2]. Harris and Stephens [3], improved the algorithm well known as the Harris Corner Detector. They detect the corners by evaluating a corner score from the second-order moment image gradient matrix. Consequently, Harris detector was made robust to scale and affine transformation [4] [5] [6]. Instead of trying to detect corners, one may use local extrema of the responses of certain filters as interest

points. In particular, many approaches aim at approximating the Laplacian of a Gaussian, which, given an appropriate normalization, was shown to be scale invariant if applied at multiple image scales [7]. Lowe in [8], [9], proposed to select the local extrema of an image filtered with differences of Gaussians, which are separable and hence faster to compute than the Laplacian. The Fast Hessian detector [10] is based on efficient-to-compute approximations to the Hessian matrix at different scales. Compactedly supported wavelets have been introduced by Daubechie in [11].Wavelets have been used for corner detection in [12]. Lowe has discussed in detail in his paper [13], scale selection of key points and scale invariance. In [14], Schmid and Mohr did an in-depth evaluation of various interest point detectors including intensity and contour based interest point detectors.

Interest points found applications in areas like image retrieval and visual tracking. Recent approaches with regards to tracking could be found in [33],[34]. In [15],[16], interest point along with texture features have been used for image retrieval. Image retrieval based on salient points have been dealt in detail by Loupias and Sebe [17],[18]. It was seen that salient points detected using multi-resolution wavelet transform performed better than classic corner detectors. Since, maximum values of wavelet coefficients gives an indication of the presence of singular points in an image, they represent the edge points of the object of interest in the image. In [19], Itti and Koch gave an important concept to form a saliency map based on combination of features that determine points that draw visual attention. Similar key point signatures based retrieval was discussed in [32]. Kadir and Brady [20] used the local entropy of regions in an image as a feature to identify salient regions. Local features from the salient point neighborhood were used for image retrieval [21], [22]. Song [22] in his paper provides a technique to use local feature like color to extract salient points using wavelet transform and Barnard detector. Liu et al [25] proposed a salient point detector based spectrum energy variation and applied it to medical images. Recently in [26],another approach namely contour based salient point detector was discussed. Reduction of salient points identified was discussed in [27]. In [23]. linear scales for wavelet base salient point maps were proposed . Multi-scale patches were formed and used for image retrieval in a method proposed in [28]. In some papers like in [29], [30], various distance metrics like KL divergence were employed for image retrieval. In [31], paper discusses an

on the fly technique to select the number of wavelet coefficients based on their distribution for image retrieval.

Our paper primarily tries to address two pertinent problems. First, the task is to find an optimum number of feature points which are just enough for effient image retrieval.The number of interest points to represent an image varies according to the nature of an image. Fewer number of points might degrade the efficiency of Image retrieval system. Conversely, if more number of interest points are used, the redundancy of points will contribute to additional time taken for image retrieval. Second task is to eliminate the false edge points.In this paper, a hybrid algorithm that extracts only the visually most significant interest points is proposed. The proposed algorithm consists of an multi-scale edge point detector and a multi-resolution wavelet transform based feature point representation. The idea is to extract only those points which are common to both these detectors and use them for retrieval purpose.

The organization of the paper is as follows . Section-II gives a detailed system overview of the proposed methodology. Steps in generating visually significant feature point maps (VSFPM) and illustrating examples with results are discussed in Section-III. Performance of the proposed algorithm against existing techniques have been discussed with comparative results in Section-IV. Also, robustness of the feature points against geometrical transformation like rotation and scaling were discussed.Image retrieval with results, distance metric employed, databases used and performance evaluation is done in Section V. Section-VI lists the conclusions and inferences of the proposed methodology.

