International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 2 Issue: 10
ISSN: 2321-8169
3061 – 3068

“Ameliorating the Saliency of Objects in digital Images”

Arvinda Sharma
M. Tech Scholar: Computer Science & Engineering Dept
Global Institutes of Management& Emerging Technology
Amritsar, Punjab, India.
arvinda11.rampal@yahoo.com

Dr. Rajiv Mahajan
Professor: Computer Science & Engineering Dept
Global Institutes of Management& Emerging Technology
Amritsar, Punjab, India
rajivmahajan08@gmail.com

Abstract**--**Visual saliency is the capability of a revelation system of human or machine to choose a sub part of information for further processing. The mechanism describes here serves as a filter to pick only the interesting information related to given errands and the tasks while ignoring inappropriate information. This work uses Random Forest to know the similarity between the image patches, apply active contour model to get the approximate contour and do dynamic thresholding segmentation .The results we get consists of many small artifacts, so to remove the low level details and to obtain the more smoothness we apply gradient minimization technique.

Keywords*- Saliency, Random Forest, Gradient minimization and Active Contour model.*

*****

I. Introduction

Object detection [1] is a technically and mathematically exigent and virtually supportive problem in the area of Computer vision. Object detection deals with identifying the existence of various objects in an image. Great success has been achieved in proscribed surroundings for object detection and recognition problem but the problem remains unsolved in many places, especially when objects are placed in capricious poses in muddled and occluded environment. As an example, it might be easy to train a domestic help robot [1] to recognize the presence of coffee machine with nothing else in the image. On the other hand, imagine the difficulty of such robot in detecting the machine on a kitchen slab that is cluttered by other utensils, gadgets, tools, etc. The probing or appreciation process in such scenario is very difficult. So far, no effective solution has been found for this problem.

Detecting saliency [2] in digital images has been an area of extensive research in the recent vision literature. Some work is done to envision the locations of human eye fixations and so provoke the primary ideology of saliency detection. Some methods of detecting saliency are bottom-up and some are top-down saliency. Top to down type uses high -level knowledge for saliency computation, calculating saliency to discover the stimuli from certain object classes. While bottom-up saliency method predicts eye fixations, so these methods are useful for high-level vision tasks, such as object recognition and localization.

Detecting saliency that is region of interest in digital images primary goal [15] as it allows proper image analysis and synthesis. Great accomplishment has been achieved at small amount of places but the problem remnants unrequited in unrestrained places; in especially when objects are placed in capricious places in jumble environment .Problems arise in discovering the important entities such as hasty object motion, varying manifestation 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.

Many object detection [1] and recognition methods can be classify into two types based on the Feature type they use in their methods. The two types are edge-based feature type and patch based feature type. It is notable that some researchers have used a combination of both the edge-based and patch-based features for object detection.

(1) Edge-based features

The methods that use edge-based feature [1] type haul out the edge map of the image and identify the features of the object in terms of edges. These are chiefly invariant to enlightenment conditions and variations in objects colors and textures. They represent the object boundaries and data efficiently in the large spatial extent of the images. Here they use the complete contour shape of the object as the feature and use the group of contour fragments as the feature of the object.

(2) Patch-based features

The second customary [1] type is the patch based feature type, which uses outer surface as cues. In this feature type, there are two main variations: (a) Patches of rectangular shapes contain the distinguishing margins describing the features of the objects. Many times these featured types are called as the local features. (b) Irregular patches in which, each patch is homogeneous in terms of intensity or texture and the change in these features are characterized by the boundary of the patches are called as the region-based features.

