Foreground Object Detection and Segmentation Based on Using Region Contrast

Samera Shams Hussein
Department of Computer Science, College of Education for pure Science (Ibn Al-haitham),
Baghdad University, Baghdad, Iraq
Email: samera.s.h@ihcoedu.uobaghdad.edu.iq

Abstract

Foreground object detection is one of the major impotent tasks in the field of computer vision, which attempt to discover important objects in still image or image sequences or locate related targets from the scene. Foreground objects detection is very important for several approaches like object recognition, surveillance, image annotation, and image retrieval, etc. In this work, we have proposed a method for detection and segmentation foreground object from image or video in case when the targets are moving or not moving. Comparisons with general foreground detectors such as background subtraction techniques our approach are able to detect important target for case the target is moving or not and can separate foreground object with high details.

Keywords: Image Segmentation, Foreground Object, Region Contrast.

كشف و تجزئة الاجسام البارزة بالأعتماد على خوارزمية تباين المنطقة

سميرة شمس حسين
قسم علوم الحاسبات، كلية التربية للعلوم الصرفة (إبن الهيثم) ، جامعة بغداد، بغداد، العراق

الخلاصة

تعد عملية كشف الأجسام البارزة من المهام الرئيسية المهمة في مجال الرؤية الحاسوبية والتي تهدف الى كشف وجود الأجسام المهمة في الصورة الثابتة او في الصور المتسلسة (الفيديو) او ايجاد الهدف المقصود في المشهد. ان كشف الأجسام البارزة لها اهميه كبيرة في العديد من المجالات مثل تمييز الأجسام، انظمة المراقبة، توضيح واسترجاع الصور، وغيرها. في هذا العمل تم اعداد طريقة لكشف وتقطيع الاجسام البارزة من الصور والفديو في حالة كان الهدف متحرك او متوقف. بالمقارنة مع طرق كشف الاجسام البارزة التقليدية مثل تقنية طرح الخلفية فأن نموذجنا المقترح يتمكن من كشف الأجسام البارزة في حالة كان الجسم متحرك او ثابت ويتمكن من فصل الجسم البارز من الخلفية مع تفاصيل عالية.

1. Introduction

Biological vision systems tend to be incredibly efficient at finding relevant targets within a scene. Most people will probably quickly and consistently spot at those important targets within the images for example in Figure 1. Certainly, to detect these types of objects from image, it commonly preselected by using a two level labelling strategy to be sure a foreground object is detached from the sense background [1].

