

Improved Salient Object Extraction using Structured Matrix Decomposition and Contour Based Spatial Prior

Ramesh Bhandari ^a, Sharad Kumar Ghimire ^b

^{a, b} Department of Electronics & Computer Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal
Corresponding Email: ^a eramesh.bhandari@gmail.com, ^b skghimire@ioe.edu.np

Abstract

Salient object detection is a useful and important technique in computer vision, which can scale up the efficiency of several applications such as object detection, object segmentation and content based image editing. In this research work, an improved technique of salient object extraction using structured matrix decomposition and contour based spatial prior is implemented. In order to assist efficient background-salient objects separation, structured matrix decomposition model with two structural regularizations namely, tree-structured sparsity-inducing regularization and Laplacian regularization are used. Tree-structured sparsity-inducing regularization captures the image structure and enforces the same object to have similar saliency values and Laplacian regularization enlarges the gap between background and the salient object. In addition, integrating general high level priors and contour based spatial prior obtained from biologically inspired model is implemented to improve the efficiency of saliency related tasks. The performance of the proposed method is evaluated on two demanding datasets, namely, ICOSEG and PASCAL-S, including single object, multiple object and 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 and similarly mean absolute error is also less by 0.019. Similarly, both visual and comprehensive evaluation using receiver operating characteristics curve and mean absolute error for other dataset also shows improved result.

Keywords

Salient Object Detection, Structured Matrix Decomposition, Contour Based Spatial Prior, Salient Object Extraction

1. Introduction

Determining visual saliency has been a fundamental research problem in vision perception for a very long period of time. It 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. Our society has become more technologically advanced than ever but our most advanced machines and computers still struggle at describing what it sees in series of photos. This is due to fact that human visual system has innate capability to extract crucial information from a scene but for machine same task is difficult. Little by little, we are giving sight to the machines or computers using state-of-the-art methods from computer vision.

Due to its wide range of applications in computer vision, such as object detection, object extraction, object recognition and automatic image editing, salient object extraction has grabbed attention over last decade. To accomplish the task of salient object detection many saliency models have been proposed. Based on whether prior knowledge is used or not, current models fall in two classes, namely, bottom-up and top-down. Bottom-up models [1] and [2] are based on low level features such as color, texture, location etc. The main downside of bottom-up methods is detected salient regions may only contain parts of salient object. On the other side, top-down salient object detection methods [3], [4] and [5] use high level human perceptual knowledge to identify potential region of salient object. These high level human perceptual knowledge are generally context, semantics and background knowledge which guide the saliency related tasks.

However, efficiency of such models suffer from diversity of object types which are encountered in real world applications. Recent development for salient object detection shows that trend is to combine bottom-up cues with top-down cues to increase the efficiency of saliency related tasks.

This research work mainly focuses on saliency related tasks, salient object detection and salient object extraction, and improved salient object extraction is implemented using structured matrix decomposition and contour based spatial prior.

2. RELATED WORK

In the past, to detect salient object many saliency models have been proposed. Based on whether prior knowledge is used or not, current models fall in two classes, namely, bottom-up and top-down. Bottom-up models [1] - [2] are based on low level features such as color, texture and location. The main limitation of these methods are that detected salient regions may only contain parts of the target objects, or be easily mixed with background. On the other hand, top-down models [3], [4] and [5] are based on high-level human perceptual knowledge, such as context, semantics and background priors, to guide the subsequent saliency computation. However, generalization and scalability of these models suffers from high diversity of object types. Fixation prediction methods proposed by L. Itti, C. Koch, and E. Niebur [1] mainly focus on high – contrast boundaries but ignores object surfaces and shapes. In contrast, methods proposed by Q. Yan, L. Xu, J. Shi, and J. Jia [4] may lacks its performance when there are no dominant objects in the scene.

Recent state-of-the-art method for saliency related task shows that trend is to combine bottom-up cues with top-down cues [6], [7], [8], [9], [10] and [11] to assist efficient salient object detection. Low-rank recovery models have shown potential for salient object detection, where matrix is decomposed into a low rank matrix representing image background and a sparse matrix identifying salient objects. Representative series of methods [6] - [11] are based on low rank matrix recovery theory [12]. Methods [6] - [10] lack performance when there are similarities between the salient objects and background or when background is complicated.

