

An Integrated Approach for Image Inpainting Based On Saliency Detection

Aswathy S ¹, Unnikrishnan A S ², Santhosh B S ²

P.G.Student, Department of Electronics & Communication Engineering, SCTCE Trivandrum, Kerala, India¹

Assistant Professor, Department of Electronics & Communication Engineering, CE Perumon, Kerala, India²

Assistant Professor, Department of Electronics & Communication Engineering, CE Perumon, Kerala, India²

ABSTRACT: Salient object detection has been a rather hot research topic recently, due to its potential applications in image compression, scene classification, image registration and so forth. The overwhelming majority of existing computational models are designed based on computer vision techniques by using lots of image cues and priors. A set of novel features including multi-scale contrast, center surround histogram and color spatial distribution to describe a salient object locally, regionally and globally is proposed. In the proposed method an integration approach is used to detect salient region based on three principles which include color contrast in a global context, spatially compact color distribution and multi-scale image abstraction. Based on the mentioned principles, region saliency analysis approach can be formulated in a unified framework. The method is fast, easy to implement and generates high quality saliency maps of the same size and resolution as the input image when compared with pixel level saliency method. Further exemplar based inpainting method is used to fill the missing information in the detected salient region. The PSNR for both region level and pixel level saliency were compared and region level saliency found to produce a better result

KEYWORDS: Saliency, Pyramid, Spatial distribution, Inpainting.

I. INTRODUCTION

Saliency is the state of quality by which it stands out relative to its neighbour. It arises from contrast between object and its neighborhood. Human pays unequal attention to what is seen in the world. When looking at some images people are usually attracted by some particular objects within the images while other seems uninteresting for them. This ability of withdrawing from something in order to deal effectively with others is called attention. Detecting these areas of attention is called saliency detection. Saliency object detection is commonly interpreted in computer vision as a process that includes two stages, detecting the most salient object and segmenting the accurate boundary of that object. In visual saliency, visual attention is guided by two components. The former one is the bottom up approach which draws attention to places that contain outliers and conspicuities in image features. The latter being top down approach will guide the attention in a task dependent manner. In bottom up saliency model [1] salient region detection is based upon biological approach which uses centre surround differences among multiscale image features such as color, intensity and orientation. In center-surround colour difference [2] a fixed neighbourhood for each pixel is directly computed. They classify the saliency as three levels: attended view, attended areas and attended points. A spectral residual approach [3] relies on frequency domain processing. In colour histogram model [4] that detects the salient region in images use only L-channel of colour so as to improve computational efficiency. In color histogram model [5] L channel of color is used to improve computational efficiency. In Frequency Tuned method [6] saliency is defined as the distance between pixel and overall mean image color.

II. SYSTEM DESCRIPTION AND IMPLEMENTATION

The system describes an integration approach based on image pyramid and region property considering multiscale representation and color feature of image. Following aspects had been considered in the paper. Image layer generation in which image pyramid and graph based super pixel generation is exploited to extract image layer. Single layer saliency cues uses location prior considering color distribution. Multi layer integration that make use of navely averaging. The block diagram of the system description is shown in Fig 1. The method proposed in the salient region detection

Department of Electrical & Electronics Engineering, College of Engineering Perumon, Kollam, Kerala – 691601, India

approach mainly consists of multiscale image abstraction and feature extraction, salient region detection in a single scale, saliency map fusion and reinforcement under multiple scales and inpainting the salient region that has detected.

