Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 571 – 580

Procedia Computer Science

International Conference on Computational Modeling and Security (CMS 2016)

Robust Pre-processing Technique Based on Saliency Detection for Content Based Image Retrieval Systems

Chesti Altaff Hussain[a]*, D. Venkata Rao[b], S. Aruna Masthani[c]

[a*]*Department of ECE, JNTU College of Engineering, Anantapuramu, chestialtaff@gmail.com*
[b]*Narasaraopet Institute of Technology, Narasaraopet,dv2002in@yahoo.com*
[c]*Department of ECE, JNTU College of Engineering, Anantapuramu, aruna_masthani@yahoo.com*

Abstract

The perceptual quality which attracts viewer attention by making the objects differ from its neighbours is said to be Visual saliency. They concentrate on the region of interest. The paper aims at improving the performance of content based image retrieval using saliency detection approach. Several methods have been developed to extract the saliency information from an image. We use the state of the art Quaternion transform for to detect the saliency. The paper focuses on the content based image retrieval systems based on scale invariant feature transform and region segmentation. Experimental results prove that the proposed technique outperforms the existing techniques and produce better retrieval results.

Peer-review under responsibility of the Organizing Committee of CMS 2016

Keywords: Visual Saliency, Perceptual quality, Segmentation, Content based image retrieval, Scale Invariant transform, Quaternion transform.

Peer-review under responsibility of the Organizing Committee of CMS 2016
doi:10.1016/j.procs.2016.05.223

1. INTRODUCTION

A computer system used to retrieve, browse and search images from large databases of digital images is known as Image retrieval system. The solution of image retrieval through the application of computer vision is known as Content Based Image Retrieval (CBIR) and also known as Query By Image Content (QBIC) and Content Based Visual Information Retrieval (CBVIR) [1]. Different applications of CBIR are Architectural and Textile industries, Engineering design, crime prevention, remote sensing systems, Military, Retail catalogues, Art collections geographical information.

To identify CBIR in this paper various textual features were used under sophisticated techniques. Similarity measurement also takes place to trace the textual features for 16 typical texture patterns. Some textual features used in this condition are contrast, coarseness, directionality, regularity, line likeness and roughness [2]. CBIR is of two types 1) offline mode, 2) online mode. In online large datasets in website and from search engines were obtained. In case of offline a large stored static database used to obtain the retrieval results. This can be explained with the following Figure 1.

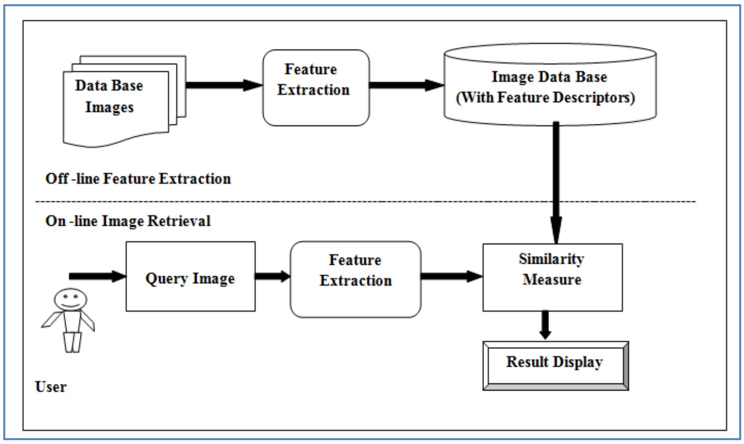


***Fig. 1.**Content based image retrieval*

In CBIR many developments were proposed and drastic need of speed and exact retrieval of images possess the necessity of exact retrieval of images based color, texture and shape. For this basis many algorithms were proposed some approaches are genetic algorithm [3], clustering and segmentation are advancing out of all other approaches comparatively.