Generally, these LR-based saliency detection methods assume that an image can be represented as a combination of a highly redundant information part (e.g., visually consistent background regions) and a sparse salient part (e.g., distinctive foreground object regions). The redundant information part usually lies in a low dimensional feature subspace, which can be approximated by a low-rank feature matrix. In contrast, the salient part deviating from the low-rank subspace can be viewed as noise or errors, which are represented by a sparse sensory matrix. Peng et. al. [11] proposes salient object detection via structured matrix decomposition to overcome these problems which shows its decent accuracy for complex scene images.

In this research, integrating contour based spatial prior with method proposed in [11] is proposed to improve the task of detection which ultimately results better for salient object extraction.

3. METHODOLOGY

We define salient object as a dominant object (e.g., people, animals, cars, flowers, or any other structures that is dominant on image) in the given image. Having an input image, it is first partitioned to perceptually homogenous elements based on low rank matrix recovery model for salient object detection [7]. After partitioning, Feature matrix (F) is computed which consists low level features and then simple linear iterative clustering (SLIC) algorithm [13] 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 [14] is implemented to merge spatially neighboring patches. After obtaining feature matrix (F) and index tree, structured matrix decomposition to decompose feature matrix (F) into low-rank part (L) and structured sparse part (S) is applied. 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

Given an input image I , it is partitioned into N patches (super pixels) $P = P_1, P_2, P_3, \dots, P_N$. For an input image shown in Figure 1 generated super-pixels using simple linear iterative clustering [13] is shown in Figure 2.



Figure 1: Input Image.

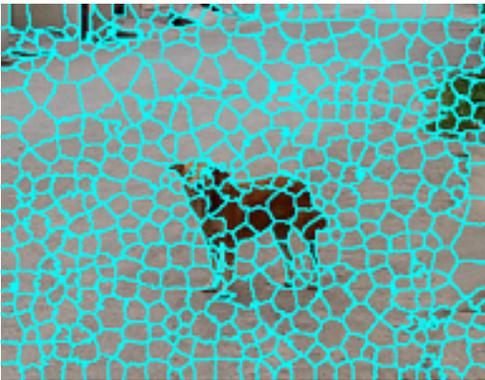


Figure 2: Generated Super-pixel for Input Image.

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 $F = f_1, f_2, \dots, f_N$. Feature matrix (F) consists of low level feature including RGB color, steerable pyramids [15] and Gabor filter [16] 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 [14] is applied to merge spatially neighboring patches. When both the feature matrix (F) and index tree are ready, structured matrix decomposition model as proposed in [11] is implemented as follows:

$$\min_{L,S} \Psi(L) + \alpha\Omega(S) + \beta\Theta(L,S) \quad s.t. \quad F = L + S \quad (1)$$

Where $\Psi(\cdot)$ is a low-rank constraint to allow identification of the intrinsic feature subspace of the redundant background patches, $\Omega(\cdot)$ is a structured sparsity regularization to capture the spatial and feature relations of patches in S , $\Theta(\cdot, \cdot)$ is an interactive regularization term to enlarge the distance between the L and S , and α, β are positive tradeoff parameters. Three interactive regularizations, Low-rank regularization for image background, Structured-sparsity regularization for salient objects and Laplacian regularization to enlarge the gap between L and S are performed as proposed in [11].

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 [7], 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 [17] 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:

$$hp = w_1 * lp + w_2 * bp + w_3 * cp \quad (2)$$

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 $w_1 + w_2 + w_3 = 1$. Contour based spatial prior obtained from biologically inspired model is also computed and later integrate contour based spatial prior with obtained high level prior (hp) in Equation (2) 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 [18] and contour based spatial prior is computed using method proposed in [19]. Efficient color boundary detection with color-opponent mechanisms imitate the working of human visual system to identify the edges from an images. This mechanism involves processing of an image information in three layers, namely, Cone layer, Ganglion/Lateral Geniculate Nucleus (LGN) layer, and Cortex layer, of our visual system.

