
ABSTRACT

Salient region detection is one of a pre-processing method which are useful for several image processing application. Image processing application includes image retrieval, object recognition, face recognition, adaptive compression and image retargeting. Thus salient region detection is detecting the more attracted regions in an image. It can be define as contrast difference between the visually attracted regions and its neighbourhood in an image. In this paper we proposed a novel method by creating a saliency map by average of the three feature maps added by colour spatial variance of an image. Main feature maps are frequency map and spectral map.

KEYWORDS: Color spatial map, Frequency map, Image processing, salient region detection, Spectral map.

INTRODUCTION

Visual saliency is a term used to identify a meaningful region in an image. This is mainly used to calculate whether an object in an image attracts viewers' attention or not. It always depends on their uniqueness or rarity. A salient object can always be identified due of many features. They can be color, edge, gradient and boundaries which are the main characteristics for visual stimuli. Visual saliency is a quality which makes some object in the scene pop out from their surroundings and immediately attract our attention. The process of reducing redundant information and detecting the most important parts of an image can significantly reduce the complexity of visual processing.

Humans always gave more attention to important regions in an image and thus effortlessly judge their importance. Detecting salient regions computationally in an image remains a significant goal, where it helps in allocating computational resources while image analysis. Salient regions are widely used in many computer vision application .They are also used as a preprocessing method for many image processing applications. They can be used in image retrieval, image resizing, image compression, object recognition, face recognition and object-of-interest image segmentation. While creating saliency map the more salient pixels will be shown in white color, less salient pixels will be shown in black color as shown in Fig.1.

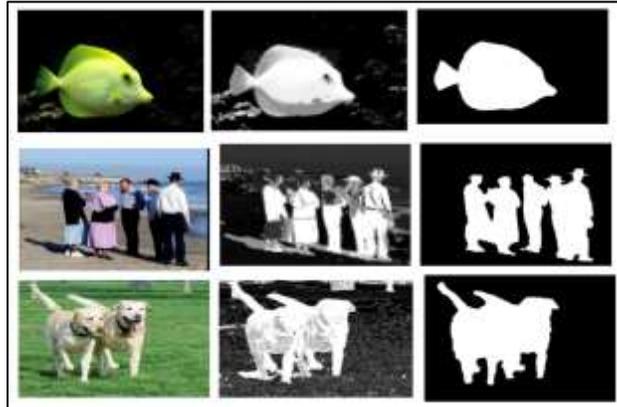


Figure 1. Left: Original images. Middle: Saliency maps generated from the proposed method. Right: Human labelled ground truths.

Saliency detection means detecting visually attracted regions in images. It is an aspect of exploring visual attention from a computer vision viewpoint. The human pays unequal attention to what is seen in the world. When looking at some images, people are usually attracted by a certain objects in an image. Other subjects appear uninteresting for them. Detecting those areas of attention is called saliency detection.

ChienChi Chen, Jian-Jiun Ding [12] proposed this method to create saliency map. This paper is based on color spatial variance, Local global contrast, Border measurement. It is based on a common assumption that, color which is extensively distributed in an image can be the background color. Based on the observation that salient objects are commonly less likely to be attached with the borders of an image. The boundary information of each segment is used to compute the salient region. Last but not least, a PCA context aware method is used for extracting local global contrast and the segment boundary is considered when calculating the saliency value. Peng Jiang, Haibin Ling [8] propose a novel salient region detection algorithm. In this paper three important visual cues are integrated namely uniqueness, focusness and objectness (UFO). Uniqueness captures the appearance derived visual contrast and focusness reflects the fact that salient regions are often photographed in focus. Objectness helps to keep completeness of detected salient regions. Zhang [3] used a global contrast based salient region detection. In this paper it evaluates global contrast differences and spatial coherence.