2. CONTENT BASED IMAGE RETRIEVAL

Content based image retrieval is to index, search, browse and retrieval relevant images from a large selection of digital image collection [4]. Globally the visual content descriptors include color, shape and texture and locally the descriptors can be derived from decomposed regions of images. Image Segmentation is helpful in pre-processing stage for CBIR algorithms. Decomposing an image into parts for further analysis and performing a change of representation are the objectives of image segmentation. In spatial segmentation, valley determination and valley growing implemented in region growing for segmented regions [5] [6].Quaternion is the generalization of complex number which has one real part and 3 imaginary parts represented in Cartesian form as:

$$A = b + ci + dj + ek \qquad (1)$$

where b, c, d, e are real numbers and the complex operators i, j and k obey the following rules:

$$i * j = k, \quad j * k = i, \; k * i = j, \; j * i = -k,$$
$$k * j = -i, \; i * k = -j \qquad (2)$$
$$i^2 = j^2 = k^2 = i * j * k = -1$$

The quaternion conjugate is given by,

$$\bar{A} = b - ci - dj - ek \qquad (3)$$

and the modulus of quaternion is given by,

$$|A| = \sqrt{b^2 + c^2 + d^2 + e^2} \qquad (4)$$

An image pixel has three components via, Red, Green and Blue which can represented using pure quaternion as

$$f(x, y) = r(x, y)i + g(x, y)j + b(x, y)k \quad (5)$$

The advantage of using quaternion based operations to manipulate color information in an image is that treating each color triple as a whole unit instead of considering each color as independent channel which results in higher color information accuracy.

3. RELATED WORK

Different approaches of saliency are proposed in this context [7].Every image $I \in R^{MxNxC}$ with at most 4 color components, i.e. $C \leq 4$ can be represented using MxN quaternion matrix

$$\boldsymbol{I}Q = \boldsymbol{I}4 + \boldsymbol{I}1i + \boldsymbol{I}2j + \boldsymbol{I}3k + \boldsymbol{I}4 + \boldsymbol{I}1i + (\boldsymbol{I}2 + \boldsymbol{I}3i)\,j \qquad (6)$$

where I_C denotes the M x N matrix of the c^{th} image channel. It is common to represent the (potential) 4th image channel as the scalar part [8] [9], because when using this definition it is possible to work with pure quaternion's for the most common color spaces such as, RGB, YUV and Lab.

Weighted Quaternion Color Components: Naturally, as done by Bian et al. [10] [11] and also related to the recent trend to learn feature dimension weights [12] [13], the relative importance of the s which is for the color space component can be modelled for the visual saliency by introducing a quaternion component weight vector $w = [\, w_1 \quad w_2 \quad w_3 \quad w_4 \quad]^T$ and adapting Eq. appropriately:

$$I_Q = w_4 \boldsymbol{I}_4 + w_1 \boldsymbol{I}_1 i + w_2 \boldsymbol{I}_2 j + w_3 \boldsymbol{I}_3 k \qquad (7)$$

In case of equal influence of each color component, i.e. uniform weights, Eq. 6 is a scaled version of Eq. 7 which is practically equivalent for our application.