In the Cone layer, three color components namely red, green and blue from an input color image is extracted and denoted by r , g and b respectively. To implement color opponent mechanisms, yellow component of a color image denoted by y is computed by taking average of red and green component as follows:

$$y = \frac{r + g}{2} \quad (3)$$

In order to obtain local color information, extracted four components from an image are subjected to Gaussian filters which simulates the receptive fields of the cones in the retina as suggested by [20] and [21]. Here, all the four components are smoothed with Gaussian filters having same standard deviation (σ). The output of the four components after simulated by Gaussian filters are output of Cone layer and are denoted by R_g , G_g , B_g , and Y_g . Now the outputs of Cone layer are passed to Ganglion/Lateral Geniculate Nucleus (LGN) layer.

In the Ganglion/Lateral Geniculate Nucleus (LGN) layer, single opponent mechanism is implemented. Ganglion and Lateral Geniculate Nucleus layers are implemented in single layer because these cells have similar receptive field properties. Response of this layer to Cone layer outputs is mathematically described as follows:

For R-G channel:

$$S(x, y) = w_1 \cdot R_g(x, y; \sigma) + w_2 \cdot G_g(x, y; \sigma) \quad (4)$$

For B-Y channel:

$$S(x, y) = w_1 \cdot B_g(x, y; \sigma) + w_2 \cdot Y_g(x, y; \sigma) \quad (5)$$

Where,

$$w_1 w_2 \leq 0 \text{ and } |w_1|, |w_2| \in [0, 1] \quad (6)$$

Here, w_1 and w_2 are connection weights from cone layer cells to Ganglion/Lateral Geniculate Nucleus (LGN)

layer. Weight w_1 and w_2 have always opposite sign.

Outputs from single opponent mechanisms i.e. (S_{rg} , S_{gr} , S_{by} and S_{yb}) of Ganglion/Lateral Geniculate Nucleus (LGN) layer is passed to Cortex layer and then double opponent mechanisms is implemented in this layer. In the cortex layer of V1, the receptive fields of color-sensitive neurons are both chromatically and spatially opponent [18]. In particular, the oriented double-opponent cells are considered to play an important role in color boundary detection [22]. In this layer boundary is detected by using set of filters having orientation $\theta \in [0, 2\pi]$. Boundary responses at each orientation is calculated using,

$$D(x, y; \theta_i) = \sum_{m, n \in N_{r+g-}} S_{r+g-}(x+m, y+n) * RF(m, n; \theta_i) + \sum_{m, n \in N_{r+g-}} S_{r-} \quad (7)$$

Here, $S_{r-g+} = -S_{r+g-}$, N_{r-g+} and N_{r+g-} are the R-off/G-on and R-on/G-off neurons in the V1 region. Where θ_i is given as:

$$\theta_i = \frac{2(i-1)\pi}{N_\theta} \quad (8)$$

And $RF(x, y; \theta_i)$ is determined by using following equations,

$$\left| RF(x, y; \theta_i) = \frac{\partial f(\tilde{x}, \tilde{y})}{\partial \tilde{x}} \right| \quad (9)$$

$$f(x, y) = \frac{1}{\sqrt{(2\pi(k\sigma^2))}} \exp\left(\frac{-(\tilde{x}^2 + \gamma^2 \tilde{y}^2)}{2(k\sigma^2)}\right) \quad (10)$$

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} x \cos(\theta) + y \sin(\theta) \\ -x \sin(\theta) + y \cos(\theta) \end{pmatrix} \quad (11)$$

γ in Equation (10) is the spatial aspect ratio of Gaussian that controls the ellipticity of receptive field. Based on physiological findings [23] and [24], value of γ is generally taken as 0.5. Similarly, product of k and σ determines size of V1 neurons in cortex layer.

After calculating response at different orientation, maximum response from each orientation is calculated as follows:

$$D(x, y) = \max\{D(x, y; \theta_i) | i = 1, 2, \dots, N_\theta\} \quad (12)$$

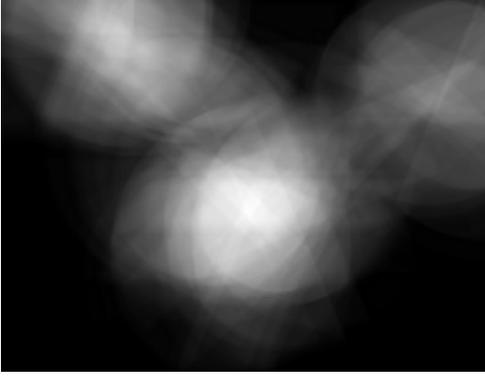


Figure 3: Computed Contour Based Spatial Prior.