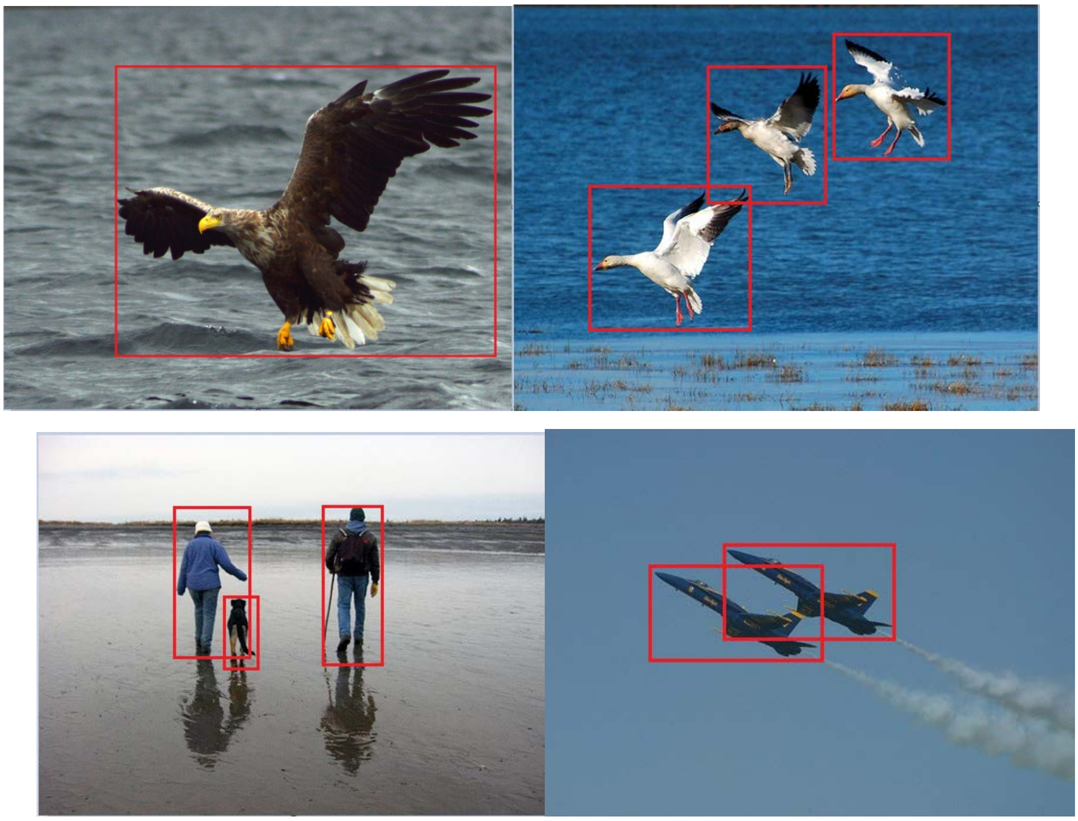


Figure 1: Sample of images having a obvious forgroubd object, where the red ectangles represent human annotation.

Recognizing these areas of prominent, or significant, within the visual field allows one to recruitment the minimal perceptual resources using an efficient way. When compared computer vision approaches with biological systems, it's certainly the computer vision techniques are far behind capability to detect foreground objects. Nevertheless, dependable foreground object detection strategies could be beneficial in numerous applications such as: object recognition, unsupervised image segmentation, and adaptive scaling and compression. One common strategy to minimize scene clutter would be to recognize foreground objects versus a static background. There are many techniques currently successful in several purposes that useful in detect moving object in the scenes such as background subtraction, but the difficult task is to detect the foreground object from moving sense or still image that can focused and segment just those important objects from other objects in background of sense [2, 3].

The automated recognition of foreground object regions in images requires a gentle breaking down the foreground from other elements of background image. This type of breaking down is a key element of several graphics and computer vision tasks. Instead of emphasizing forecasting human fixation points (an alternative significant research path of visual attention modelling), foreground region detection strategies focus on regularly featuring existing foreground object regions, and as a result benefiting numerous applications, such as object recognition, object level image manipulation, object of-interest, internet visual media retrieval, image segmentation, content aware image editing and adaptive compression, [4].

Extraction of salient or foreground objects in a scene relates to appropriate object retrieval and image segmentation. Apparently, reliable foreground evaluation is frequently achievable without the need of actual scene knowing. Foreground, considered a bottom-up procedure that derives from visual surprise, rarity or distinctness, and is often related to variants in image characteristics like gradient, color, boundaries and edges. Visual foreground objects are investigated throughout numerous disciplines such as computer vision, cognitive psychology, and neurobiology. According to observed reaction along biological methods, the human attention theories hypothesize that the techniques of human vision system basically areas of an image in detail, along with leaving the remaining around unprocessed. Earlier work suggests two steps of visual consideration: a quick, data driven, bottom-up, pre-attentive, foreground object extraction; then, slow, top-down, goal driven, task dependent, foreground extraction [5].

This paper focused to detect the important objects and separate it from moving and still sense with high details in order to increase the accuracy of recognition process. We have been proposing an approach for detect and segment foreground object depending on region Contrast. Since the theory of object based attention is recommending to start using the complex image segmentation towards proto-objects. Even though, common object segmentation is a difficult task, estimated at describing significant segments within an image via feature grouping is probably achievable. We then determine foreground for each proto-object coupled with describe the foremost salient one.