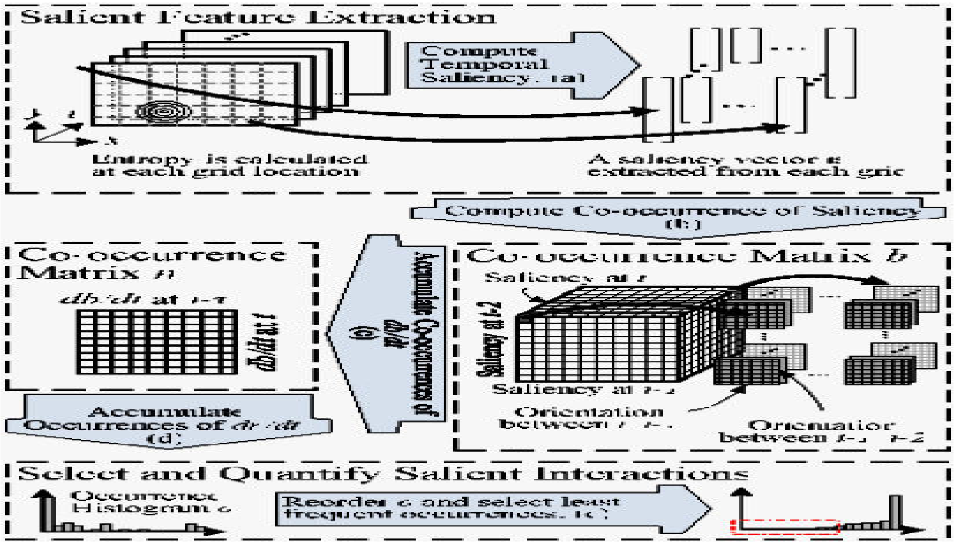


***Fig. 2.**: Diagrammatic representation of Saliency*

I. ALGORITHM: IMAGE SEGMENTATION

1. Read the input image

2. At the corresponding scale, calculate the local minima values.

3. Then start the region growing algorithm which in turn has two steps – Valley determination and valley growing.

Valley Determination

3.1.1. Initially, the local minima’s are considered to be the base of region growing

3.1.2. For the pixels present the mean (μ) and standard deviation (σ) is to be calculated

3.1.3. Set a threshold TH at

3.1.4. If the minima values are less than the threshold value then these pixels are

3.1.5. The value a is chosen from the values [–0.6, –0.4, –0.2, 0, 0.2, 0.4] which

3.1.6. The regions are grown from the valley points.

3.1.7. Remove "holes" in the valleys.

3.1.8. Average the local pixel values in the remaining un-segmented part of the region technique. around.$TH = \mu H + \alpha\sigma H$ called valley points. gives most of the valley points.

3.1.9. Calculate local pixel values for the remaining pixels at the next smaller scale

3.1.10. Grow the remaining pixels one by one at the smallest scale. To form growing areas below the average connect pixels. If a growing area is adjacent to one and only one valley, it is assigned to that valley. to more accurately locate the boundaries. Repeat steps 3.1.7 and 3.1.8. in a buffer Pixels at the valley boundaries are stored. Each time, the pixel with the minimum local pixel value is assigned to its adjacent "valley" and the buffer is updated till all the pixels are classified.

4. Finally, the resulted regions are merged based on their color similarity using Euclidean distance measure.

SIFT algorithm consists of four major steps they are

- Scale-space extreme detection
- Key point localization
- Orientation assignment
  Key point descriptor

II. ALGORITHM FOR SCALE INVARIANT FEATURE TRANSFORM SALIENCY ALGORITHM [3]

1: for each pixel of the current image frame do

2: Calculate the hue-saturation distribution,b_{HS} for all spatio-temporal scales that are considered;

3: for each S_t do

4: for each S_s do

5: Calculate the entropy$H_D(x)$ of the hue-saturation distribution ;

6: end

7: end

8: $\hat{S}_s$ Calculate the peaks in$H_D(x)$ across spatial scales where S_t = 0;

9: Calculate the inter-scale saliencyW_D;

10: for each peak in the feature vector $\hat{S}_s$ do

11: $\hat{S}_t$: Calculate the peaks in$H_D(x)$ over temporal scales where S_s=$\hat{S}_s$;

12: if Peaks in $\hat{S}_t$ exist, then

13: Calculate the spatial saliency,$Y_D([\hat{S}_s \quad 0]^T, X)$and at$S_T = 0$;

14: if there is a peak over spatial scales in $H_D(x)$ at S_T=$\hat{S}_T$ andS_s=$\hat{S}_s$ then

15: Calculate the spatial saliency, $Y_D([\hat{S}_s \quad \hat{S}_T]^T, X)$at S_T=$\hat{S}_T$ andS_s=$\hat{S}_s$;

16: if $Y_D([\hat{S}_s \quad \hat{S}_T]^T, X)$>$Y_D([\hat{S}_s \quad 0]^T, X)$ then

17: Discard this peak;

18: end

19: end

20: end

21: end

22: for All the spatio-temporally salient peaks that are left, do

23: Calculate the inter-scale saliency W_D over temporal scales at the current salient spatial scale. Calculate the temporal saliency at this point.