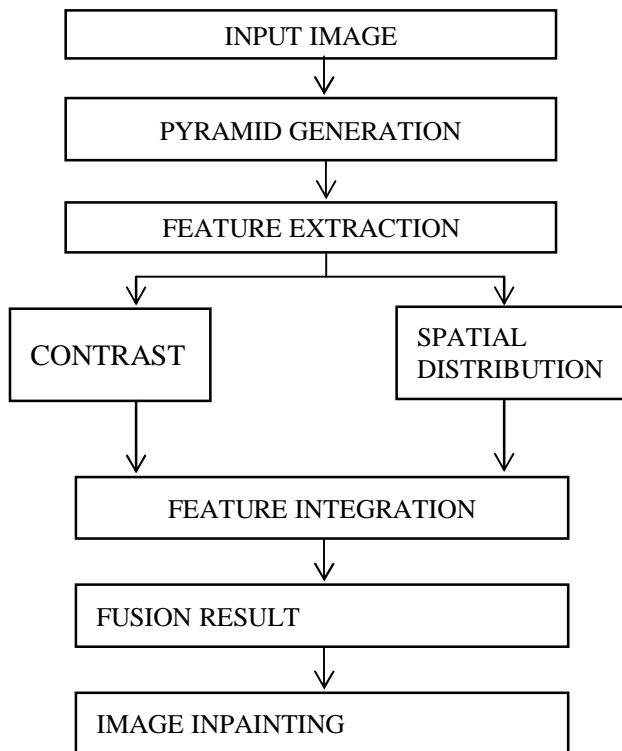


Fig.1 Block diagram of the system

1) Multiscale Image Abstraction and Feature Extraction

Pixel level saliency analysis approach in [7] which cannot handle the natural image with complex background. The above problem is solved by using multi scale image level representation. Here image is divided into non overlapped regions and the saliency analysis is based upon these regions.

1.1) Pyramid Generation

Eventhough convolution deals with most image analysis systems, it creates problem when working with multi resolution image representation. The image pyramid is a data structure designed to support efficient scaled convolution through reduced image representation. It consists of a sequence of copies of an original image in which both sample density and resolution are decreased in regular steps. Reduced resolution level of pyramid is obtained through highly efficient iterative algorithm. Zero level of pyramid G_0 is equal to the original image. This image is low pass filtered and subsampled by a particular factor to obtain the next pyramid level. The process repeats to generate remaining pyramid levels. Thus the process can be considered as a standard reduce operation and the general expression for l^{th} level of a pyramid is given below

$$G_i = REDUCE [G_{l-1}] \quad (1)$$

A Gaussian pyramid is used for constructing image pyramid. It is constructed by convolving the original image with a set of Gaussian like weighting function and then subsampling. The image pyramid thus provides a convenient multi resolution format that mirrors the multiple scales of processing in the human visual system.

1.2) Graph Based Image Segmentation

Let $G = (V, E)$ be an undirected graph with vertices $v_i \in V$, set of elements to be segmented and edges $(v_i, v_j) \in E$ corresponding to pairs of neighboring vertices. Each edge has a corresponding weights which is a non-negative measure of dissimilarity between two pixel connected by that edge. In graph based approach a segmentation S is a portion of V into components $C \in S$ corresponds to a connected components in a graph. The edges between two vertices in the same component should have relatively low weights and edges between vertices in different components should have higher weights.

2) Salient Region Detection in Single Scale

To detect visual salient part of images non overlapped regions were introduced in the saliency analysis task that should be represented by reasonable features like color, edge, texture, size, shape and so on. The important clues in saliency detection task such as region color contrast and region color spatial distribution had been extracted.

2.1) Region Color Contrast

In the work proposed on [6],[8] global color contrast is used for salient region detection that separates large scale objects from its surroundings. as this method does not consider spatial relationship of pixel in images the final saliency map can only highlight part of salient object. This is resolved by using Region Color (RC) contrast based approach which considers spatial relationship among non-overlapped regions. Region contrast is a contrast analysis method to integrate spatial relationship into region level contrast computation. Contrast analysis method is used to overcome the computation expensive while introducing spatial relationship for computing pixel level contrast. In RC first segment the input image into region and then compute color contrast at region level using color histogram. Define saliency for each region as weighted sum of region contrast to all other regions in the image. Finally set the weight according to the spatial distance with farther regions being assigned smaller weights. For salient analysis Lab color space is used that utilizes average region color to represent a region. The color of region R_i and the region center $C(R_i)$ in the m^{th} layer of image pyramid is represented by:

$$F(R_i) = \frac{\sum P_f}{\text{Num}(R_i)} \quad (2)$$

where P is the pixel in the region R_i whereas P_f represents the coordinate vector of pixel and $\text{Num}(R_i)$ denotes the number of pixel in region R_i . These non-overlapped regions are introduced in image saliency task to reduce the computational complexity. By integrating the element of color contrast, spatial distance and the size of region, the saliency of the region is defined as

