



# A robust RGBD saliency method with improved probabilistic contrast and the global reference surface

Surya Kant Singh<sup>1</sup> · Rajeev Srivastava<sup>1</sup>

Accepted: 15 December 2020 / Published online: 6 January 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

## Abstract

The human attention mechanism inspires salient object detection. Most of the saliency methods work on 2D perception mechanisms, while human attention systems work on 3D perception mechanisms. This proposed method makes use of depth information from RGBD to robustly and correctly detect the salient object in a complex and clutter background. The saliency of regions related to object border increases in Poisson probabilistic contrast space while distinguishing the conspicuous object in a complex and clutter background. This process produces a global concave reference surface. This global reference plane integrated with intra-regional spatial, structural, color, and depth information detects the salient object correctly. Background estimation and central saliency integration thoroughly remove the background. This algorithm generates a robust conspicuous object. The experimental result presented here shows that the proposed method performs better in comparison to the recent, highly referenced and closely related fourteen state-of-the-art methods, and the three publicly available complex RGBD datasets and six evaluation parameters.

**Keywords** Salient object detection · Poisson probabilistic modeling · Probabilistic contrast · Background elimination · Global saliency · Regional saliency · Depth saliency

## 1 Introduction

Visual saliency can be defined as the highlighting of the most prominent object from the cluttered and complicated backgrounds. Initially, the problem of saliency originated from the field of neuroscience, psychology and computer vision applications. The applications of salient object detection are diverse and applicable in attention of neurobiology [1], automatic action recognition [2], saliency-aware video object segmentation [3,4], video salient object detection [5,6], corner detection [7], etc.

The computational model of saliency is mainly categorized into three domains. The initial computational model is based on a bottom-up and low-level feature without any training and learning. And the most dominating feature

in this model is contrast, which is defined as pixel-level contrast [8], patch-level contrast [9], region-level contrast [10], multi-scale contrast [11], center-surround contrast [12] color contrast and spatial contrast [13], etc. However, there are some other highly reported features in recent literature like center prior [14], background connectivity [15], surroundness [16], depth [17,18], superpixel-based [19] and objectness [20,21].

The second computational model is top-down learning-based and dependent on high level semantic [22], contextual [23,24] and structural features [25]. These features are used in training and learning with manually annotated ground-truth data. The next recent computation model is a hybrid model that integrates the low level and high-level features. Most of the models vary on proposing the integration strategies to increase the robustness in saliency detection.

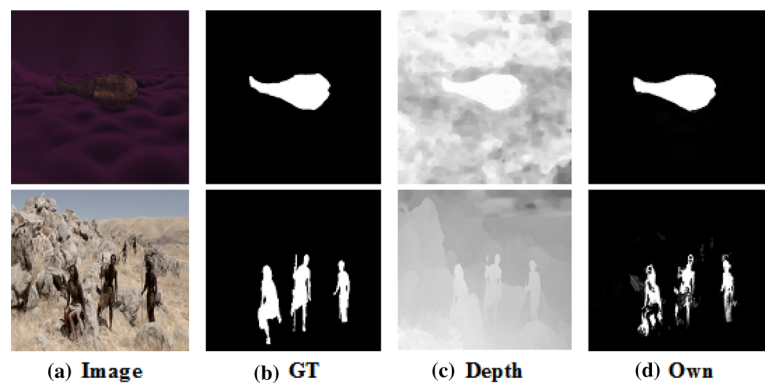
In particular, we define three challenges in the model of saliency computation: (1) interior saliency discrepancy, (2) exterior saliency discrepancy and (3) object border discrepancy. Next, we discuss these challenges and our motivation to mitigate these challenges in holistic integration-based approaches. Interior saliency discrepancy is defined as removing the saliency of the salient region, which is simi-

✉ Surya Kant Singh  
suryakantsingh20@gmail.com

Rajeev Srivastava  
rajeev.cse@iitbhu.ac.in

<sup>1</sup> Department of Computer Science and Engineering, Indian Institute Of Technology, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India

**Fig. 1** Saliency detection in complex and clutter background with minimizing discrepancies in the interior, exterior and border regions saliency



lar to the background regions. Exterior saliency discrepancy is explained as increasing the saliency of non-salient regions, which are identical to prominent object regions. Object border discrepancy is defined as destroying the structure of the object border regions.

The global contrast and other low-level features-based saliency models [12,13,26,27] are computationally efficient and produce full-length saliency. This model generates interior, exterior and border region saliency discrepancies. Most of the models fail in complex and clutter background. The central saliency-based cellular automata are optimal and highly referenced in the literature, which minimizes the interior saliency. In contrast, these models produce exterior as well as border region discrepancy. Therefore, this model is used in the saliency enhancement stage in the proposed method. Some other recent saliency models measure the backgrounds to remove the exterior saliency discrepancy. These approaches minimize the exterior saliency but do not reduce the interior and border saliency discrepancy (Fig. 1).

All these 2-D models generally fail in low depth images. The depth clues in RGBD saliency give an extra space to increase the saliency in low depth 3D images. However, it has discriminative power against the complex and clutter background. Therefore, to capture the low depth features, the RGBD saliency model is used. RGBD saliency model captures low-level depth information but is itself not sufficient for salient object detection. Therefore, Cheng et al [17] used color and structural-based regional saliencies [13] to exploit the regional features. But this model produces saliency, which minimizes exterior saliency discrepancy and improves the interior saliency discrepancy but fails in border region discrepancy. Zhu et al. [28] used dark channel, central channel and other purifying saliency features along with depth features. This model produces better saliency than other RGBD model. This model minimizes both interior and exterior saliency discrepancy, although this model fails in border region discrepancy.

To address the limitations as mentioned above, the global concave topographical saliency is used as a reference surface

in the proposed method. This surface is used as a reference surface for minimizing all these said discrepancies. The main contributions of the proposed method for addressing the discrepancies mentioned above are as follows:

- We are proposing for the first time a global concave saliency-based reference plane in RGBD saliency computation.
- In this paper, a novel global concave reference surface is proposed by the addition of DoG-based contour and improved probabilistic contrast (IPC)-based saliency to minimize the border region discrepancies.
- The integration of spatial, regional, color and depth saliencies into a global concave reference surface is proposed to minimize the interior saliency discrepancy.
- The integration of spatial, regional, color and depth saliencies into Gaussian-based background elimination model is proposed to minimize exterior saliency.

Global concave-based saliency is vital in providing a reference plane for regional saliencies integration, because this reference surface has enhanced saliency on object boundary through the DoG based contour. This enhanced boundary with improved probabilistic contrast preserves the border region saliency during regional saliency integration. This integration proves the discriminating characteristics through the results obtained. This novelty has added a different space attention rule for salient object detection.

This paper is divided into six sections. The detailed survey of closely related literature is studied in Sect. 2. In Sect. 3, the proposed method is adequately defined and explained. The experiments and result analysis with state-of-the-art methods are demonstrated in Sect. 4. The conclusion and the future scope are presented in Sect. 5. The theoretical principle for normalization of global probabilistic contrast is presented in Sect. 6.

## 2 Related works

Recently, a significant advancement in the saliency computation model to predict the salient object accurately has been witnessed. Various saliency computation models have been reported in the literature. These models have a significant contribution toward accuracy and robustness. In these models, global prior, central prior, background prior, connectivity prior and depth prior are highly reported features to enhance the saliency computation. The global prior is mainly designed by pixel-level contrast method [12,29,30]. There are some other global contrast-based methods like histogram-based [13], region-based contrast [31], directional contrast-based [27], distance-based model [28] which have been studied and investigated.

