Level set method with Retinex-corrected saliency embedded for image segmentation

Dongmei Liu1,2 | Faliang Chang2 | Huaxiang Zhang1 | Li Liu1

1 School of Information Science and Engineering, Shandong Normal University, Jinan, China
2 School of Control Science and Engineering, Shandong University, Jinan, China

Abstract
It can be a very challenging task when using level set method segmenting natural images with high intensity inhomogeneity and complex background scenes. A new synthesis level set method for robust image segmentation based on the combination of Retinex-corrected saliency region information and edge information is proposed in this work. First, the Retinex theory is introduced to correct the saliency information extraction. Second, the Retinex-corrected saliency information is embedded into the level set method due to its advantageous quality which makes a foreground object stand out relative to the backgrounds. Combined with the edge information, the boundary of segmentation will be more precise and smooth. Experiments indicate that the proposed segmentation algorithm is efficient, fast, reliable, and robust.

1 | INTRODUCTION

Image segmentation is a process of decomposing an image into multiple areas, which aims to facilitate the extraction of information and interpretation of image contents. It is a fundamental but important task in image analysis and computer vision applications, e.g. object extraction [1], content-based image retrieval [2] and medical image analysis [3]. Over the past decades, a plethora of methods for image segmentation have been proposed, such as clustering [4], region growing [5] and curve evolution [6]. In recent years, active contour method (ACM) [6–8] has been one of the most extensively applied methods for image segmentation. The basic idea is to evolve a curve to approach object and finally converges to object’s edge by minimizing its energy functional.

Among energy minimization approaches, the level set method (LSM) [9] has been proved an effective way to handle curve evolution. The existing LSM-based image segmentation models can be categorized into two classes: region-based models and edge-based models. The region-based level set models define a region descriptor (e.g. intensity, colour, texture or motion) to guide the motion of the level set function (LSF) to identify each region of interest. One of the early region-based models is proposed by Mumford and Shah [10] which approximates the image using a smooth function inside each region. Chan–Vese (CV) model [6] is one of the most popular region-based models where the image is approximated using a constant function. Some local fitting models have been proposed later on, such as local binary fitting (LBF) [11], local image fitting (LIF) [12], Markov random field (MRF) embedded [13] etc. Li et al. [11] proposed a region-based ACM (LBF) that draws upon intensity information in local regions at a controllable scale. Zhang et al. [14] presented a level set method for image segmentation with intensity inhomogeneity by exploiting local image region statistics, in which they defined a mapping from the original image domain to another domain so that intensity probability model is more robust to noise while suppressing the intensity overlapping to some extent. Ding et al. [15] presented an active contour model using local pre-fitting energy for fast image segmentation, and the core idea of local pre-fitting energy is to define two pre-fitting functions by computing average image intensities locally before the evolution of curve. The edge-based level set models often use an image gradient to control the motion of the active contours and stop the contour on the desired object’s boundaries. Li et al. [16] proposed an distance regularization level set evolution model (DRLSE) to solve the periodical re-initialization problem. Wang et al. [17] introduced an enhanced distance regularized level set evolution completely free of the re-initialization procedure based on analysing these recent regularization models. Liu et al. [18] proposed a
weighted edge-based level set method based on multi-local statistical information to better segment synthetic and real images that have added different types and levels of noise. The region-based models usually rely on the statistical information of the image, which are sensitive to the intensity inhomogeneity of the image. The edge-based models depend on the object edge information, which suffer to boundary leakage problem with weak edges.

Recently, image saliency detection, which aims at locating important and semantic areas or objects from a complicated circumstance, has been beneficial for computer vision applications including object recognition [19], image and video compression [20], object detection [21], to name a few. In terms of the used models, saliency detection methods can be categorized into two classes: top-down methods (task-driven) and bottom-up methods (stimuli-driven). Top-down models [22, 23] are driven by task, which use prior knowledge to find information relevant to the ongoing tasks or goals. These methods have achieved encouraging performance, but it is expensive to generate pixel-wise training data [24]. By comparison, bottom-up methods [25–34] are more extensively studied and used. These methods are usually triggered by stimulus such as colour, intensity, orientation, shape etc.

