DOI: 10.24996/ ijs.2017.58.4A.18

Iraqi Journal of Science

ISSN: 0067-2904

Foreground Object Detection and Separation Based on Region Contrast

Samera Shams Hussein
Department of Computer Science, College of Education for Pure Science, Baghdad University, Baghdad, Iraq

Abstract

Foreground object detection is one of the major important 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, a proposed method has been presented for detection and separation foreground object from image or video in both of moving and stable targets. 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 Separation, Foreground Object, Region Contrast.

كشف وفصل الاجسام البارزة بالاعتماد على خوارزمية تباين المنطقة

سميرة شمس حسين
قسم علوم الحاسبات، كلية التربية للعلوم الصرفة ، جامعة بغداد، بغداد، العراق.

الخلاصة

تعد عملية كشف الأجسام البارزة من المهام الرئيسية المهمة في مجال الرؤيه الحاسوبية والتي تهدف الى كشف وجود الأجسام المهمة في الصورة الثابتة او في الصور المتسلسلة (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 تصميم طريقة لكشف وتجزئة 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 بتفاصيل 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].

Email: samera.s.h@ihcodeu.uobaghdad.edu.iq

Figure 1- Sample of images having an obvious forgroubd object, where the red rectangles 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 Separation, 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 separation , content aware image editing and adaptive compression, [4].

Extraction of salient or foreground objects in a scene relates to appropriate object retrieval and image Separation. 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 it 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 processing the important area of an image with leaving the remaining around unprocessed [5].

This paper focused to detect the important objects and separate them from moving and still sense with high details in order to increase the accuracy of recognition process. An approach for detect and segment (separate) foreground object has been proposed depending on region contrast. Since the theory of object based attention is recommending to start using the complex image separation towards proto-objects. Even though, common object separation is a difficult task, estimated at describing significant segments within an image via feature grouping is probably achievable. Finally, determines the foreground for each proto-object coupled with describes 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.

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 separation inside each segment. In same year, 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. The approach has been proven to acquire highly good results whenever annotated image data is obtainable.

Van de Sande et al. (2011) [11], presented modify hierarchical separation 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 separation method has been used and all segments all over the entire structure are considered. 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, this 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 color, 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 [12].

In summary, it can pointed from previous strategy, in order to get a good foreground separation results, the proposed model should be pass to several processing steps t: image Preprocessing, image separation, extracting features from image, creating index tree, then decomposed to organized components structured.

For purpose of image separation, the SUPERPIXEL algorithm is one of the effective algorithms for foreground image separation. It group pixels into perceptually meaningful atomic regions which can be used to replace the rigid structure of the pixel grid. They capture image redundancy, provide a convenient primitive from which to compute image features, and greatly reduce the complexity of subsequent image processing tasks. One of research that based on SUPERPIXEL algorithm is the simple linear iterative clustering (SLIC) that done by Achanta et al. [13], which adapts a k-means clustering approach to efficiently generate superpixels. SLIC algorithms are over-segment the image to many super pixels (patches). Every patch has been represented via a feature vector, and every one of these feature vectors constitutes the feature matrix. In this work for feature extraction, the Gabor texture method [14] is one good strategy that help in extraction effective area, which used color cues to obtain the color sub-saliency map of an image. Based on the responses of a Gabor filter, texture features are extracted from the image in order to obtain the texture sub-saliency map. Then, the color and texture sub-saliency maps are combined to obtain the final saliency map. The results of our experiments showed that the proposed method outperforms other state-of-the-art methods.

In order to create index tree, there are several approaches that can used in that purpose, one of them is the “graph based image separation” that proposed by Felzenszwalb [15]. The graph-based representation of the image combines spatial neighboring patches based on their affinity, which generates a granularity-increases separations sequence. The method measures the evidence for a boundary between two regions by comparing two quantities: one based on intensity differences across the boundary, and the other based on intensity differences between neighboring pixels within each region. Intuitively, the intensity differences across the boundary of two regions are perceptually important if they are large relative to the intensity differences inside at least one of the regions.

The final process is to reconstruction the image to generate output result that produce foreground object, this need to decompose feature matrix to structured-sparse component and a low-rank component. A “Structured Low-Rank Matrix Factorization” [16] has been used to that purpose. This method is using a projective tensor norm, which includes classical image regularizes such as total variation and the nuclear norm as particular cases. The method uses local minimum of the factorization that is sufficient to find a global minimum of the product, offering the potential to solve the factorization using a highly reduced set of variables.

