



# Performance Analysis of Structured Matrix Decomposition with Contour Based Spatial Prior for Extracting Salient Object from Complex Scene

Ramesh Bhandari<sup>1</sup>, Sharad Kumar Ghimire<sup>2</sup>

Department of Electronics and Computer Engineering

Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal

Corresponding author: <sup>1</sup>eramesh.bhandari@gmail.com and <sup>2</sup>skghimire@ioe.edu.np

Received: Dec 20, 2018

Revised: April 3, 2019

Accepted: April 5, 2019

---

**Abstract:** Automatically extracting most conspicuous object from an image is useful and important for many computer vision related tasks. Performance of several applications such as object segmentation, image classification based on salient object and content based image editing in computer vision can be improved using this technique. In this research work, performance of structured matrix decomposition with contour based spatial prior is analyzed for extracting salient object from the complex scene. To separate background and salient object, structured matrix decomposition model based on low rank matrix recovery theory is used along with two structural regularizations. Tree structured sparsity inducing regularization is used to capture image structure and to enforce the same object to assign similar saliency values. And, Laplacian regularization is used to enlarge the gap between background part and salient object part. In addition to structured matrix decomposition model, general high level priors along with biologically inspired contour based spatial prior is integrated to improve the performance of saliency related tasks. The performance of the proposed method is evaluated on two demanding datasets, namely, ICOSEG and PASCAL-S for complex scene images. For PASCAL-S dataset precision recall curve of proposed method starts from 0.81 and follows top and right-hand border more than structured matrix decomposition which starts from 0.79. Similarly, structural similarity index score, which is 0.596654 and 0.394864 without using contour based spatial prior and 0.720875 and 0.568001 using contour based spatial prior for ICOSEG and PASCAL-S datasets shows improved result.

**Keywords:** Salient object detection, structured matrix decomposition, contour based spatial prior, salient object extraction

---

## 1. Introduction

Extracting proper salient object from complex scene has been a fundamental research problem in vision perception for a very long period of time. Visual data is being generated in every minute. Growth in image data has led to a new challenge for processing these images and extracting proper information from it. Human visual system has tremendous capacity to extract information of interest from a given scene using their perceptual and cognitive knowledge. But, one peculiar problem for computer vision algorithms is to detect and extract object of interest automatically from a given complex scene. Saliency alludes to the recognition of crucial visual information for

further processing. Identifying and segmenting the most conspicuous object from the scene referred to as salient object detection and salient object extraction is an important branch of visual saliency.

Visual saliency finds its applications in the field of computer vision for several tasks such as object detection, object extraction, object recognition and automatic image editing. For the accomplishment of this task, several salient object detections models have been proposed. On the basis of use of prior knowledge, current methods fall into two classes, namely, bottom-up and top-down method. However, due to diversity of object types encountered in real world applications, performance of these models suffer giving part of salient object or mixing the salient object with background of an image. Recently many models have been proposed to combine both bottom-up cues with top-down cues to increase the performance of saliency related tasks. This research work mainly focuses on analyzing performance of structured matrix decomposition and contour based spatial prior for extracting salient object form complex scene.

## **2. Literature Review**

Based on whether prior knowledge is used or not, different methods for saliency related task have been proposed in the past. Current models mainly fall in two classes, namely, bottom-up and top-down model. Low level features such as color, texture and location are used in bottom-up models. Research work [6] and [10] are based on bottom-up approach. However, these methods suffer from problem that detected salient regions may only contain parts of the target objects, or be easily mixed with background and resulting background part in the detected object. On the other hand, methods [11], [25] and [5] uses high level knowledge, such as context, semantics and background priors for computing saliency and are referred to as top-down models. However, generalization and scalability of both top-down and bottom-up models suffer from diversity of object in the complex scene.