$$S_{color} = \sum \psi(R_i, R_k) * \phi(R_i, R_k) * \text{Num}(R_k) \quad (3)$$

where

$$\psi(R_i, R_k) = 1 - \text{Dist Positions}(R_i, R_k)$$

and

$$\phi(R_i, R_k) = \| F(R_i) - F(R_k) \|$$

Department of Electrical & Electronics Engineering, College of Engineering Perumon, Kollam, Kerala – 691601, India

Num(R_k) is the no: of pixel in the region R_k and $\psi(R_i, R_k)$ is the distance between R_i and R_k and $\phi(R_i, R_k)$ represents the color distance between R_i and R_k . Dist positions (R_i, R_k) denotes the spatial distance between R_i and R_k .

2.2) Spatial Distribution of Color

By only calculating the color contrast term natural image with complex texture cannot be analyzed. Based on image pixel, authors in [9] proposed a novel salient feature which exploits the spatial distribution of color. Initially all region colors are represented by Gaussian mixture model (GMM) $\{\pi_c, \mu_c, \Sigma_c\} = 1$, where $\{\pi_c, \mu_c, \Sigma_c\}$ are the weight, the mean color and the covariance matrix of the c^{th} Gaussian component respectively. K-mean clustering is used to initialize the parameters of GMMs and exploit expectation maximization (EM) algorithm to train the GMMs. A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. (EM) is a well-established maximum likelihood algorithm for fitting a mixture model to a set of training data. Estimation step compute the expected value of latent variable and maximization step updates the parameter of model based on latent variable. Probability of a region belonging to the c^{th} clustered component is given by,

$$P(c|R_i^m) = \frac{\pi_c N(R_i^m | \mu_c, \Sigma_c)}{\sum_c \pi_c N(R_i^m | \mu_c, \Sigma_c)} \quad (4)$$

The horizontal spatial variance $D_x(c)$ of the c^{th} clustered component is given by

$$D_x(c) = \frac{1}{\Gamma_c} \sum_{i=1}^{k_m} P(c|R_i^m) * |C_x(R_i^m) - \theta_x(c)|^2 \quad (5)$$

$$\theta_x(c) = \frac{1}{\Gamma_c} \sum_{i=1}^{k_m} P(c|R_i^m) * C_x(R_i^m) \quad (6)$$

Similarly vertical spatial variance $D_y(c)$ can be defined in the same way and thus the total spatial variance of the c^{th} clustered component is defined as

$$D_{full}(c) = D_x(c) + D_y(c) \quad (7)$$

High variance indicates that the region color distributed in the whole image and it is less likely belonging to the salient object. Based on three variances, the salient features of region color spatial distribution is defined by,

$$S_{distribution(x)}(R_i^m) = 1 / \exp(\sum_{c=1}^6 P(c|R_i^m) * D_x(c)) \quad (8)$$

$$S_{distribution(y)}(R_i^m) = 1 / \exp(\sum_{c=1}^6 P(c|R_i^m) * D_y(c)) \quad (9)$$

$$S_{distribution(full)}(R_i^m) = 1 / \exp(\sum_{c=1}^6 P(c|R_i^m) * D_{full}(c)) \quad (10)$$

If spatial distribution features are properly used it will generate a high quality saliency detection result.

3) Feature Integration

In feature integration theory features are registered early, automatically, and in parallel, while objects are identified separately" and at a later stage in processing. The theory has been one of the most influential [psychological models](#) of human visual [attention](#). It has two stages pre-attentive and focused attention stage. In pre-attentive stage the object is analyzed for details such as shape, color, orientation and movement and is processed in different areas of the brain. During second stage the individual features of an object are combining to perceive the whole object. To combine the individual features attention is required and a master map location is used to select those features. Authors in [10] found that using linear combination methods good features becomes worst after combination. Thus a nonlinear approach is adopted to uniformly highlight the salient objects and thereby reduce the influence of complex background texture. The fusion result of salient feature is given by,

$$Saliency(R_i^m) = S_{color}(R_i^m) * (\eta + S_{distribution}(R_i^m)) \quad (11)$$

