IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

Detection of Salient Object in a Static Image

K.P. Kamble[1], M.S. Ghute[2], Mohsina Anjum[3]

Assistant Prof, Electronics and Telecommunication, YCCE, Nagpur, India[1,2]

Assistant Prof, Electronics and Telecommunication, Anjuman COE, Nagpur, India [1]

Abstract: To use the human visual attention concept to maximum benefit in the domain of image processing and thus extract important or the most prominent object in a static image. The research up till now in this area has been limited to indicating just the Focus of Attention (FOA) and not the entire region of interest. The salient object detection problem for static images is used to formulate salient object detection as one which separates a salient object from the background. In this paper a bottom up approach to compute the saliency map analogous to one calculated by human brain. The mathematical model is prepared, based on study of anatomy of physiology of visual system primates. Low level features are extracted based on color, contrast and intensity. After normalization and linear combination, a master map or a saliency map is computed to represent the saliency of each image pixel. Finally, the image is segmented out from the background. About 30 images are selected from a large database of images from Microsoft Asia which contains a salient object or a distinctive foreground object.

Keywords: salient object, static image, FOA, normalization, segmentation.

I. INTRODUCTION

In electrical engineering and computer science, image processing is any form of signal processing for which the input is an image such as photographs or frames of video; the output of image processing can be an image or a set of characteristics or parameters related to the image. Most image processing techniques involved treating the image as 2D signal and applying standard signal processing techniques to Visual attention is supposed to be driven by low-level stimulus in the scene, such as intensity, contrast, and motion. The visual saliency approach is one of the best approach.Visual saliency approach as defined by Koch and Ullman in 1985 said that the brain calculates an explicit saliency map of the visual world. Their definition of saliency relied on centre-surround principles. Points in the visual scene are salient if they differ from their neighbours. Visual saliency is based upon a model that simulates the elements of a visual scene that are likely to attract the attention of human observers. Given an image or video sequence, the model computes a saliency map, which topographically encodes for conspicuity (or ``saliency'') at every location in the visual input. In this paper, the object is extracted from an image which draws maximum visual attention of the observer. The final result obtained will be the prominent object in the image while the background pixels would be assigned a low score (black background).

II. ALGORITHM

In algorithm consider an image to detect a salient object from the image. This salient object will be extracted from the image using the feature maps. Here three methods to determine the feature maps are considered. The methods are Thresholding, multi-scale contrast, colour spatial distribution. The results obtained from all these three methods are combined to form a saliency map. This saliency map will provide the required region of interest in order to detect the salient object.

The following operations are performed on the input image in order to obtain the required feature maps.
(1)Thresholding: OSTU's method is used for Thresholding method which uses histograms for its computational analysis which is easily obtainable in 1D array. This method maximizes between class variance.

(2)Multi-scale contrast: A contrast map is calculated at every level of dyadic pyramid and final feature map is computed by across- scale addition of contrast maps.

(3)Colour spatial distribution: This method finds the Euclidean distance between the Lab pixel vectors in a Gaussian filtered image with the average Lab vector for the input image.

Three feature maps Thresholding, multi-scale contrast, colour spatial distribution are normalized and combined to form a saliency map. This saliency map detects the most prominent object in an image. The above mentioned feature maps are explained in detail in the following sections.