Combining both low-level feature and high-level feature for detecting salient object efficiently from images is new trend. Recent state-of-the-art methods [29, 20, 24, 14, 22, 3, 17] are based on both top-down cues and bottom-up cues to assist efficient salient object detection and extraction. Structured matrix decomposition method [17] based on low-rank matrix recovery models has shown decent accuracy for detecting salient object, where matrix is decomposed into a low rank matrix representing image background and a sparse matrix identifying salient objects. Representative series of methods [29, 20, 24, 14, 22, 3, 17] are based on low rank matrix recovery theory [4]. Methods [29, 20, 24, 14, 22] lack performance when there are similarities between the salient objects and background or when background is complicated. To improve performance, structured matrix decomposition with another cue known as contour based spatial prior is introduced in research work [3]. In this research work, performance of structured matrix decomposition along with contour based spatial prior is analyzed for complex scene datasets to detect and extract salient object from the given complex scene.

## **3. Methodology**

Visual saliency is the quality of a region in an image to be noticeable from the rest of the part of an image and grab our attention. In other words, salient objects are identified as a dominant object in the given image. For example, people, animals, cars, flowers, or any other structures that is dominant in the given image and grabs our attention. Given and input image, perceptually homogenous regions are found based on low rank matrix recovery models [20]. Having found

perceptually homogenous elements, matrix of low level features ( $F$ ) is constructed and then simple linear iterative clustering (SLIC) algorithm [18] is performed to over-segment the image to generate super-pixels. On the other hand, high level features like - color prior, location prior, background prior, contour based spatial prior is computed and index tree to encode structure information is constructed. Graph based segmentation algorithm [8] is implemented to merge spatially neighboring patches. Feature matrix is then decomposed into low rank part ( $L$ ) and sparse part ( $S$ ) using structured matrix decomposition model. From an input image, contour based spatial prior is computed and integrated to the structured matrix decomposition by multiplying each component in feature matrix ( $F$ ) to guide matrix decomposition. After decomposition, saliency map is calculated by transferring structured sparse part ( $S$ ) from feature domain to spatial domain. Finally, with the help of original image salient object is extracted by making decision whether a pixel corresponding to original image falls in detected salient object or not.

### 3.1 Structured Matrix Decomposition

Input image  $I$  is first partitioned into  $N$  different super-pixels  $P = \{P_1, P_2, P_3 \dots, P_N\}$ . For an input image shown in Fig. 1, super-pixels generated using simple linear iterative clustering algorithm [13] is shown in Fig. 2.



Fig. 1: Input Image

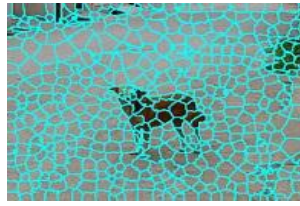


Fig. 2: Generated Super-pixel for Input Image

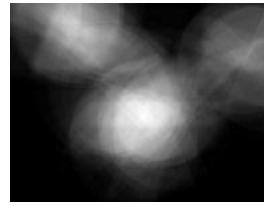


Fig. 3: Computed Contour Based Spatial Prior

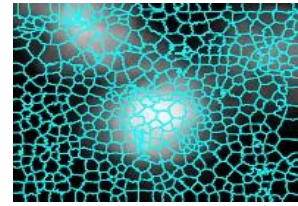


Fig. 4: Mapped N patches over contour based spatial prior

For each super pixel  $P_i$  a  $D$  dimensional feature (here  $D = 53$ ) vector is extracted and denoted as  $f_i$ . Feature vector forms a matrix representation of  $I$ , denoted as  $F = \{f_1, f_2, f_3 \dots, f_N\}$ . Feature matrix ( $F$ ) consists of low level feature including RGB color, steerable pyramids [21] and Gabor filter [9] to construct 53-dimensional feature representation. On top of super-pixels, an index tree is constructed to encode structure information via hierarchical segmentation. Then graph based segmentation algorithm [8] is applied to merge spatially neighboring patches. When both the feature matrix ( $F$ ) and index tree are ready, structured matrix decomposition model is implemented as proposed in [17].