Saliency information is sensitive to human visual system, and it can be taken into account to yield promising results in image segmentation. In this paper, a new hybrid model for image segmentation which combines the saliency information with the level set segmentation method is proposed. The saliency information of an image is extracted after Retinex-based inhomogeneity correction and embedded into the region-based level set function, and then the evolution of the curve is driven by an improved level set equation combining both region and edge information.

The remainder of this paper is summarized as follows. Related works and motivation of the proposed algorithm is presented in Section 2; In Section 3, the Retinex-corrected saliency extraction and the proposed level set method are described in detail. Experimental results and discussions are given in Section 4. The paper is then concluded in Section 5.

2 BACKGROUND AND MOTIVATION

In this section, some back ground knowledge of related works are first briefly introduced: the Retinex theory, the saliency detection model, the CV model and DRLSE model. And the motivation of this work is given in Section 2.4.

2.1 Retinex theory

The Retinex theory was first proposed by Land and McCann [35]. This theory attempts to explain colour constancy phenomenon of the human visual system and compensates illumination effects in images. It is now widely used in tone mapping [36], image enhancement [37], image segmentation [38] etc.

The Retinex theory assumes that an observed image $I_{obs}(x, y)$ can be decomposed into an illumination component $L(x, y)$ and a reflectance component $R(x, y)$, such that:

$$\begin{aligned}
I_{obs}(x, y) &= L(x, y)R(x, y) \\
\end{aligned} \tag{1}$$

Reflectance describes the nature property of objects, which is considered to be consistent under any lightness conditions. The illumination represents the various lightness on objects. The primary goal of Retinex-based algorithm is to obtain reflectance component $R(x, y)$ from observed image, which is an ill-posed problem. Multi-scale Retinex with colour restoration (MSRCR) [39] is one of centre/surround approaches to solve this problem, and it is widely used to achieve simultaneous dynamic range compression, colour consistency and lightness rendition. Mathematically,

$$\begin{aligned}
R_{msrcr}(x, y) &= C_i(x, y) \sum_{n=1}^{N} w_n \log |I_{msrcr}(x, y)| \\
&- \log |F(\sigma_n) * I_{obs}(x, y)| \\
C_i(x, y) &= \rho \log |yI_{obs}(x, y)| - \log \left( \sum_{n=1}^{N} I_{obs}(x, y) \right) \\
\end{aligned} \tag{2}$$

where $N$ is the number of scales, $w_n$ and $F(\sigma_n)$ are respectively the weight and Gaussian function associated with the $n$th scale, and a log operation is applied to simplify the calculation. $C_i$ is the $i$th band of colour restoration function, where $\rho$ is a gain constant, and $y$ controls the strength of the non-linearity.

2.2 Saliency detection methods

Saliency detection aims to locate the important and conspicuous areas or objects of natural scenes. Itti et al. [25] introduced a biologically inspired saliency model by evaluating the contrast of low-level features (colour, orientation of edges and intensity). FT [27] is a frequency-tuned approach of finding salient objects in images using low level features of colour and luminance. Cheng et al. [29] introduced a regional contrast (RC) based salient object detection algorithm, which simultaneously evaluates global contrast differences and spatial weighted coherence scores. Ishikura et al. [31] present a saliency detection technique, which can maintain high detection performance on images containing objects of various sizes and salient sub-objects by using a combination of multiscale perceptual-colour-difference extrema and a measure of global saliency based on directional-rarity and colour-rarity. Wu et al. [34] proposed a novel propagation model via deformed smoothness constraint to deal with the problem that some object regions with low contrast to background will miss during saliency detection. Some saliency detection models [41–43] have been proposed based on deep learning [44]. Li et al. [41] presented a multi-scale cascade network to identify the most visually conspicuous objects in an image. Luo et al. [42] introduced a webly-supervised learning method utilizing large amounts of web data for salient object detection.

Figure 1 shows results of some saliency detection models: Itti (IT) model [25], graph-based (GB) model [26], frequency-tuned
(FT) model [27], context-aware (CA) model [40], wavelet transform saliency detection (WTS) model [28] and non-subsampled contourlet transform saliency detection (NSCTS) model [30]. The pixels with higher values in bright of these typical results of saliency detection models are salient.

In comparison, the FT method is chosen in this work, because it outputs full resolution saliency maps with well-defined boundaries of salient objects and is simple to implement, which is computationally efficient.

