SALIENCY BASED SEGMENTATION OF SATELLITE IMAGES

Ashu Sharma^a, J.K. Ghosh^b

^aResearch Scholar, Geomatics Section, Civil Engineering Department, Indian Institute of Technology, Roorkee, Uttarakhand, India –

ashu.mini123@gmail.com

^b Associate Professor, Civil Engineering Department, Indian Institue of Technology, Roorkee, Uttarakhand, India –

gjkumfce@yahoo.com

KEY WORDS: psychovisual, saliency, segmentation, satellite image

ABSTRACT:

Saliency gives the way as humans see any image and saliency based segmentation can be eventually helpful in Psychovisual image interpretation. Keeping this in view few saliency models are used along with segmentation algorithm and only the salient segments from image have been extracted. The work is carried out for terrestrial images as well as for satellite images. The methodology used in this work extracts those segments from segmented image which are having higher or equal saliency value than a threshold value. Salient and non salient regions of image become foreground and background respectively and thus image gets separated. For carrying out this work a dataset of terrestrial images and Worldview 2 satellite images (sample data) are used. Results show that those saliency models which works better for terrestrial images are not good enough for satellite image in terms of foreground and background separation. Foreground and background separation in terrestrial images is based on salient objects visible on the images whereas in satellite images this separation is based on salient area rather than salient objects.

1. INTRODUCTION

Satellite image interpretation is necessary part for further planning in civil engineering based applications. Many computer based applications provides different type of algorithms which can be helpful in image interpretation but expert human image interpreter can only be able to interpret an image at its best. If such an algorithm can be developed which can mimic the human way of image interpretation then huge reduction in cost and time can be done for the civil applications. Therefore psychovisual image interpretation is needed to interpret an image as human do. Image segmentation is a key step in image interpretation and it is typically defined as exhaustive partitioning of an input image into regions, each of which is considered to be homogeneous with respect to some image property of interest like intensity, color or texture etc (Jain 2013). In saliency based image segmentation, saliency computes the most attentive location on the basis of human vision system which will give the foreground of image and rest of the area will be as background. The more saliency model is closer to human vision mechanism the more the probability will be to extract all the salient objects needed for image interpretation. Thus saliency based segmentation can be eventually helpful in psychovisual image interpretation.

There are many saliency models are available and they even perform well on terrestrial images (Tavakoli et al 2011), (Riche et al 2013), (Technion et al 2010) (Achanta et al 2008). The efficiency of these models is calculated on the basis of ground truth images. In the ground truth images objects presented in image becomes foreground (1s) and rest part become background (0s). Here foreground and background separation is precisely done on pixel basis. But with the satellite image this case is different. For satellite images there are numerous objects are presented in image and all (or some of them) may be required for image interpretation. So for such cases a human labeled pixel wise precise foreground and background reference image can't be prepared until target object is not defined. Therefore saliency based segmentation for satellite images is a better way to segment a satellite image specially when target object is not defined. The whole idea behind is that even humans also perceive on the basis of those object which catches attention the most within the area of vision, so if the most attentive locations as per human vision can be extracted from satellite image then the whole image can be given as input for final image interpretation in an human inspired way.

For saliency based segmentation first there is need to understand how and where humans generally look at. This can be computed by different available visual saliency models which resembles the human quality of prioritizing the incoming stimuli from a scene and focus on those parts (Riche et al 2013). If image is segmented on the basis of this saliency then there is only need to concentrate over a limited area of image. Although many saliency models are available and some of which have even used saliency based segmentation (Hou et al 2007) (Achanta et al 2009) but these are performed over terrestrial image only. In this work different saliency models are used in association with single segmentation algorithm and these models are tested for satellite images. It is not necessary that the saliency model which gives better result for certain data set gives the same for other type of data set. Saliency for satellite images plays differently than any other data set used such as indoor or outdoor images. Till now so far on the basis of literature review done, any of introduced saliency models have neither used satellite image for measuring saliency nor for segmentation.

Keeping in view the above idea this paper demonstrates the implementation of saliency model based segmentation on a set of satellite images. Performance of the same models is also judged on terrestrial image dataset with respect to the reference images given along with dataset. Results of satellite images have been discussed on the basis of capability for image interpretation from objects or area extracted. It means the goodness of a model is compared on the basis that the objects or area extracted are enough for image interpretation or not.