II. Random forest

Random Forests are made up of multiple independently formed decision trees. A Random Forest[4] is a classifier that consists of a collection of tree-structured classifiers $h(x, \Theta k)$ $k=1, 2,. . .$ where the Θ k are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x. The area of expertise of this combination is that each decision tree is made from a random vector of parameters [1]. Random Forest generates an ensemble of decision trees. Among different approaches that aims at forming ensembles of diverse classifiers, those using randomization to produce diversity have proven to be particularly efficient, as for bagging or random subspace methods. To classify a new object from an input vector, the input vector is run down each of the trees in the forest. Each tree provides a classification votes for the class. Interesting features of Random Forests [3] include: (i) their computational efficiency in both training and classification, (ii) their probabilistic output, (iii) the seamless handling of a large variety of visual features e.g. color, texture, shape, depth etc, and (iv)the inherent feature sharing of a multi-class classifier. The core idea behind Random Forest is to alleviate such problems by (i) injecting randomness into the training of the trees, and (ii) combining the output of multiple randomized trees into a single classifier.

III. Active contour model

Image segmentation [5] is a primary imaging problem that requires a resourceful modeling of image and texture features. Many existing active contour methods types segment an image according to either edge or region information. The evolution of segmenting curve is driven by the minimizing some energy that taken into account this given data. Active contour models are of two types:

(a) Edge-based active contours [5]

Edge-based active contour models use some edge detector and evolve the segmenting curve towards sharp gradients of pixel intensity. These local models are perceptive to the noise that affects the results of edge detectors. To avoid the difficult problem of edge localization, region based active contour models incorporate more global information to obtain segmented regions with homogeneous features.

(b) Region-based active contours:

To avoid the difficult problems that occurred with edge based methods, region based active contour models are employed and to incorporate more global information to obtain segmented regions with homogeneous features these region based are employed. In region based active contour model, the image is approximated using a smooth function inside each region. Many variants of this preliminary model have been projected that allow simpler and more efficient Implementation details. This type of active contour model are level set and level set free type. Level set methods have burden of re –initializations which induce large overhead and high cost. So here is some advantage of level set free time active contour.

Re-initialization [6] is a technique that sporadically re-initializes the function to a signed distance function during the evolution. It has been extensively used as a statistical therapy for maintaining stable curve evolution and ensuring usable results. However, re-initialization can cause dissimilarity between the theory of the level set method and its implementation. Moreover, it has a disadvantageous side effect of moving the zero level set away from its original location. The problem arises that when and how to apply the re-initialization. So there occurs a technique that forces the level set function to be close to function and it eliminates the need of the costly re-initialization procedure.

The variation level set formulation presented here has three main advantages over the traditional level set formulations. First, a significantly larger time step can be used for numerically solving the evolution PDE and therefore speeds up the curve evolution. Second, the level set function could be initialized as functions that are computationally more efficient to generate than the signed distance function. Third, the proposed level set evolution can be implemented using simple finite difference scheme, instead of complex upwind scheme as in traditional level Set formulations. However, like most of classical snakes and active contour models, the model proposed in relies on edge indicator functions or gradient-based stopping functions, depending on the image gradient, to stop the curve evolution. These models can detect only objects with edges defined by gradient. The discrete gradient modules can have moderately small local maximums on the object edges and then the stopping function is far from zero on the edges, and the curve may pass through the boundary. Also, the local maximums of the discrete gradient modules in the textured or noisy regions can be very close or equal to those of the object edges. Therefore, the evolving curve may stop before getting the object boundaries. With level set free active contours we have advantages like only one global initialization of the evolving curve can be performed to detect image salient objects. Second, we obtain a fast and re-initialization free active contour model which is robust to the textured and noisy regions. These also operate well when no object is detected in the image.

IV. *Segmentation*

Image segmentation, is necessary step in image analysis, object representation, visualization, and many other images processing tasks. Segmentation divides an image into its ingredient regions or objects. Segmentation accuracy helps in evaluating the eventual success or failure of computerized analysis procedure. Segmentation algorithms are based on one of two basic properties of intensity values discontinuity and similarity. First category is to partition an image based on abrupt changes in intensity, such as edges in an image. Second category is based on partitioning an image into regions that are similar in accordance to some previously defined criteria. Histogram thresholding and slicing techniques are used to segment the image. They may be applied directly to an image, but can also be combined with pre- and post-processing techniques.

