



A Scrutiny of detecting Salient Objects in Digital Images

ARVINDA SHARMA

M.Tech Scholar,

Global Institutes of Management and Emerging Technology,
Amritsar, Punjab - 143501

Abstract:-This paper presents a review of various object detection techniques. Object Detection will localize the objects in the digital image or in the video. It is beneficial when the object is not in the centre of image or is of size different from the other object's size. Object detection is useful for face detection, detecting objects in cctv camera and in vehicular adhoc networks. This paper explores the various techniques for efficiently extracting the object from digital images ends up with the short comings of the existing techniques.

Keywords-object detection, random forest, frequency, tree sharing and global contrast.

I. INTRODUCTION

Object detection [1] is a process of discovering the important objects in the digital images and videos. Detecting such an object is pre-processing stair of segmentation. There are large numbers of methods that segment the multiple objects the scene. Each method has its own benefits and drawbacks. It will discount the effects of background from image and detect the salient objects and saliency is focus of attention in a scene. Saliency maps [2] compute saliency value for each pixel. The salient object is always different from its neighbouring perspective, placed near the centre of the image and has a sharp closed margin.

Detecting such salient image regions remains a significant goal, as it allows proper image analysis and synthesis. Great achievement has been achieved in few places but the problem remains unanswered in unrestrained places, in especially when objects are placed in capricious places in jumble environment .Problems in discovering the important entities arise hasty object motion, changing appearance patterns of an object and a background, object-to-object occlusions and object-to-background occlusions. Image segmentation has been the subject of active research in computer vision and image processing. A large body of work on geometric active contours (i.e.) active contours implemented via level set methods, has been proposed to address a wide range of image segmentation problems.

Various Salient Object detection techniques are as follows:

- Salient Object detection via random forest
- Sharing of trees among random forest
- Global contrast based salient region detection
- Frequency tuned saliency region detection

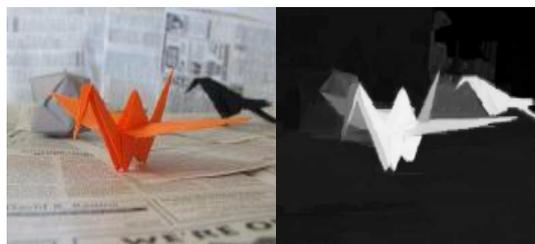


Fig 1: Salient Object Detection [2]

II. ACTIVE CONTOUR MODELS

ACM [3] is to evolve a curve under some constraints to extract the desired object. ACM are also known as snakes. It is a rubber band [15] of any shape that is deform with time and try to get as close as possible to object contour. These are the energy minimizations [15] that will consider only the pixels or patches that lie on the boundary of object. ACMs [3] can be categorized into two types: edge-based models and region-based models. Region-based ACMs have many advantages over edge-based ones. First, region-based models utilize the statistical information inside and outside the contour to control the evolution, which are less sensitive to noise and have better performance for images with weak edges or without edges. Second, they are significantly less sensitive to the location of initial contour and then can efficiently detect the exterior and interior boundaries simultaneously.

III. RANDOM FOREST

Random forest [4] is an ensemble classifier consisted of a number of decision trees. Since the performance of the trees is highly related to the connection among each model in it, the trees are often constructed with some randomization and randomization comes from two points: Sub sampling the training data and each tree are grown with different data, for each internal node, selecting some attributes for split. Besides, each internal node contains a best split of training data.