2.3 | Traditional level set models for image segmentation

The level set method was first introduced by Osher et al. [9] for capturing moving fronts, and it is an effective approach to handle curve evolution. In two-dimension image segmentation, the LSM represents a closed curve \( C(t) = \{(x, y)|\phi(x, y, t) = 0\} \) as the zero level set of a level set function (LSF) \( \phi(x, y, t) \) in three dimensions.

The curve moves toward its interior or exterior normal direction and stop at the edge of object. The level set equation (LSE) which defines the evolution function can be written as:

\[
\frac{\partial \phi}{\partial t} = |\nabla \phi| F
\]

where \( F \) is the speed of evolving the curve for solving the differential equation, \( \nabla \) is the gradient operator and \( t \) is the optimization step.

Chan et al. [6] proposed a classic region-based active contour model which used the mean values of different regions to guide contour evolution, and the level set formulation is expressed as follows:

\[
E^{\text{CV}} = \mu \int_{\Omega} \delta(\phi)|\nabla \phi|dxdy + \nu \int_{\Omega} H(\phi)dxdy + \lambda_1 \int_{\Omega} (|I(x, y) - e_1|^2 H(\phi))dxdy + \lambda_2 \int_{\Omega} (|I(x, y) - e_2|^2 (1 - H(\phi)))dxdy
\]

where \( \mu \geq 0, \nu \geq 0, \lambda_1 > 0, \lambda_2 > 0 \) are fixed parameters; the first and second integral items are the length of the curve \( C \) and the area of the region inside \( C \) respectively; the constants \( e_1 \) and \( e_2 \) are the averages of image \( I \) inside \( C \) and outside \( C \) respectively; \( H \) is the Heaviside function: \( H(\phi) = 1 \), when \( \phi(x, y) \geq 0 \); \( H(\phi) = 0 \), when \( \phi(x, y) < 0 \).

Li et al. [16, 45] proposed an edge-based level set active contour model with a regularization term, which is used to solve the periodic re-initialization problem, and the energy functional is expressed as follows:

\[
E^{\text{DRLSE}} = \lambda \int_{\Omega} g \delta(\phi)|\nabla \phi|dxdy + \alpha \int_{\Omega} gH(-\phi)dxdy + \beta \int_{\Omega} \frac{1}{2}(|\nabla \phi| - 1)^2dxdy
\]

where \( g = \frac{1}{1 + |\nabla \phi|^2} \) is an edge detector depending on the gradient of the image and \( C_\nu \) is a Gaussian kernel with standard deviation \( \sigma \).

2.4 | Motivation

The CV model [6] is a region-based level set model, which is less sensitive to noise and the initial placement of evolving contour, but may not work well for segmenting some images with intensity inhomogeneity. The DRLSE model [16] is an edge-based level set model with a regularization term to avoid periodical re-initialization, but may have problems to segment images in the presence of noise and weak object edges.

The human visual system can effectively select important information from complex scenes for further processing. Inspired by this biological capability, visual saliency detection, which can locate important and semantic areas or objects from an complicated circumstance, can be taken into account in image segmentation to yield promising results. Qin et al. [46] combined the region saliency based on entropy rate superpixel with the affinity propagation clustering algorithm, and use random walks method to obtain the segmentation results. Zhi et al. [8] presented a two-stage level set evolution protocol with both saliency map and colour intensity, but this work doesn’t consider the problem of intensity inhomogeneity in saliency detection and segmentation. Wang et al. [47] extracted saliency map to determine the initial contour, and the level set function is obtained by computing the local average Renyi entropy and Canny operator.

Therefore, motivated by aforementioned works, a synthetical level set model based on the combination of edge information and region information is proposed, wherein the region-based information is constituted by Retinex-corrected saliency information and colour intensity. Saliency information embedded into the energy functional can make the foreground objects more distinct to be segmented, with Retinex corrected, it can also address challenges presented by intensity inhomogeneity. The combination of region information and edge information can make the segmentation insensitive to the initializations, robust and fast.
3 | PROPOSED MODEL

The overview of the proposed image segmentation model is illustrated in Figure 2. The saliency information of the input image is firstly generated after Retinex inhomogeneity correction. Then the saliency information is embedded into the region-based level set function. The final energy functional contains both region and edge information and the segmentation is guided by minimizing the proposed energy functional.