4) Saliency Map Fusion and Reinforcement

The **Saliency Map** is a topographically arranged map that represents visual saliency of a corresponding visual scene. They combine different visual features into a single topographically oriented map and integrate the normalized information from the individual feature maps into a global measure of conspicuity. Saliency map was designed for covert selective attention. Here most salient location is obtained by computing the maximum position established in the topographical map of saliency by Winner take all mechanism. Thus combining image pyramid property difference between salient region and background region was increased by suppressing background region. For generating saliency map, spatial scales are created using Open CV function. Saliency map fusion result is given by,

$$MSaliency(I) = \bigoplus_{m=1}^M Saliency(I^m) \quad (12)$$

where I is the input image, I^m is the saliency detection result of the mth layer and M*Saliency (I) represents integration result under multiple scales.

In reinforcement a new image center is defined as

$$Saliency X = \frac{\sum_{i=1}^N P_i(x) * MSaliency(P_i)}{\sum_{i=1}^N MSaliency(P_i)} \quad (13)$$

$$Saliency Y = \frac{\sum_{i=1}^N P_i(y) * MSaliency(P_i)}{\sum_{i=1}^N MSaliency(P_i)} \quad (14)$$

where N is the number of pixels in the image, p_i is the ith pixel combining the weighted salient image center. The final pixel level saliency is given by,

$$MRS(I_p) = MSaliency(I_p) * \exp\left(-\frac{D_p^2}{\sigma^2}\right) \quad (15)$$

D_p is the Euclidian distance between pixel I_p and the weighted salient image center.

5) Inpainting of Salient Object

Inpainting is the technique used to reconstruct lost or deteriorated parts of images and videos.. Exemplar based inpainting, [1] one of the robust inpainting technique is used to fill the missing portion in a salient region In inpainting the missing portion is being filled and usually the missing portion is selected manually. In this work saliency generated region is the missing portion and hence the computational time for obtaining the target region in the mentioned inpainting technique is reduced. The algorithm explaining the inpainting technique is mentioned in Fig 2

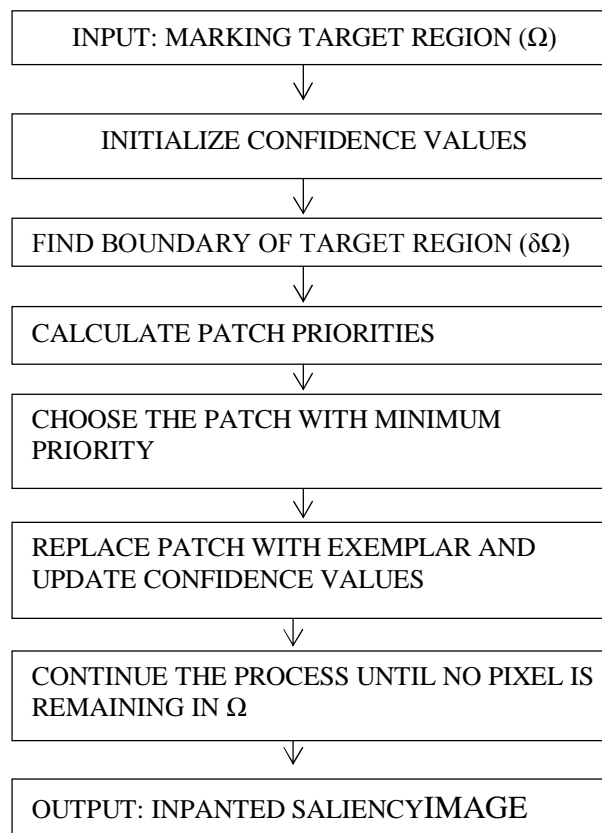


Fig.2 Inpainting Algorithm

III.RESULTS AND DISCUSSION

In order to validate the performance of the proposed method different image of resolution 512×512 pixels shown in Fig.3 is used. The experiment is implemented in MATLAB platform