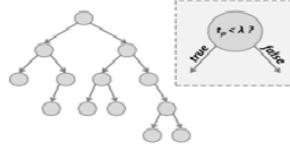


Fig2: Decision trees [4]

IV. TECHNIQUES OF SALIENT OBJECT DETECTION

A. SALIENT OBJECT DETECTION VIA RANDOM FOREST

It is a patch-based approach [5] to detect the salient object through its approximate contour. First, we use the global patch rarity to capture the approximate contour of the salient object by constructing the random forest using the equation [5],

$$t_n(s_n; h_1, h_2) = \begin{cases} p_i \in s_l, & \text{if } d_i(h_1, h_2) \leq \theta_{h_1, h_2} \\ p_i \in s_r, & \text{otherwise} \end{cases}$$

Where s_l and s_r are the patch sets contained in node's left and right child, $\theta_{h_1, h_2} = \frac{1}{|s_n|} \sum_{p_i \in s_n} d_i(h_1, h_2)$, $|s_n|$ is the cardinality of s_n . After the forest is built, we use it to measure the rarities of patches and compute similarities among them, and then detect the contour through equation[5]:

$$cS(p_i) = \frac{1}{\sum_{k=0}^n |L_k|} \cdot w(x_i, x_c),$$

The above equation is used to evaluate the rarity of p_i , where $|L_k|$ is the number of patches contained in L_k . $w(x_i, x_c)$. Through level set Active Contour method, we discover the salient object according to the detected contour C . We partition these patches into two subsets, inside and outside, respectively. The formal definition can be written as [5]:

$$p_i \in \begin{cases} S_{in}, & \text{if } \frac{|p_i \cap C|}{p_i^2} > \lambda, \\ S_{out}, & \text{otherwise,} \end{cases}$$

Where $|p_i \cap C|$ denotes the number of pixels of contained inside the extracted contour, λ is a constant, and its value varies between 0 and 1 according to the dataset. The p_i and p_j are similar patches if they belong to leaf node and then we measure the contrasts between the inner patches and

The outer patches, aiming to suppress the patches of the inner part similar to the outer patches while highlight the outer patches similar to the inner patches. Through the contrasts among patches, our method highlights the whole object uniformly. Finally, we refine the local map using graph-cut based segmentation [5]. This approach performs well not only for the single-object case, but also for the multiple-object case.



Fig3: (a) input image (b) object detection [5]

B. FREQUENCY TUNED SALIENT REGION DETECTION

This method [6] highlights the whole salient objects, by establishing the well-defined borders of salient objects, and eliminating high frequencies arising from noise. In it, let W_{LC} [6] be the low frequency cut-off value and W_{HC} be the high frequency cut-off value. To highlight large salient objects [6], we need to consider very low frequencies from the original Image, i.e. W_{LC} has to be low. In order to have well defined boundaries, we need to retain high frequencies from the original image, i.e. W_{HC} has to be high. However, to avoid noise, coding artifacts, and texture patterns, the highest frequencies need to be disregarded.

Saliency map S as for an image I of width W and height H pixels can thus be formulated as [6]:

$$S(x, y) = |I_U - I_{Whc}(x, y)|$$

Where I_U is the arithmetic mean pixel value of the image and $I_{Whc}(x, y)$ is the Gaussian blurred version of the original image to eliminate fine texture details as well as noise and coding artifacts. The criterion of difference is used as we are interested only in magnitude of the differences. So the method of finding the saliency map S for an image I of width W and height H pixels can thus be formulated to use features of color and luminance; we rewrite it as [6]:

$$S(x, y) = ||I_U - I_{Whc}(x, y)||$$

Fig4 :(a) Original image (b) saliency map [6]

C. GLOBAL CONTRAST BASED SALIENT REGION DETECTION

In Global contrast based saliency [7] the HC method is fast and generates results with full precision, the RC method generates saliency maps at the cost of reduced computational efficiency and at the same time evaluate global contrast differences and spatial coherence. This method is efficient, and generates full resolution saliency maps.



➤ Histogram based contrast

This method is used to define saliency values for image pixels using color statistics of the input image. The saliency value of a pixel I_k in image I is defined [7] as-

$$S(I_k) = \sum_{I_i \in I} D(I_k, I_i)$$

Where $D(I_k, I_i)$ is the color distance metric between pixels I_k and I_i in the L*b*c color space.

The terms with the same color value C_j are grouped together; we get saliency value for each color as [7],

$$S(I_k) = S(c_i) = \sum_{j=1}^n f_j D(c_i, c_j)$$

Where c_i is the color value of pixel I_k , n is the number of distinct pixel colors, and f_j is the probability of pixel color c_j in image I.