BACKGROUND

The framework in this paper includes two main features as a base, L_0 smoothing filter and color spatial variance. Boundary information is computed using L_0 smoothing filter and segmentation. L_0 smoothing filter mainly reduce redundant information and thus preserving the prominent set of edges in an image. The L_0 smoothing filter is a sparse gradient counting scheme in an optimization framework to restrict the discrete number of intensity changes among neighboring pixels, which is mathematically related to the L_0 norm for the pursuit of information sparsity. It also preserves main color changes and strengthens the edges by increasing gradient of transition. Fig. 2 shows that the structure of the smoothed result is comparing with the original structure. The textures of background is actually smoothed, while the edges between them are actually sharpened.

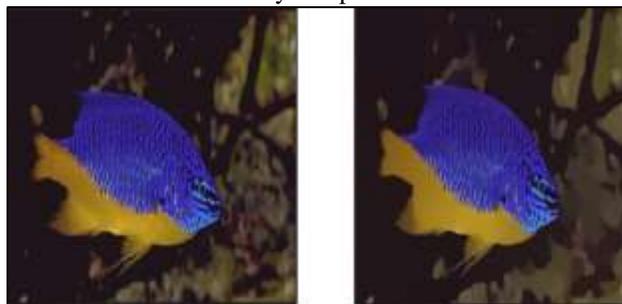


Figure 2. A demonstration of the L_0 smoothing filter: Left: The original image. Right: The smoothed image using L_0 smoothing filter.

For a given input image I , S denotes the computed result, and the gradient $\nabla S_p = (\partial_x S_p, \partial_y S_p)^T$ is used to calculate the colour difference between the neighbouring pixels along the x and y directions for each pixel, p . The gradient measure is defined as

$$C(S) = \# \{p \mid |\partial_x S_p| + |\partial_y S_p| \neq 0\} \quad (1)$$

Equation (1) determines the number of non-zero magnitudes $\partial_x S_p, \partial_y S_p$ in I . S is computed by

$$\min_S \{ \sum_p (S_p - I_p) + \lambda \cdot C(S) \} \quad (2)$$

Here λ is recommended to be 0.01 and it shows the degree of smoothness in an image.

PROPOSED METHOD

Extracting salient regions from an image is very difficult if it contain multi salient objects. Existing methods of saliency detection generally have many disadvantages like low resolution, poorly defined borders and some may be highly expensive to compute. Some method will not uniformly highlight whole salient regions, thus may produce higher saliency values at object edges. This project proposed a method that is salient region detection based on average saliency map. To detect salient region from an image extract four features, namely global contrast, frequency weight map, boundary information and color spatial variance of the image. Global contrast find distinction between regions, boundary information used to extract object boundary information, spatial weight used to find spatial weight of an image. color spatial variance used to highlight entire region.

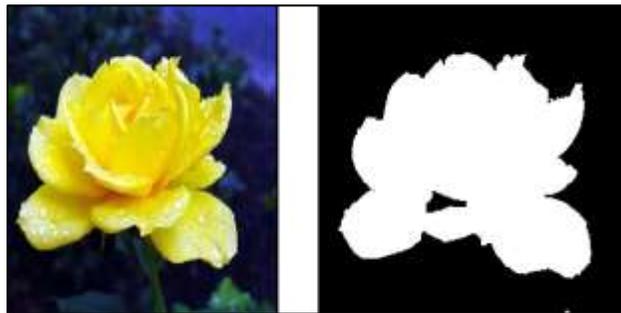


Figure 3: Segmentation Result

In the proposed saliency framework, the input image is passed into four phases and produce four saliency map. First an input image is passed into L_0 smoothing filter. Next step is segmentation. In segmentation grey level threshold of an image is taken. Using that grey level threshold, image is segmented. After segmentation, compute boundary weight of the salient objects. Once get boundary weight then calculate saliency map using boundary weight.

APPROACH

A. Saliency map using boundary weight

Boundary of a salient object can be determined by using L_0 Smoothing filter and image segmentation. A smoothed image can be segmented perfectly than a normal image. Filtered image can be segmented efficiently and thus segments can be separated in a proper way. Currently there is no best algorithm for every kind of normal images. Segmentation helps in simplifying and representing the whole image into a more simpler and meaningful regions which helps to analyse the image easily. One of the method of image segmentation is thresholding method.