2. System Overview

Figure 1 shows the broad framework for the Image indexing and retrieval system used in our paper.
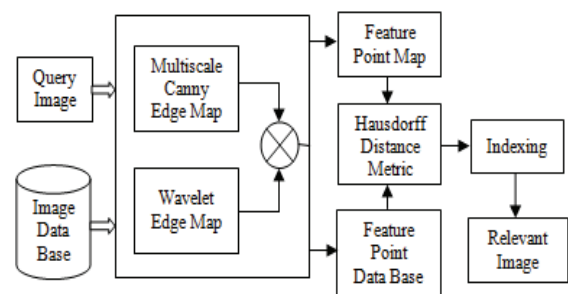

Figure 1: VSFPM based Image Retrieval System

3. Feature Point Representation

Image pre-processing is the initial step before representing it by a feature point map. Usually noisy images are treated to image smoothing prior to further processing. One should bear in mind that pre-processing step also depends on the type of application. In some cases, a sub-image or a region of interest may well be identified based on low-level features such as color, gray level or texture.

3.1 Feature Point Representation (Wavelet based method)

The wavelet transform is a multi-resolution representation of a signal that expresses variation of the signal at different scales. Daubechies proposed wavelets, that are also orthogonal and compactly supported [11]. Given an image signal, A 2-D discrete wavelet transform is defined as:

$$\varphi_{jk}(x) = 2^{-j/2} \varphi (2^{-j}x-k), \qquad (1)$$

$$\psi_{jk}(x) = 2^{-j/2} \psi (2^{-j}x-k) \qquad (2)$$

where $\varphi(x)$ and $\psi(x)$ are the scaling function and wavelet function respectively , and $\{ \varphi_{jk}(x)\} \{\psi_{jk}(x)\}$ are the two orthogonal function basis sets.

In this paper, we want to represent the high frequency information of the given image at finer scales.The dyadic wavelet transform uses the dyadic scale 2^j, where $j \in Z$ is the scaling. The wavelet transform gives the information about the variations in the signal at different scales. The idea is that the local absolute maximum values of the wavelet coefficients at a coarse resolution corresponds to a region with high global variation of a signal. So significant points can be detected by searching for coefficients at finer resolutions.

If we want to locate the edges of a 2D signal, we should consider implementing the approximation of the canny edge algorithm using wavelets. The image signal is convolved with a dilated wavelet to implement the wavelet transform.

The functions $\psi_s^1(x,y)$ and $\psi_s^2(x,y)$ are the wavelet functions in our paper defined as

$$\psi_s^1(x,y) = \frac{\partial \theta_s(x,y)}{\partial x} \quad \text{and} \quad \psi_s^2(x,y) = \frac{\partial \theta_s(x,y)}{\partial y}$$

where s = 2^j, j€ Z and $\boldsymbol{\theta}_s$ is the 2D Gaussian smoothing function at different scales.

The wavelet transform of a 2-D image thus obtained has three detail images. We denote these detail images as cH (containing horizontal information in high frequency), cV (containing vertical information in high frequency), and cD (containing diagonal information in high frequency). The gradient information is obtained by taking the sum of these detail images. The local maxima of the gradient magnitude at various scales is then given by taking the modulus of the wavelet transform [35] at a particular scale **s,**

$$M_s(x,y) = ||s\nabla(f * \psi_s)(x,y)||, \qquad (3)$$

After obtaining the local modulus maxima information, we compute the edge image at scale s by thresholding the magnitude function

$$W_s(x,y) = \begin{cases} 1, & |M_s(x,y)| \geq \tau \\ 0, & |M_s(x,y)| < \tau \end{cases} \qquad (4)$$

where the threshold has been selected as

$$\tau = avg(|M_s(x,y)|) \\ + \alpha(s)[\max(|M_s(x,y)| - avg(|M_s(x,y)|]] \qquad (5)$$
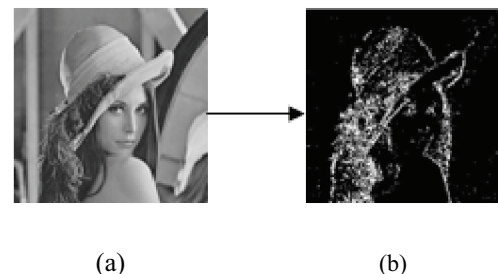

(a)   (b)