There are various global contrast-based methods reported in the literature review. But, here, we are going to use relevant, recent and widely reference methods. A recent global contrast-based method uses the color histogram HC [13] of quantized color channels. In this method, global color contrast is computed in the most frequently used color in the form of the histogram. This algorithm has failed in complex images where regions have a similar color histogram. This method uses color-based regional saliency (RC) [13] to address this issue. In many complex image sets, it increases the interior as well as the exterior saliency discrepancy simultaneously. Therefore, this method used a saliency cut algorithm based on a computationally expensive graph cut algorithm for the segmentation. These methods have common limitation to produce saliency with backgrounds. Hang et al. [27] proposed the minimum directional global contrast based on spatial distribution. Minimum directional contrast (MDC) [27] is the minimum value of directional contrast (DC). Background region shows low MDC value, while the foreground pixels represent the high MDC value. This method destroys the border of a salient object in complex and clutter backgrounds.

The major drawback of the above methods is that it has failed in producing the saliency in an image having a salient object that has multiple color regions and low depth feature. It has also failed in producing saliency where there is a structural similarity between prominent and non-salient regions. But, the main advantage of these methods is to produce a full-length global saliency. It uniformly distinguishes the salient object from the background. To overcome these limitations, most of the methods use some other supplementary methods. MDC [27] uses boundary connectivity to measure the background regions and differentiate it from the foreground region, having a salient object. Background prior-based methods minimized the exterior saliency discrepancy while increasing interior saliency discrepancy. Some early works use center prior [32] based on object size and location similar to the Gaussian fall-off map. In this feature, the

central region assigns higher saliency gradually to other border regions. Usually, these cues are used as weights [33,34] or a feature in learning-based methods [35]. These methods reduce the interior saliency discrepancy while increasing the exterior saliency discrepancy.

To overcome the limitation of low depth in sophisticated image, recently depth features are being used to improve the saliency computations. Cheng et al. [17] compute saliency by the integration of using both color contrast and depth contrast features. Peng et al. [36] use a fusion-based model to combine the RGB saliency with depth-based saliency. Geng et al. [37] proposed the salient object detection in the stereo images based on depth-based saliency. Recently, depth cue is combined by Zhu et al. [28] with regional saliency, dark saliency and center saliency. In this method, the author uses the dark channels prior, center and depth to increase the robustness. As per these results, depth saliency is a valuable feature compared to other visual features in improving the robustness of saliency computation.

Addressing the above-described limitations related to global, regional as well as background prior, a global concave topographical reference surface is used to initialize the saliency computation. This surface improves saliency in border regions. The intra-regional distance-based saliency and spatially weighted saliency are integrated into the global reference surface to increase the interior saliency. The regional color saliency and depth saliency are integrated using Gaussian function into a well-defined global reference plane to enhance the object's interior region saliency and minimizes the exterior saliency. Further integration of the center saliency uniformly removes the background and highlights the prominent object in complex and challenging images.

Recently, deep learning-based methods exploit the learning concept from CNNs to improve the results at the next level. The representative model of CNN in RGBD saliency is proposed by Liangqiong et. al [38]. In this deep fusion DF [38] model, CNN is used to integrate different low-level features into hierarchical saliency. A adaptive fusion AF-Net-based [39] saliency model is used to combine the results of two-stream CNN by using switch maps. In this adaptive fusion, the loss function is composed of three different components like saliency supervision, switch map supervision and edge-preserving constraints. These loss functions guided the training of the network in an end-to-end manner. But this model is failed in the complex and cluttered background image. A conditional variational autoencoder (VAE)-based probabilistic RGBD model UC-Net [40] is proposed to produce multiple saliency maps for each input image. This model adds different dimensions to improve the performance of salient object detection methods. In JL-DCF [41] proposed the joint learning (JL) and a densely cooperative fusion (DCF) based on a shared Siamese backbone network. DCF module is used to find the complementary feature among

RGB and depth features, while the JL module is used to learn saliency features from coarse deep level to final image levels. This model uses the middle fusion strategy. This model produces better saliency maps in complex and cluttered background images. In D3Net [42], three-stream separate network for RGBNet, RGBDNet and DepthNet is proposed with a depth, depurator unit to learn the modalities specific information. This unit identified and removed the low-quality depth maps, and it is used for effective multi-modal fusion to achieve robust SOD and improved the performance.

The next representative module ICNet [43] is proposed interactive and adaptive model to learn high-level features between RGB and depth modalities. A cross-modal depth-weighted combination block is used to enhance the saliency. This network exploits the complementarity cross modalities to improve the performance. A recent region-wise attention is proposed as complementary interaction module SSF [44] to supplement rich boundary information for each modality to learn and fuse cross-modal features. The main objective of proposed model SSF is to effectively find the complementarity features in cross-modal while minimizing the negative effects introduced by low-quality depth maps to enhance the performance.

### 3 The proposed method

#### 3.1 Initialization through global concave surface (GCS)

The global concave surface is used to initialize saliency computation. This surface is composed of improved Poisson probabilistic contrast (IPC) and DoG-based contour-enhanced global surface.

##### 3.1.1 Improved Poisson probabilistic contrast (IPC)

Poisson probabilistic modeling is used as *pdf* of the image plane. This *pdf* is used to compute the probabilistic contrast. Generalized Poisson distribution is a better choice because it has characteristics of convergence of information divergence into a concave shape topographical surface. The probabilistic distribution  $\phi$  with mean  $\mu$  is defined over color planes  $c = [l, a, b]$  for input original image  $I_0$  in CIE-LAB space. It is defined as:

$$\phi^c(I_0, \mu) = \frac{e^{-\mu} \mu^{I^c}}{I^c!} \quad (1)$$

The improved Poisson probabilistic contrast (IPC) is the addition of contour enhanced global surface with Poisson probabilistic contrast [45]. It is Poisson-based global contrast with normalized likelihood surround symmetry. The

improved Poisson probabilistic contrast (IPC) is formulated using color chrominance channel  $c = [a, b]$  and luminance channel  $l$  in CIE – LAB space for input image  $I_0$ . It is defined as follows:

$$S_{IPC} = \sum_{c \in a, b} \underbrace{\|I_0^c - \phi^c I_0^c\|}_{\text{Chrominance Contrast}} + \underbrace{\|I_0^l - \phi^l I_0^l\|}_{\text{Luminance Contrast}} \quad (2)$$

where  $N_{\text{Coff}}$  is called the normalization coefficient. This coefficient is used to normalize the Poisson probabilistic luminance contrast. It is defined as:

$$N_{\text{Coff}} = \left( \frac{1}{2} p_k^2 + \frac{1}{6} p_k^3 + \frac{1}{3} p_k^4 \right) * \gamma \quad (3)$$

where  $\gamma = \sqrt{R_\sigma} - \sqrt{R_\mu}$  is formulated to approximate the divergence effect of uneven luminance into chrominance planes.  $R_\sigma$  is the relative contrast of variance between the luminance plane and chrominance plane.  $R_\mu$  is the measure of the relative mean contrast between luminance mean, and chrominance mean.  $P_k$  is the  $k$ th point probability in Poisson probabilistic space.  $N_{\text{Coff}}$  is visualized as the region of uneven distribution of luminance. The mathematical formulation of  $N_{\text{Coff}}$  and corresponding proof are properly described in Sect. 6.