A global threshold value is used to turn a colour image into a binary image. So first is a global threshold of an image is selected. The pixel value greater than the global threshold is set as 1 and pixel values less than threshold is set as 0. So a binary image is obtained after segmentation. In the binary image the object area will be white in colour and the background region will be black in colour. Segmented image is shown in fig 3.

Basic global thresholding contain following steps:

- 1) Initial estimate is selected for threshold value T
- 2) Image is segmented using T .
- 3) Get two group of pixel values P_1 and P_2 . In P_1 grey values $> T$ and P_2 grey values $\leq T$.
- 4) Compute average grey level values μ_1 and μ_2 for pixels in P_1 and P_2 .

5) Compute a new threshold value T_2 .

$$T_2 = \frac{1}{2} (\mu_1 + \mu_2)$$

6) Repeat from 2 to 4 until difference in T_1 successive iteration is less than predefined parameter T_2 .

The grey level threshold value of an image is taken in segmentation. Image is segmented using that threshold value. Then compute boundary weight of the salient object.

$$BW3 = (BW1 - BW2) \quad (3)$$

BW1 is the segmented image and BW2 is the image that is eroded by structuring element. BW3 is border of an image. First saliency map is calculated using boundary weight.

$$S1 = \exp(-BW3 / \sigma) \quad (4)$$

B. Saliency map using global contrast

Spectral residual saliency map does not use prior knowledge of targets, it is independent of features and categories. First analysing the log-spectrum of an image, then extract the spectral residual of an image in spectral domain, and last will construct the corresponding saliency map in spatial domain. From the perspective of information theory, effective coding decompose the image information $H(\text{Image})$ into two parts:

$$H(\text{Image}) = H(\text{Innovation}) + H(\text{PriorKnowledge}) \quad (5)$$

$H(\text{Innovation})$ is the novelty part, and $H(\text{Prior Knowledge})$ is the redundant information that can be suppressed by a coding system. Main steps for computing saliency map are:

1. First apply FFT on the image
2. Extract amplitude spectrum and phase spectrum
3. Apply Gaussian kernel on the amplitude spectrum then get a smoothed spectrum
4. Subtract smoothed spectrum from amplitude spectrum to get the spectral residual of an image
5. Apply inverse FFT to get saliency map
6. First apply fast Fourier transform on the image.

$$\text{FFT} = \text{fft2}(\text{image}) \quad (6)$$

7. Extract amplitude spectrum and phase spectrum.

$$\text{LogAmplitude} = \log(\text{abs}(\text{FFT})) \quad (7)$$

$$\text{myPhase} = \text{angle}(\text{FFT}) \quad (8)$$

8. Apply filter on the amplitude spectrum then get a smoothed spectrum.

$$\text{Smoothened} = \text{imfilter}(\text{LogAmplitude}, \text{fspecial}(\text{average}, 3); \text{replicate}) \quad (9)$$

9. Subtract smoothed spectrum from amplitude spectrum to get the spectral residual of an image. The spectral residual contains the innovation of an image. It serves like the compressed representation of a scene.

$$\text{SpectralResidual} = \text{LogAmplitude} - \text{smoothened} \quad (10)$$

10. Apply inverse FFT on spectral residual and phase spectrum to get saliency map. Using Inverse Fourier Transform, construct the output image in spatial domain called the saliency map

$$S2 = \text{abs}(\text{ifft2}(\exp(\text{SpectralResidual} + \text{Phase})))^2 \quad (11)$$

This model highlight salient object boundaries and saliency covers unknown features such as curve, density, intersection.