Figure 2: (a) Original Image   (b) Wavelet based Feature points

3.2 Multiscale Edge Map

The multiscale edge map is formed by finding the edges using the canny edge algorithm [36] and sub-sampling the resultant image. Let the edge map be $E_s(x,y)$.
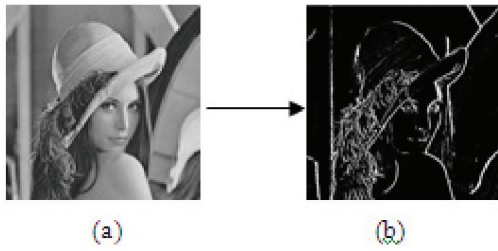

Figure 3: (a) Original Image (264x 400)   (b) Edge map at higher scale (66 x 100)

3.3 Selection of Significant Feature Points

It can be seen from figure 4, that the images have been represented by feature point maps obtained using wavelet points only and using our fusion map. The redundancy observed in wavelet edge maps have been greatly reduced and the points in the fusion map seem to be nicely distributed. The selection of significant feature points is based on the logic that the feature points common to both these detectors must be stable and robust as they have survived operations like sub-sampling over multiple scales. It is therefore sensible to use both these edge maps for image indexing. This technique has various advantages. Firstly the redundant feature points to represent the image in wavelet domain is greatly reduced. Second, the false edge points that might have crept in due to sub-sampling of spatial edge map is removed. And lastly, optimum number of visually significant feature points have been extracted. The spatial location of these feature points are later used for faster image querying and indexing. The feature point map thus generated is given by

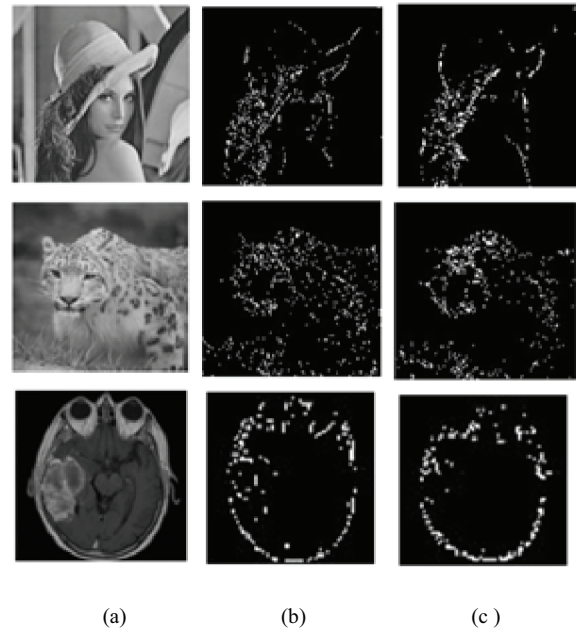$$G_s(x,y) = W_s(x,y) . E_s (x,y) \qquad (6)$$

Figure 4: Examples of feature point maps (a) Original images (b) Proposed feature point maps (c) Wavelet representation

Figure 5 illustrates the generation of the feature point map with visually significant feature points.
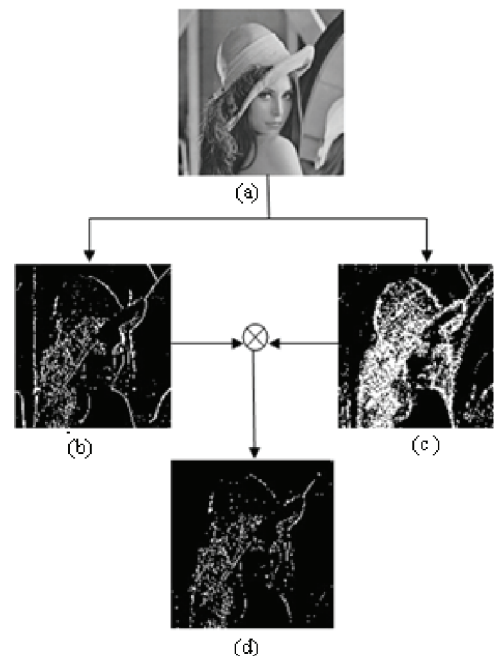

Figure 5: Selection of visually significant Interest points (a)Original Image (b) Canny edge map (c) Wavelet edge map (d) Visually significant feature point map (VSFMP)