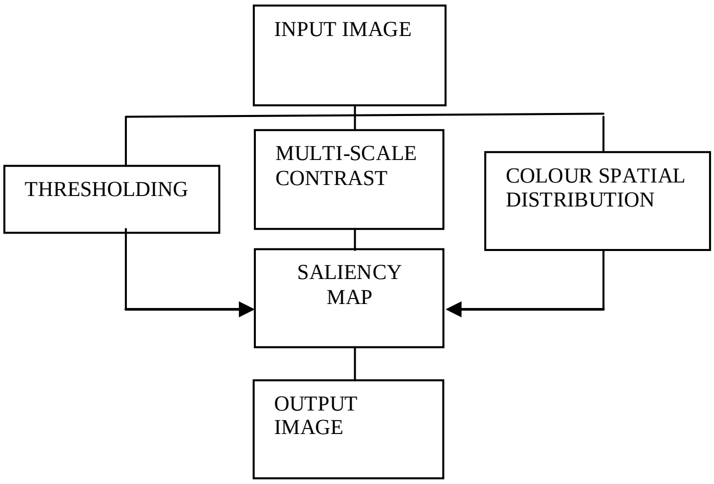

Fig. 1. Flow chart for Proposed Algorithm

1.Thresholding

A feature map based on intensity values is calculated by thresholding image by Otsu Method. This method assumes that the image to be threshold contains two classes of pixels (e.g. foreground and background) and then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal.

In Otsu's method , the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t)$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variances of these classes.

Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance:

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = w_1(t)w_2(t)\left[\mu_1(t) - \mu_2(t)\right]^2$$

Which is expressed in terms of class probabilities w_i and class means μ_i which in turn can be updated iteratively.

Algorithm of thresholding:

1. Compute histogram and probabilities of each intensity level
2. Set up initial $w_i(0)$ and $\mu_i(0)$
3. Step through all possible thresholds $t = 1,2, \ldots$ maximum intensity

a. Update w_i and μ_i

b. Compute $\sigma_b^2(t)$

4. Desired threshold corresponds to the maximum $\sigma_b^2(t)$

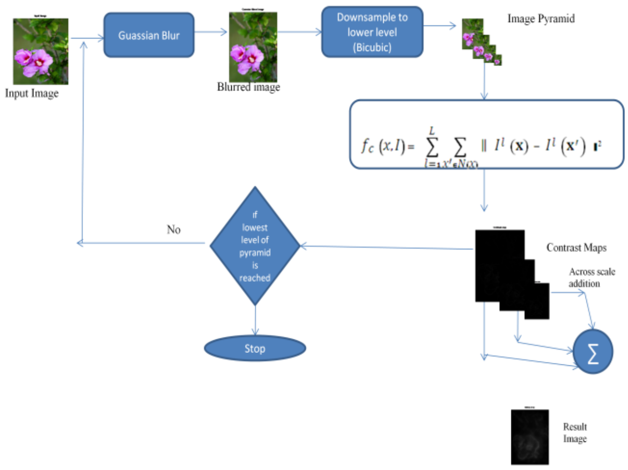

Fig.2 Flow Chart of MSCM

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

IMAGE PYRAMIDS

The data structure used to represent image information can be critical to the successful completion of an image processing task. One structure that has attracted considerable attention is the image pyramid. This consists of a set of lowpass or bandpass copies of an image, each representing pattern information of a different scale. A variety of pyramid methods is developed for image data compression, enhancement, analysis and graphics. There are two main types of pyramids; lowpass pyramids and bandpass pyramids. A lowpass pyramid is generated by first smoothing the image with an appropriate smoothing filter and then sub-sampling the smoothed image, usually by a factor of two along each coordinate direction. This smoothed image is then subjected to the same processing, resulting in a yet smaller image.

As this process proceeds, the result will be a set of gradually more smoothed images, where in addition the spatial sampling density decreases level by level. If illustrated graphically, this multi-scale representation will look like a pyramid, from which the name has been obtained. A bandpass pyramid is obtained by forming the difference between adjacent levels in a pyramid, where in addition some kind of interpolation is performed between representations at adjacent levels of resolution, to enable the computation of pixel-wise differences.

In the early days of computer vision, pyramids were used as the main type of multi-scale representation for computing multi-scale image features from real-world image data. Today, this role has been taken over by scale space representation, motivated by the more solid theoretical foundation, the ability to decouple the subsampling stage from the multi-scale representation, the more powerful tools for theoretical analysis as well as the ability to compute a representation at any desired scale, thus avoiding the algorithmic problems of relating image representations at different resolution. Nevertheless, pyramids are still frequently used for expressing computationally efficient approximations to scale-space representation.