C. Saliency map using frequency weight map

The saliency map S for an image I with width W and height H pixels can be formulated as:

$$S(x, y) = | I_{\mu} - I_{whc}(x, y) | \quad (12)$$

where I is the arithmetic mean pixel value of the image and I_{whc} is the Gaussian blurred version of the original image to eliminate noise, texture details as well as coding artifacts. To extend Equation 10 to use features of colour and luminance, rewrite it as:

$$S_3(x, y) = \| I_\mu - I_{whc}(x, y) \| \quad (13)$$

where I is the mean image feature vector, I_{whc} is the corresponding image pixel vector value in the Gaussian blurred version of the original image, and $\|$ is the L2 norm which is the Euclidean distance. A Spatial weight map is shown in fig 5. The resulting saliency maps are better suited to salient object segmentation. It uniformly highlight the salient regions with sharpened boundaries. A saliency map using frequency weight map is shown in fig 6.

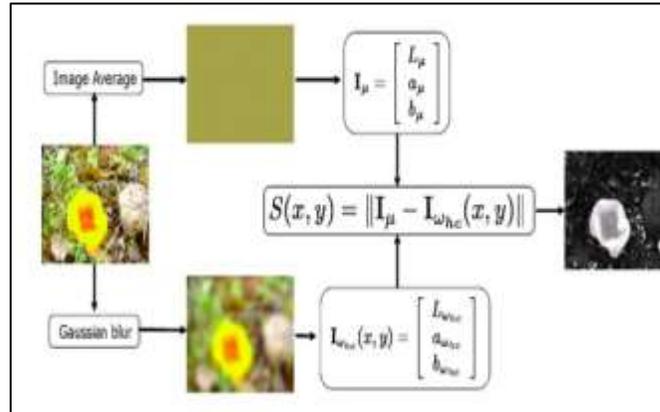


Figure 5: Spatial weight map representation



Figure 6: Saliency map using frequency weight map

D. Saliency map using color spatial variance

The Color spatial variance based on GMM, GMM stands for Gaussian Mixture Model, which is one of the most used methods for clustering. Color spatial variance is a widely used global feature that compares with the human visual system. If a color is extensively distributed within an image, it can be assumed as a background color. In other words, a specific color having a smaller spatial variance will attract greater attention, and thus is more likely to be part of the salient object. Saliency map using color spatial variance is shown in fig 7. Each pixel in the GMM can be represented as

$$P(c|I_{x,y}) = \frac{W_c N(I_{x,y} | \mu_c \Sigma_c)}{\sum_c W_c N(I_{x,y} | \mu_c \Sigma_c)} \quad (14)$$

Where W_c , μ_c , Σ_c are the weight, mean, and covariance of the c^{th} component. $N()$ is the Gaussian model, and $I(x, y)$ is the pixel at the coordinate (x, y) . In this framework, the maximum likelihood is used instead of probability models.

$$C^* I_{(x,y)} = \text{argmax}_c (c|I_{(x,y)}) \quad (15)$$

where C is the set of all components. The horizontal and vertical spatial positions are $v_h(c^*)$ and $v_v(c^*)$ of each color component can then be further simplified as

$$v_h(c^*) = \frac{1}{N} \sum_{I_{x,y} \in C^*} |x - \mu_x(c^*)|^2 \quad (16)$$

$$v_v(c^*) = \frac{1}{N} \sum_{I_{x,y} \in c^*} |y - \mu_y(c^*)|^2 \quad (17)$$

where μ_x and μ_y are the means of the x- and y-coordinates, respectively, and N is the number of pixels with a maximum likelihood of c^* . The color spatial variance of the c^* th component can be defined as being the maximum of the x- and y-coordinate variances:

$$V(c^*) = \max(v_h(c^*), v_v(c^*)) \quad (18)$$

Finally, the global feature map is determined from

$$\text{Color}_{\text{sal}}(I_{(x,y)}) = V(c^*I_{x,y}) \quad (19)$$



Figure 7: Saliency map using color spatial variance

E. Final Saliency Map Computation

Finally, the saliency map is defined as being the average of the three feature maps added by the color spatial variance,

$$S = \frac{1}{3}(S1 + S2 + S3) + \text{Color}_{\text{sal}} \quad (20)$$

Using this proposed method, it can detect multi salient objects and detect the salient region if the object is far from the centre of the image, border of the image. This proposed system used for improving the performance of the saliency detection. Because of the four feature maps that are combined, the proposed algorithm is more robust than other methods and has similar saliency values in the entire salient region. That will uniformly highlight entire salient region and emphasize the large salient object. It also provide well-defined boundaries of salient objects. And give full resolution saliency map efficiently.