##### 3.1.2 Contour-based global surface

The DoG filter efficiently approximates the Laplacian of Gaussian. This edge detection method is widely reported in the literature [46,47]. The global concave topographical surface is computed by the addition of contour based global surface and improved Poisson probabilistic. This contour-based surface is defined by difference of Gaussian  $DoG(x, y)$  of the input image  $I_0(x, y)$ , as defined in Eqs. 4 and 5. Here,  $\sigma_1$  and  $\sigma_2$  are the standard deviation, where  $\sigma_1 > \sigma_2$ . The  $DoG(x, y)$  is defined as follows:

$$\begin{aligned} DoG(x, y) &= \frac{1}{2\pi} \left[ \frac{1}{\sigma_1^2} e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2^2} e^{-\frac{(x^2+y^2)}{2\sigma_2^2}} \right] \\ &= G(x, y, \sigma_1) - G(x, y, \sigma_2) \end{aligned} \quad (4)$$

This contoured surface is generated by the integration of multiple edges. For integration of the range of edges, let  $\varphi = \sigma_1/\sigma_2$ . So, all the edges over DoG will have the standard deviations in the ratio  $\varphi$ . It is defined as:

$$S_{ES} = \sum_{i=0}^{N-1} G(x, y, \varphi^{i+1}, \sigma) - G(x, y, \varphi^i, \sigma) \quad (5)$$

This  $S_{ES}$  surface integrates a  $N - 1$  number of edge surfaces into the first edge surface to enrich the boundary of the object. The initial  $GCS$  is a simple addition of edge-enhanced global contour surface and enhanced Poisson probabilistic contrast surface. This surface has characteristics of enriched object boundary [7,48]. Therefore, it is used as a reference plane for other saliencies integration and background removal, which is defined as follows:

$$S_{GCS} = S_{IPC} + S_{ES}. \tag{6}$$

### 3.2 Regional contrast integration into GCS

Initial GCS, global concave topographical saliency, is used as a reference plane. It maximizes the information of the object and reduces the saliency of background regions.

### 3.3 Regional saliency integration with GCS

The regional features are integrated using the region descriptor based on the K-mean algorithm in the initial image plane  $I_0$ . This process uses  $K$  color-based region. The same regional descriptor is used for all regional saliencies into initial GCS topographical saliency. The spatial regional saliency  $SS_i$  is defined as regional density. It is defined as a ratio of number of pixel(NOP) in the region  $i$  and total pixels(TP) in image as  $SS(i) = NOP(r_i)/TP$ . The spatial depth saliency into depth  $I_d$  space is defined as follows:

$$S_{SD}(r_k) = \sum_{i=1, i \neq k}^K SS_i e^{\frac{Dis_0(r_k, r_i)}{\sigma^2}} Dis_d(r_k, r_i) \tag{7}$$

where  $Dis_d(r_k, r_i) = |r_k, r_i|$ ,  $| - |$  is the Euclidean distance between region  $i$  with central region  $k$  in depth space. Similarly, regional color saliency in color space is defined as:

$$S_{color}(r_k) = \sum_{i=1, i \neq k}^K SS_i e^{\frac{Dis_0(r_k, r_i)}{\sigma^2}} Dis_{color}(r_k, r_i) \tag{8}$$

where  $Dis_0(r_k, r_i)$  is the spatial saliency and  $\sigma$  is controlling parameter while  $Dis_{color}(r_k, r_i)$  is regional color saliency based on the Euclidean distance between central region  $k$ th and  $i$ th region in  $L * a * b$  color space. Similarly, regional probabilistic contrast in GCS space is defined as:

$$S_{GC}(r_k) = \sum_{i=1, i \neq k}^K SS_i e^{\frac{Dis_0(r_k, r_i)}{\sigma^2}} Dis_{GCS}(r_k, r_i) \tag{9}$$

$Dis_d(r_k, r_i)$ ,  $Dis_{color}(r_k, r_i)$  and  $Dis_{GCS}(r_k, r_i)$  are depth, color and probabilistic contrast-based regional saliencies, respectively. These saliencies minimize the interior discrepancy with regional weighting parameter like  $SS_i$ ,  $Dis_0(r_k, r_j)$ ,  $Dis_d(r_k, r_i)$ ,  $Dis_{color}(r_k, r_i)$  and  $Dis_{GCS}(r_k, r_i)$ . These regional saliencies integrations into  $GCS$  space increase some background saliency also. In this integrations, the exterior saliency discrepancy increases, which removes in next background estimation model.

#### 3.3.1 Background estimation model

In the saliency computation domain, it is the assumption of center prior and background prior hypothesis [17] that the salient object is mostly located in the center of the image. So, the integrating factor of these regional saliencies assigns more weight to the central regions and less weight to the border regions. This background estimation is approximated with the normalized Gaussian function. Let  $Pos_k$  define  $k$ th region and  $Pos_c$  define central region respectively. The weight-based integrating factor ( $WF_c$ ) is calculated as:

$$WF_c(r_k) = \frac{Gaussian(Dis_d(Pos_k - Pos_c))}{N_k} GW(D_k) \tag{10}$$

where  $N_k$ , denotes the number of pixels in  $k^{th}$  region. In this equation,  $Dis_d$  is the distance in Euclidean space and  $GW(D_k)$  is depth weight, which is calculated as:

$$GW(D_k) = max(D - D_k)^{\frac{1}{(max(D) - min(D))}} \tag{11}$$

Finally, the normalized Gaussian function is used to integrate color, depth, spatial and probabilistic saliencies. This integration increases the interior saliency and minimizes the background saliency, which is defined as:

$$S_G = Gaussian(S_{GD}(r_k) + S_{color}(r_k) + S_{GC}(r_k)) WF_c. \tag{12}$$

### 3.4 Saliency enhancement

The hypothesis of biologically plausible architecture [49] describes the phenomena of centralization of the object always towards the center of the image. The center prior is also preferred because of the mindset characteristics of the photographer. Central bias is preferred in saliency detection and enhancement [50,51]. Therefore,  $S_{Cen}$  is used in saliency



enhancement, which is based on the central saliency BSA [32] algorithm. This algorithm is used to remove the edge effect and minimize the exterior saliency discrepancy while enriching the interior saliency. Final saliency is the simple addition of central saliency  $S_{Cen}$  and  $S_G$

$$S = (S_G + S_{Cen}). \quad (13)$$

### 3.5 The proposed algorithm

The proposed method is summarized in the following algorithm. The two algorithms, Algorithms 1 and 2, are used to describe the sequential steps as follows.

---

**Algorithm 1:** To generate the global concave reference plane using improved Poisson-based probabilistic contrast (IPC) and Gaussian-based contour

---

**Input** : Input original image  $I_0 = \{l, a, b\}$  in CIE-LAB color format.  $\phi$  is Poisson-based  $PDF$ ,  $\mu$  is mean of respective color planes, and  $\sigma$  is variance.