As described above, the primary contributions of this method can be summarized as follows:

- Retinex theory is introduced to correct image inhomogeneity and enhance saliency detection, which shows beneficial effect and promote the performance of the proposed model.
- The improved saliency information is embedded as the region information of the energy function, which makes the foreground objects more distinct and the segmentation results more reliable.
- With saliency information embedded, a synthetic level set model based on the combination of region information and edge local information is proposed. Experiments demonstrate its effectiveness and robustness in image segmentation.

3.1 | Retinex-corrected saliency detection

The FT method proposed by Achanta et al. [27] is employed to extract the saliency information in this work. It can output full resolution saliency maps with well-defined boundaries of salient objects, which is easy to implement and computationally efficient. But it does not work well under the circumstance of the illumination problems such as poor illumination or occurrence of shadow like some other saliency detection models. In addition, intensity inhomogeneity is one of significant challenges to some region-based LSM image segmentation. In this paper, we use Retinex to ameliorate intensity inhomogeneity of the image and to suppress the influence of illumination on saliency detection. The Retinex theory originally dealt with colour constancy, which in human vision ensures that the perceived colours of objects remain relatively constant under varying illumination conditions. The Retinex theory can be used to correct intensity inhomogeneity when illumination is recovered and removed from the image.

The saliency detection of this paper consists of the following steps:

1. Convert the input image $I_{\text{in}}$ from RGB colour space to CIELAB colour space.
2. Apply MSR method to the L-channel of the input image to estimate and remove the luminance component in this channel, while the colour channels A and B remain unchanged: $I = (R_{\text{msrcr}}(I_{\text{in}}), I_{\text{A,in}}, I_{\text{B,in}})$.
3. Calculate the saliency map of the improved image $I$ of step (2) by FT model:

$$S(x,y) = ||I_m - I_{\omega}(x,y)||$$  \hspace{1cm} (7)

where $I_m$ is the arithmetic mean pixel value of the image $I$, $I_{\omega}(x,y)$ is the corresponding image pixel vector value in the Gaussian blurred version (using a $5 \times 5$ separable binomial kernel, to eliminate fine texture details as well as noise and coding artefacts) of the original image, and $||\cdot||$ is the $L_2$ norm.

Three examples are given to show the effect of Retinex-correction in saliency detection in Figure 3. The saliency maps are converted into heat maps to get a better visual display. The Retinex method enhances the contrast between the salient objects and background, and ameliorate intensity inhomogeneity of the image. In the example of boat, the boat and people on the boat of Figure 3d show higher saliency than that of Figure 3b while the saliency of the object reflection in the water much lower, which is beneficial in the later segmentation process (Figure 7f). In the second example, some texture details around the tiger are eliminated and the contrast of tiger and background is enhanced, meanwhile, the entire tiger is intensity homogeneous. In the third case, the saliency map after Retinex-corrected of the entire object is more homogeneous than that of before Retinex-corrected. With Gaussian filter operation in Retinex, the background noise is also suppressed. The quantitative comparison between saliency detection with and without Retinex-correction is carried out on dataset MSRA1000 [27]. The average accuracy, average recall and F-Measure are shown in Table 1, and it can be seen that the corrected saliency detection shows better performance.
3.2 Saliency-embedded region energy

In light of the above discussion, we can take the Retinex-corrected saliency information into account to improve segmentation results, which can make an object stand out relative to its neighbors with well-defined boundaries and saliency homogeneity. Thus, under the framework of CV models, the new saliency-embedded region energy arises:

\[
E^{SE} = \mu \int_{\Omega} \delta(\phi)|\nabla \phi|dxdy + \nu \int_{\Omega} H(\phi)dxdy + \lambda_1 \int_{\Omega} |I(x,y) - c_1|^2 H(\phi)dxdy + \lambda_2 \int_{\Omega} |I(x,y) - c_2|^2 (1 - H(\phi))dxdy + \lambda_3 \int_{\Omega} |S(x,y) - s_1|^2 H(\phi)dxdy + \lambda_4 \int_{\Omega} |S(x,y) - s_2|^2 (1 - H(\phi))dxdy
\]

where \(\lambda_3, \lambda_4 > 0\), \(S\) is the saliency map generated from Section 3.1, \(s_1\) and \(s_2\) are the averages of saliency map \(S\) inside and outside contour \(C\) respectively.