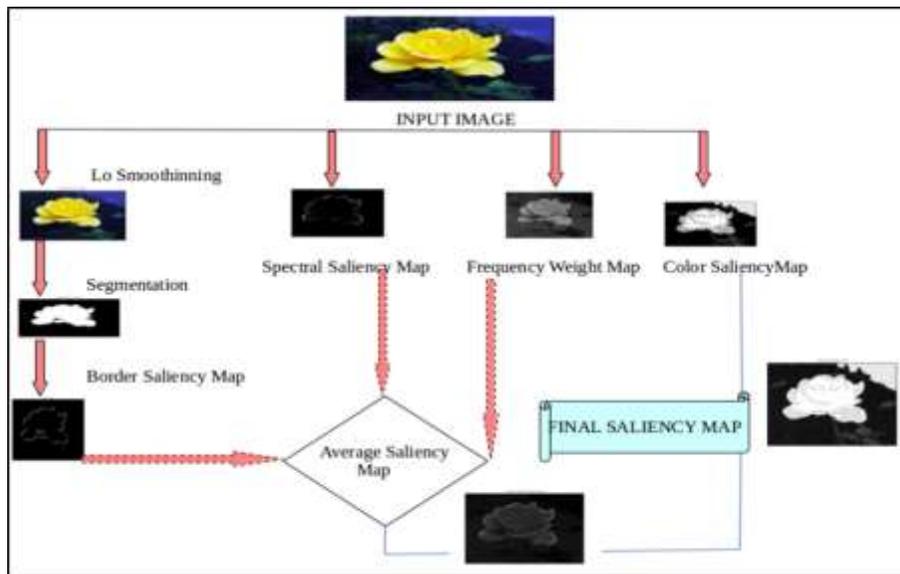


Figure 7a: Framework of our proposed methodology

Second saliency map is obtained by spectral residual. To get spectral residual map, first apply FFT on image. Then compute log spectrum and smoothed spectrum. Subtract smoothed spectrum from log spectrum. Finally apply inverse FFT to get a second saliency map. To compute frequency weight map. First compute arithmetic mean pixel value of the image. Apply Gaussian kernel on the image. Find Euclidian distance of the difference between arithmetic mean pixel value of the image and blurred version of the image.

Third saliency map is obtained using frequency weight map. Take the average of these three saliency map to get an average saliency map. Next compute color spatial variance. The color spatial variance is utilized to calculate the degree of color distribution. When a color is extensively distributed in an image, it may be the background color. Otherwise it is a salient object. That is lower spatial variance provide more attention, so it may be a salient object. So using color spatial variance another saliency map is obtained.

Finally saliency map is obtained by average saliency map is added with color spatial map. The final saliency map is very efficient to highlight the salient region. It detect salient region border of the image, in the center of the image, or even far from the center of the image. This saliency map offers many advantages over existing methods such as uniformly highlighted salient regions with well-defined boundaries, high efficiency and full resolution. Fig 7a shows the proposed method.

EXPERIMENTAL RESULTS

The proposed method was evaluated on MSRA dataset with a ground truth in the form of accurate human-marked labels for the salient regions. The saliency map is detected and are correctly localized. The success rate of saliency detection is 100 percent. The proposed technique is tested on MATLAB R2012a, to justify the effectiveness of the proposed algorithm. The proposed framework is compared with five state-of-the-art saliency detection methods: SR [4], AC [5], FT [6], CA [7], PH [13].