The organisation of this paper is started with introduction followed by brief details about saliency models and segmentation algorithm used in this work. After this section methodology and implementation details are given. Further results have been discussed followed by conclusion of the work done.

2. BACKGROUND THOERY

Our different saliency models in association with SLIC segmentation algorithm have been used in this work. Description of saliency models is given followed by segmentation algorithm.

2.1 Saliency by Sparse Sampling & Kernel Density Estimation (SS&KD):

This center surround saliency model is proposed by Tavakoli et al. in 2011 (Tavakoli 2011) in which it is hypothesized that there exists a local window which is divided into a center which contains an object and a surround. Saliency belonging to center in this model utilizes Bayes's theorem. Then multi scale measure is done by changing the radius and number of samples. Here the radius is "size scale" denoted by r and the number of samples as "precision scale" denoted by n. Saliency S(x) of a pixel at different scales is calculated by the average taken over all scale:

$$S(x) = \frac{1}{M} \sum_{i=1}^{M} S_{r_i}^{n_i}(x)$$
(1)

where M = number of scales,

 $S_{r_i}^{n_i}(x) = i^{th}$ saliency map calculated at a different scale using the equation (2.1.2).

$$S_{r}^{n}(x) = A_{c} * [P_{r}^{n}(1|f,\bar{x})]^{\alpha}$$
(2)

where $A_c = a$ circular averaging filter,

* = convolution operator,

 $[P_r^n(1|f,\bar{x})]^{\alpha}$ = calculated by using Bayes's theorem and $\alpha \ge 1$ is an attenuation factor which emphasizes the effect of high probability areas. Work flow diagram is given in Figure 1.



Figure 1. Work Flow diagram of SS7KD Saliency Model

2.2 Multi Scale Rarity-based Saliency (RARE2012):

This model is a 'multi-scale rarity-based saliency detection' and it is also called RARE2012 (Riche et al 2013). There are three main steps of this bottom up saliency model. In first step lowlevel features such as color and medium-level orientation features get extracted. Color transformation is done to obtain a maximum color features decorrelation. Afterwards, a multiscale rarity mechanism is applied as a feature is salient only in specific context. Therefore the mechanism used for multi-scale rarity allows detecting both locally contrasted and globally rare regions in the input image. Finally, rarity maps are fused into a single final saliency map. The flow chart of this model is shown in Figure 2.



Figure 2 Work Flow diagram of RARE2012 Saliency Model

2.3 Saliency by Low level Feature Contrast (Achanta 08):

This method is based on 'Local Contrast' proposed by Achanta (Achanta et. al. 2008). In this method salient regions are identified as the local contrast of an image region with respect to its neighborhood at various scales. It is evaluated as the distance between the average feature vector of the pixels of an image sub-region with the average feature vector of the pixels of its neighborhood. This allows obtaining a combined feature map at a given scale by using feature vectors for each pixel, instead of combining separate saliency maps for scalar values of each feature. At a given scale, the contrast based saliency value $c_{I,j}$ for a pixel at position (i, j) in the image is determined as the distance D between the average vectors of pixel features of the inner region R1 and that of the outer region R2 as:

$$c_{i,j} = D\left[\left(\frac{1}{N_1}\sum_{p=1}^{N_1} V_p\right), \left(\frac{1}{N_2}\sum_{q=1}^{N_2} V_q\right)\right]$$
(3)

where NI = number of pixels in RI

N2 = number of pixels in R2

v = vector of feature elements corresponding to a pixel.

D = a Euclidean distance if v is a vector of uncorrelated feature elements, and it is a Mahalanobis distance (or any other suitable distance measure) if the elements of the vector are correlated.

In this work, the CIELab color space has been used, assuming RGB images, to generate feature vectors for color and luminance. Since perceptual differences in CIELab color space are approximately Euclidian, *D* in Equation 2.3.2 is:

$$D_{i,j} = ||v_1 - v_2|| \tag{4}$$

where $v_1 = [L1; a1; b1]^T$ and $v_2 = [L2; a2; b2]^T$ are the average vectors for regions R1 and R2, respectively.