**Output:** Global concave topographical reference saliency surface  $S_{GCS}$

- 1 Compute  $\delta_a \leftarrow \tan^{-1}(\sigma_a/\sigma_l)$
  - 2 Compute  $\delta_b \leftarrow \tan^{-1}(\sigma_b/\sigma_l)$
  - 3 Compute  $\gamma \leftarrow |\sqrt{R_\sigma} - \sqrt{R_\mu}|$
  - 4 Compute  $p_k \leftarrow \tan(45^\circ - \delta_a)\tan(45^\circ - \delta_b) \quad // \text{ for } k^{th} \text{ image}$
  - 5 Compute  $N_{Coff} = (\frac{1}{2}p_k^2 + \frac{1}{6}p_k^3 + \frac{1}{3}p_k^4) * \gamma$
  - 6 Compute  $S_{IPC} \leftarrow (I_0^l - \phi^l I_0^l)^2 N_{Coff} + (I_0^a - \phi^a I_0^a)^2 + (I_0^b - \phi^b I_0^b)^2$
  - 7 Compute  $S_{ES} = \sum_{i=0}^{N-1} G(x, y, \varphi^{i+1}, \sigma) - G(x, y, \varphi^i, \sigma)$
  - 8 Finally Compute  $S_{GCS} \leftarrow S_{IPC} + S_{ES}$
- 

---

**Algorithm 2:** Regional depth, color and spatial saliencies integration in global concave topographical saliency  $S_{GCS}$

---

**Input** : Input initial global concave saliency  $S_{GCS}$  and original input color image  $I_0$  into CIE-LAB space

**Output:** Final saliency  $S$

- 1 **for**  $k \leftarrow 1$  **to**  $K$  *color region do*
  - 2   Compute the regional depth saliency  $S_{SD}(r_k)$  by Eq. 7
  - 3   Compute the regional color saliency  $S_{color}(r_k)$  by Eq. 8
  - 4   Compute the regional spatial saliency  $S_{GC}(r_k)$  by Eq. 9
  - 5   Compute the regional Gaussian background elimination weight  $WF_c(r_k)$  by Eq. 10
  - 6   Integrate all regional saliency  $S_G = \text{Gaussian}(S_{GD}(r_k) + S_{color}(r_k) + S_{GC}(r_k))WF_c$  by Eq. 12
  - 7 **end**
  - 8 Compute the central saliency  $S_{Cen}$
  - 9 Compute final saliency  $S \leftarrow (S_{Cen} + S_G)$  by Eq. 13
- 

## 4 Experiment and result analysis

### 4.1 Dataset

The proposed method is extensively evaluated on three publicly available complex datasets for salient object detection. The first dataset, RGBD-1000 or NLPR [36], contains 1000 images, which includes a complex background, very similar to the conspicuous object. Each image has a resolution of  $640 \times 480$ . The second dataset is PU-80 or SSD [28], which contains low depth images having multiple similar objects and confusing backgrounds with a resolution of  $960 \times 1080$ . This dataset is designed with a complex scene to make it a computational challenge for salient object detection. Next most preferred dataset is NJUD-1985 [52]. It has 1985 stereo-image pairs. These images have been collected from different sources like some from the internet and some from 3D movies.

### 4.2 Evaluation metrics

To evaluate the performance of the proposed method with other state-of-the-art methods, we used six performance metrics (1) F-measure (2), precision–recall curve (PR curve), (3) receiver operating characteristic (ROC curve), (4) mean absolute error (MAE), (5) E-measure ( $E_\psi$ ) and (6) S-measure.

#### 4.2.1 F-measure

The comprehensive evaluation of the proposed method is demonstrated with F-measure. This metric is used to compute the relevancy of parameters like precision and recall. In this metric, precision and recall are combined as weighted harmonic, which is defined as follows:

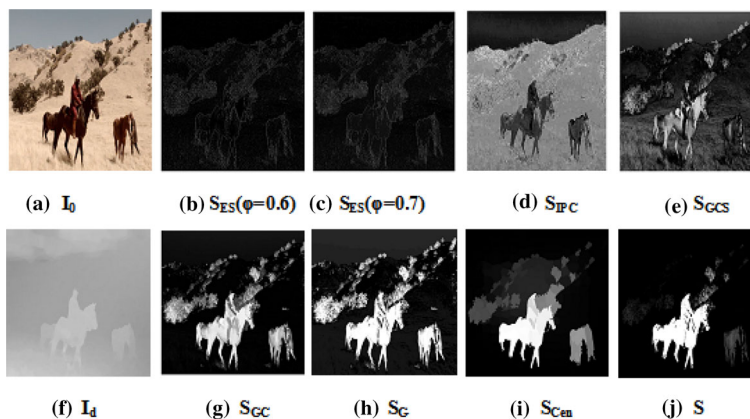
$$\text{F-measure} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}} \quad (14)$$

We use  $\beta^2 = 0.3$  in Eq. 20 F-measure for uniform comparison because the same value is preferred in the majority of saliency methods.

#### 4.2.2 Precision–recall curve (PR curve)

The most widely used metric in saliency is precision and recall curve for a fair evaluation. The precision describes the correct prediction of salient points. The recall represents the accurate detection of the percentage of the salient pixels. We use  $S_{map}$ , and corresponding binary masks  $B_{mask}$ , in the saliency evaluation. Precision and recall are defined as:

**Fig. 2** Visual demonstration of the contribution of each step in the proposed algorithm



$$\text{Precision} = \frac{|S_{\text{map}} \cap B_{\text{mask}}|}{|S_{\text{map}}|} \tag{15}$$

$$\text{Recall} = \frac{|S_{\text{map}} \cap B_{\text{mask}}|}{|B_{\text{mask}}|} \tag{16}$$

In Eqs. 15 and 16,  $|-|$  represents the intersection between binary mask  $B_{\text{mask}}$  and saliency map  $S_{\text{map}}$ . For the accurate analysis of these metrics, bipartite saliency  $S_{\text{map}}$  is used. In this evaluation, multiple fixed thresholds are used, which change from 0 to 255. Precision and recall are computed on each threshold and are combined to form a precision–recall (PR) curve.

### 4.2.3 Receiver operating characteristic (ROC curve)

The ROC curve demonstrates the graphical representation between true positive rate (TPR) and false positive rate (FPR). In this computation, multiple fixed thresholds are used, which range from 0 to 255.

$$\text{TPR} = \frac{|S_{\text{map}} \cap B_{\text{mask}}|}{|S_{\text{map}}|} \tag{17}$$

$$\text{FPR} = \frac{|S_{\text{map}} \cap B_{\text{mask}}|}{|S_{\text{map}} \cap B_{\text{mask}}| + |\overline{S_{\text{map}}} \cap \overline{B_{\text{mask}}}|} \tag{18}$$

In Eqs. 17 and 18,  $S_{\text{map}}$ ,  $B_{\text{mask}}$ ,  $\overline{S_{\text{map}}}$  and  $\overline{B_{\text{mask}}}$  represent true salient points, true ground-truth points, false salient points and false ground-truth points, respectively.

### 4.2.4 Mean absolute error (MAE)

The mean absolute error (MAE) is the preferred metric in successive steps validation and demonstration of successive steps contribution. MAE is defined in normalized range  $[0, 1]$  saliency map  $S_{\text{map}}$  and the ground-truth binary mask  $B_{\text{mask}}$ ,

which is defined as follows:

$$\text{MAE} = \frac{1}{n} \sum_{k \in n} (S_{\text{map}}(k) - B_{\text{mask}}(k)). \tag{19}$$

### 4.2.5 E-measure ( $E_{\psi}$ )

E-measure is recently defined as enhanced alignment measure, and the detailed definition and formulation are available here [53]. This measure is based on cognitive vision studies. It uses image-level statistics (mean) and local level pixel matching information. To demonstrate a comprehensive evaluation, we use maximum value of E-measure.