2. Previous Works

A different method has been provided for foreground and silence object detection approaches, which have been aimed to locating specific categories such as tables, cars, persons, airplane, etc.

Walther and Koch (2006) [6], specify proto objects as peaks spatial format of the foreground map. Basically, their method suggests that the proto object in image has a set of pixels that is determined by a continuing four linked neighborhood of the peak with foreground over a specific threshold. Hence, in their method, the majority of salient points are determined based on the spatial-based model, following that the foreground is distributing on the area around them, which mean that there are proto objects has been obtained from the foreground map.

Liu et al. (2007) [7], consider color spatial distribution, center-surround color histograms, and multi-scale contrast to evaluate pixel foreground. For the localization step, overall characteristics are combined in a depending random field causing a binary label map that isolates the foreground object away from the background. This technique proves a good performance. Nevertheless, it involves learning the Conditional Random Fields "CRF" that is generally computationally high priced.

Valenti et al. (2009) [8], presented a salient object detector in real time. With their approach, pixel foreground has been computed as a linear mix of three characteristics: curvedness, rarity and isocentricity of color edges. This method demonstrates edges and centers of the image constructions. To be able to distribute foreground objects inside connected regions, the researcher attempts average values of the foreground map and graph based segmentation inside each segment. In a localization step, effective sub-window search has been used.

Bruce and Tsotsos (2009) [9], outline that foreground depending on optimum information simples. They calculated the Shannon self-information via using the possibilities of the local image content within a patch provided the content of the entire image. Patches having unusual content tend to be more informative, and therefore salient.

Marchesotti et al. (2009) [10], suggest a detector for salient object that is depending on the hypothesis that images with a similar visual appearance will probably have salient objects with similar characteristics. In order to determine foreground inside a target image, the researchers train a classifier for the almost all identical images, with presented ground truth bounding boxes close to foreground objects. A 2 class problem for classification has been considered: non-salient class includes the background and the salient class contains salient objects. Every patch in the target image is categorized for being non-salient/salient. To be able to discover a salient object, the classifier output has been used to initialize the algorithm of iterative graph-cut. Therefore, the segment, which provides coverage for many of the salient pixels, is described to be salient object. The approach has been proven to acquire highly good results whenever annotated image data is obtainable. Nevertheless, the researchers additionally demonstrated how the technique is greatly dependent upon the grade of the retrieval step where the almost all related images are extracted.

Van de Sande et al. (2011) [11], has been modify hierarchical segmentation to locate beneficial candidates of object locations. Their work is dependent on a couple key suggestions. Firstly, objects could be of any size that can appear at any scale. As a result, a hierarchical segmentation method has been used and all segments all over the entire structure are considered. Secondly, to be able to take into account various object performances and image conditions, the outcome of numerous, complementary segmentations is integrated. This approach has shown itself powerful task for object localization.

3. Foreground Detection Theory

A different method has been provided for foreground and silence object detection approaches, which have been aimed to locating specific categories such as tables, cars, persons, airplane, etc.

There's two popular hypotheses for human visual attention: object-based and spatial-based attention. For the spatial-based hypothesis, interest is over a zoom lens or a spotlight that shifting our focus from one spatial spot to one other to sample surrounding. Consequently, all visual content inside a fovea-sized region close to those locations has been processed. On the other hand, the object-based attention hypothesis proposes that attention is in fact focused upon objects or what known as proto-objects, which is defined as a visual information unit that could be converted to an object-part or a plausible object. The idea suggests that in early pre-attentive stage, the visual system pre-segments a complicated scene into proto-objects Therefore, a foreground object has been identified mainly via the structure of the foreground map. Nevertheless, these map doesn't consider explicitly the information with regards to objects of an image. Thus, many of the modern foreground object detectors follow the hypothesis of spatial-based attention [5].