We further extend the structured matrix decomposition based salient object detection to integrate high level priors and contour based spatial prior. Three types of high level priors, namely, location prior, color prior and background priors are used. Specifically, the location prior is generated by Gaussian distribution based on the distance of the pixels from an image center and it is denoted by  $lp$ . The color prior used here is same as [20], which measures human eye sensitivity to red and yellow color. Computed color prior is denoted by  $cp$ . The background prior calculates the probabilities of image regions connected to image boundaries [28] and is denoted by  $bp$ . These three high level priors are combined by taking weighted sum to get high level prior denoted by  $hp$  as follows:

$$hp = w_1 * lp + w_2 * bp + w_3 * cp \quad (1)$$

where  $w_1$ ,  $w_2$  and  $w_3$  are weight given to each prior. Values of  $w_1$ ,  $w_2$  and  $w_3$  lies between 0 and 1, and satisfies the following relations:

$$w_1 + w_2 + w_3 = 1 \quad (2)$$

Contour based spatial prior obtained from biologically inspired model is also computed and later integrated with obtained high level prior ( $hp$ ) in Equation (1) to get final high level prior map ( $fp$ ).

### 3.2 Contour Based Spatial Prior

To estimate contour based spatial prior, we first estimate the edge response and corresponding orientations by using, biologically inspired method, efficient color boundary detection with color-opponent mechanisms proposed in [16, 26] and contour based spatial prior is computed using method proposed in [13, 27, 23]. Efficient color boundary detection with color-opponent mechanisms imitates the working of human visual system to identify the edges from an image. This mechanism involves processing of image information in three layers, namely, Cone layer, Ganglion/Lateral Geniculate Nucleus (LGN) layer, and Cortex layer, of our visual system. Obtained contour based spatial prior denoted by  $cbsp$  for input image shown in Fig. 1 is shown in Fig. 3.

### 3.3 Integrating Contour Based Spatial Prior

Since contour based spatial prior is determined pixel wise while high level priors are computed for super pixels. To integrate contour-based spatial prior ( $cbsp$ ) with high level prior ( $hp$ ) obtained earlier in Equation (1), we map generated  $N$  patches (super-pixels) over contour based spatial prior and average value of all the pixels value within each patch is calculated to obtain final contour based spatial prior. Final contour based spatial prior is denoted by  $fcbsp$ . Mapped  $N$  patches (super-pixels) over contour based spatial prior for input image shown in Fig. 1 is shown in Fig. 4. After finding final contour based spatial prior ( $fcbsp$ ), it is combined with previously obtained high level prior ( $hp$ ) by taking weighted sum to get final high level prior denoted by  $fp$  as follows:

$$fp = m_1 * hp + m_2 * fcbsp \quad (3)$$

where,  $m_1$  and  $m_2$  are weight given to each prior. Values of  $m_1$  and  $m_2$  lies between 0 and 1, and satisfies  $m_1 + m_2 = 1$ . Final high level prior,  $fp \in [0,1]$  for each patch  $P_i$  indicates the likelihood that  $P_i$  belongs to a salient object based on high level information. This prior is encoded into structured matrix decomposition by multiplying each component in feature matrix ( $F$ ) to guide matrix decomposition.

### 3.4 Salient Object Detection and Extraction

After decomposition of feature matrix ( $F$ ) into low-rank part ( $L$ ) and structured sparse part ( $S$ ), saliency map is calculated by transferring structured sparse part ( $S$ ) from feature domain to spatial domain as in [17]. For an input image shown in Fig. 1 detected salient object i.e. saliency map is shown in Fig. 5.