Final Saliency map is calculated as sum of saliency values across the scales *S* as per following equation:

$$m_{i,j} = \sum_{S} c_{i,j} \tag{5}$$

Here $m_{i,j} = a$ element of combined saliency map M at pixel value (I,j).

Work flow diagram of this model is shown in the following Figure 3.



Figure 3. Work Flow diagram of Saliency by Low level Feature Contrast

2.4 Frequency-tuned Salient Region Detection (Achanta 09):

This method is also proposed by Achanta et al (Achanta 2009) and exploits features of color and luminance. It is simple to implement but computationally efficient as compared to Achanta 2008. In this method finding the saliency map S for an image I of width W and height H pixels can be formulated as:

$$\mathbf{S}(\mathbf{x},\mathbf{y}) = |\mathbf{I}_{\mu} - \mathbf{I}_{\omega hc}(\mathbf{x},\mathbf{y})| \tag{6}$$

where, I_{μ} = arithmetic mean pixel value of the image $I_{\omega hc}$ = Gaussian blurred version of the original image

 I_{ohc} is used to eliminate fine texture details as well as noise and coding artifacts. The norm of the difference is used because the main focus is only in the magnitude of the differences. This is computationally quite efficient.

To extend the above equation to use features of color and luminance, rewrite it as:

$$\mathbf{S}(\mathbf{x}, \mathbf{y}) = \| \mathbf{I}_{\mu} - \mathbf{I}_{\omega h c}(\mathbf{x}, \mathbf{y}) \|$$
(7)

where, I_{μ} = mean image feature vector,

 $I_{\omega hc}(x,y)$ = corresponding image pixel vector value in the Gaussian blurred version (using a 5 X 5 separable binomial kernel) of the original image,

|| = L2 norm.

Using the Lab color space, each pixel location is an $[L,a,b]^T$ vector, and the L2 norm is the Euclidean distance. The work flow diagram is shown in Figure 4.



Figure 4. Work Flow diagram of Frequency-tuned Saliency Detection

2.5 Segmentation algorithm

For segmentation SLIC (Simple Linear Iterative Clustering) algorithm is used, proposed by Achanta et al. (Achanta 2010). SLIC is a simple and efficient method to decompose an image in visually homogeneous regions. It is based on a spatially localized version of k-means clustering in which each pixel is associated to a feature vector $\Psi(x,y)$ of image I(x,y):

$$\Psi(\mathbf{x},\mathbf{y}) = \begin{bmatrix} \lambda_{\mathbf{x}} \\ \lambda_{\mathbf{y}} \\ \mathbf{I}(\mathbf{x},\mathbf{y}) \end{bmatrix}$$
(8)

The coefficient λ balances the spatial and appearance components of the feature vectors.

$$\lambda = \frac{\text{regularizer}}{\text{regionsize}}$$
(9)

SLIC takes two parameters: the nominal size of the regions (superpixels) **regionSize** and the strength of the spatial regularization **regularizer**. SLIC starts by dividing the image domain into a regular grid with M×N tiles, where:

$$\mathbf{M} = \begin{bmatrix} \frac{\text{imageWidth}}{\text{regionSize}} \end{bmatrix}, \qquad \mathbf{N} = \begin{bmatrix} \frac{\text{imageHeight}}{\text{regionSize}} \end{bmatrix}$$
(10)

Then a region (superpixel or k-means cluster) is initialized from each grid center. In order to avoid placing these centers on top of image discontinuities, the centers are then moved in a 3×3 neighbourohood to minimize the edge strength. Then the regions are obtained by running k-means clustering, started from the centers.

$$C = \left\{ \psi(x_i, y_j), i = 0, 1, \dots, M - 1 \, j = 0, 1, \dots, N - 1 \right\}$$
(11)

K-means uses the standard LLoyd algorithm alternating assigning pixels to the closest centers a re-estimating the centers as the average of the corresponding feature vectors of the pixel assigned to them. The only difference compared to standard k-means is that each pixel can be assigned only to the center originated from the neighbour tiles.

After k-means has converged, SLIC eliminates any connected region whose area is less than **minRegionSize** pixels. This is done by greedily merging regions to neighbour ones: the pixels p are scanned in lexicographical order and the corresponding connected components are visited. If a region has already been visited, it is skipped; if not, its area is computed and if this is less than **minRegionSize** its label is changed to the one of a

neighbour region at p that has already been visited. The working flow of this segmentation algorithm is given in the Figure 5.