Now output of each channel after taking maximum at N_θ orientation are represented by D_{rg} , D_{gr} , D_{by} and D_{yb} . These outputs are normalized linearly. Final response of cortex layer is calculated by taking maximum response at each channel, which is done as follows:

$$r(x,y) = \max\{D_{c_i}(x,y;\theta_i) | c_i \in D_{rg}, D_{gr}, D_{by}, D_{yb}\} \quad (13)$$

$r(x,y)$ is the output of color opponent mechanisms for color boundary extraction. $r(x,y)$ gives the detected color boundary and these are further processed using non-maximum suppression for thinning edge.

Edges obtained after color opponent mechanisms are subjected to non-maximum suppression for thinning edge and thresholded to find dominant edges only. In this research work threshold value 0.33 is used. After obtaining dominant edge response and corresponding orientations, for each edge pixel, Average Edge Response (AER) is calculated in the left and right half disk around it with disk radius $d_r = \min(W,H)/3$. Where W and H are width and height of given image. Then we carry out simple voting to compute rough spatial weights of saliency S_e . Voting here is: all of the pixels within half disk having higher average edge response between two half disks are voted 1, and the pixels in the other half are voted 0.

To obtain final contour based spatial prior, Center-bias weighting [25], [26] modeled by Gaussian masks with standard deviation $\sigma_c = d_r$ is also considered and this saliency is represented by S_c . Using saliency measure S_e and S_c , contour based spatial prior denoted by $cbasp$ is obtained by:

$$cbasp = S_e + S_c \quad (14)$$

Obtained contour based spatial prior for input image shown in Figure 1 is shown in Figure 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 ($cbasp$) with high level prior (hp) obtained earlier in Equation (2), 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 Figure 1 is shown in Figure 4.

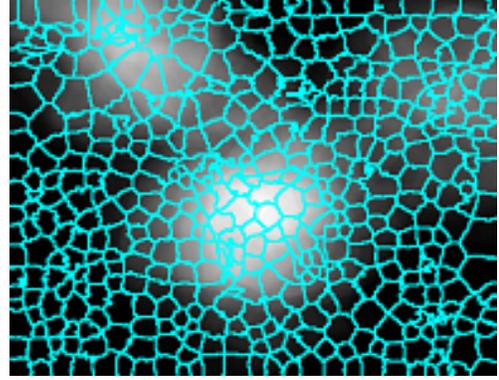


Figure 4: Mapped N patches over contour based spatial prior.

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:

$$fp = m_1 * hp + m_2 * fcbsp \quad (15)$$

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 the structured matrix decomposition by multiplying each component in feature matrix (F) to guide matrix decomposition.



Figure 5: Saliency Map.

3.4 Salient Object Detection

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 [11]. For an input image shown in Figure 1 detected salient object i.e. saliency map is shown in Figure 5.



Figure 6: Extracted Salient Object.

3.5 Salient Object Extraction

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 Figure 1 using structured matrix decomposition with contour based spatial prior (SMD

with CBSP) is shown in Figure 6.

4. EXPERIMENT

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 various scenes. Experimental analysis on ICOSEG [27] and PASCAL-S [28] datasets is performed to evaluate metrics like receiver operating characteristic (ROC) curve, precision recall (PR) curve and Mean Absolute Error (MAE) for comprehensive evaluation.

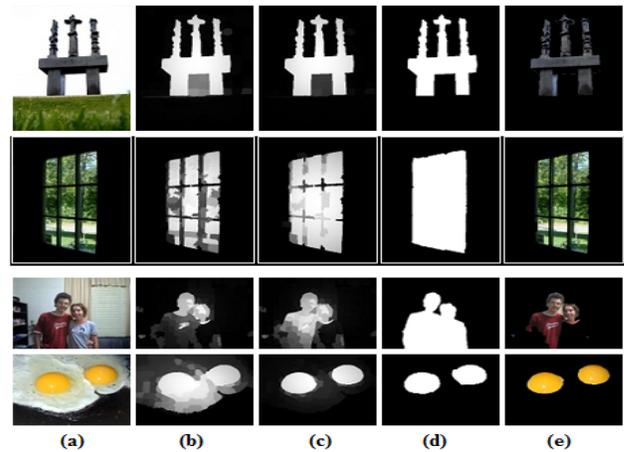


Figure 7: Output Images (a) Original images (b) Detected images by structured matrix decomposition (SMD). (c) Detected images by SMD with CBSP. (d) Ground Truth (e) Extracted images using SMD with CBSP.