24: end

25: end

Fig. 3. Sample images from Corel Database

Fig. 4. Saliency detection output

The figure 2 the saliency detection output of the images used. This will enhance the pre-processing stage of feature extraction by including only the prominent features of the object that is present in the ROI. The background in complex images will not be considered for feature extraction.

4. RESULT ANALYSIS

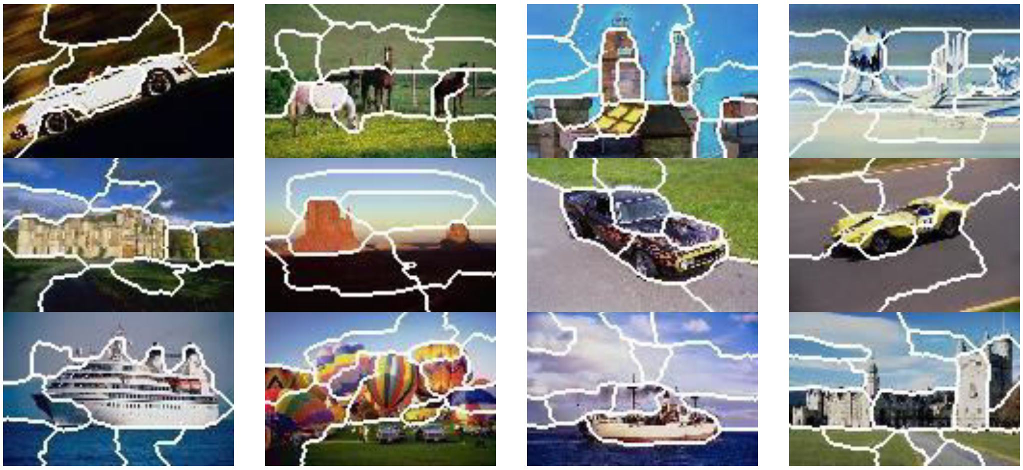


Fig. 5. Output of region segmentation

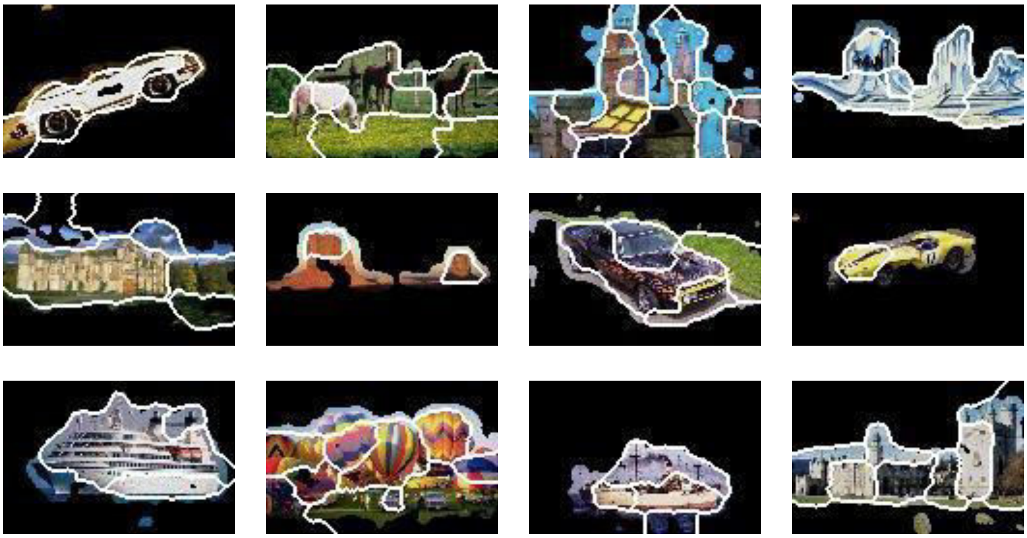


Fig. 6. Output of region segmentation after saliency extraction

As a result of the saliency detection, the number of regions detected in the image is reduced prominently. This will increase the efficiency of the retrieval as false regions will not be included.