4. Experimental Results

4.1 Performance of Proposed Technique against existing Algorithms

We have taken the Fox image (Figure 6a) considered in the paper by Loupias and Sebe [1],[2]. Figure 6c represents the interest points extracted in the wavelet domain. Two-level wavelet decomposition is applied over the image. Resultant edge map by our technique is shown in figure 6b.
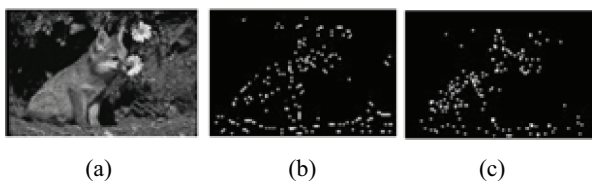

(a)   (b)   (c)

Figure 6:(a) Original image (b) Proposed method (171 salient points) (c) Loupias method [17] (171 feature points)

It can be observed that although Loupias method of salient point extraction has an advantage of flexibility in selecting the number of salient points, it is not possible to know the optimum number of points for image indexation. One can also see that the fur of the fox in the image is also identified by points apart from the body contour. The distribution of feature points is better compared to the Loupias technique for the same number of feature points.

Our next comparision is on medical images like x-rays. Figure 7 shows that our technique gives appreciable results. Ony 45 points have been used to represent the x-ray image. method identifies salient regions based on local entropy. While Itti and Koch's method is generating a saliency map based on neurobiological concept of visual attention. Liu's paper on medical images have been used for comparing our proposed algorithm.
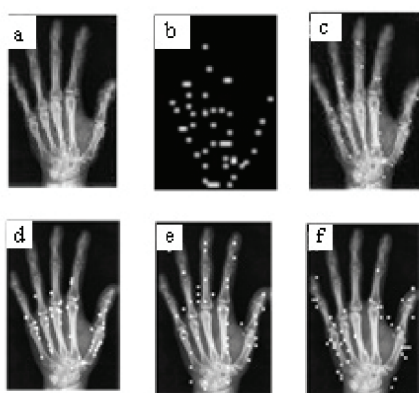

Figure 7: (a) X-ray Image (b) Proposed method (c) Superimposed image (d) Kadir's method[20] (e) Itti's method[19] (f) Liu's method [25]

As a third example, we have compared our proposed algorithm with two other methods. Song's method exploited local features like color from identified wavelet points while in Tsai's method, the image is divided into subimages and saliency map is formed for subimages based on the wavelet coefficients. We have taken 110 salient points for this example. As can be seen in figure 8b, the points are distributed. There is no grouping of points as in figures 8e and 8f.
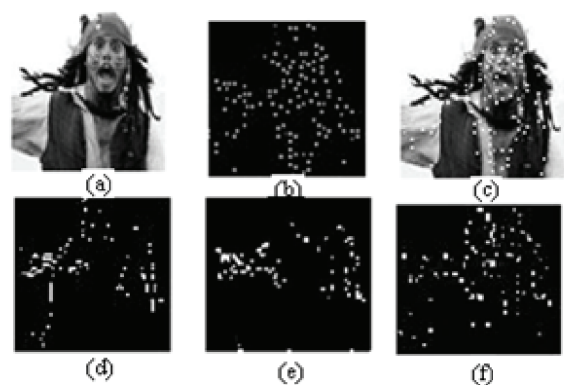

Figure 8: (a) Original Image (b)Proposed method (c) Superimposed image (d) Loupias method[17] (e)Song's method[22] (f) Tsai's method [27]

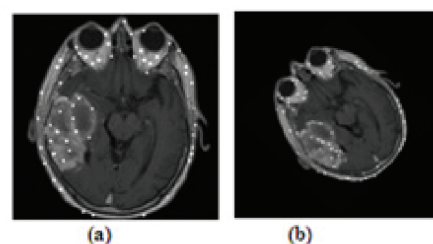### 4.2 Robustness to Geometric Transformations