Fig. 5: Saliency Map

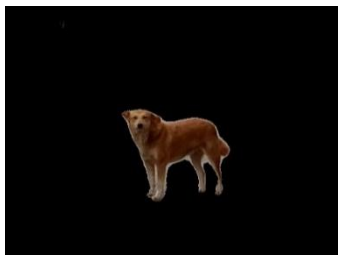


Fig. 6: Extracted Saliency Image

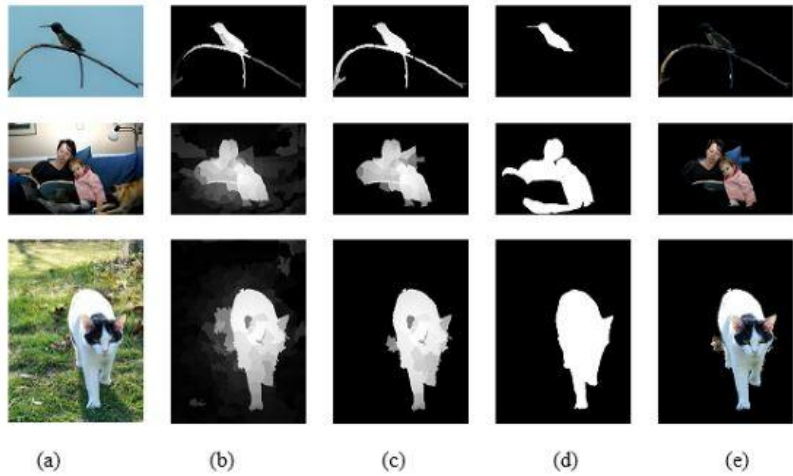


Fig. 7: Output Images (a) Original images (b) Detected images by structured matrix decomposition (SMD). (c) Detected images by SMD with CBSP. (d) Ground truth (e) Extracted images using SMD with CBSP

Finally, salient object is extracted by making decision whether a pixel corresponding to original image falls in detected salient object (saliency map) or not. This can be done by taking each pixel from original image and checking whether that pixel belongs to detected salient region or not. If that pixel belongs to detected salient region, then value of that pixel is taken otherwise some standard color which will be background color in extracted image is set. Extracted salient object from an input image shown in Fig. 1 using structured matrix decomposition with contour based spatial prior (SMD with CBSP) is shown in Fig. 6.

## 4. Results and Discussion

### 4.1 Experimental Setup

To evaluate the performance of proposed method to detect and extract salient object from an image, series of experiments are conducted using different images and standard datasets involving complex scenes. Experimental analysis on ICOSEG [2] and PASCAL-S [15] datasets, which mainly consists complex scene, is performed to evaluate metrics like precision recall (PR) curve and Structural Similarity Index (SSIM) score for comprehensive evaluation. Different parameters for the implementation of this research work are set as follows. While computing contour based spatial prior, in color opponent mechanisms value of sigma ( $\sigma$ ) is set to 1.5, cone input weights are set to -0.6 and 1, and number of orientation for color opponent mechanisms are set to 8. Similarly, value of  $\gamma$  is chosen to be 0.5 based on physiological findings [7] and [19]. Similarly, while integrating high level priors and contour based spatial prior values of weight  $w_1$ ,  $w_2$  and  $w_3$  prior are set to 1/3 and value of weights  $m_1$  and  $m_2$  while integrating high level prior with final contour based spatial prior is set to 1/2.

### 4.2 Experimental Results

To validate the effectiveness of proposed method, first, visual analysis between the output of structured matrix decomposition (SMD) and structured matrix decomposition with contour based

spatial prior (SMD with CBSP) is performed for various complex scene from ICOSEG [2] and PASCAL-S [15] datasets in Fig. 7. The detection and extraction results for three images are shown. For these three examples, output of proposed method is closer to ground truth as compared to output of structured matrix decomposition model.