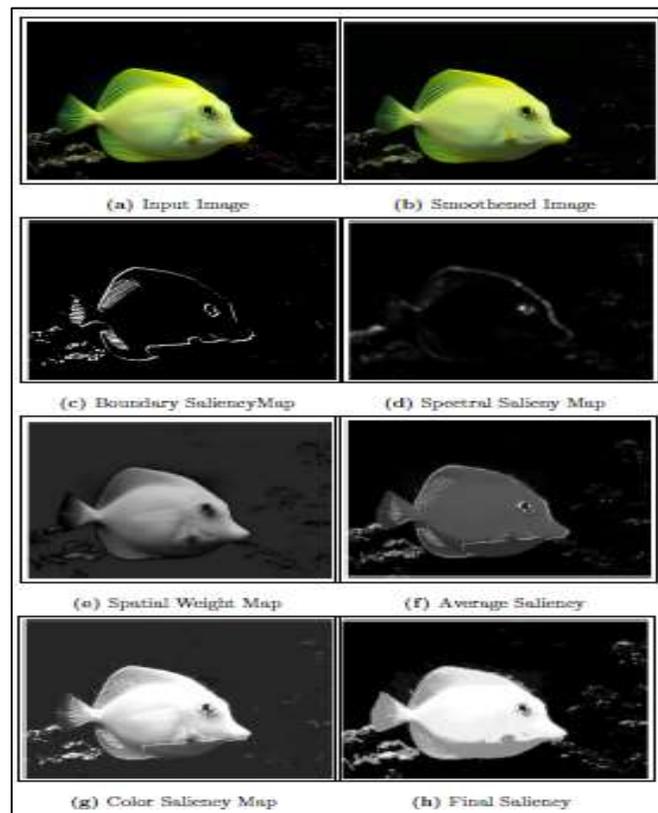


Figure 8: Result 1- One salient object

In phase of L_0 smoothing filter set as $2e-2$ and Parameter that controls the rate is 2. Therefore the smoothing image preserve the edges and provides a perfect smoothed image. In phase of saliency map using boundary weight set as σ

= 0.5. The exponential function is used for visual enhancement. In frequency spatial weight map using a 5 * 5 separable binomial kernel. The parameters used in the system, colour components c are 0.01, 5 respectively.

In Fig 8 contain only one salient object while Fig. 9 contain two salient objects. Both the cases proposed method can efficiently detect the salient region. Visual comparison of saliency maps are given in Fig 11. It is evident from the figure that proposed method outperforms state-of-the-art methods for salient region detection. The programming and simulation of the processes as well as analysis of the results were implemented in Mat lab. For the other methods namely SR, FT, AC and CA, used the author's implementations.

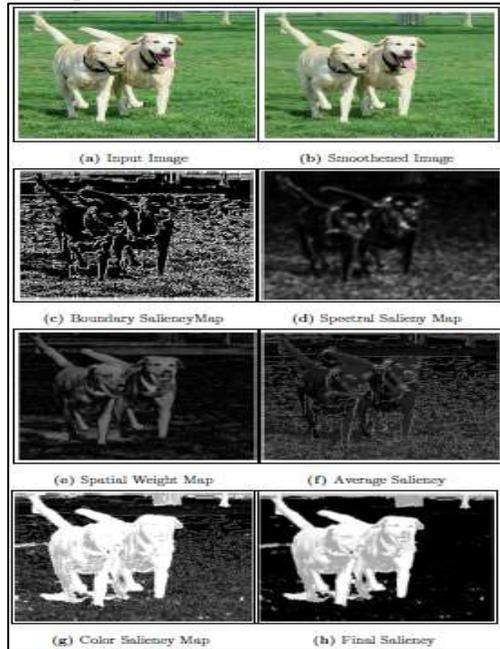


Figure 9: Result 2- Two salient object



Figure 10: Result 3-Multi-salient-objects

A. Experiment on single salient object

First verify the performance of the method for single saliency detection on the MSRA-B dataset. In Fig. 8 contain only one single salient object. So based on color contrast, border information, global contrast saliency map can easily detect a salient object from the image. The Fig 8 shows that it is efficiently output the saliency map and also provide a more clearer saliency map which exclude the non-salient region.