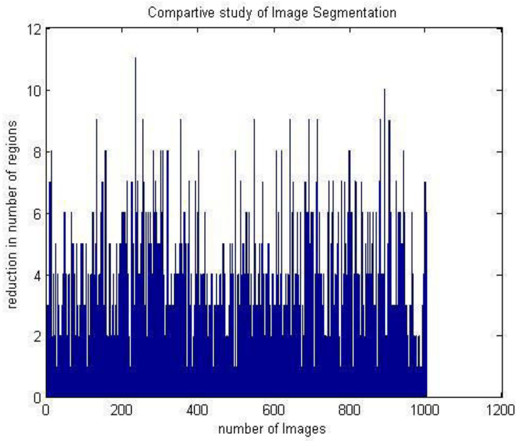


Fig. 7. Plot representing the reduction in number of regions detected.

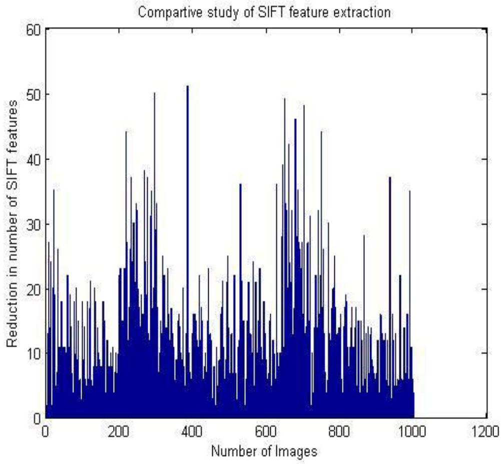


Fig. 8. Plot representing the reduction in SIFT features detected

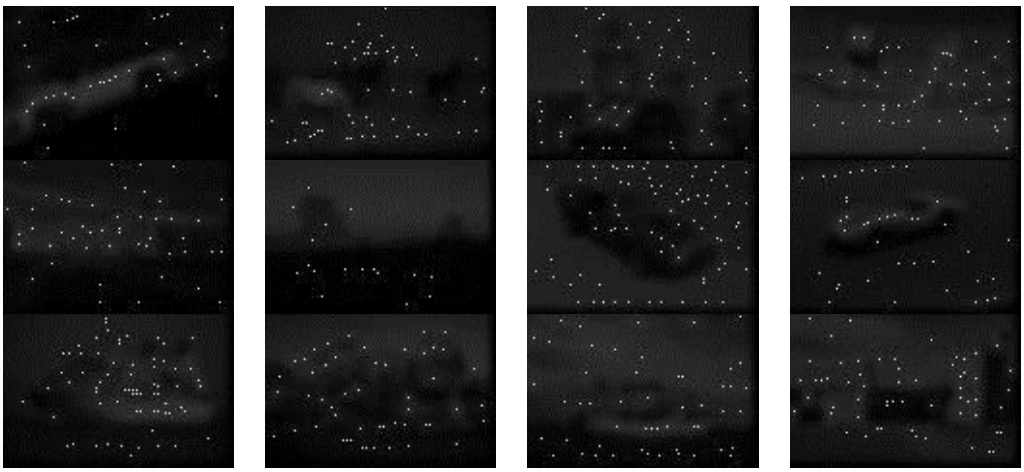
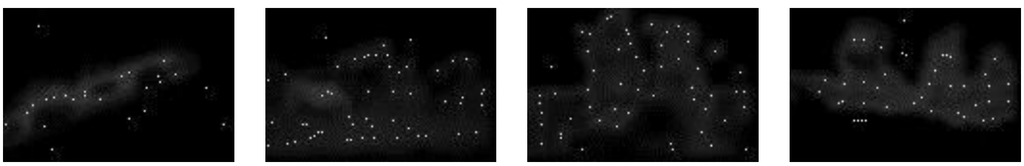