Fig.3 Test images obtained from dataset MSRA 5000

Region level saliency for the test images are shown in Fig 4



Fig.4 Region level saliency for tested image

Exemplar based inpainting done in bungee image is shown in detail from Fig. 5



Fig. 5.Original Image

The pyramid layer generation of the input (bungee) image is shown in Fig.6

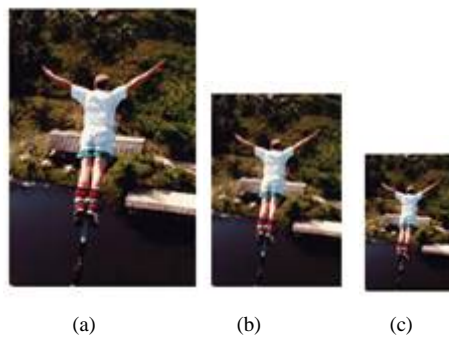


Fig.6 Generated pyramid for three levels for Image 1 (a) Level 1 (b) Level 2 (c) Level 3

The Region Color contrast for multilayer image is shown in Fig.7

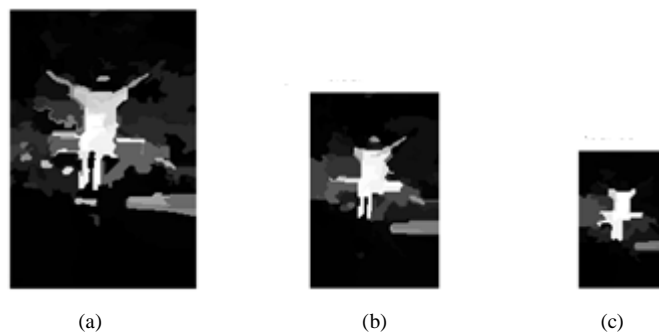


Fig.7 Region Color contrast for multilayer Image 1 (a) Level 1 (b) Level 2 (c) Level 3

The spatial distribution for multilayer image is shown in Fig.8

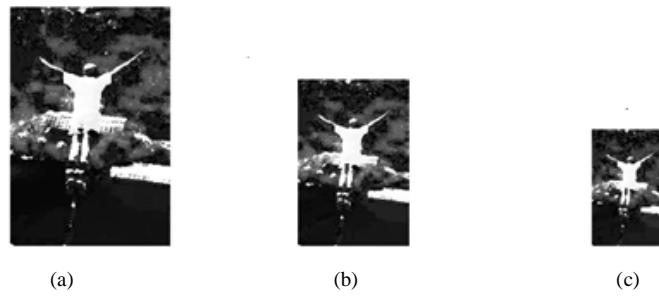


Fig.8 Spatial distribution for multilayer Image 1 (a) Level 1 (b) Level 2 (c) Level 3

The Feature Integration Result is shown in Fig.9.

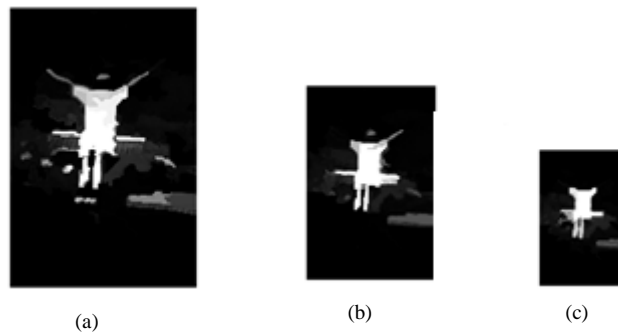


Fig.9 Feature Integration Result for Image 1 (a) Level 1 (b) Level 2 (c) Level 3

The Fusion Result for bungee Image is shown in Fig.10.



Fig.10. Region level Saliency Fusion result

Department of Electrical & Electronics Engineering, College of Engineering Perumon, Kollam, Kerala – 691601, India

The target region and segmentation map for inpainting is taken from the fusion result and is shown in Fig.11.