B. Experiment on Two salient objects.

In the Fig.9 contain two salient objects. The existing work is not efficiently output the saliency map. The saliency map was not uniformly highlight the entire region, and also boundaries are not clear. To avoid these problems proposed saliency map consider color spatial variance and border measurement.

Table 1 shows the computational time. The time taken by proposed method is higher than SR and FT but it is lower than CA, PH. That means proposed method provide accurate saliency map within a short time.

Table 1: The time consumed by each saliency detection methods

Method	Time in seconds
SR[4]	0.013
CA[7]	33.662
AC[5]	0.042
FT[6]	0.045
PH[13]	1.13
Proposed method	0.058

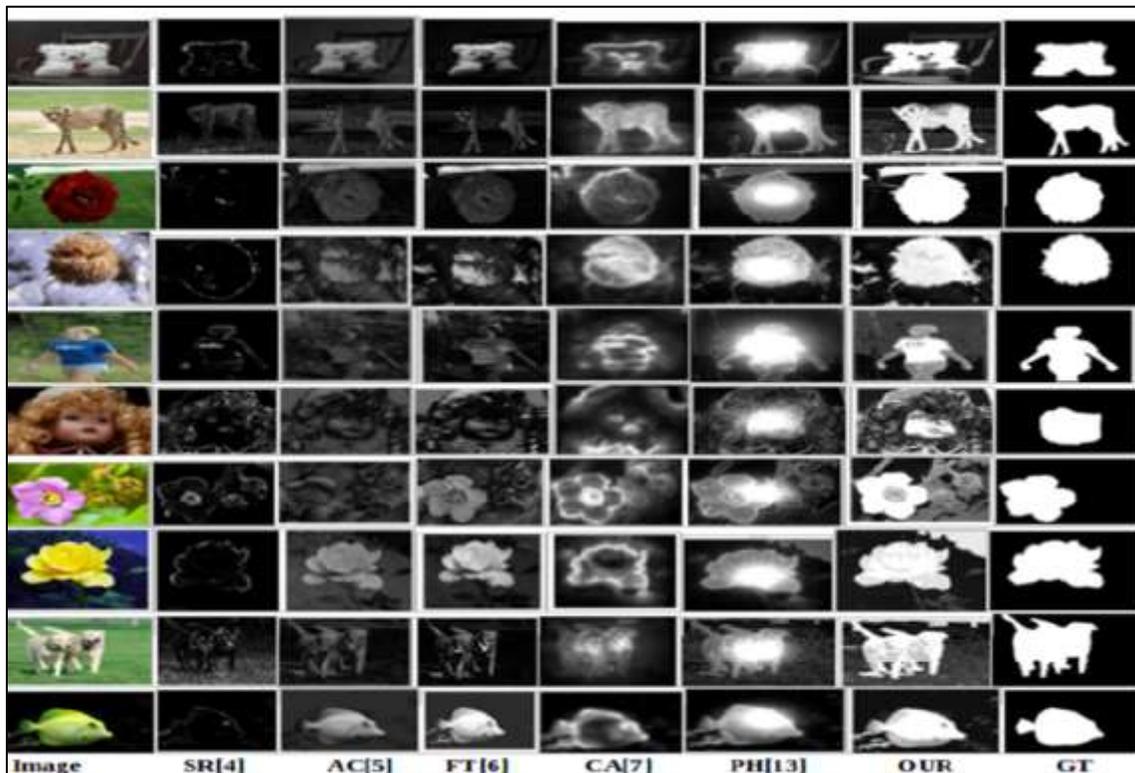


Figure 11: Saliency maps computed by different state-of-the-art methods and with proposed method

C. Experiment on Multi salient objects.

It is very difficult to detect salient objects in particular image when it contain more than two objects. The saliency map creation is very rare in the case of previous works. If it occur, then wide range of problems like ill-defined boundaries, not uniformly highlight the entire salient region etc. happens. The proposed system is efficient to handle multi saliency in different kinds of images. The Fig.10 shows experimental result of multi saliency Table 1 SR shows much lesser time compared to our proposed system but SR is a basic default method which gives more blurred and less quality images. So comparing to those our proposed method gives more efficient and clear images with 0.058 seconds.