Figure 9. (a) Feature point image (b) Rotation by 45º

The feature points have been shown to be robust to geometrical transformations like rotation (figure 9). We have extracted the feature points from the image under different angles of rotations. 12 Steps of 30º have been taken to repeat the experiment and the graph has been plotted as seen in figure 10. The radius of error is very small for any particular interest point. So, the feature points can be a good representation of the local structures of the brain.
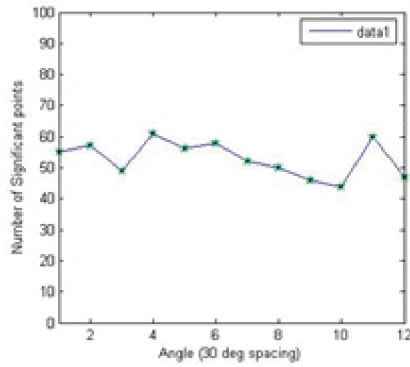
Fig.10. Repeatability rate of feature points (rotation)

The feature points have been extracted for different Gaussian blurs as shown in figure 11. We have taken scales at multiples of $\sqrt{2}$. Repeatability rates shown in figure 12 indicates that the feature points are robust to scales upto σ $=2\sqrt{2}$.

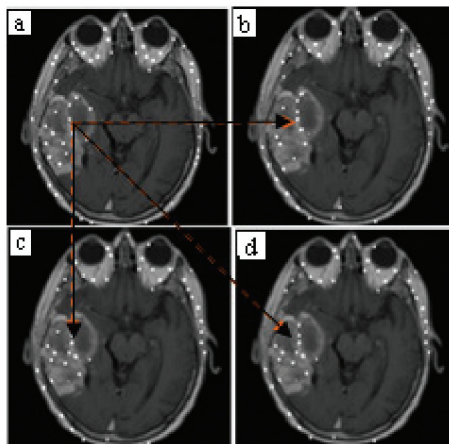Figure 11 (a) Feature Point Image(52 points/) (b) Gaussian blur (σ=1)(48 points) (c) Gaussian blur (σ =1.707)(46 points) (d) Gaussian blur (σ =2)(49 points)
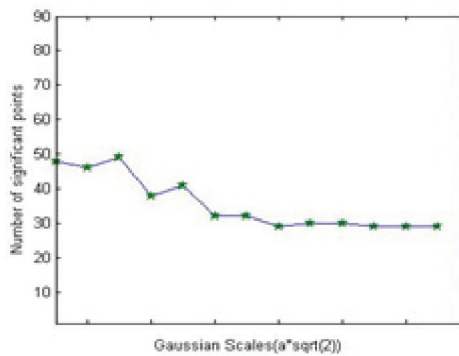

Figure 12. Repeatability rate (Gaussian blurs)

5. Image Retrieval

5.1 Distance Measure

The Hausdorff distance is a nonlinear operator which measures the mismatch of the two sets. Therefore, it can also be used as a measure for similarity between two images if images are represented by feature point sets. In our paper, we measure the distance of one point of the query image that is nearest to any point of the database image and vice versa. Spatial location of feature points is used to calculate the distance. Then the Hausdorff distance stored will be the maximum of these distances calculated for all points in the query image. Given two finite point sets A = { a_1, a_2,……a_m} and B = { b_1, b_2, …..b_n } , the Forward Hausdorff distance , H_D between them is defined as

$$H_D \ = \max_{a_i \in A}(D_{a_i}B) \qquad (7)$$

where $D_{a_i}B$ is the distance metric from point a_i to set B expressed as

$$D_{a_i}B = \ \min_{b_j \in B} D_{a_i b_j} \qquad (8)$$

Therefore, the most similar images will correspond to minimum of the Hausdorff distances calculated and stored for the images in the database.

5.2 Retrieval Results

We conducted our experiments on broadly three different datasets. We used medical images like 2D MRI scans of brain lesions and x-rays. We have used outdoor scenes like lakes, roads and houses and clubbed them as category 'Natural Scenery'. We have also tested our algorithm on a third category from a known image database like the COREL which has been widely used for CBIR evaluation purposes. We randomly selected 5 sample images from each category and used them as query images to test the retrieval accuracy. Query image was included in the database. Similarity metric used is Forward Hausdorff as it takes into account the spatial information of the feature points. Comparision of top 5 retrieval results using wavelet points only and feature points using proposed method have been shown.