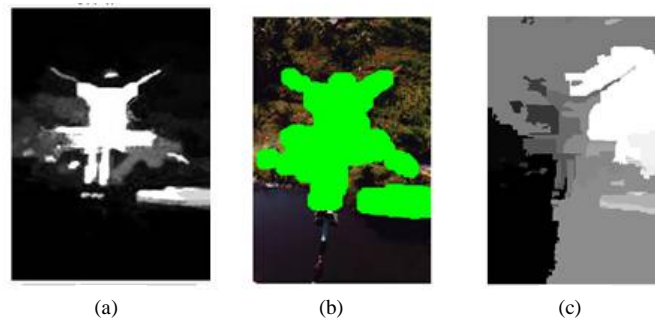


Fig.11. (a) Saliency region, (b) Target selection (c) Segmentation map

The inpainted result for bungee image after exemplar based inpainting technique is shown in Fig. 12



Fig.12 Inpainted Image

The comparison of Peak Signal to Noise Ratio(PSNR) of region level saliency and pixel level saliency after inpainting are shown in TABLE I.

SALIENCY TECHNIQUE	PSNR(dB)
REGION LEVEL SALIENCY	37.7090
PIXEL LEVEL SALIENCY	32.4571

TABLE 1 COMPARISON OF PSNR (dB) FOR REGION LEVEL SALIENCY AND PIXEL LEVEL SALIENCY

IV. CONCLUSION

In this approach initially image pyramid is build and the two salient feature of the image pyramid in each layer is calculated. Salient features are combined using nonlinear approach. Saliency map of single scale is generated and fuse the map of multiple levels and combine salient image center to obtain region level saliency. Further Exemplar based image inpainting was applied to both pixel level saliency and region level saliency. The Peak Signal to Noise Ratio (PSNR) was found to be better for region level saliency technique

REFERENCES

- [1] O. Le Meur, J. Gautier, and C. Guillemot, "Exemplar-based inpainting based on local geometry," in Proc. 18th IEEE Int. Conf. Image Process., Sep. 2011, pp. 3401–3404.
- [2] Ma, Y.F., Zhang, H.J., "Contrast-based image attention analysis by using fuzzy growing," ACM Int. Conf. Multimedia, USA, 2003, pp. 374–381.
- [3] Hou, X.D., Zhang, L.Q., "Saliency detection: a spectral residual Approach," IEEE Conf. Computer Vision and Pattern Recognition, USA, 2007, pp. 1–8.
- [4] Zhai, Y., Shah, M., "Visual attention detection in video sequences using spatiotemporal cues," ACM Int. Conf. Multimedia, USA, 2006, pp. 815–824.
- [5] Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X., Hu, S.M., "Global contrast based salient region detection," IEEE Conf. Computer Vision and Pattern Recognition, USA, 2011, pp. 409–416.
- [6] Achanta, R., Hemami, S, Estrada, F, Susstrunk, S, "Frequency-tuned salient region detection". IEEE Conf. Computer Vision and Pattern Recognition, USA, 2009, pp. 1597–1604.
- [7] Ren, X., Jitendra, M., "Learning a classification model for segmentation," IEEE Conf. Computer Vision (ICCV), China, 2003, pp. 10–17.
- [8] Zhai, Y., Shah, M., "Visual attention detection in video sequences using spatiotemporal cues," ACM Int. Conf. Multimedia, USA, 2006, pp. 815–824.
- [9] Liu, T., Yuan, Z., Sun, J., Wang, J., Zheng, N., Tang, X., Shum, H.Y., "Learning to detect a salient object," IEEE Trans. Pattern Anal. Mach. Intell., 2011, 33, (2), pp. 353–367.
- [10] Gopalakrishnan, V., Hu, Y., Rajan, D., "Salient region detection by modeling distributions of color and orientation," IEEE Trans. Multimedia, 2009, 11, (5), pp. 892–905.
- [11] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk. "Salient region detection and segmentation", In ICVS, pages 66–75. Springer, 2008.
- [12] D.J.Heeger And J.R. Bergen. Pyramid-Based Texture Analysis/Synthesis. In proceedings Of ACM Conf. Comp. Graphics (Siggraph), Volume 29, Pages 229{233, Los Angeles, Ca, 1995. ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 12, December 2013 All Rights Reserved © 2013 IJARCET 3214