### 4.2.6 S-measure

S-measure [54] is a recent metric which is used to compare the structural similarity and dissimilarity. This metric computes region-aware  $S_r$  and object-aware  $S_o$  structural similarity between computed saliency map and ground-truth map. This metric is defined as:

$$S_{\text{measure}} = \alpha S_o + (1 - \alpha) S_r \tag{20}$$

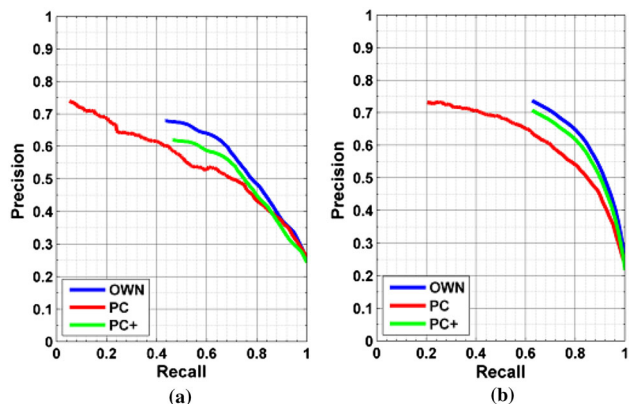
where  $\alpha \in [0, 1]$  is set to 0.5.

## 4.3 Parameters and constraints selection

A set of extensive experiments are performed to evaluate the final value of the various parameters and constraints. These experiments are performed on RGB-1000 and PU-80 or SSD dataset. These extensive experiments have been done to finalize the value and range of the following parameters. The outer border of the salient object is enclosed with reference contour by combining the result of applying several DoG base edges. The range of  $\sigma_1$  and  $\sigma_2$  varies to keep the value of  $\varphi = \sigma_1/\sigma_2$  constant at 1.6. The controlling parameter  $\sigma^2$  is 0.4 in Eqs. 4 and 5.

**Table 1** Stepwise mean absolute error in the proposed method—GCS

Data – Set	Initial $S_{GCS}$	Final $S_G$	$S_{Cen}$	$S$
PU-80	0.5875	0.3605	0.1593	0.0716
RGBD-1000	0.3540	0.2451	0.1097	0.0581

**Fig. 3** Comparison of probabilistic contrast-based method using PR curve on (1) PU-80 (2) RGB-1000 dataset, respectively

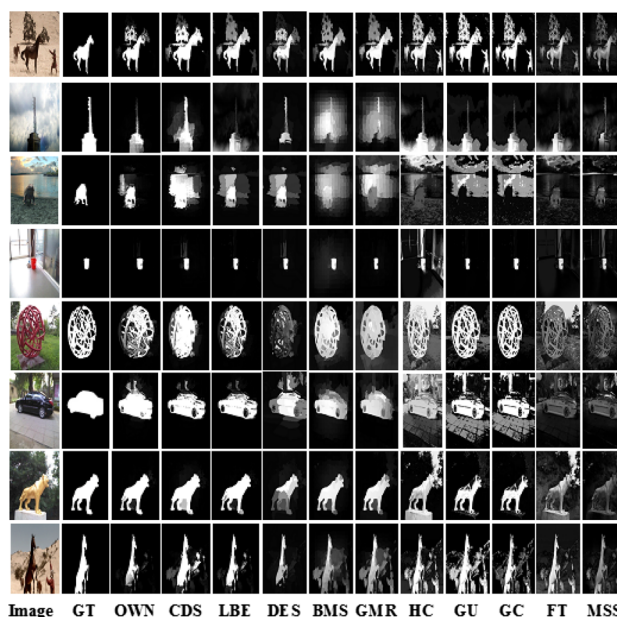
#### 4.4 Successive steps validation

Successive steps of the proposed method GCS are validated on the publicly available complex image dataset PU-80 or SSD and RGBD-1000 or NLPR having depth information. Validation of each step is essential to demonstrate the contributions in saliency. In complex and clutter background images, the salient object cannot be separable by a single-stage algorithm. The visual contribution of each step is shown in Fig. 2. The validation of effectiveness is measured through MAE (mean absolute error), which is shown in Table 1 and PR curve in Fig. 3. The result is shown in Table 1, which validates each step of GCS on PU-80 and RGBD-1000 datasets. This result demonstrates the effectiveness of each step.

#### 4.5 Comparative analysis

The extensive experiment is performed on using three RGBD benchmark datasets having images with complex and clutter background. This result analysis of the proposed method is evaluated through a visual qualitative scale (Figs. 2, 4) and quantitative scale (Figs. 3, 5, 6; Tables 1, 2). Our proposed method is initialized with global probabilistic contrast and difference of Gaussian-based contour model. Therefore, in these evaluations, the top five global contrast-based methods like MDC [27], HC [13], GC [10], MSS [26] and FT [12] are selected.

These methods are selected in the result analysis are based on highly referenced, computationally fast, recent and closely related to our proposed method. The proposed

**Fig. 4** Visual comparison of saliency of proposed method with other state-of-the-art methods

method GCS is also compared with efficient graphical model GMR [55] and cellular automata-based central saliency model BMS [32]. The proposed method is also compared with top, efficient and recent RGBD-based models like DCMC [56], LBE [57], CDS [28] and DES [17] and two recent deep learning-based methods such as DF [38] and AFNet [39] in Table 2. In this evaluation, some other state-of-the-art methods like GU [31], RC [13], RBD [15] and MST [58] are also compared.

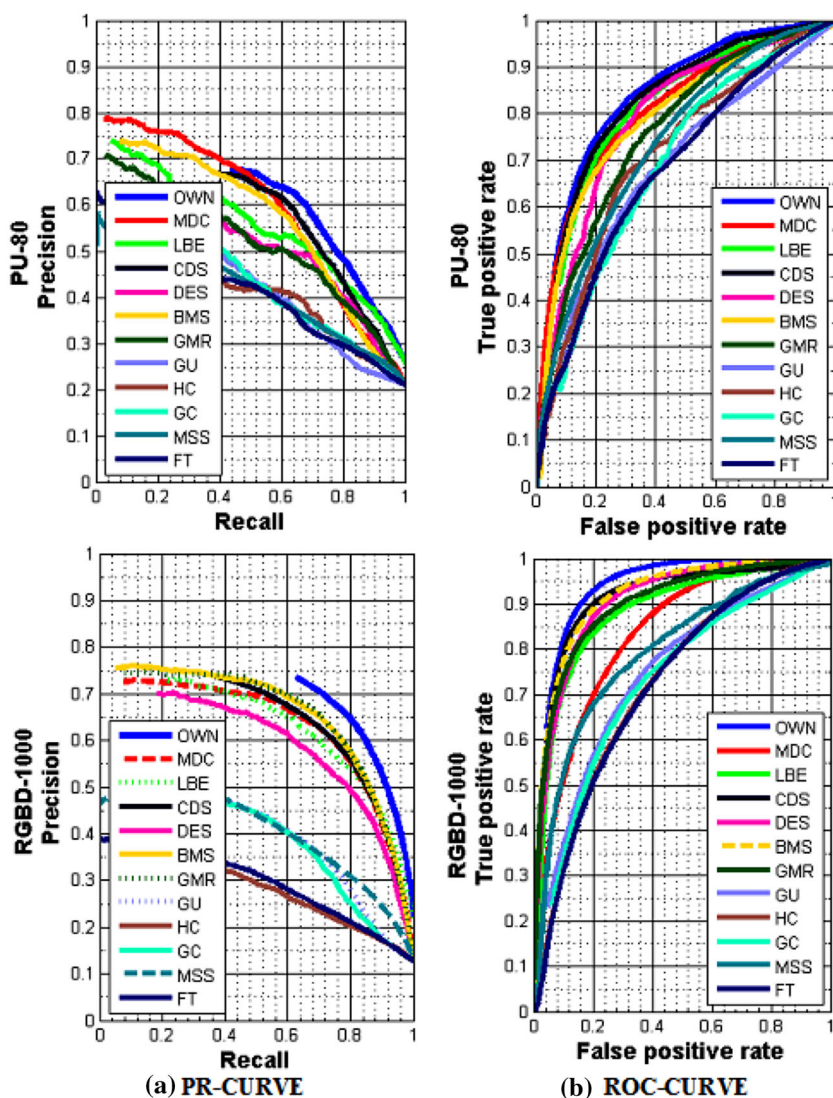
The qualitative analysis is demonstrated through the visual saliency map, which is shown in Fig. 4. In this observation, the global contrast-based methods produce full-length saliency. The global contrast-based techniques—FT, GC, MDC, HC and MSS, highlight some background similar to the salient regions and suppress some interior saliency, which has similar characteristics. So, these methods produce interior and exterior saliency discrepancy. To remove the backgrounds, HC and RC used saliency cut, FT used mean shift, MSS used graph-cut algorithm. However, these methods enhance the computation cost. MDC uses a marker-based watershed segmentation algorithm to separate the salient object. This method destroys some structural information like shape and border regions, which is shown in Fig. 4.