4. Proposed Method

The proposed method of foreground object detection is based on following algorithm:

Foreground Object Detection Algorithm

Input:	RGB Image
Output:	Grayscale foreground object

Step 1: Start.
Step 2: Request Images from folder.
Step 3: Convert to gray scaled.
Step 4: Set the maximum possible width of image frame edge (if found) for top, bottom, left and right of image then evaluate the edge density to removing it.
Step 5: Segmented image by applying simple linear iterative clustering (SLIC) super-pixels algorithm.
Step 6: Create index tree (along with super-pixel process), to encode construction info through hierarchical separation.

i. Calculate appreciation of each surrounding patch through get the first and second order reachable matrix.
ii. Applying the "graph based image separation" algorithm.
iii. 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.
iv. Got fine-coarse hierarchical separation from the input image.

Step 7: Extract features from image

i. Portioned input image into small and perceptually homogeneous features
ii. Extract the low-level features that includes Gabor filter, steerable pyramids and RGB color
iii. Generate a dimension feature description.

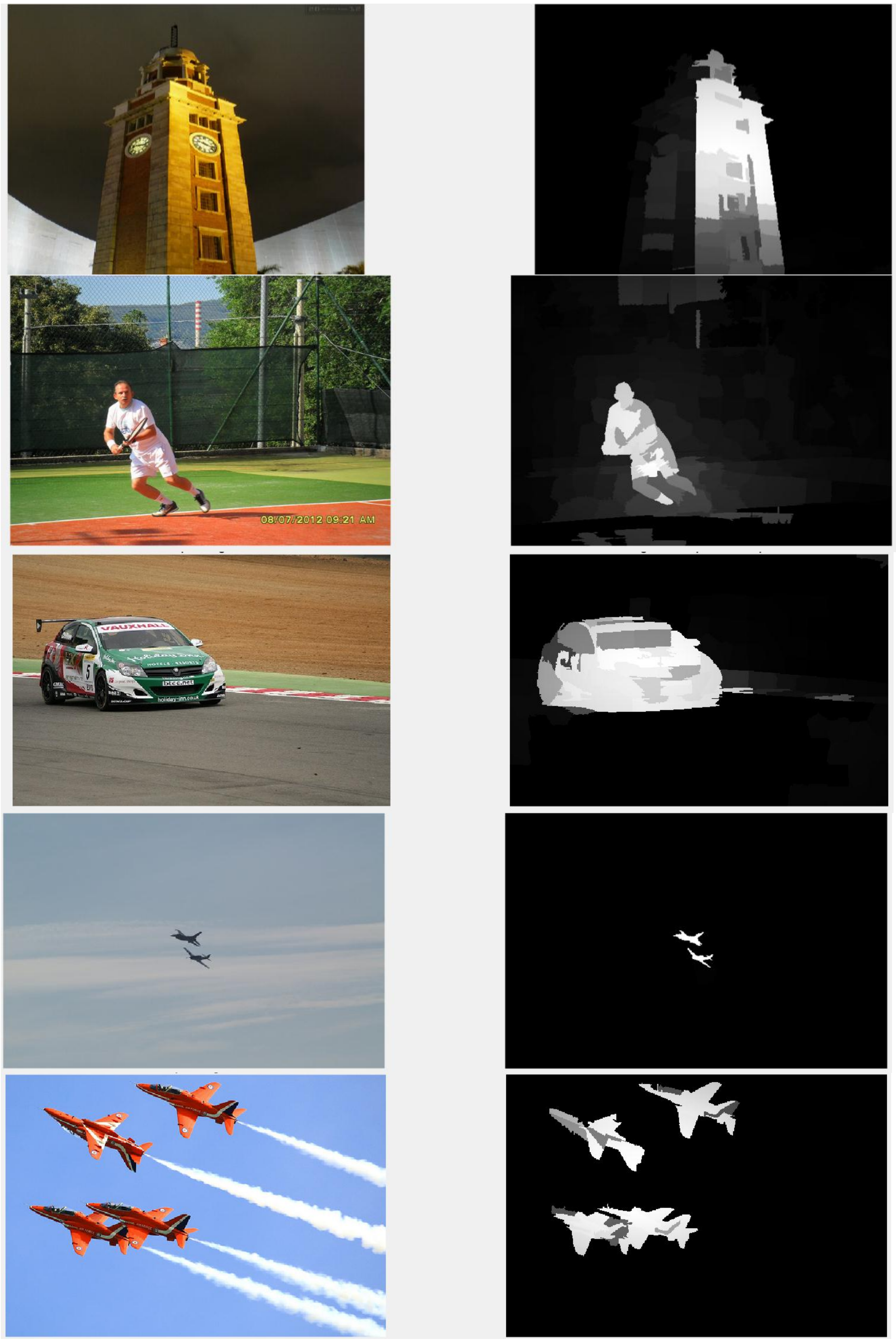
Step 8: Apply "Structured Low-Rank Matrix Factorization", to decompose feature matrix to structured-sparse component and a low-rank component.

i. Collectively imposing the Laplacian regularization and structured-sparsity,
ii. Input feature matrix is decomposed straight to organized components structured-sparse component and a low-rank component .