➤ Color Space Smoothing

We use a smoothing procedure to reduce noisy results [7], and for this we replace the saliency value of each color by the weighted average of the saliency values of similar colors (measured by L*a*b distance), refine the saliency value of color c by [7],

$$S'(c) = \frac{1}{(m-1)T} \sum_{i=1}^m (T - D(c, c_i)) S(c_i)$$

Where $T = \sum_{i=1}^m (D(c, c_i)) S(c_i)$, is the sum of distances between color c and its m nearest neighbors c_i and the normalization factor comes from $\sum_{i=1}^m (T - D(c, c_i)) = (m-1)T$.

➤ Region based contrast

We first segment the input image into regions using a graph based image segmentation method. Then we build the color histogram for each region.

For a region r_k we compute its saliency value by measuring its color contrast to all other regions in the image [7],

$$S(r_k) = \sum_{r_k \neq r_i} w(r_i) D_r(r_i, r_k)$$

Where $w(r_i)$ is the weight of region r_i and $D_r(\cdot)$ is the color distance metric between the two regions.



Fig5: (a) Input Image (b) Saliency maps (c) Region of interest [7]

D. SHARING OF TREES AMONG RANDOM FOREST

Such method [8] improves the efficiency of the system in testing phase and save memory by reducing the total number of trees. Here the random forest is used for object detection task. In this method, when the trees are shared among different classes of objects, it is called as tree sharing. Then, the weight of each tree needs to be decided. After deciding the tree weights we can remove the less used trees according to the system budget. Finally, the testing result is the combination of the responses from all remaining trees with the decided weights. After that, tree weights are decided as by defining problem such as [8]: Given m concepts and construct n trees for each of them. We denote the j -th tree of the i -th concept t_{ij} . For each concept, we would like to decide the weights of total mn trees.

We denote the weights for i -th concept by a $m \times 1$ vector w^i . We will obtain with rough minimizing a loss function $loss^i(w)$

$$w^i = \arg \min_w loss^i(w)$$

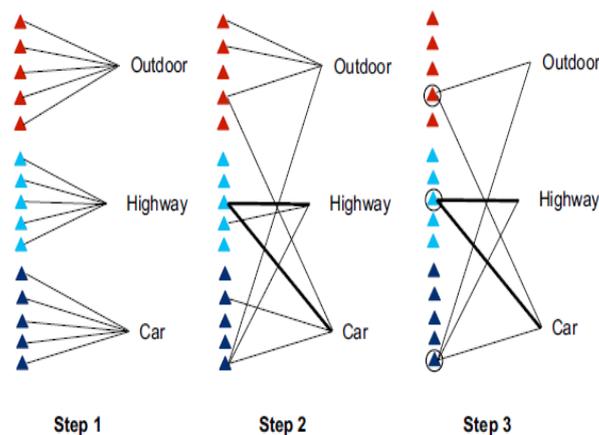
The loss functions is defined as [8]: For the i -th concept, we define the scoring function to be

$$F_w^i(img) = Rsp^i(img) * w,$$

Where $Rsp^i(img)$ a $1 \times mn$ vector is formed by testing on all mn trees, and img is the test image. Since the trees are binary classifiers here, the mn elements are either 0 or 1. Suppose we have p image $I = \{img_1^i, img_2^i, img_3^i, \dots, img_p^i\}$ and their relevance scores $r = \{r_1^i, r_2^i, \dots, r_p^i\}$ or it concept.

Finally the tree is reduced as [8]: To reduce the tree number, we need to measure the significance of each tree from the view of whole system. After deciding the weights, we get w_1, w_2, w_m . We gather these weight vectors to form a $m \times m$ matrix $W = [w_1, w_2, w_m]$. Each row of this matrix is consisted of the weights of each tree on all concepts.

Since the absolute value [8] of the weight can be considered as how important the tree is for a specific concept, we take the L1-norm of each row vector as the significance of each tree. Then, we sort the trees by their significances and leave the most important ones according to the system budget.