The marker-based watershed segmentation algorithm produces multiple markers, in which some are related to background and others to the objects.

Cellular automata-based central saliency creates a saliency map with no interior saliency discrepancy. Therefore, this algorithm used in saliency enhancement. All the above methods fail in producing saliency in low depth images. The



**Fig. 5** Quantitative comparison of proposed method on RGBD saliency map with PR curve and ROC curve **a** RGBD-1000, **b** PU-80



proposed method, GCS, minimizes the limitations mentioned above and builds a robust salient object, which reduces the interior saliency discrepancy and altogether remove the backgrounds in low depth and sophisticated image. Through visual comparisons, the proposed method can detect single, multiple and complex images precisely. Through all these observations, our proposed method GCS performs better than other state-of-the-art methods.

The proposed method GCS is compared with fourteen state-of-the-art top-performing methods. Eleven state-of-the-art methods are using PR curve, ROC curve and F-measure (Figs. 5, 6), while the two deep-based learning methods and four recent RGBD-based methods are compared using S-measure, E-measure( $E_\psi$ ) and MAE in Table 2. The proposed method outperforms on the recall axis while maintaining the same level of precision, which is visible in PR curve in Figs. 3 and 5. These characteristics demonstrate the robustness of the

proposed method GCS with better saliency maps with other state-of-the-art methods.

#### 4.6 Comparison with RGBD deep learning-based methods

The proposed method is based on a global topographical surface. It is not related to the learning-based method. However, the proposed method is compared with two published RGBD deep learning-based methods, DF [38] and AFNet [39]. In this evaluation, recent metrics like S-measure, E-measure ( $E_\psi$ ) and MAE have been used to compare the quantitative performance. In this evaluation, we use the same learning and testing pattern as used in DF [38]. This evaluation is shown in Table 2. The deep learning-based methods improved the performance significantly of SODs methods recently over non-deep learning-based methods because deep learning-based method learned the structures, semantic, object characteris-

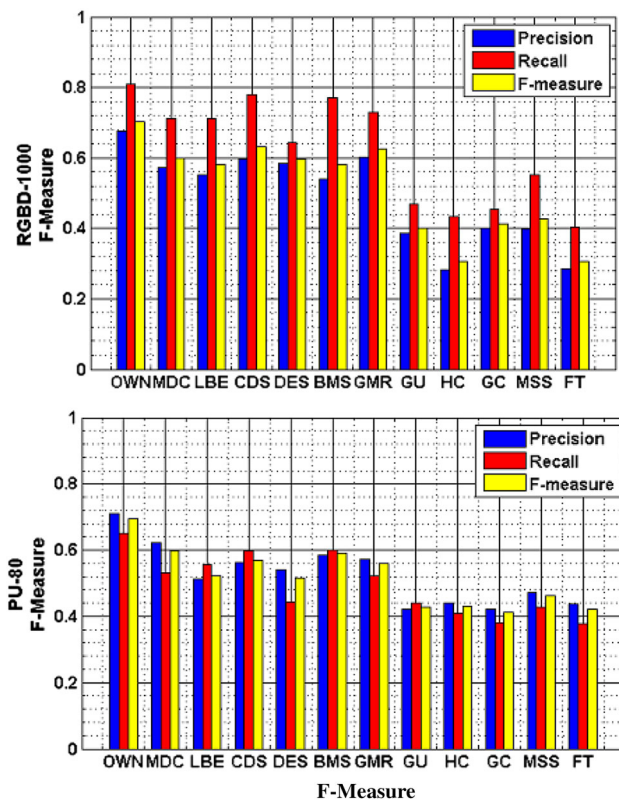


Fig. 6 Quantitative comparison of proposed method on RGBD saliency map with F-measure a RGBD-1000, b PU-80

tics, low-level local features and high-level global features to correctly identify the salient points. The proposed method used probabilistic global contrast to generate the reference surface for multiple integrations of regional saliency integration. Deep-based learning method used multi-stage, object semantics-based learning to improve the performance. The proposed method generates the saliency with comparable in results with two recent deep learning-based methods to show the robustness of the proposed method.

### 5 Conclusion

This paper uses an additional parameter depth to increase the robustness in saliency detection in complex and clutter background. In this method, an innovative and robust approach of global concave topographical surface (GCS) is prepared for regional features integration. This surface designs with the difference of Gaussian-based contours. So, this reference plane is used to minimize the border region discrepancies. This integration works efficiently and effectively in regional saliencies integration to reduce the interior, exterior and regional saliency discrepancies. The robustness of GCS increases the preservation of the structure, shape and border-related information in saliency estimation. These regional saliencies integrations remove the interior saliencies discrepancies. Gaussian weighted background estimation and central saliency integration remove the exterior saliency discrepancies. Finally, all these integrations into the global concave surface increase robustness and help in achieving state-of-the-art results. The improvements in robustness through adding some complimentary deep features can act as a guiding beacon for future work on this framework.

### 6 Proof of theoretic principle for normalization of global probabilistic contrast

The normalization of luminance plane over chrominance plane in Poisson distribution is defined as maximum likelihood estimation. The measure of uneven distribution is defined in terms of the influence of region of surround symmetry. The characteristics of surround symmetry regions and its information divergence are proved, formulated and described by H.Perter [59]. In this paper, the topographical

Table 2 Quantitative comparison of our proposed method with state-of-the-art RGBD deep learning-based saliency methods, DF [38] and AFNet [39], and traditional methods, DES [17], LBE [57], DCMC [56] and CDS [28], and methods on three datasets

Metric	RGBD-1000 or NLPR [36]			NJU2K [52]			PKU-80 or SSD [28]		
	$S_\alpha \uparrow$	$MAE \downarrow$	$E_\psi^m \uparrow$	$S_\alpha \uparrow$	$MAE \downarrow$	$E_\psi^m \uparrow$	$S_\alpha \uparrow$	$MAE \downarrow$	$E_\psi^m \uparrow$
OWN	0.776	0.0581	0.877	0.758	0.137	0.732	0.946	0.076	0.976
AFNet [39]	0.779	0.058	0.879	0.775	0.100	0.853	0.714	0.118	0.807
DF [38]	0.802	0.085	0.880	0.763	0.141	0.864	0.747	0.142	0.828
CDS [28]	0.782	0.098	0.884	0.744	0.160	0.833	0.741	0.0958	0.801
DCMC [56]	0.802	0.085	0.880	0.763	0.141	0.864	0.704	0.169	0.786
LBE [57]	0.762	0.081	0.855	0.695	0.153	0.803	0.621	0.278	0.736
DES [17]	0.702	0.125	0.700	0.713	0.189	0.754	0.602	0.243	0.705

concave surface lemma [59] is used to derive the normalization coefficient. Suppose a pixel has Poisson probability  $p_i$ , then  $p_i$  is approximate as probability measure for the maximum-likelihood estimation. [45]. Hence, the maximum bond of similarity of region or pixel's symmetric surround is approximated in terms of the total variation in Poisson distribution [60]. This normalized likelihood luminance plane is used to measure the global probabilistic contrast by subtracting the maximum likelihood from image planes in CIE-LAB space rather than mean of image planes.