The requirement for the distance metric to be used is the visual similarity of the retrieved results to the query image. Here, the set of feature points selected are visually significant because they have a large magnitude response and correspond to finer details in the image like edges and texture. Wavelet transform method has been used as it packs much of the information content of an image into few coefficients of large amplitude while the rest of the coefficients are small. Extracting these coefficients by judicious thresholding is sufficient to represent most of the relevant image information. In the feature maps, one can see the grouping of wavelet coefficients with high amplitude in textured regions. Such redundancy could be reduced for faster retrieval. Similarly the implementation of modulus edge map has edge points which are of different amplitudes. The points with higher magnitude response would survive the sub-sampling operation. Further reduction is done by picking only those points which are common to both the wavelet map and the modulus map. The end result has only few points but visually significant.

Hausdorff metric was in many ways a natural choice as the feature points can be used as local descriptors. The spatial information of the feature point set could be utilized for efficient retrieval. It is also practical as we now represent the image with a limited number of feature points. The two point sets for comparision would be that of the query image and that of the image from the database respectively. Minimum hausdorff distance implies perceptually most similar image to the query image. Importantly, since the number of feature points will vary for the query image and the database image, Hausdorff distance is a good choice. It does not need the length of the feature vector to be fixed. Point sets can have different number of feature points.

We define precision as the ratio of the number of relevant images retrieved to the number of returned images. Average precision is defined for all image classes. As shown in figure 13, average precision improves with our proposed method.
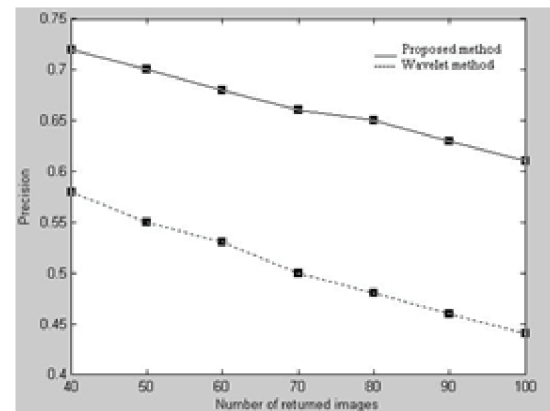

Fig 13: Average precision Vs Number of returned images using proposed method and wavelet method

Some preprocessing of the images was done for images before the implementation of the algorithm. For example, in the case of 'flowers' category, color channel was used for improved retrieval results. In case of MRI scans, histogram processing of the image was initially done. A thresholding was done on the image by carefully computing a threshold, $\tau_1 = 1.2*(\mu + \sigma)$ for better representation of the brain lesion where μ and σ are mean and standard deviation of the image respectively. Figure 14 illustrates the retrieval results for different categories of images.
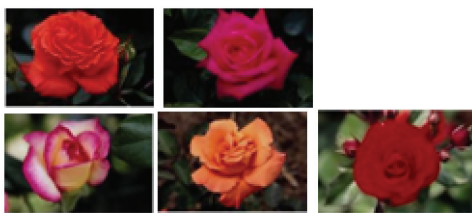
Figure 14: (a)Flowers - Query image with top 5 Retrieval results using wavelet based points only
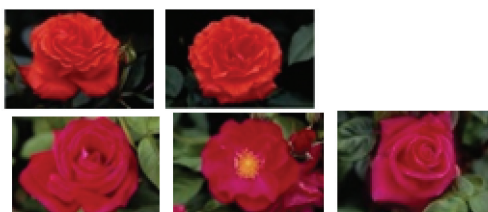

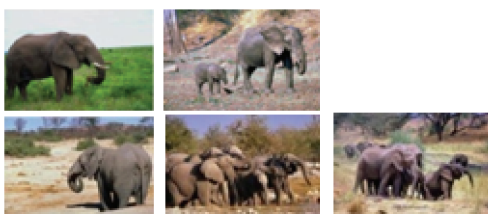Figure 14: (b)Flowers – Query image with top 5 Retrieval results using significant feature points