V. *Gradient Minimization*

L_0 smoothing [8] results in global small-magnitude gradient removal. This method suppresses low-amplitude details and it retains and sharpens salient edges. These methods after doing the segmentation, in the proposed work sharpen the major edges. It will characterize and enhance fundamental image constituent's i.e. salient edges, and in the meantime eliminate the insignificant details. Algorithmically, we propose a sparse gradient counting scheme in an optimization framework. Its contribution is to limit the no of intensity changes in neighboring pixels. The qualitative effect of our method is to thin salient edges, which makes them easier to be detected and more visually distinct. Our enhanced edges are generally in line with the original ones. This method can have can also profit edge extraction, a fundamentally important operator, by effectively removing part of noise, unimportant details, and even of slight blurriness, making the results immediately usable in image abstraction and pencil sketch production. Our method is applicable to detail enhancement based on separating layers, and possibly to HDR tone mapping after parameter tuning. User interaction can be performed more efficiently on our Edge-enhanced images after removing low-amplitude structures. User interaction can be performed more efficiently on our edge-enhanced images after removing low-amplitude structures.

VI. *Related Study*

In recent years, several kinds of object-detection techniques prior have been studied such as given as: Mo Dai et al. [10] has presented method for grey level thresholding in which it considers the gray level value of an edge pixel as the ideal threshold for its neighborhood and a dynamic threshold function will be constructed by a dilation process from these primitive thresholds. The gray level value of a pixel found on the edge can be considered as ideal threshold and then a dilation process is used for assigning an appropriate value to threshold. Once the threshold image has been constructed then it is easy to obtain a binary image.Shuze Du et al. [11] has detected the salient object from the view of the object contour. It has used the random forest to measure patch rarities and compute similarities among patches. The approximate contour of the salient object is extracted based on this rarity map by using an active contour model. Then has a local saliency map as obtained by the similarities of patches inside the contour and those outside and then the local map is refined through image segmentation. Tony Chan et al. [7] have used an active contour model based on Mumford shah techniques and level set methods. The model automatically detects the interior contours starting with only the initial curve and the initial curve does not necessarily started around the objects to be detected. It has a different active contour model which performs its function that it detects its contour with or without gradient for instance objects with very smooth boundaries or even with discontinuous boundaries. This model is not dependent on the gradient of image for smoothing process.F.Schroff1 et al. [3] has presented Random Forest classifier. It has made three type of contribution such as it apparently show quite dissimilar classifiers that is nearest neighbor matching to text on class histograms can be mapped onto a Random Forest architecture. It also show that the ability of Random Forests to combine multiple features leads to a further increase in performance when color, filter banks, and HOG features are used simultaneously. Miyoun Jung et al. [5] have used a new class of active contour models that unifies patch processing and piecewise regular image models. It makes use of nonlocal comparisons between pairs of patches within the segmented regions and had illustrated the superiority of our models over existing active contour models. Due to the local homogeneity property, such segmentation model is able to detect regions with smoothly

ISSN: 2321-8169
3061 – 3068

spatially varying features and segment separated objects with different features with a level set function. Here the instantiation is done on basis of intensity, color, texture, or motion information; by designing appropriate metrics between patches such as the $L2$ norm for piecewise smooth features.Huaizu Jiang et al.[9]has used an automatic salient object segmentation algorithm which integrates both bottom-up salient stimuli and object-level shape prior, i.e., a salient object has a well-defined closed boundary. Saliency map is computed based on multi-scale super pixels, which proves to significantly enhance saliency, through context analysis. Zuhua Liu et al. [13] we propose a novel algorithm for object detection and localization based on structural model, it considers random forest as class classifier and for each leaf node of the decision tree and the offsets of the local features reach the node along with their class information should be recorded. Qiong Yan et al. [12] presented a multi-layer approach to analyze saliency cues. The final saliency map is produced in a hierarchical model. Different from varying patch sizes or downsizing images, the scale-based region is finding saliency values optimally in a tree model.

VII. *Proposed Technique*

In order to obtain the Objective, following number of steps is used.