Let us define the Poisson probability distribution  $\phi(\mu)$  with mean  $\mu$ . Let  $P$  and  $Q$  be probability measures on  $\{0, 1, 2, 3, \dots, N\}$  with point probabilities as  $p_i$  and  $q_i$  in terms of pixel values in CIE-LAB color image where  $i = \{0, 1, 2, 3, \dots, N\}$  and  $N = 255$ . Then, the total variation between the distributions is defined as:

$$\|P - Q\| = \sum_{i=0}^N |p_i - q_i| \quad (21)$$

The divergence of information or region of similarity for creating the contrast is defined as:

$$D(P \parallel Q) = \sum_{i=0}^N p_i \log \frac{p_i}{q_i} \quad (22)$$

**Lemma 1** *The bond of maximum divergence in total variation in Poisson distribution is used as normalization coefficient of luminance plane over chrominance plane.*

**Proof** The bond of maximum divergence in total variation is defined as a region of influence of the uneven distribution of luminance over chrominance planes around the points in Poisson distribution space. Consider the  $X_1, X_2, X_3, \dots, X_N$  is a sequence of image planes, defined as independent Bernoulli probabilistic distribution, where  $P_k = P(X_k = 1)$  and  $\mu = \sum^n p_k$ .

$$\begin{aligned} D(X_k) &= (1-p_k) \ln \left( \frac{1-p_k}{\exp(-p_k)} \right) + p_k \ln \left( \frac{1-p_k}{\exp(-p_k)} \right) \\ &= (1-p_k) \ln(1-p_k) + p_k \\ &\leq (1-p_k) \left( -p_k - \frac{p_k^2}{2} - \frac{p_k^3}{3} \right) + p_k \\ &= \left( \frac{1}{2} p_k^2 + \frac{1}{6} p_k^3 + \frac{1}{3} p_k^4 \right) \end{aligned} \quad (23)$$

□

Other characteristics to measure the divergence of information in Poisson distribution are duly proved and discussed in various lemmas and theorems by H.Perter [59].

## Compliance with ethical standards

**Conflict of interest** We declare that we have no conflict of interest.

## References

- Itti, L., Rees, G., Tsotsos, J.K.: Neurobiology of Attention. Elsevier, Amsterdam (2005)
- Sharma, S., Kiros, R., Salakhutdinov, R.: Action recognition using visual attention. arXiv preprint [arXiv:1511.04119](https://arxiv.org/abs/1511.04119) (2015)
- Wang, W., Shen, J., Yang, R., Porikli, F.: Saliency-aware video object segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(1), 20–33 (2017)
- Wang, W., Shen, J., Shao, L.: Consistent video saliency using local gradient flow optimization and global refinement. *IEEE Trans. Image Process.* **24**(11), 4185–4196 (2015)
- Fan, D.P., Wang, W., Cheng, M.M., Shen, J.: Shifting more attention to video salient object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8554–8564 (2019)
- Wang, W., Shen, J., Shao, L.: Video salient object detection via fully convolutional networks. *IEEE Trans. Image Process.* **27**(1), 38–49 (2017)
- Zhang, W.C., Shui, P.L.: Contour-based corner detection via angle difference of principal directions of anisotropic gaussian directional derivatives. *Pattern Recogn.* **48**(9), 2785–2797 (2015)
- Cheng, M.M., Mitra, N.J., Huang, X., Torr, P.H., Hu, S.M.: Salient object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1**, 1–1 (2014)
- Margolin, R., Tal, A., Zelnik-Manor, L.: What makes a patch distinct? In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1139–1146 (2013)
- Perazzi, F., Krähenbühl, P., Pritch, Y., Hornung, A.: Saliency filters: contrast based filtering for salient region detection. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 733–740 (2012)
- Shi, J., Yan, Q., Xu, L., Jia, J.: Hierarchical image saliency detection on extended cssd. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(4), 717–729 (2016)
- Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition, 2009: cvpr 2009, IEEE*, pp. 1597–1604 (2009)
- Cheng, M.M., Mitra, N.J., Huang, X., Torr, P.H., Hu, S.M.: Global contrast based salient region detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(3), 569–582 (2015)
- Kienzle, W., Franz, M.O., Schölkopf, B., Wichmann, F.A.: Center-surround patterns emerge as optimal predictors for human saccade targets. *J. Vis.* **9**(5), 7–7 (2009)
- Zhu, W., Liang, S., Wei, Y., Sun, J.: Saliency optimization from robust background detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2814–2821 (2014)
- Zhang, J., Sclaroff, S.: Exploiting surroundedness for saliency detection: a boolean map approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(5), 889–902 (2016)