Figure 5. Working flow of SLIC algorithm

3. DATA USED AND METHODOLOGY

For implementation 2 set of images are used: one set of images are terrestrial outdoor images taken from dataset used by (Hou et al 2007), and another set of image used are sample natural color satellite images of worldview-2. There are 3 images (viz. img1, img2 and img3) results have been taken from first dataset to show in this paper. Among these images img1 shows band 1 and band 2 (band red and green respectively from visible range of EM spectrum) shows a strong correlation while less correlation with band 3 (blue band). Similar correlation is seen in between the band 1 and 2 of other images also shown in Figure 8(e) and 9(e) which shows the redundancy of data in band 1 and 2. All these images show a wide range of DN values which signifies no atmospheric effect. In img3 band 1 is bimodal and gives peaks at 15 and 147 DN value. First peak is because of wide area of sky of blue color in image and second peak is due to the land. Mean of the image img1 ranges from 60 to 80 nearly for all bands, similarly for img2 and img3 this range is 60 to 105 and 90 to 130 respectively. Standard deviation also for these bands is also within the range of 30 to 50. This dataset has been chosen because reference image for segmentation provided with this dataset is human labeled and hand labelers concentrate only on the edges between the foreground and the background. So this type of segmentation more resembles to human vision as when human see some object in an image then not only that object with crisp boundary comes within vision but the whole specific area comes within the vision range. There are three types of human inspired segmentation reference images are available but in this paper only those reference images are used which are having at most number of object.

The satellite image used in this work both shows a high correlation between all three bands. Standard deviation is in range of 40 to 50 only. Mean value ranges from 105 to 120 for satellite image 1 and from 75 to 95 (nearly) for satellite image 2. Resolution of the used sample satellite imagery is 0.5 meters.

A threshold based hybrid methodology, inspired by (Achanta 2008), is used for each saliency model for segmenting input image. The idea behind the methodology used is to calculate average saliency for each segment in segmented image and then extracting only those segments which are having higher saliency than threshold value. For implementing the above idea saliency map of input image is calculated and segmentation is done separately by SLIC algorithm. Then both the outputs (segments

from segmented image and saliency value of each pixel from saliency map) are combined for final output generation. If there are k segments in a segmented image then the average saliency for that segment is:

$$s_k = \frac{1}{|r_k|} \sum_{i,j \in k} sm_{i,j} \tag{12}$$

where, $sm_{i,j}$ = pixel value of the saliency map for the segment *k* which average saliency is to be calculated

Implementation as per methodology is done first for terrestrial images having mean saliency of saliency map as a threshold value and their performance are measured. Then the same method is used for satellite images keeping the threshold value same as mean and visually performance is measured. The flow chart of methodology is shown in Figure 6.

Figure 6. Methodology Work Flow Diagram



4. RESULTS AND DISCUSSION

The results for 3 images from first type of dataset used which consist of terrestrial images are shown in Figure 7, 8 and 9. First image used in Figure 7 shows that SS&KD (Figure 4.1(b)) and RARE2012 (Figure 7(c)) both the models cover almost all the important objects that are required to describe the scene. Rare2012 do highlights other small objects (other small animals) but covers the area of building which is behind tree, whereas SS&KD doesn't remove the tree but do omits the small animals (white color animal in left of the image Figure 7(b)).

Now if result of Achanta 08 and 09 models are to be considered then in Figrue 7 (d) & (e), very less information is available to describe the scene. Area near by the tree which is masked by these models, create vague impression in results which will eventually hard to deal at the time of image interpretation. In other models (Figure 7(b) & (c)) objects like building and other animals are clearly and fully visible whereas this is not the case with Achanta 08 and 09 based models as these two models are better for one object image.

Similar types of results are found for other images also from the dataset which is shown in Figure 8 and Figure 9. one things comes out form these results that even Achanta 08 and 09

models are not that much better performing in the case of foreground and background separation but it delineates crisp boundary of objects while other models based segmentation do include the background along with salient objects.