Begin: Proposed(Input_image)

Step 1-Input the image to detect salient objects. Salient objects are the objects that represent the region of interest.

Step 2-Apply Random forest Classifier

We will form patches of the input image and apply random forest, to measure the patch rarities and patch similarity. Patches of rectangular shape represent the characteristic boundaries describing the features of the objects and Irregular patches are those in which, each patch is homogeneous in terms of intensity or texture and the change in these features are characterized by the boundary of the patches. These features are commonly called the region-based features. We then use random forest classifier and go through each pixel by using the patches. As we go through each of the pixels, we will add probability of pixels to get the overall probability of the entire image pixels. Hence we get the saliency map formed as a result of this step.

$$t_n(s_n; h_1, h_2) = \begin{cases} p_i \epsilon s_l, & if \;\; d_i(h_1, h_2) \leq \theta_{h1,h2} \\ p_i \epsilon s_r, & otherwise \end{cases}$$

Where s_l and s_r are the patch sets contained in node's left and right child,

$\theta_{h1,h2} = \frac{1}{|s_n|} \sum \forall p_i \epsilon s_n d_i(h_1, h_2)$,$|s_n|$ 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 equation2:

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

The above equation is used to evaluate the rarity of p_i, where $|L_k|$ is the number of patches contained in $L_k . w(x_i, x_c)$.

Step3-Apply ACM

We then apply active contour model to extract the approximate contour of object. The patch rarity we get from the previous step will be used to get the contour. Active contour is a circle that varies by the energy function. These are the level set active contours whose idea is to find a contour that segments image into two areas one inside and other outside the contour by approximating the image intensities in two regions. Then the histogram is plotted which gives exact measures of each pixel, hence it is useful in segmentation. Segmentation is done using the dynamic thresholding through equation.

$$g(x, y) = \begin{cases} 1, & if \; f(x, y) \geq T \\ 0, & otherwise \end{cases}$$

Where g(x, y) is threshold version of f (x, y) at some global threshold T which is computed at run time.

Step 4- Perform optimization

It is done by removing low amplitude structures from the base layer in a controllable degree. In 2D image representation, we denote by *I* the input image and by *S* the computed result. The gradient ∇Sp = (∂x Sp, ∂y Sp) T for each pixel *p* is calculated as color difference between neighboring pixels along the *x* and *y* directions. Our gradient measure is expressed as:

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

It counts *p* whose magnitude |*∂x Sp*|+|*∂y Sp*| is not zero. With this definition, *S* is estimated by solving

$$\min S\{\textstyle\sum_p (S_p - I_p) power\ 2 + \lambda . C(S)$$

In practice, for color images, the gradient magnitude |*∂ Sp*| is defined as the sum of gradient magnitudes in RGB. The term Σ (*S*−*I*) power 2 constrains image structure similarity. Such step is done because it has profit of edge extraction, a fundamentally important operator, by effectively removing part of noise, unimportant details, and even of slight blurriness, making the results immediately usable in image abstraction. Here only the salient edges are preserved. This step is useful as it provides the abstraction feature which means hide the irrelevant details and presenting only the important details. We get the final output.

Return: Object detected image

ISSN: 2321-8169
3061 – 3068

VIII. *Experimental Results*

As shown in figures, we are comparing the existing and proposed technique results.

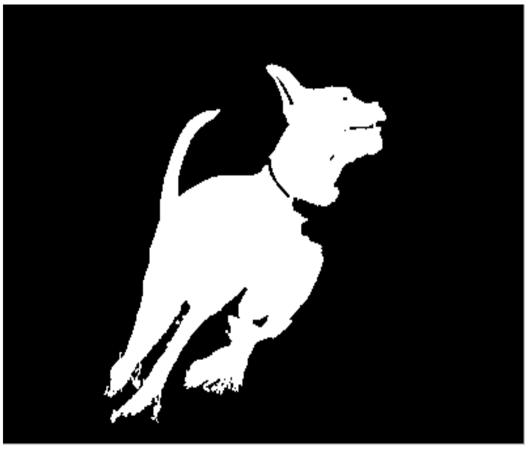
Input Image

RF Image

Old work

Detected object in old work

Proposed Work



Detecting object in new work

International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 2 Issue: 10
ISSN: 2321-8169
3061 – 3068