Figure 14: (c) Elephants - Query image with top 5 Retrieval results using wavelet based points only
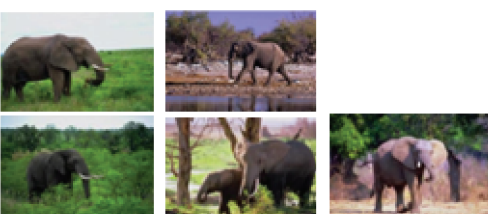

Figure 14: (d) Elephants – Query image with top 5 Retrieval results using significant feature points

Figure 14: (e) Natural Scenes - Query image with top 5 Retrieval results using wavelet based points only

Figure 14 : (f) Natural Scenes – Query image with top 5 Retrieval results using significant feature points

Slice 14   Slice 19

Slice 15   Slice 16   Slice 13
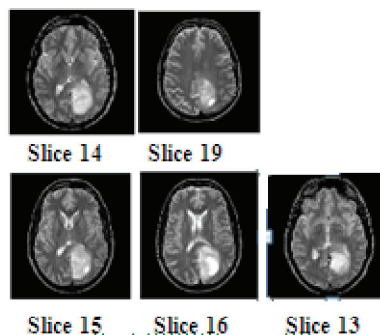
Figure 14: (g) MRI scan - Query image with top 5 Retrieval results using wavelet based points only

Slice 14   Slice 18

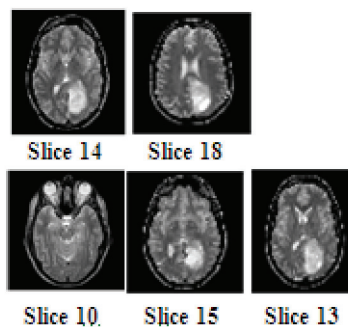Slice 10   Slice 15   Slice 13

Figure 14: (h) MRI scan – Query image with top 5 Retrieval results using significant feature points

6. Conclusions

We presented a visually significant feature point detector based on combining wavelet method and edge maps. The feature points are located in visually significant regions (both edges and textured region). There is always a need of speedy but efficient retrieval algorithm. We tried to bring a novel method to tackle this problem of image retrieval. Sparse representation of image information coupled with a distance metric which exploits the spatial information of these point sets have been employed for faster computation and better retrieval results.

Acknowledgements

References

1. R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: ideas, influences ,and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, Article 5, 2008.
2. Moravec, H. (1980). Obstacle avoidance and navigation in the real world by a seeing robot rover (Technical Report CMU-RI-TR- 80-03). Robotics Institute, Carnegie Mellon University.
3. C. Harris and M. Stephens, '' A Combined Corner and Edge Detector '', Proc. of 4th Alvey Vision Conference, 1988, pp. 147-151.
4. Shi, J., & Tomasi, C. (1994). Good features to track. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR'94)* (pp. 593–600).
5. Mikolajczyk, K., & Schmid, C. (2001). Indexing based on scale invariant interest points. In *Proceedings of the IEEE intl. conference on computer vision (ICCV'01)* (Vol. 1, p. 525).
6. Mikolajczyk, K., & Schmid, C. (2004). Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, *60*(1), 63–86.
7. Lindeberg, T. (1994). Scale-space theory: A basic tool for analyzing structures at different scales. *Journal of Applied Statistics*, *21*(2),224–270.
8. Lowe, D. G. (1999). Object recognition from local scale-invariant features.In *Proceedings of the IEEE intl. conference on computer vision (ICCV'99)*, Corfu (pp. 1150–1157).
9. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*(2), 91–110.
10. Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, *110*, 346–359.
11. I. Daubechies, '' Ortho-normal Bases of Compactly Supported Wavelets '', Communications on Pure and Applied Mathematics, 1988, Vol. 41, pp. 909-996.
12. C.-H. Chen, J.-S. Lee and Y.-N. Sun, '' Wavelet Transformation for Gray-level Corner Detection '', Pattern Recognition, 1995, Vol. 28, No. 6, pp. 853-861.
13. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. journal of Computer Vision,* 2004.
14. Schmid, C., Mohr, R., & Bauckhage, C. (2000). Evaluation of interest point detectors. *International Journal of Computer Vision*, *37*(2), 151–172.
15. S. Bres and J.-M. Jolion, '' Detection of Interest Points for Image Indexation '', 3rd Int. Conf.on Visual Information Systems, Visual99, Amsterdam, The Netherlands, June 2-4 1999,pp. 427-434.
16. C. Wolf, J.-M. Jolion, W. Kropatsch and H. Bischof, '' Content Based Image Retrieval Using Interest Points and Texture Features '', ICPR, 2000.
17. Etienne Loupias, Nicu Sebe, 'Wavelet-based Salient Points: Applications to Image Retrieval Using Color and Texture Features', Lecture Notes in Computer Science
18. Q. Tian, N. Sebe,M.S. Lew,E. Loupias,T. S. Huang, " Image Retrieval Using Wavelet-Based Salient Points", *Journal of Electronic Imaging, Special Issue on Storage and Retrieval of Digital Media*, pp. 835-849, Vol. 10(4), October Issue, 2001
19. L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual-attention for rapid scene analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20(11), pp. 1254-1259,1998.
20. T. Kadir and M. Brady, "Scale, Saliency and Image Description", *International Journal of Computer Vision*, Vol. 45(2), pp. 83-105, 2001.
21. H. Zhang, R Rahmani, S R Cholleti, S A Goldman, "Local Image Representations using Pruned Salient Points with Applications to CBIR," ACM, October 23-27, 2006, Santa Barbara, California ,USA.
22. H. Song, B. Li and L. Zhang, "Color salient points detection using wavelet," *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, June 21–23, 2006, pp. 10298–10301.
23. Alexandra Teynor,Hans Burkhardt," Wavelet-based Salient Points with Scale Information for Classification", IEEE 2008
24. Minakshi Banerjee, Malay K. Kundu, Pradipta Maji," Content-based image retrieval using visually significant