The idea of many foreground and silence object detection algorithms are able to go back towards the feature integration theory that posits that varieties of interest are dependable with regard to joining different features into knowingly experienced wholes. Later, a model of computational attention has been built based on a biologically credible architecture. It symbolizes the input image via the colour, orientation, and intensity channels, as well as can determine foreground maps by making use of center surround differences that are used together to make a final foreground map. In recent times, many research has been developed to design numerous foreground features characterizing foreground objects or regions [5, 12].

The majority of research generally follow the contrast (or center-surround difference) framework. In center-surround theory, the color histograms, calculated to present the center and the surround, which are utilized to find center surround dissimilarity. The information hypothesis standpoint is subjected to provide a mathematical formula, calculating the center surround divergence depending on feature information. The framework “center surround difference" could be examined to determine the foreground from region-based image information. The variance between straightaway neighboring regions and color histogram region are utilized to analyze the foreground rating. The global contrast approaches, computing the foreground map by comparing all area with each other's, intend to immediately calculate the global uniqueness. Using the regional contrast, feature color uniqueness and spatial distribution are brought to calculate the foreground ratings of regions. The foreground map is made by propagating the regions of foreground rates to the pixels. Other models have been also suggested for foreground computation, for example the center-bias i.e. the foreground object, commonly is founded on the center of an image. Object prior, such as auto-context cue, concavity context, the background prior and connectivity prior are usually researched for foreground computation. Example-based methods, seeking identical images of the input have been made for foreground object detection. The graphical model has been presented to merge general objectness and visual foreground together to identify objects. A method of low rank matrix recovery has been proposed for foreground object detection. A top-down method through joint dictionary learning and conditional random fields is introduced. Stereopsis is utilized for foreground analysis. Moreover, spectral analyzing for the frequency domain has been used to identify salient regions. In addition, there are many projects immediately checking if an image window includes an object. The universal objectness measure is based on mixing several image cues to evaluate the chance that a window includes an object. There is besides the previous methods there is attracts a lot of interests for independent object detection category that utilizes window symmetry, edge distribution and super pixel boundary integral to describe objectness, which trained to rate several object window candidates. Foreground object detection by composition has been determines when the content inside a window could be composed by neighbor regions. The approach of random forest regression has been used to immediately regress the object rectangle out of the foreground map. Another visual foreground research direction is eye fixation prediction. The latest research for detect foreground/silence objects are using statistically modeling, isocentric color and curvedness, making use of multitask sparsity pursuit, depth cues, implementing image histogram, combing top-down and bottom-up features, quaternion-based spectral analysis, task-specific visual attention, exploring patch rarities, and so on. There are several other foreground object detections, such as context-aware foreground detection that attempting to discover the image regions that present the scene [12].

4. Proposed Method

Our method of foreground object detection consists of following four steps: abstraction of an image, Tree structure, matrix breaking down and assignment foreground object.

1- Image Preprocessing: In this part, we first converted image to grayscale then remove image frame by detect if there is any frame edge for top, bottom, left and right of image. This done by set the maximum possible width of frame then evaluate the edge density, then removing farm if it founded.

2- Image Segmentation: In this part we have been segmented image by used SLIC (simple linear iterative clustering) superpixels algorithm based on worked done by Achanta et al. [13]. The SLIC algorithm are over-segment the image to a many super pixels (patches). Every patch has been represented via a feature vector, and every one of these feature vectors constitute the feature matrix.

3- Extract features from image: In this part, an input image is portioned into small and perceptually homogeneous features. We tend to first extract the low-level features that includes Gabor filter as described in [14], steerable pyramids and RGB color to generate a dimension feature description.

4- Create index tree: Together with super pixels, an index tree is created to encode construction info through hierarchical segmentation. As a result, we initially calculate the appreciation of each surrounding patch by get the first and second order reachable matrix.