Fig. 9. Output of SIFT feature Extraction



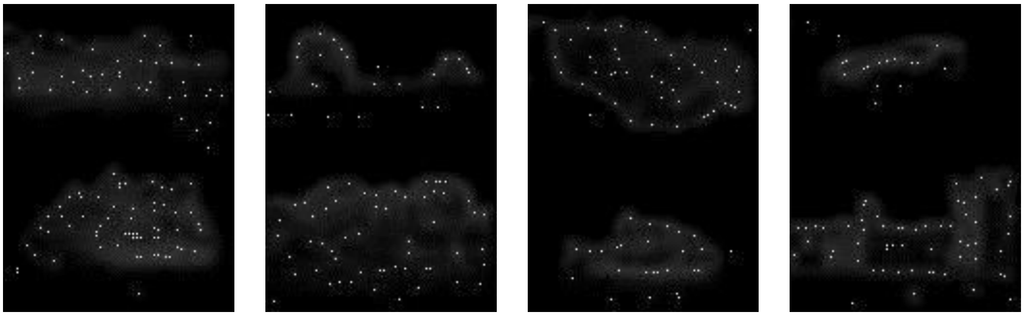


Fig. 10. Output of region segmentation after saliency extraction

Figure7 and Figure 8 depicts the comparison of number of SIFT features detected. Figure 9 and Figure 10 show the SIFT feature output. The SIFT features extracted are only from the ROI thus increasing the accuracy of the image retrieval system. At the time of matching of the SIFT features, since only the ROI ones are present, better retrieval results.

5. CONCLUSION

The paper presents an improved pre-processing technique based on image saliency detection using state of the art quaternion transform. As the saliency output produces only the regions in the ROI, pre-processing techniques like region segmentation and Scale Invariant Feature Transform extract accurate features by eliminating false ones. By proposing SIFT saliency transformation technique helps in improving the accuracy of image retrieval system. The experimental results performed on the Corel Database have has proved that the performance of the image retrieval system with complex background images has been improved.

References

1. Eakins, John, Graham, Margaret, 'Content Based Image Retrieval', University of Northumbria at Newcastle. Retrieved 10-03-2014.
2. Tamura, Hideyuki, Mori, Ymawaki, Shunji, Takashi, 'Textual Features corresponding to visual perception', Vol 8, Issue 6, Systems Man and Cybernetics, IEEE transactions, 2007.
3. Chih-Chin Lai, Ying-Chuan Chen," Color Image Retrieval Based on Interactive Genetic Algorithm", Next Generation Applied Intelligence,Lecture Notes in Computer Science Volume 5579, 2009, pp 343-349.
4. Datta, Ritendra; Dhiraj Joshi, Jia Li, James Z. Wang (April 2008). "Image Retrieval: Ideas, Influences, and Trends of the New Age". ACM Computing Surveys 40 (2): 1–60.
5. Sangwine. S, Ell, 'Colour image lters based on hypercomplex convolution'. IEEE Proc. Vision, Image and Signal Processing 147 (2000) 8993.
6. Mustafa Ozden, EdizPolat," A color image segmentation approach for content-based image retrieval", Pattern Recognition 40 (2007) 1318–1325.
7. Guo, C., Zhang, L, 'A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression'. IEEE Trans. Image Process. 19 (2010) 185 – 198.
8. Bian, P., Zhang, L. 'Biological plausibility of spectral domain approach for spatiotemporal visual saliency'. In advances in Neural Information Processing Systems. (2009)
9. Zhao, Q., Koch, C. 'Learning a saliency map using _xated locations in natural scenes'. Journal of Vision 11 (2011) 1_15
10. Judd, T., Ehinger, K., Durand, F., Torralba, A. 'Learning to predict where humans look'. Proc. Int. Conf. Comp. Vis. (2009).
11. X. Hou and L. Zhang. "Saliency detection A spectral residual approach". In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pages 1–8. 2007.

12. Ling Shao, Timor Kadir and Michael Brady" Geometric and Photometric Invariant Distinctive Regions Detection" Information Sciences. 177 (4):1088- 1122, 2007.
13. J. Sivic and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," Proc. IEEE Int'l Conf. Computer Vision, pp. 1470-1477, 2003.