INTERPOLATION

Generating Image Pyramids require many image resizing. Image resizing is done by estimating neighboring pixels. The method used for estimation of pixels is Interpolation. In the mathematical subfield of numerical analysis, interpolation is a method of constructing new data points within the range of a discrete set of known data points.

In engineering and science one often has a number of data points, as obtained by sampling or experimentation, and tries to construct a function which closely fits those data points. This is called curve fitting or regression analysis. Interpolation is a specific case of curve fitting, in which the function must go exactly through the data points.

A different problem which is closely related to interpolation is the approximation of a complicated function by a simple function. Suppose the function is too complex to evaluate efficiently. Then pick a few known data points from the complicated function, creating a lookup table, and try to interpolate those data points to construct a simpler function. Of course, when using the simple function to calculate new data points , the same result is not received as when using the original function, but depending on the problem domain and the interpolation method used the gain in simplicity might offset the error.

Interpolation as stated earlier is used for estimation of neighboring pixels. Three type of interpolation is generally used in Image Processing.
1. Nearest neighbor Interpolation
2. Bilinear Interpolation
3. Bicubic Interpolation

BICUBIC INTERPOLATION

Bicubic interpolation is an extension of cubic interpolation for interpolating data points on a two dimensional regular grid. The interpolated surface is smoother than corresponding surfaces obtained by bilinear interpolation or nearest-neighbor interpolation. Bicubic interpolation can be accomplished using either Lagrange polynomials, cubic splines, or cubic convolution algorithm.

Colour Spatial distribution

Colour feature map is constructed on the basis of colour features of Image. Study of visual system of primates show that colour which is spread more is least concentrated by primates. Hence the more the colour is spread in image least is its possibility of being colour of Salient object.

Following flowchart shows algorithm for calculating Colour Feature Map.

DOI10.17148/IJARCCE.2017.6737

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

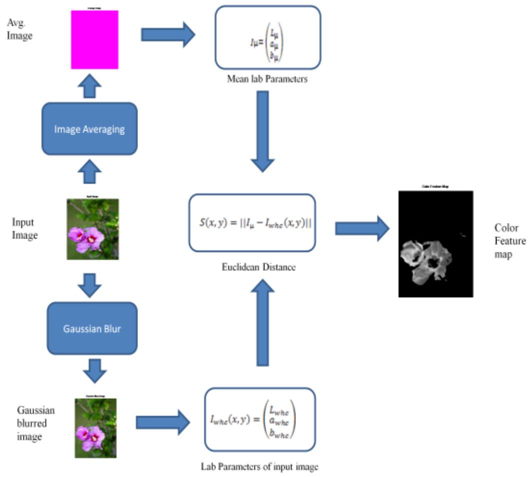

Fig.3 Flow Chart for colour feature map

III. RESULT

Fig 4(a) original image **(b)** blurred image **(c)** colour feature map

Saliency map computation

It is seen that all the three feature map are not taken in equal weights as each feature map do not contribute to the target result with the same efficiency. So with a research based trial and error method we assigned weights to these feature maps so that satisfactory results are obtained.

$$\mathrm{SM} = 0.2\mathrm{N}(\mathrm{MSCM}) + 0.7\mathrm{N}(\mathrm{CM}) + 0.1\mathrm{N}(\mathrm{TM})$$

Where,

SM = final saliency map

N(MSCM) = Normalized value of multi-scale contrast feature map.

N(CM) = Normalized value of colour spatial distribution feature map

N(TM)= Normalized value of threshold feature map.

For this purpose the feature map obtained by Thresholding was assigned (0.1) weight while that of Multi-scale contrast was assigned (0.2) weight. It was seen that the colour spatial distribution provided us with the best of the results hence was given the weight of (0.7).