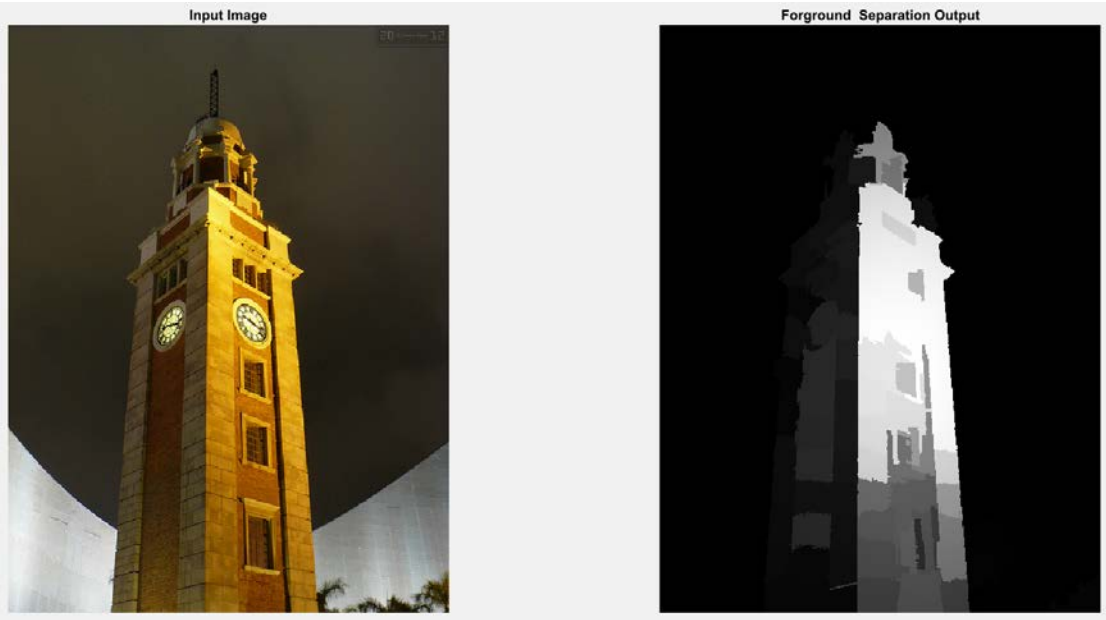
After that, we use the algorithm “graph based image segmentation” that described in [15] to combine spatial neighboring patches based on their affinity. The graph based image segmentation algorithm generates a granularity-increases segmentations sequence. In every granularity layer, segments are corresponding to nodes in the equivalent layer within the index tree. Especially, the granularity is controlling by an affinity threshold. At last, we get a fine-coarse hierarchical segmentation from the input image.

5- Structured matrix decomposition: after extracted features and created the index tree, we used “Structured Low-Rank Matrix Factorization” [16], to decompose feature matrix to structured-sparse component and a low-rank component. After collectively imposing the Laplacian regularization and structured-sparsity, the input feature matrix is decomposed straight to organized components structured-sparse component and a low-rank component.
6- Post-processing: Following decomposing, we proceed the outcome from the feature to some preprocessing algorithms in order to get improvements for foreground object. Depending on the structured matrix, we specify the function of straightforward foreground estimation for each patch.

After combining all patches together and executing contextbased propagation, we obtained the final forground map of the input imag.

5. Results

This part shows a number of chosen results verifying the usage of the foreground object detection approach, it is useful to analyze how our approach could be successfully extract important object from background. We have been select different cases such as birds, human, cars, jet fighters, etc.