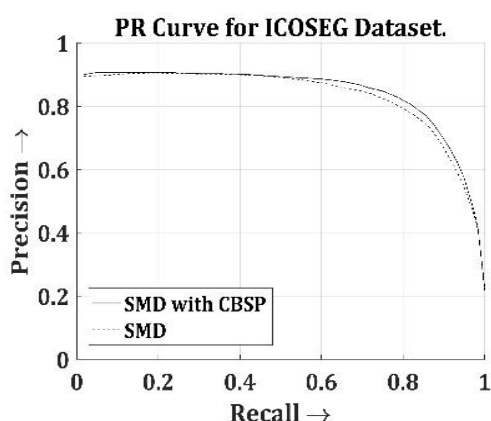


Fig. 8: PR curve for ICOSEG [20] Dataset

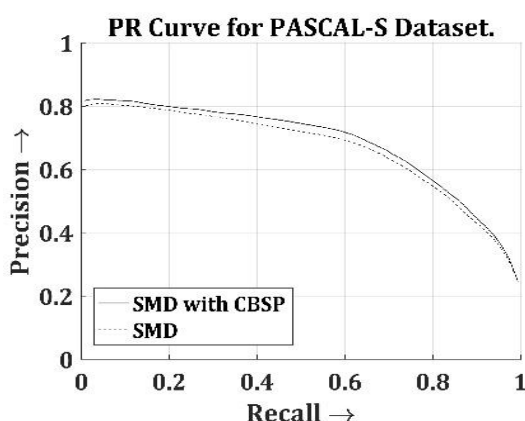


Fig. 9: PR curve for PASCAL-S [8] Dataset

Similarly, for comprehensive analysis, quantitative comparison on two different datasets using different evaluation metrics like precision-recall (PR) curve and structural similarity index (SSIM) is performed. Precision-recall (PR) curve is also plotted for ICOSEG [2] dataset using SMD and SMD with CBSP shown in Fig. 8 and similarly for PASCAL-S [15] dataset shown in Fig. 9. For both datasets PR curve obtained using SMD with CBSP follows the right-hand border and the top border more than using SMD indicating SMD with CBSP has better accuracy. As complementary to PR curves, structural similarity index (SSIM) is evaluated and tabulated in Table 1 for quantitative analysis. It determines the similarity between the ground truth and saliency map. For the both datasets SSIM is more using SMD with CBSP than SMD indicating detected and extracted salient object is much closer to ground truth.

Table 1: Comparison of Structural Similarity Index (SSIM)

Datasets	Structural Similarity Index (SSIM)	
	SMD	SMD with CBSP
PASCAL-S	0.596654	0.720875
ICOSEG	0.394864	0.568001

## 5. Conclusion

This research work aims to analyze the performance of structured matrix decomposition model with contour based spatial prior obtained from biologically inspired framework for extracting salient object from complex scene. The proposed method has decent accuracy for detecting and extracting salient object from complex scene and shows improvement over structured matrix decomposition model without using contour based spatial prior. Its effectiveness is demonstrated by comprehensive experiments using precision-recall curve and structural similarity index on two



widely used datasets targeting complex scene. In addition, its performance is also presented for visual analysis on both datasets. Additionally, method to integrate contour based spatial prior to the structured matrix decomposition model is also presented. Exploring more robust high-level prior may merit further study for saliency related tasks.