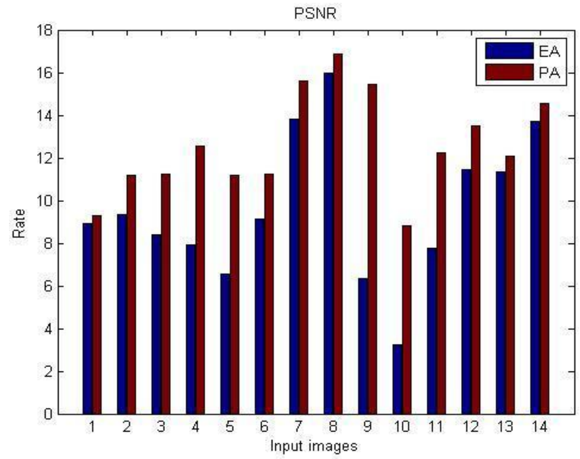
IX. *Performance Analysis*

(a) PSNR

Table 1 shows the result of peak signal to noise ratio value for multiple images. The values of existing and proposed method show the difference between the values for the same image. Here in the table values of proposed technique increased from existing to proposed ones.

Table1: PSNR ANALYSIS

Images	Existing Technique	Proposed Technique
1.jpg	8.9431	9.2944
2.jpg	9.3322	11.1689
3.jpg	8.3843	11.2246
4.jpg	7.9284	12.5382
5.jpg	6.5732	11.1723
6.jpg	9.1499	11.2353
7.jpg	13.7970	15.6141
8.jpg	15.9897	16.8833
9.jpg	6.3288	15.4501
10.jpg	3.2191	8.7923
11.jpg	7.7490	12.2287
12.jpg	11.4589	13.5260
13.jpg	11.3282	12.0671
14.jpg	13.7150	14.5485



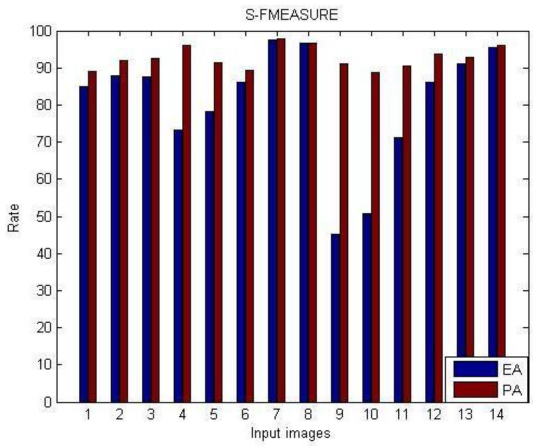
Graph1

Graph1 shows the input images across x-axis and rate across the y-axis and the bar graph shows the variation of values of proposed and existing technique.

b) *Sf-measure*

Table2: SF-measure

Images	Existing Technique	Proposed Technique
1.jpg	84.8098	88.9057
2.jpg	87.9034	92.0342
3.jpg	87.6319	92.5395
4.jpg	73.1100	96.1679
5.jpg	78.1099	91.4308
6.jpg	86.2531	89.4296
7.jpg	97.5646	97.7103
8.jpg	96.6943	96.7022
9.jpg	45.1349	91.0010
10.jpg	50.7232	88.6161
11.jpg	71.1265	90.4369
12.jpg	86.1233	93.8309
13.jpg	91.1786	92.7388
14.jpg	95.4378	96.1920