ISSN: 2414 – 7419

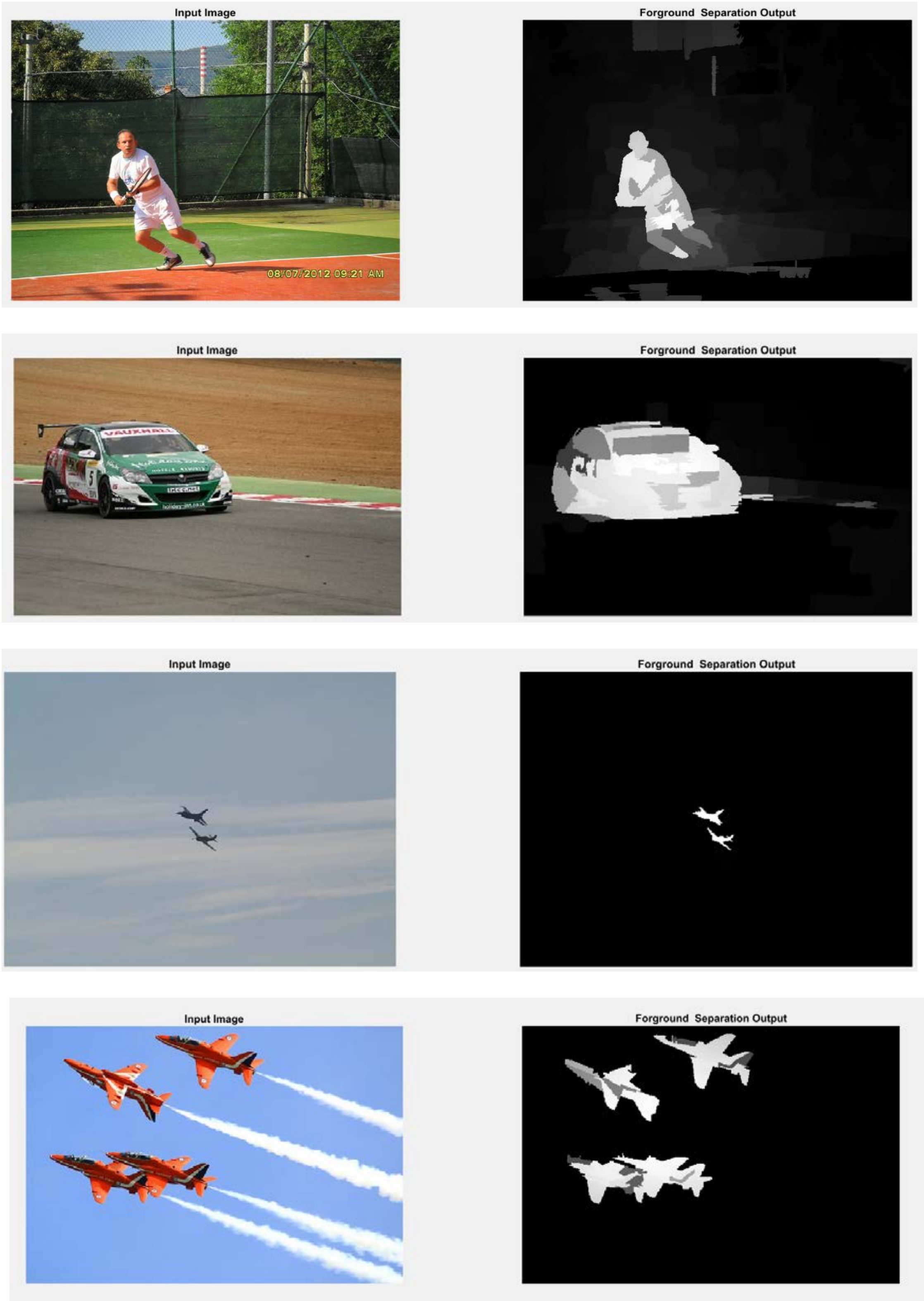


Figure 2: Prossessing Results. Left: origional image , Right: forground object detect

As shown from figure, it’s clear that this method efficiently removes foreground object from background with high details that can be detect easily and accurately with many recognition techniques.