Figure 7. (a) original image (b) SS&KD, (c) Rare2012, (d) Achanta 08, (e) Achanta 09, (f) human labeled



Figure 8. (a) original image, (b) SS&KD, (c) Rare2012, (d) Achanta 08, (e) Achanta 09, (f) human labeled



Figure 9. (a) original image, (b) SS&KD, (c) Rare2012, (d) Achanta 08, (e) Achanta 09, (f) human labeled

One more important point about these results is Achanta 08 model based segmentation extract the flying object at the upper right corner (Figure 8 (d)) which other models are failed to extract at this threshold level even that object is not much distinguishable but may be eventually helpful at the time of image interpretation

The above results shows that Achanta 08 model based segmentation able to extract even some small objects (which may not be much salient). The performance of each of these models on the first type of dataset has been analyzed by calculating running time and F-1 score with reference of human labeled images given along with dataset. Based on the

performance measured by F-1 score it can be said that SS&KD and RARE2012 performs well for terrestrial images with respect to the human labeled image used as reference for checking. Whereas segmentation based on Achanta 08 and 09 could not perform well for the same. Only those areas which are having high intensity values are extracted from these two models (Achanta 08, 09). SS&KD performed well with image where main objects in image are in center (Figure 7(b)). Rare2012 shows comparatively low performance than SS&KD as it uses low level color feature and then orientation, therefore even extracting the area as a salient object which is not even an object in the image (bottom left corner of image in Figure 9(c)) The following table (Table 1.) shows F-1 scores calculated for each of these 3 images:

	Running Time (secs)			F-1 score		
Model	IM1	IM2	IM3	IM1	IM2	IM3
Name						
SS&KD	0 10	7.06	6.01	0.6620	0.506	0.644
based	0.40	7.90	0.01	≈ 0.7	≈ 0.5	≈ 0.6
Rare2012	7 806	0.60	10.12	0.5503	0.518	0.5905
Based	7.800	9.09	10.12	≈ 0.6	≈ 0.5	≈ 0.6
Achanta	56	7 1 07	0.20	0.2302	0.258	0.225
08 based	5.0	/.18/	9.20	≈ 0.2	≈ 0.3	≈ 0.2
Achanta	56.06	71 12	74 49	0.3762	0.452	0.419
09 based	50.00	/1.12	/4.40	≈ 0.4	≈ 0.4	≈ 0.4

 Table 1. Comparative analysis of Saliency based Segmentation

 Models for terrestrial data

The same models are implemented for satellite images from worldview -2 of Washington, D.C.; June 8, 2011 and Madrid, Spain; February 7, 2011. The result of first image after implementing the four mentioned saliency model based segmentation is shown in the following figure 10. The threshold value for these results is mean of the complete saliency map. In results some interesting pattern objects are completely removed by the SS&KD based models and unnecessary part of roads are extracted. Because of using center surround method by SS&KD model it leaves the salient objects lying in corner or boundary area. Because of this reason two visual attention grabbing objects at lower portion of image are removed completely Figure 10 (b). Even for this method if the threshold value increases then also only the center portion of image will be enhanced and again corner area will be extracted.

RARE2012 (Figure 10 (c)) also gives considerable results as it also extracts the major highlighted portion of imagery. At this threshold value maximum road side trees also extracted. But if it is compared with result of Achanta 08 and 09 then it can be said that roads need not to be extracted completely as this type of information can be bet by another marks like zebra crossing on road which always have high luminance and always grab our attention very easily. Thus it extracts unnecessary parts. Also a part at upper right corner is also completely removed by this model whereas in Achanta 09 based segmentation it is clear and in Achanta 08 it is having some trace upto some extent. The results from Achanta 08 and 09 based models gives almost all the necessary objects required to interpret that image and the area which is not completely extracted e.g. roads that can also be interpreted on the basis of linear car like objects and zebra crossing over it. Trees near by the road is also gets extracted by these two models.

Now such type of image is having more 0s, thus redundant values. So now for interpretation less part of image can be taken for consideration and not all image is required until the target object or application is regarding a specific object. On the basis of above discussion it can be said that segmentation based on SS&KD model and RARE2012 models do not give better result as compared to what is got by Achanta 08, 09.