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 (σ) 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 γ is chosen to be 0.5 based on physiological findings [23] and [24].

Similarly, while integrating high level priors and contour based spatial prior values of weight w_1 , w_2 and w_3 while finding high level 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 in Figure 7. The detection and extraction results for four images are shown. For these four examples, output of proposed method is closer to ground truth as compared to output of structured matrix decomposition model.

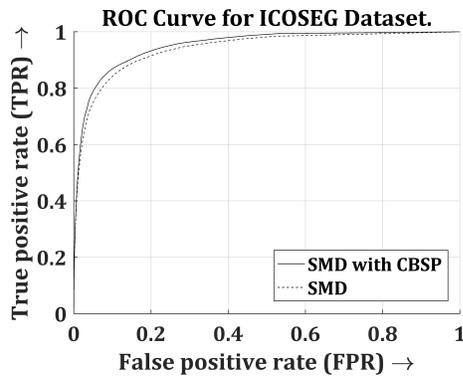


Figure 8: ROC curve for ICOSEG [27] Dataset.

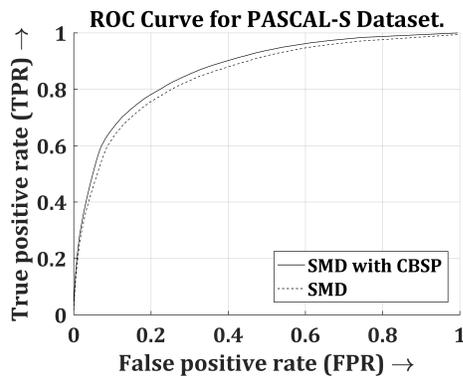


Figure 9: ROC curve for PASCAL-S [28] Dataset.

Similarly, for comprehensive analysis, quantitative comparison on two different datasets using three different evaluation metrics like receiver operating characteristics (ROC) curve, precision-recall (PR) curve and mean absolute error (MAE) is performed. ROC curve for ICOSEG [27] dataset using SMD and SMD with CBSP shown in Figure 8 and similarly for

PASCAL-S [28] dataset shown in Figure 9. For both datasets ROC curve obtained using SMD with CBSP follows the left-hand border and the top border of the receiver operating characteristics space more than using SMD indicating SMD with CBSP has better accuracy.

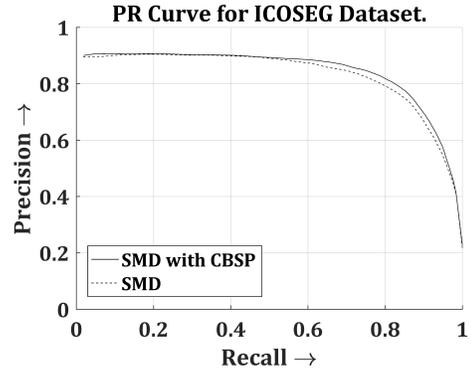


Figure 10: PR curve for ICOSEG [27] Dataset.

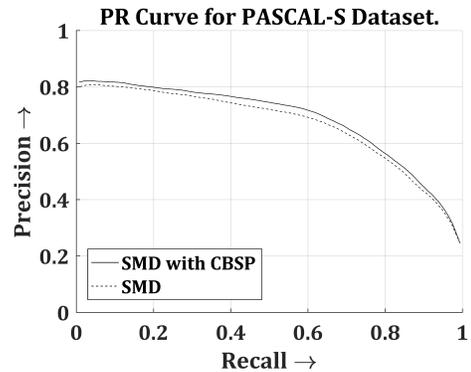


Figure 11: PR curve for PASCAL-S [28] Dataset.

Table 1: Comparison of Mean Absolute Error

Datasets	Mean Absolute Error	
	SMD	SMD with CBSP
PASCAL-S	0.208472	0.189488
ICOSEG	0.138161	0.118017

In addition to ROC curves, precision-recall (PR) curve is also plotted for ICOSEG [27] dataset using SMD and SMD with CBSP shown in Figure 10 and similarly for PASCAL-S [28] dataset shown in Figure 11. For both datasets PR curve obtained using SMD with CBSP follows the right-hand border and the top border more than using SMD.