Step 9: Proceed the outcome from the feature to some preprocessing algorithms in order to get improvements for foreground object. Depending on the structured matrix, it would specify the function of straightforward foreground estimation for each patch.
Step 10: Combining all patches together and executing context based propagation
Step 11: Obtained the final foreground map of the input image.
Step 12: End.

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. Many different cases have been used for test, such as birds, human, cars, jet fighters, etc. to investigate the accuracy of separation in different situation, as illustrated in Figure- 2.



(a) **(b)**

Figure 2- Prossessing Results. (a): origional image , (b): forground object detection

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, a model that dependent on the theory of object detection based attention and computes foreground of segments for image has been proposed. This model can automatically detect the distribution of foreground objects that can be found inside the connected regions. Compared with learning based separation, our method doesn't depend on any learning strategy that needs data based to identify objects. On the other hand, this model is wholly depending on the object-based consideration hypothesis and extracts 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 knowledge, at first, apply this model to task of foreground object detection. To determine integrated foreground, it should 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, the visual color and word histograms have utilized for it. Finally, the foreground has been demonstrated 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.

References

1. Daniilidis, K., Maragos, P. and Paragios, N. **2010**. Computer Vision -- ECCV 2010. 11th European Conference on Computer Vision, Heraklion, Crete, Greece, Proceedings, Part V. , 282–295.
2. Rahtu, E., Kannala, J., Salo, M. and Heikkila, J. **2012**. Segmenting Salient Objects from Images and Videos. *Power Electronics, IEEE Transactions on*, **27**(12): 4858-4867.
3. Marchesotti, L., Cifarelli; C. and Csurka, G. **2009**. A framework for visual saliency detection with applications to image thumbnailing. In: ICCV, 2232-2239.
4. Zhaoyu, P., Pingping, L. and Changjiu, L. **2009**. Automatic Detection of Salient Object based on Multi-features. *Second International Symposium on Intelligent Information Technology Application IEEE*, ISBN: 978-0-7695-3497-8,437-441.
5. Cheng, M., Mitra, N. J., Huang, X., Torr, P. and Hu, S. **2015**. Global Contrast Based Salient Region Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**(3):1-14.
6. Walther, D. and Koch, C. **2006**. Modeling attention to salient proto-objects. *Neural Networks*, **19**(9):1395–1407.
7. Liu, T., Sun, J., Zheng, N.N.,Tang, X. and Shum, H.Y. **2011**. Learning to detect a salient object. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, **33**(2): 353-367.
8. Valenti, R., Sebe, N. and Gevers, T. **2009**. Image saliency by isocentric curvedness and color. *International Conference on Computer Vision*, 2185-2192.
9. Bruce, N. and Tsotsos, J. **2009**. Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, **9**(3): 1-24.
10. Marchesotti, L., Cifarelli, C. and Csurka, G. **2009**. A framework for visual saliency detection with applications to image thumbnailing. *International Conference on Computer Vision*, 2232-2239.
11. Van de Sande, K., Uijlings, J.R., Gevers, T. and Smeulders, A. **2011**. Segmentation as Selective Search for Object Recognition. *International Conference on Computer Vision*, 1879-1886.
12. Yanulevskaya, V., Uijlings, J. and Geusebroek, J. M. **2012**. Salient object detection: from pixels to segments. *Image and Vision Computing*,1-28.
13. Achanta, R., Shaji A., Smith K., Lucchi, A., Fua P. and Susstrunk, S. **2010**. SLIC Superpixels. EPFL Technical, Report 149300, 1-15.
14. Zujovic, J. **2011**. Perceptual Texture Similarity Metrics", Phd Thesis, Northwestern University.
15. Felzenszwalb, P. and Huttenlocher, D. **2004**. Efficient graph-based image Segmentation. *IJCV*, **59**(2): 167-181.
16. Haeffele, B.D., Young, E.D. and Vidal, R. **2014**. Structured Low-Rank Matrix Factorization: Optimality, Algorithm, and Applications to Image Processing. Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland USA, 4108-4117.