Figure 10. (a) original image, (b) SS&KD, (c) Rare2012, (d) Achanta 08, (e) Achanta 09

As these two models (viz. Achanta08, 09) based segmentation have performed better for satellite images then it is again tested for different threshold value for same and for other satellite image. This time threshold value is taken as 'mean/2' and implemented for both satellite images. The result of the implementation of this threshold value is shown in figure 11. After decreasing the threshold value some other less salient areas have been extracted after segmentation, which gives better understanding for image. Small trees on road side are also visible at this threshold value. In second satellite image also almost all important features are visible (e.g. upper left corner in Figure 11 (f)).



Figure 11. Image 1: (a) original image, (b)-(c) thresh >=mean/2 for Achanta 08, Achanta 09 respectively Image 2 : (d) original image (e) –(f) Achanta 08 at Thresh >= mean and >= mean/2 respectively (g)-(h) Achanta 09 atThresh >=mean and >= mean/2 respectively

One more advantage is seen for Achanta 08 and 09 models over other two models used in this work is that even shadow of the tower is darker than the roads but still after segmentation roads are omitted but shadow still remains which is necessary information for interpretation of the towers.

The output of saliency based segmentation has been compared with other segmentation techniques. Multiresolution segmentation is applied on the satellite image of Washington and results are generated at different values of the parameters. These results are shown in Figure 12. Any segmentation techniques, generally, divides the image into parts or some regions. The multiresolution segmentation technique used here for comparison is a bottom up region merging technique and a local optimization procedure. The results of this segmentation show that image is divided into regions based on the parameters used.



Figure 12. Segmentation Result at (a) scale=10, shape=0.7 compactness=0.7, and (b) scale=25, shape=0.9 compactness=0.7

As aim of segmentation in this work is to separate an image into such foreground and background in which foreground is based on human perception or in similar way as human prioritize a scene which comes into their vision, such segmentation is not possible with traditional segmentation techniques which only make parts or region in image based on different parameters. Such segmentations can be used as intermediate step in saliency based segmentation but the parameters should be chosen as per the requirement of satellite image. For example for, here the satellite image used for implementing multiresolution segmentation is of urban area, therefore shape parameter is of higher importance as mostly objects in urban area are manmade and therefore having proper geometrical shape (except trees on road side). Similarly if the satellite image is of natural landscape then shape parameter will have weightage.

5. CONCLUSION, LIMITATIONS AND FUTURE SCOPE

This paper has presented and evaluated four models to visualize saliency based segmentation for high resolution satellite images. The focus of the work is to segment the satellite image from human vision point of view which is brought by the use of saliency models for segmenting the high resolution satellite image. From the results discussed above it can be said that for satellite image interpretation Achanta 08 based segmentation model has given the better results than as compared to other saliency based models used in the work. SS&KD based model mainly concentrates in the center of the image and thus looses the information content at the corners of the image. Rare2012 performed somewhat better than SS&KD as it includes the corner highlighted value. Rare2012 uses low level features color for rarity map calculation; therefore it highlights the colors with

much luminance in image while areas having colors with less illumination become non salient or lesser salient. Frequency tuned Achanta 09 based segmentation model do perform better than the above two discussed. By using local contrast Achanta 08 based model extracts the most information. For example all major building's top portions, shadow of towers, trees at the road side etc all are extracted by this model which is necessary for scene interpretation.

The satellite image used is of urban area and having mainly manmade objects in the image e.g. buildings, roads, cars. Therefore maximum objects in the image is having regular boundary. Therefore using SLIC segmentation it gives a neat boundary of objects extracted. The regulizer parameter of SLIC helped in keeping the object boundary so that all the segments extracted are either a part of object or object itself but no segment comes in between the boundary or sharp change of pixel values. For natural scene image the same may not perform well because of irregular boundaries. From the results it can be inferred that those saliency based segmentation which works efficiently for terrestrial images are not good enough for satellite image. As for terrestrial image even if it is a complex image then also training for satellite image will be different for satellite image for developing intelligent systems.

Another important conclusion about precise boundary of objects in satellite images is the segmentation algorithm used.. In this way one very much important concluding remark is for satellite images saliency is not same as we generally define for other images always and also if saliency based segmentation is done for satellite images then with less information other opt out values can also be inferred.

Limitation noticed of the work done is that the result of final segmentation is dependent on the quality of saliency calculation of saliency model. If the saliency model cannot mimics well the human way of prioritizing the stimuli then we may loose some important objects while interpreting the satellite image as this happened with satellite image segmentation doen with SS&KD saliency based model (Figure 10(b)). Even this model performs well on terrestrial images but huge area is left and only area in center is considered. In this way we have loosed some important and quite salient building structure at the corners and also the trees on the road side.