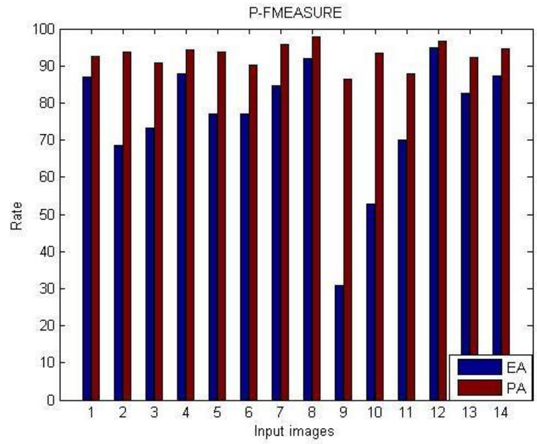


Graph2

c) *Pf-measure*

Table3: pf measure

Images	Existing Technique	Proposed Technique
1.jpg	86.9504	92.4829
2.jpg	68.4613	93.6551
3.jpg	73.2038	90.6570
4.jpg	87.7882	94.2021
5.jpg	77.0621	93.8004
6.jpg	77.1097	90.2449
7.jpg	84.7488	95.7362
8.jpg	92.0196	97.8829
9.jpg	30.7352	86.5112
10.jpg	52.7567	93.3616
11.jpg	69.9013	87.9232
12.jpg	94.7993	96.5854
13.jpg	82.5776	92.1399
14.jpg	87.2068	94.5913



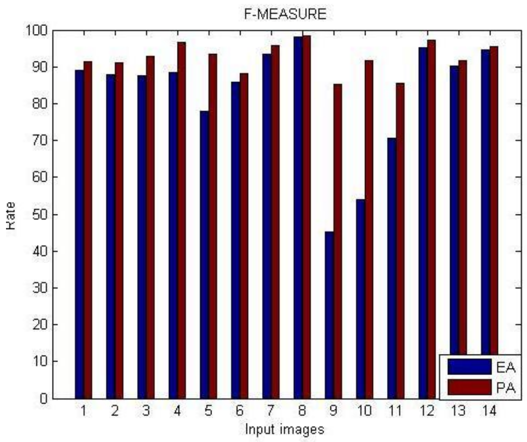
Graph3

d) *F-measure*

The F-Measure computes average of the information retrieval precision and recall metrics. Precision also known as positive predictive value is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance. The computed values are between 0 and 1 and a larger F-Measure value indicates higher classification/clustering quality. Table 4 indicates increase in the proposed technique values from existing technique.

Table 4: F-measure

Images	Existing Technique	Proposed Technique
1.jpg	89.1679	91.2824
2.jpg	87.9007	91.1051
3.jpg	87.6673	92.8359
4.jpg	88.5155	96.5463
5.jpg	77.8620	93.5095
6.jpg	85.8279	88.1328
7.jpg	93.2867	95.8525
8.jpg	97.9568	98.3158
9.jpg	45.1239	85.2527
10.jpg	53.8897	91.7950
11.jpg	70.6737	85.6694
12.jpg	95.2118	97.1659
13.jpg	90.2882	91.6181
14.jpg	94.6672	95.6051



Graph4

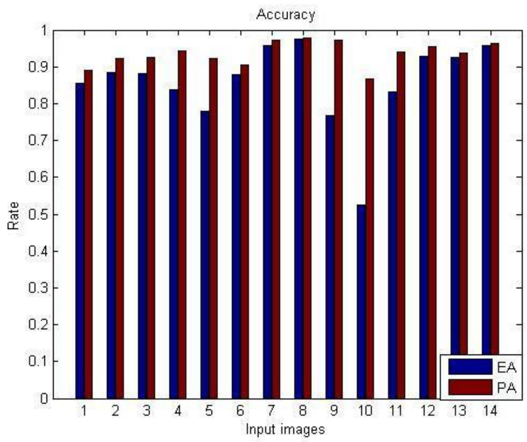
Graph4 shows the input images across x-axis and rate across the y-axis and the bar graph shows the variation of values of proposed and existing technique, which therefore shows the results increase in propose technique values of salient object detection.

e) Accuracy

It is defined as the extent to which we are close to the actual value by maintaining the good quality of digital images.Table1 indicates accuracy value of existing and proposed technique, and shows the continuous increase in accuracy value of proposed method.