As complementary to ROC and PR curves, Mean Absolute Error (MAE) is evaluated and tabulated in Table 1 for quantitative analysis. It determines the mean difference between the ground truth and saliency map in pixel level. For the both datasets MAE is less using SMD with CBSP than SMD.

Conclusion

This research work aims to improve the performance of structured matrix decomposition model for saliency related tasks by integrating contour based spatial prior obtained from biologically inspired framework. The proposed method can detect and extract salient object and shows improvement over structured matrix decomposition model. Its effectiveness is demonstrated by various experiments on two widely used 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] L. Itti, C. Koch, and E. Niebur. A model of saliency based visual attention for rapid scene analysis. *IEEE TPAMI*, 1998.
- [2] M. Cheng, G. Zhang, N. J. Mitra, X. Huang, and S. Hu. Global contrast based salient region detection. *CVPR*, 2011.
- [3] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li. Automatic salient object segmentation based on context and shape prior. *BMVC*, 2011.
- [4] Q. Yan, L. Xu, J. Shi, and J. Jia. Hierarchical saliency detection. *CVPR*, 2013.
- [5] M. Cheng, J. Warrell, W. Lin, S. Zheng, V. Vineet, and N. Crook. Efficient salient region detection with soft image abstraction. *ICCV*, 2013.
- [6] J. Yan, M. Zhu, H. Liu, and Y. Liu. Visual saliency detection via sparsity pursuit. *IEEE SPL*, 2010.
- [7] X. Shen and Y. Wu. A unified approach to salient object detection via low rank matrix recovery. *CVPR*, 2012.
- [8] C. Lang, G. Liu, J. Yu, and S. Yan. Saliency detection by multitask sparsity pursuit. *IEEE TIP*, 2012.
- [9] W. Zou, K. Kpalma, Z. Liu, and J. Ronsin. Segmentation driven low-rank matrix recovery for saliency detection. *BMVC*, 2013.
- [10] Chang Tang, Jin Wu, Changqing, Pichao Wang, and Wanqing Li. Salient object detection via weighted low rank matrix recovery. *LSP*, 2017.
- [11] Houwen Peng, Bing Li, Haibin Ling, Weiming Hu, Weihua Xiong, and Stephen J. Maybank. Salient object detection via structured matrix decomposition. *TPAMI*, 2016.
- [12] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *J. ACM*, 2011.
- [13] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 2012.
- [14] P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *IJCV*, 2004.
- [15] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. *ICIP*, 1995.
- [16] H. G. Feichtinger and T. Strohmer. Gabor analysis and algorithms: theory and applications. *Springer*, 1998.
- [17] W. Zhu, S. Liang, Y. Wei, and J. Sun. Saliency optimization from robust background detection. *CVPR*, 2014.
- [18] K. Yang, S. Gao, C. Li, and Y. Li. Efficient color boundary detection with color-opponent mechanisms. *IEEE CVPR*, 2013.
- [19] Kai-Fu Yang, Hui Li, Chao-Yi Li, and Yong-Jie Li. A unified framework for salient structure detection by contour-guided visual search. *IEEE TIP*, 2016.
- [20] G. Papari, P. Campisi, N. Petkov, and A. Neri. A biologically motivated multiresolution approach to contour detection. *EURASIP Journal on Applied Signal Processing*, 2007.
- [21] N. Petkov and M. A. Westenberg. Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition. *Biological Cybernetics*, 2003.
- [22] E. N. Johnson, M. J. Hawken, and R. Shapley. The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience*, 2001.
- [23] B. R. Conway. Advances in color science: from retina to behavior. *The Journal of Neuroscience*, 2010.
- [24] R. Shapley and M. Hawken. Color in the cortex—single- and double-opponent cells. *Vision Research*, 2011.
- [25] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. *IEEE ICCV*, 2009.
- [26] B. W. Tatler. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *J. Vis.*, 2007.
- [27] D. Batra, A. Kowdle, D. Parikh, J. Luo, and T. Chen. Interactively co-segmenting topically related images with intelligent scribble guidance. *IJCV*, 2011.
- [28] Y. Li, X. Hou, C. Koch, J. Rehg, and A. Yuille. The secret of salient object segmentation. *CVPR*, 2014.