Other limitation of the work done is still the threshold value is negotiable. As increasing the threshold will increase only that area extraction which was priory less salient. This will increase only some number of objects in the image but how many objects are necessary and sufficient for complete image interpretation that is still variable from image to image in terms of resolution, viewing angle, objects present in the image etc.

The future scope of this work can be suggested as saliency based segmentation for satellite image can be helpful in psychovisual satellite image interpretation as it separated the foreground and background on the basis of human vision system and ultimately can be helpful in many other civil applications in which complete interpretation of a high resolution satellite image is required. Ability of intelligent image interpretation systems can be increased by giving training to system about where to look and what objects are necessary to interpret an image in a way as human mind can interpret. In this way if segmenting an image in a way of only concentrating image objects cannot give much better result as satellite image generally have multiple objects and almost every object may or may not contribute in image interpretation. So in this way if techniques regarding imitating human vision system of prioritizing the objects is used then it may be helpful in image interpretation as human mind.

REFERENCES

Achanta Radhakrishna, Estrada Francisco, Wils Patricia, and Susstrunk Sabine, 2008. "Salient Region Detection and Segmentation", International Conference on Computer Vision Systems (ICVS '08), Vol. 5008, Springer Lecture Notes in Computer Science, pp. 66-75,

Achanta Radhakrishna, Hemami Sheila, Estrada Francisco, and Süsstrunk Sabine, 2009. "Frequency-tuned salient Region Detection", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 1597 - 1604.

Achanta Radhakrishna., Shaji A., Smith K., Lucchi A., Fua P. and Süsstrunk S., 2010. "SLIC Superpixels", Technical Report, EPFL (Ecole polytechnique fédérale de Lausanne).

Gonzalez R. C., and Woods R. E., 2008. "Digital Image Processing," Addison-Wesley Publishing Company, III edition, Asia.

Hou X. and Zhang L., June 2007. "Saliency Detection: A Spectral Residual Approach", Computer Vision and Pattern Recognition (CVPR '07).

Jain A. K., 2013. Fundamentals of Digital Image Processing, PHI Learning Pvt Ltd.

Judd Tilke, Ehinger Krista, Durand Fredo, Torralba Antonio, 2009. "Learning to Predict Where Humans Look", International Conference on Computer Vision (ICCV).

Karakis, S., Marangoz, A. M., Buyuksalih, G,2006. Topographic Mapping from Space (with Special Emphasis on Small Satellites) ISPRS Archives-VolumeXXXVI-1/W41, "Analysis Of Segmentation Parameters In Ecognition Software Using High Resolution Quickbird Ms Imagery", Ankara, Turkey, <u>http://www.isprs.org/proceedings/XX</u> XVI/1-W41/default.aspx, (February 14-16, 2006)

Kiefer R. W., Lillesand T. M., and Chipman J. W., 2009. "Remote Sensing and Image Interpretation", John Willey and Sons, V edition, University of Wisconsin, Madison, chapter-4,.

Riche Nicolas, Mancas Matei, Duvinage Matthieu, Mibulumukini Makiese, Gosselin Bernard, Dutoit Thierry, July 2013. "RARE2012: A Multi-Scale rarity-Based Saliency Detection with Its Comparative Statistical Analysis", Signal Processing: Image Communication at Science Direct, Volume 28, Issue 6, pp 642-658 28.

Stiefelhagen R. and Ekenel HK, 2011. 'Content-Based Image and Video Retrieval', Lectures in the summer semester, Computer Vision for Human Computer Interaction Lab.

Tavakoli Hamed Rezazadegan, Rahtu Esa, and Heikkila Janne, 2011. "Fast and Efficient Saliency Detection Using Sparse Sampling and Kernel Density Estimation", Scandinavian Conference on Image Analysis (SCIA), Lecture Notes in Computer Science 6688, pp. 666–675.

Technion Stas Goferman, Technion Lihi Zelnik-Manor, Technion Ayellet Tal, June 2010. "Context-Aware Saliency Detection", Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, ISSN: 1063-6919, pp 2376 – 2383.