point features, Fuzzy Sets and Systems, Elsevier, March 2009.

25. Liu W, Tong QY , " Medical image retrieval using salient point detector" , IEEE proceedings, 27th Annual international Conference of Engineering in Medicine and Biology Society, pp 6352-6355,2005.

26. R. da S. Torres, A.X. Falca˜o, "Contour salience descriptors for effective image retrieval and analysis", Elsevier, Image and Vision Computing 25 (2007) 3–13.

27. Yao-Hong Tsai , "Salient Points Reduction for Content-Based Image Retrieval", World Academy of Science, Engineering and Technology 25 2009.

28. Paolo Piro, Sandrine Anthoine, Eric Debreuve, and Michel Barlaud, "Sparse Multiscale Patches for Image Processing", Springer Verlag LNCS 5416, p. 284-304, 2009.

29. M. Do and M. Vetterli. Wavelet based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. TIP, 11:146–158, 2002.

30. Q. Zhang and E. Izquierdo. optimizing metrics combining low-level visual descriptors for image annotation and retrieval", ICASSP, Toulouse, France, 2006.

31. Sandrine Anthoine, Eric Debreuve, Paolo Piro, Michel Barlaud, "Using neighborhood distributions of wavelet coefficients for on-the-fly, multiscale-based image

retrieval" , Proc. of WIAMIS'08, 07- 09/05/2008, Klagenfurt, Austria, pp. 38-41.

32. Calonder, M., Lepetit, V., & Fua, P. (2008). Keypoint signatures for fast learning and recognition. In *Proceedings of the 11th European conference on computer vision (ECCV'08)*, Marseille, France.

33. Carrera, G., Savage, J., & Mayol-Cuevas, W. (2007). Robust feature descriptors for efficient vision-based tracking. In *Proceedings of the 12th Iberoamerican congress on pattern recognition* (pp. 251–260).

34. Steffen Gauglitz ,Tobias Höllerer , Matthew Turk , " Evaluation of Interest Point Detectors and Feature Descriptors for Visual Tracking" , International Journal of Comput erVision , Springer, March 2011.

35. Zhu-lihua, Ji-xiaoping. "The application of wavelet modulus maxima algorithm in the image retrieval".SCI-tech information development and economy. 2007, vol.17(3): 169-17

36. J.F. Canny, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986 ,vol.8, pp.679-698