Step 1: Construct forests
Step2: Decide the weights
Step 3: Reducing the total tree numbers

Fig 6: Sharing of trees [8]

V. CONCLUSION

Object Detection is a computer technology related to image processing that deals with detecting instances of semantic objects such as humans, cars, buildings in digital images and videos. In this paper we review the existing techniques of salient object detection

The existing methods don't consider many of the factors like complexity of background, some are optimal only for natural scenes, many methods have overhead of reinitializing the energy functions, some have reduce the large number of trees, often results in poor performance i.e. will leave the most important one's from the view of whole system, some may output the object on basis of contrast which may undergo error if noise is present in the images.

So these methods still have difficulties of highlighting the entire object when the inner region of the object is inhomogeneous. To reduce the problems of existing literature a new technique will be proposed in near future that will consider all the weaknesses of various methods and will produce better results.

In near future we can introduce the new level set free active contour method. This method reduces overhead of cost and reinitialisations of energy functions and can include dynamic thresholding method because it remains ineffective even in case of presence of noise.

ACKNOWLEDGMENT

I would like to express my truthful thanks to Dr. Rajiv Mahajan (H.O.D.), Computer Science, Global Institutes of Management and Emerging Technology, Amritsar, Punjab for his guidance, motivation and productive suggestions that help us in the preparation of this paper.

REFERENCES

- [1] Cheng, Ming-Ming, et al. "Global contrast based salient region detection." *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on. IEEE, 2011
- [2] Jiang, Huaizu, et al. "Automatic salient object segmentation based on context and shape prior." *BMVC*. Vol. 3.No. 4. 2011
- [3] Zhang, Kailua, et al. "Active contours with selective local or global segmentation: A new formulation and level set method." *Image and Vision Computing* 28.4 (2010): 668-676.
- [4] Gall, Juergen, NimaRazavi, and Luc Van Gool. "An introduction to random forests for multi-class object detection." *Outdoor and Large-Scale Real-World Scene Analysis*. Springer Berlin Heidelberg, 2012.243-263.
- [5] Du, Shuze, and Shifeng Chen. "Salient Object Detection via Random Forest." (2013): 1-1
- [6] Achanta, Radhakrishna, et al. "Frequency-tuned salient region detection." *Computer Vision and Pattern Recognition, 2009.CVPR 2009*.IEEE Conference on.IEEE, 2009.
- [7] Cheng, Ming-Ming, et al. "Global contrast based salient region detection." *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on.IEEE, 2011.
- [8] Chiu, Tzu-Hsuan, et al. "Sharing the trees among random forests for effective and efficient concept detection." *Multimedia Signal Processing (MMSp)*, 2012 IEEE 14th International Workshop on. IEEE, 2012.
- [9] Yanulevskaya, Victoria, Jasper Uijlings, and Jan-Mark Geusebroek. "Salient object detection: From pixels to segments." *Image and Vision Computing* 31.1 (2013): 31-42.
- [10] Prasad, Dilip K. "Survey of the problem of object detection in real images." *International Journal of Image Processing (IJIP)* 6.6 (2012): 441.
- [11] Zhang, Chao, et al. "Contour detection via random forest." *Pattern Recognition (ICPR)*, 2012 21st International Conference on. IEEE, 2012
- [12] Ksantini, Riadh, BoubakeurBoufama, and Sara Memar. "A new efficient active contour model without local initializations for salient object detection." *EURASIP Journal on Image and Video Processing* 2013.1 (2013): 1-13.
- [13] Chen, Qiang, et al. "Efficient maximum appearance search for large-scale object detection." *Computer Vision and Pattern Recognition (CVPR)*, 2013 IEEE Conference on.IEEE, 2013.
- [14] Feng, Jie, et al. "Salient object detection by composition." *Computer Vision (ICCV)*, 2011 IEEE International Conference on. IEEE, 2011.
- [15]]http://en.wikipedia.org/wiki/Active_contour_model