CONCLUSION

This project proposes a novel framework for salient region detection method by integrating four features namely boundary information, global contrast, frequency weight map and color spatial variance. Using this proposed method, it can detect multi salient objects and detect the salient region if the object is in border of the image, centre of the image, or even far from the centre of the image. L0 smoothing filter reduce noise and other redundant information and increase both the accuracy and efficiency of saliency detection. Furthermore saliency map uniformly highlight whole salient regions. Emphasize the large salient object. Provide well-defined boundaries of salient objects. Efficiently output full resolution saliency map. Disregard high frequencies. Our framework was tested on MSRA dataset. Experimental result shows that the proposed method produce better performance compared to the state-of-the-art of methods.

REFERENCES

- [1] Achanta.R, Estrada.F, Wils.P, and Ssstrunk. S, "Salient region detection and segmentation, in Computer Vision Systems" (Lecture Notes in Computer Science), vol. 5008.New York, NY, USA: Springer-Verlag, 2008, pp. 6675.
- [2] Achanta. R. Hemami . S, Estrada. F, and Susstrunk S., "Frequency-tuned salient region detection", in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp.1597160
- [3] Cheng. M. M, Zhang.G. X, Mitra .N. J, Huang .X., and Hu S. M, " Global contrast based salient region detection", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 409416.
- [4] Goferman S, Lihi Z. M., and Tal. A, "Context-aware saliency detection", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 23762383.
- [5] Hou.X , and Zang.L, Saliency detection: "A spectral residual approach", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2007, pp. 18.
- [6] Harel J, Koch C, and Perona P, "Graph-based visual saliency, in Advance in Neural Information Processing Systems", vol. 19. New York, NY, USA: Springer-Verlag, 2006, pp. 545552.
- [7] Itti L, Koch C, and Niebur E, "A model of saliency-based visual attention for rapid scene analysis", IEEE Trans. Pattern Anal. Mach. Intell, vol. 20, no. 11, pp. 12541259, Nov. 1988
- [8] Jiang .P. et al, "Salient Region Detection by UFO: Uniqueness, Focusness and Objectness", In Proc. ICCV, 2013
- [9] Liu .T, Yuan. Z, Sun J, Wang J, Zheng N, Tang X, and Shum H, "Learning to detect a salient object", IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 2, pp. 353367, Feb. 2011.
- [10] Ma Y.-F, and Zhang H. J, "Contrast-based image attention analysis by using fuzzy growing", in Proc. 11th Int. Conf. Multimedia, 2003, pp. 374381.
- [11] Perazzi F, Krahenbuhl P, Pritch Y, and Hornung A, "Saliency filters: Contrast based filtering for salient region detection", in CVPR, 2012, pp. 733740K.
- [12] Wu P. H, Chen, C. C, Ding. J. J., Hsu .C. Y. and Huang. Y. W., "Salient Region Detection Improved by Principle Component Analysis and Boundary Information", Image Processing, IEEE Transactions on, 2013, 22(9): 3614-3624.
- [13] Xu L, Lu C, Xu Y, and Jia J, "Image smoothing via L0 gradient minimization", ACM Trans. Graph., vol. 30, no. 5, p. 174, 2011. 2000.
- [14] Yafeng Zheng, Qiaorong Zhang, HuiminXiao, "Salient region detection based on automatic feature selection", Proceedings on Third International Symposium on Intelligent Information Technology Application, pp.232.
- [15] <http://mmcheng.net/msra10k/>