## References

- [1] Achanta R, Shaji A, Smith K, Lucchi A, Fua P and Susstrunk S (2012), Slic superpixels compared to state of-the-art superpixel methods, *IEEE TPAMI*, **34(11)**:2274-2282.
- [2] Batra D, Kowdle A, Parikh D and Luo Jand Chen T (2011), Interactively co-segmenting topically related images with intelligent scribble guidance, *IJCV*, **93(3)**: 272-292.
- [3] Bhandari R and Ghimire SK (2017), Improved Salient Object Extraction using Structured Matrix Decomposition and Contour Based Spatial Prior, *Proceedings of IOE Graduate Conference*, Kathmandu, Nepal.
- [4] Candes E, Li X, Ma Y and Wright J (2011), Robust principal component analysis? *J. ACM*, **58(3)**, Article No. 11.
- [5] Cheng M, Warrell J, Lin W, Zheng S, Vineet V and Crook N (2013), Efficient salient region detection with soft image abstraction, *International Conference on Computer Vision*, Sydney, Australia.
- [6] Cheng M, Zhang G, Mitra NJ, Huang X and Hu S (2011), Global contrast based salient region detection, *Conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, USA.
- [7] Conway BR (2010), Advances in color science: from retina to behavior, *The Journal of Neuroscience*, **30(45)**: 14955 – 14963.
- [8] Felzenszwalb P and Huttenlocher D (2004), Efficient graph- based image segmentation, *International Journal of Computer Vision*, **59(2)**: 167-181.
- [9] Feichtinger HG and Strohmer T (1998), *Gabor analysis and algorithms: theory and applications*, Springer.
- [10] Itti L, Koch C and Niebur E (1998), A model of saliency based visual attention for rapid scene analysis, *IEEE TPAMI*, **20(11)**: 1254-1259
- [11] Jiang H, Wang J, Yuan Z, Liu T, Zheng N and Li S (2011), Automatic salient object segmentation based on context and shape prior, *Proceedings of the British Machine Vision Conference*, University Dundee, UK.
- [12] Johnson EN, Hawken MJ and Shapley R (2001), The spatial transformation of color in the primary visual cortex of the macaque monkey, *Nature Neuroscience*, **4(4)**: 409-416.
- [13] Judd T, Ehinger K, Durand F and Torralba A (2009), Learning to predict where humans look, *IEEE International Conference on Computer Vision*, Kyoto, Japan.
- [14] Lang C, Liu G, Yu J and Yan S (2012), Saliency detection by multitask sparsity pursuit, *IEEE TIP*, **21(3)**: 1327-1338.
- [15] Li Y, Hou X, Koch C, Rehg J and Yuille A (2014), The secret of salient object segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, USA.
- [16] Papari G, Campisi P, Petkov N and Neri A (2007), A biologically motivated multiresolution approach to contour detection, *EURASIP Journal on Applied Signal Processing*, Vol. 2007, Article Id 071828.
- [17] Peng H, Li B, Ling H, Hu W, Xiong W and Maybank SJ (2016), Salient object detection via structured matrix decomposition, *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, **39(4)**: 818-832.
- [18] Petkov N and Westenberg MA (2003), Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition, *Biological Cybernetics*, **88(3)**: 236-246.

- [19] Shapley R and Hawken M (2011), Color in the cortex—single-and double-opponent cells, *Vision Research*, **51(7)**: 701-717.
- [20] Shen X and Wu Y (2012), A unified approach to salient object detection via low rank matrix recovery, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, USA.
- [21] Simoncelli EP and Freeman WT (1995), The steerable pyramid: A flexible architecture for multi-scale derivative computation, *International Conference on Image Processing*, Washington, DC, USA.
- [22] Tang C, Wu J, Changqing, Wang P and Li W (2017), Salient object detection via weighted low rank matrix recovery, *Signal Processing Letters (LSP)*, **24(4)**:490-494.
- [23] Tatler BW (2007), The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions, *J. Vis*, **7(14)**/4: 1-17.
- [24] Yan J, Zhu M, Liu H and Liu Y (2010), Visual saliency detection via sparsity pursuit, *IEEE Signal Processing Letters (SPL)*, **17(8)**: 739-742.
- [25] Yan Q, Xu L, Shi J and Jia J (2013), Hierarchical saliency detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, OR, USA.
- [26] Yang K, Gao S, Li C and Li Y (2013), Efficient color boundary detection with color-opponent mechanisms, *IEEE Conference on Computer Vision & Pattern Recognition, Portland, OR, USA*.
- [27] Yang KF, Li H, Li CY and Li YJ (2016), A unified framework for salient structure detection by contour-guided visual search, *IEEE Transactions on Image Processing (TIP)*, **25(8)**: 3475-3488.
- [28] Zhu W, Wei Y, Liang S and Sun J (2014), Saliency optimization from robust background detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA.
- [29] Zou W, Kpalma K, Liu Z and Ronsin J (2013), Segmentation driven low-rank matrix recovery for saliency detection, *Proceedings of the British Machine Vision Conference (BMVC)*, Bristol University, UK.