6. Conclusion

In this work we have been proposed a model that dependent on the theory of object detection based attention and computes foreground of segments for image. This model can automatically detect the distribution of foreground objects that can be found inside the connected regions. Compared, our method doesn't depend on any learning; therefore, it doesn't need image retrieval and annotation. On the other hand, we wholly depend on the object-based consideration hypothesis and extract foreground objects straight from the image via feature grouping. That permits us to evaluate foreground at the level of proto object. Towards the best of our knowledge, we have been first apply it to task of foreground object detection. To determine integrated foreground, we determine the object details rarity. Automatically, the image locations that deviate via the remainder of an image needs to be foreground. To reflect the content of image, we have been employed visual color and word histograms. We're demonstrated this foreground depending on these features outperforms standard information maximization foreground and regular spectral recurring foreground for that task of detection foreground object. The test provide that this method is able to separate target foreground object from sense with high accuracy and good details that can be recognized effectively by using recognition techniques such as neural network.

7. Reference

[1] Daniilidis, K.; Maragos, P.; Paragios, N.; "Computer Vision -- ECCV 2010," 11th European Conference on Computer Vision, Heraklion, Crete, Greece, Proceedings, Part V. , September 5-11, pp.282–295, 2010.

[2] Rahtu, E.; Kannala, J.; Salo, M.; Heikkila, J.; " Segmenting Salient Objects from Images and Videos," Power Electronics, IEEE Transactions on , vol.27, no.12, pp.4858,4867, Dec. 2012.

[3] Marchesotti, L., Cifarelli; C., Csurka, G.; "A framework for visual saliency detection with applications to image thumbnailing,". In: ICCV, pp.2232-2239, 2009.

[4] Zhaoyu, P.; Pingping, L.; Changjiu, L., "Automatic Detection of Salient Object based on Multi-features," Second International Symposium on Intelligent Information Technology Application. IEEE, ISBN: 978-0-7695-3497-8, pp.437-441, January 2009.

[5] Cheng, M.; Mitra, N. J.; Huang, X.; Torr, P.; Hu, S.; " Global Contrast Based Salient Region Detection,". IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 37, Issue: 3, ISSN:0162-8828, pp.1-14, 2015.

[6] Walther, D.; and Koch, C.; "Modeling attention to salient proto-objects". Neural Networks, 19(9), pp.1395–1407, 2006.

[7] Liu, T.; Sun, J.; Zheng, N.N.; Tang, X.; Shum, H.Y.; "Learning to detect a salient object. In: IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol.33, No.2, pp.353-367, February 2011.

[8] Valenti, R.; Sebe, N.; Gevers, T.; "Image saliency by isocentric curvedness and color,". International Conference on Computer Vision, pp.2185-2192, 2009.

[9] Bruce, N.; Tsotsos, J.; "Saliency, attention, and visual search: An information theoretic approach,". Journal of Vision, 9(3):5, pp.1-24, 2009.

[10] Marchesotti, L., Cifarelli, C.; Csurka, G.; "A framework for visual saliency detection with applications to image thumbnailing," In: International Conference on Computer Vision, pp.2232-2239, 2009.

[11] Van de Sande, K.; Uijlings, J.R.; Gevers, T.; Smeulders, A.; "Segmentation as Selective Search for Object Recognition,". International Conference on Computer Vision, pp.1879-1886, 2011.

[12] Yanulevskaya, V.; Uijlings, J.; Geusebroek, J.M.; "Salient object detection: from pixels to segments", Preprint submitted to Image and Vision Computing, pp.1-28, October 2012.

[13] Achanta, R.; Shaji A., Smith K.; Lucchi, A.; Fua P., Susstrunk, S.; "SLIC Superpixels", EPFL Technical, Report 149300, pp.1-15, June 2010.

[14] Zujovic, J.; “Perceptual Texture Similarity Metrics”, Phd Thesis, Northwestern University, August 2011.

[15] Felzenszwalb, P.; Huttenlocher, D.; “Efficient graph-based image segmentation,” IJCV, Volum 59, No.2, pp.167-181, 2004.

[16] Haeffele, B.D.; Young, E.D.; Vidal, R.;“ Structured Low-Rank Matrix Factorization: Optimality, Algorithm, and Applications to Image Processing,” Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland USA, pp.4108-4117, 2014.