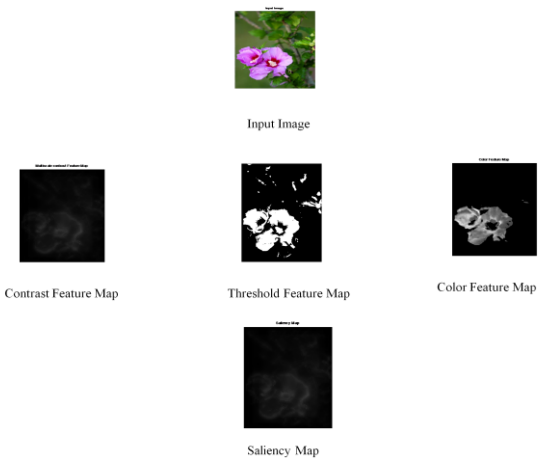

Fig 5 (a) original image (b)contrast feature map(c) threshold feature map **(d)** colour feature map (e)saliency map

Hence the saliency map obtained has the feature maps in combination of 10% of Thresholding feature map, 20% of multi-scale feature map and 70% of colour spatial distribution feature map. The saliency map thus obtained is then multiplied with the original image so that the salient object is distinctly extracted from the image. Thus the result obtained is to extract the most salient object from the image.

IV. CONCLUSION

Bottom up approach of visual attention in primates. Given an input image, this system attempt to predict which location in the image will automatically and unconsciously draw your attention towards them.

In this biologically-inspired system, an input image is decomposed into a set of ``feature maps'' which extract local spatial discontinuities in the modalities of color, intensity and contrast. All feature maps are then combined into a unique scalar ``saliency map'' which encodes for the salience of a location in the scene irrespectively of the particular feature which detected this location as conspicuous.

Feature map obtained by Thresholding was assigned (0.1) weight as it was giving the least efficient result, while that of Multi-scale contrast was assigned (0.2) weight. MSCM provided us with fairly good results with the region of interest being highlighted with high score pixels. It was seen that the colour spatial distribution provided us with the best of the results as the extraction of salient object was almost accurate hence was given the weight of (0.7).

This algorithm is capable of detecting multiple salient objects provided that it is distinct and prominent in the image and is found to be detecting the most salient object in an image with fair accuracy

ACKNOWLEDGMENT

Completing a task is never a one man effort. It is often the result of invaluable contribution of number of individuals in a direct or indirect manner.

The motivating factor for this project was the inspiration to me by my husband **Mr. P.B. Kamble** he has given many valuable suggestions and encouraging generously throughout.

I wish to convey our gratitude to my family member. With their encouragement has been a constant source of inspiration for me.

REFERENCES

[1] Tie Lie, Zejian Yuan, Jian Sun, Jingdong Wang, Nanning Zheng, "Learning to Detect Salient Object" IEEE Trans. on pattern analysis and machine intelligence, vol. X, no. Xx, December 2009

[2] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. on PAMI, vol. 20, no. 11, pp. 1254–1259, 1998

[3]. L. Elazary and L. Itti, "Interesting objects are visually salient," Journal of Vision, vol. 8, pp. 1–15, 2008.

[4] R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk, Frequency-tuned Salient Region Detection, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[5] Rafael C. Gonzalez , Richard E. Woods, Steven L. Eddins "Digital image processing using MATLAB"

BIOGRAPHIES

Mrs. K.P.Kamble holds a B.E. degree in 2002 from Nagpur University and M.E. degree in 2004 from S.R.T.M.U. NandedIn Electronics Engineering. She has presented papers in international conference and international journal. Her research focuses on Digital Image Processing, Digital Signal Processing and Fuzzy Logic and Neural Network

Mrs M.S.Ghute holds a B.E. degree in 2005 and M.Tech. degree in 200 8 Electronics Engineering. She has presented papers in international conference and international journal. Her research focuses on Digital Image Processing, ,Digital Signal Processing.