Table 5: Accuracy

Images	Existing Technique	Proposed Technique
1.jpg	0.8562	0.8892
2.jpg	0.8834	0.9236
3.jpg	0.8828	0.9246
4.jpg	0.8389	0.9443
5.jpg	0.7799	0.9237
6.jpg	0.8784	0.9053
7.jpg	0.9583	0.9725
8.jpg	0.9748	0.9795
9.jpg	0.7671	0.9715
10.jpg	0.5235	0.8679
11.jpg	0.8321	0.9401
12.jpg	0.9285	0.9556
13.jpg	0.9263	0.9379
14.jpg	0.9575	0.9649



Graph 5

Graph5 shows the input images across x-axis and rate across the y-axis and the bar graph shows the variation of values of proposed and existing technique, which therefore shows the results increase in propose technique values of salient object detection in terms of accuracy rate.

(x)Conclusion

Salient object detection has attracted a lot of interest in computer vision as it provides fast solutions to several complex processes. Firstly, it detects the most salient and attention-grabbing object in a scene, and then it segments the whole extent of that object. The previous approach had used normal thresholding which in turn also faces problems if due to noise more of the pixels are added or removed, and resulting image has much of the artifacts which don't properly detect the salient object. So our proposed work overcome all these deficiency .The proposed work has used dynamic thresholding which will calculate global thresholding for each pixel at run time and in final step perform smoothening to remove low level details through use of gradient minimization.

References

[1] Prasad, Dilip K. "Survey of the problem of object detection in real images."*International Journal of Image Processing (IJIP)* 6.6 (2012): 441.

[2] Arvinda Sharma "A Compendium of Salient Object Detection" International Journal of advanced research in computer science and software engineering, IJARCSSE in Volume 4, Issue 8, 2014.

[3] Schroff, Florian, Antonio Criminisi, and Andrew Zisserman. "Object class segmentation using random forests." (2008).

ISSN: 2321-8169
3061 – 3068

[4] Kulkarni, V. Y., and P. K. Sinha. "Pruning of random forest classifiers: a survey and future directions." *Data Science & Engineering (ICDSE),2012 International Conference on*. IEEE 2012.

[5] Jung, Miyoun, Gabriel Peyré, and Laurent D. Cohen. "Nonlocal active contours." *SIAM Journal on Imaging Sciences* 5.3 (2012): 1022-1054.

[6] Ksantini, Riadh, Boubakeur Boufama, 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.

[7] Chan, Tony, and Luminita Vese. "An active contour model without edges."*Scale-Space Theories in Computer Vision*. Springer Berlin Heidelberg, 1999. 141-151.

[8] Xu, Li, et al. "Image smoothing via L_0 gradient minimization." *ACM Transactions on Graphics (TOG)* 30.6 (2011): 174.

[9] Jiang, Huaizu, et al. "Automatic salient object segmentation based on context and shape prior." BMVC. Vol. 3. No. 4. 2011

[10] [10] Dai, Mo, et al. "Image segmentation by a dynamic thresholding using edge detection based on cascaded uniform filters." *Signal processing* 52.1 (1996): 49-63.

[11] Borji, Ali, Dicky N. Sihite, and Laurent Itti. "Salient object detection: A benchmark." *Computer Vision–ECCV 2012*. Springer Berlin Heidelberg, 2012. 414-429.

[12] Du, Shuze, and Shifeng Chen. "Salient Object Detection via Random Forest." (2013): 1-1

[13] Yan, Qiong, et al. "Hierarchical saliency detection." *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013.

[14] Liu, Zuhua, and Huilin Xiong. "Object detection and localization using random forest." *Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference on*. IEEE, 2012.

[15] Arvinda Sharma"A scrutiny of detecting salient object in digital images". International Journal of advanced research in computer science and software engineering, IJARCSSE in Volume 4, Issue 5, 2014.

[16] R. Gonzalez and R. Woods *Digital Image Processing*, Addison-Wesley Publishing Company, 1992, pp 518 - 548.