17. Cheng, Y., Fu, H., Wei, X., Xiao, J., Cao, X.: Depth enhanced saliency detection method. In: Proceedings of International Conference on Internet Multimedia Computing and Service, ACM, p. 23 (2014)
18. Tang, Y., Tong, R., Tang, M., Zhang, Y.: Depth incorporating with color improves salient object detection. *Visual Comput.* **32**(1), 111–121 (2016)
19. Zhou, X., Wang, Y., Zhu, Q., Xiao, C., Lu, X.: Ssg: superpixel segmentation and grabcut-based salient object segmentation. *Visual Comput.* **35**(3), 385–398 (2019)
20. Zhang, Q., Lin, J., Li, W., Shi, Y., Cao, G.: Salient object detection via compactness and objectness cues. *Visual Comput.* **34**(4), 473–489 (2018)
21. Zhong, G., Liu, R., Cao, J., Su, Z.: A generalized nonlocal mean framework with object-level cues for saliency detection. *Visual Comput.* **32**(5), 611–623 (2016)
22. Kapoor, A., Biswas, K., Hanmandlu, M.: An evolutionary learning based fuzzy theoretic approach for salient object detection. *Visual Comput.* **33**(5), 665–685 (2017)
23. Zhao, R., Ouyang, W., Li, H., Wang, X.: Saliency detection by multi-context deep learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1265–1274 (2015)
24. Li, X., Li, Y., Shen, C., Dick, A., Van Den Hengel, A.: Contextual hypergraph modeling for salient object detection. In: 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, pp. 3328–3335 (2013)
25. Li, Z., Lang, C., Feng, S., Wang, T.: Saliency ranker: a new salient object detection method. *J. Vis. Commun. Image Represent.* **50**, 16–26 (2018)
26. Achanta, R., Süssstrunk, S.: Saliency detection using maximum symmetric surround. In: 2010 17th IEEE International Conference on Image Processing (ICIP), IEEE, pp. 2653–2656 (2010)
27. Huang, X., Zhang, Y.J.: 300-fps salient object detection via minimum directional contrast. *IEEE Trans. Image Process.* **26**(9), 4243–4254 (2017)
28. Zhu, C., Li, G., Wang, W., Wang, R.: An innovative salient object detection using center-dark channel prior [c]. In: IEEE International Conference on Computer Vision Workshop (ICCVW) (2017)
29. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: CVPR'07: IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1–8 (2007)
30. Lv, Q., Wang, B., Zhang, L.: Saliency computation via whitened frequency band selection. *Cogn. Neurodyn.* **10**(3), 255–267 (2016)
31. Cheng, M.M., Warrell, J., Lin, W.Y., Zheng, S., Vineet, V., Crook, N.: Efficient salient region detection with soft image abstraction. In: 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, pp. 1529–1536 (2013)
32. Qin, Y., Lu, H., Xu, Y., Wang, H.: Saliency detection via cellular automata. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 110–119 (2015)
33. Zhu, C., Li, G.: A three-pathway psychobiological framework of salient object detection using stereoscopic technology. In: The IEEE International Conference on Computer Vision (ICCV) (2017)
34. Cheng, G., Han, J., Zhou, P., Xu, D.: Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection. *IEEE Trans. Image Process.* **28**(1), 265–278 (2019)
35. Zhu, C., Li, G., Guo, X., Wang, W., Wang, R.: A multilayer back-propagation saliency detection algorithm based on depth mining. In: International Conference on Computer Analysis of Images and Patterns, Springer, pp. 14–23 (2017)
36. Peng, H., Li, B., Xiong, W., Hu, W., Ji, R.: Rgb-d salient object detection: a benchmark and algorithms. In: European Conference on Computer Vision, Springer, pp. 92–109 (2014)
37. Niu, Y., Geng, Y., Li, X., Liu, F.: Leveraging stereopsis for saliency analysis. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 454–461 (2012)
38. Qu, L., He, S., Zhang, J., Tian, J., Tang, Y., Yang, Q.: Rgb-d salient object detection via deep fusion. *IEEE Trans. Image Process.* **26**(5), 2274–2285 (2017)
39. Wang, N., Gong, X.: Adaptive fusion for rgb-d salient object detection. *IEEE Access* **7**:55(284), 277–55 (2019)
40. Zhang, J., Fan, D.P., Dai, Y., Anwar, S., Saleh, F.S., Zhang, T., Barnes, N.: Uc-net: uncertainty inspired rgb-d saliency detection via conditional variational autoencoders. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8582–8591 (2020)
41. Fu, K., Fan, D.P., Ji, G.P., Zhao, Q.: JI-dcf: Joint learning and densely-cooperative fusion framework for rgb-d salient object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3052–3062 (2020)
42. Fan, D.P., Lin, Z., Zhang, Z., Zhu, M., Cheng, M.M.: Rethinking rgb-d salient object detection: models, data sets, and large-scale benchmarks. *IEEE Trans. Neural Netw. Learn. Syst.* **49**(4), 1115–1133 (2020)
43. Li, G., Liu, Z., Ling, H.: Icnnet: Information conversion network for rgb-d based salient object detection. *IEEE Trans. Image Process.* **29**, 4873–4884 (2020)
44. Zhang, M., Ren, W., Piao, Y., Rong, Z., Lu, H.: Select, supplement and focus for RGB-D saliency detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3472–3481 (2020)
45. Singh, S.K., Srivastava, R.: A novel probabilistic contrast-based complex salient object detection. *J. Math. Imaging Vis.* **61**(7), 990–1006 (2019)
46. Wang, L., He, L., Mishra, A., Li, C.: Active contours driven by local gaussian distribution fitting energy. *Sig. Process.* **89**(12), 2435–2447 (2009)
47. Grigorescu, C., Petkov, N., Westenberg, M.A.: Contour and boundary detection improved by surround suppression of texture edges. *Image Vis. Comput.* **22**(8), 609–622 (2004)
48. Malik, J., Belongie, S., Leung, T., Shi, J.: Contour and texture analysis for image segmentation. *Int. J. Comput. Vis.* **43**(1), 7–27 (2001)
49. Kourtzi, Z., Kanwisher, N.: Representation of perceived object shape by the human lateral occipital complex. *Science* **293**(5534), 1506–1509 (2001)
50. Zhang, L., Tong, M.H., Marks, T.K., Shan, H., Cottrell, G.W.: Sun: a bayesian framework for saliency using natural statistics. *J. Vis.* **8**(7), 32–32 (2008)
51. Li, Y., Hou, X., Koch, C., Rehg, J.M., Yuille, A.L.: The Secrets of Salient Object Segmentation, Georgia Institute of Technology (2014)
52. Ju, R., Ge, L., Geng, W., Ren, T., Wu, G.: Depth saliency based on anisotropic center-surround difference. In: 2014 IEEE International Conference on Image Processing (ICIP), IEEE, pp. 1115–1119 (2014)
53. Fan, D.P., Gong, C., Cao, Y., Ren, B., Cheng, M.M., Borji, A.: Enhanced-alignment measure for binary foreground map evaluation. arXiv preprint [arXiv:1805.10421](https://arxiv.org/abs/1805.10421) (2018)
54. Fan, D.P., Cheng, M.M., Liu, Y., Li, T., Borji, A.: Structure-measure: a new way to evaluate foreground maps. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4548–4557 (2017)
55. Zhang, L., Yang, C., Lu, H., Ruan, X., Yang, M.H.: Ranking saliency. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(9), 1892–1904 (2017)
56. Cong, R., Lei, J., Zhang, C., Huang, Q., Cao, X., Hou, C.: Saliency detection for stereoscopic images based on depth confidence anal-



- ysis and multiple cues fusion. *IEEE Signal Process. Lett.* **23**(6), 819–823 (2016)
57. Feng, D., Barnes, N., You, S., McCarthy, C.: Local background enclosure for RGB-D salient object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2343–2350 (2016)
  58. Tu, W.C., He, S., Yang, Q., Chien, S.Y.: Real-time salient object detection with a minimum spanning tree. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2334–2342 (2016)
  59. Harremoës, P.: Binomial and poisson distributions as maximum entropy distributions. *IEEE Trans. Inf. Theory* **47**(5), 2039–2041 (2001)
  60. Serfling, R.J., et al.: A general poisson approximation theorem. *Ann. Probab.* **3**(4), 726–731 (1975)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Rajeev Srivastava** is working as a Professor and Head in the Department of Computer Science and Engineering, Indian Institute of Technology (BHU), IIT (BHU), Varanasi, Uttar Pradesh, India. He received his B.E. in Computer Engineering from Gorakhpur University, India, his M.E. degree in Computer Technology and Applications and a PhD degree in Computer Engineering both from the University of Delhi, Delhi, India. He has around 21 years of teaching and research experience. He

has around 120+ research publications to his credit. He has also edited four books in the areas of image processing and computer vision published from internationally reputed publishers. His research interests include computer vision and machine learning.



**Surya Kant Singh** received the B. Tech degree in Computer Science and Engineering from Uttar Pradesh Technical University, Lucknow, in 2008. He received the M.E. degree in Computer Science and Engineering from the PEC University of Technology, Chandigarh, India, in 2012. He is currently pursuing PhD in Computer Science Department, with Indian Institute of Technology (Banaras Hindu University) Varanasi, Uttar Pradesh, India. He is working as Assistant Professor in LNJPIT

Chapra. He worked as Assistant Professor in GLA University Mathura in Computer Science and Engineering from 2008 to 2010 and 2012 to 2017. His research interests include algorithm designing, computer vision, salient object detection, pattern recognition image analysis machine and deep learning.