The further and detailed discussion on the Retinex-correction saliency effect on the level set method is given in Section 4.1.

3.3 Improved LSM

The CV model often fails to segment images with intensity inhomogeneity or whose objects have similar colour information with background, and the contour can be trapped in a local minimum. The saliency information corrected by Retinex method is embedded into the level set function to solve these problems, and then combing the edge-based DRLSE model because it has a penalty term to solve the LSF’s periodic re-initialization problem, and the segmentation results of this model often have smooth boundaries. The final energy functional is written as follows:

\[
E = \eta E^{SE} + \kappa E^{DRLSE}
\]

\[
= \eta \left[ \mu \int_{\Omega} \delta(\phi)|\nabla \phi|dxdy + \nu \int_{\Omega} H(\phi)dxdy \right] + \lambda_1 \int_{\Omega} |I(x,y) - c_1|^2 H(\phi)dxdy + \lambda_2 \int_{\Omega} |I(x,y) - c_2|^2 (1 - H(\phi))dxdy + \lambda_3 \int_{\Omega} |S(x,y) - s_1|^2 H(\phi)dxdy + \lambda_4 \int_{\Omega} |S(x,y) - s_2|^2 (1 - H(\phi))dxdy + \kappa \left[ \int_{\Omega} g(\phi)|\nabla \phi|dxdy + \alpha \int_{\Omega} gH(\phi)dxdy \right] + \beta \int_{\Omega} \frac{1}{2}(|\nabla \phi| - 1)^2dxdy
\]

where \(H\) is the Heaviside function, and it is usually replaced by its smooth approximation function:

\[
H_\varepsilon(\varepsilon) = \frac{1}{2} \left( 1 + \frac{2}{\pi} \arctan \left( \frac{\varepsilon}{\varepsilon} \right) \right), \quad \varepsilon \to 0
\]

The minimization of the energy functional is an unconstrained optimization problem, so it can be minimized with respect to the fitting variables \(c_1, c_2, s_1, s_2\) and \(\phi\), and the steepest descent method is used to minimize them.

Keeping \(\phi\) fixed and minimizing the energy functional with respect to \(c_1, c_2, s_1\) and \(s_2\), we can obtain:

\[
c_1 = \frac{\int_{\Omega} I(x,y)H_\varepsilon(\phi)dxdy}{\int_{\Omega} H_\varepsilon(\phi)dxdy}, \quad (11)
\]

\[
s_1 = \frac{\int_{\Omega} S(x,y)H_\varepsilon(\phi)dxdy}{\int_{\Omega} H_\varepsilon(\phi)dxdy}, \quad (12)
\]
if \( \int_{\Omega} H_k(\phi) \, dx \, dy > 0 \), i.e. the curve \( C \) has a non-empty interior in \( \Omega \), and

\[
eq \frac{\int_{\Omega} I(x,y)(1 - H_k(\phi)) \, dx \, dy}{\int_{\Omega} (1 - H_k(\phi)) \, dx \, dy},
\]

(13)

\[
eq \frac{\int_{\Omega} S(x,y)(1 - H_k(\phi)) \, dx \, dy}{\int_{\Omega} (1 - H_k(\phi)) \, dx \, dy},
\]

(14)

if \( \int_{\Omega} (1 - H_k(\phi)) \, dx \, dy > 0 \), i.e. the curve \( C \) has a nonempty exterior in \( \Omega \).

Keeping \( \epsilon_1, \epsilon_2, \epsilon_1 \) and \( \epsilon_2 \) fixed, and minimizing the energy functional with respect to \( \phi \). According to the Euler-Lagrange equation \( \partial E / \partial \phi = 0 \), The evolution equation of level set function \( \phi \) is computed as follows:

\[
\frac{\partial \phi}{\partial t} = - \frac{\partial E}{\partial \phi} \]

\[
= \eta \delta_\varepsilon(\phi) \left[ \mu \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_1 (1 - \epsilon_1)^2 
+ \lambda_2 (1 - \epsilon_2)^2 - \lambda_3 (\delta - \eta)^2 + \lambda_4 (\delta - \eta_2)^2 \right] 
+ \kappa \left\{ \lambda \delta_\varepsilon(\phi) \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + \alpha \sigma_\varepsilon(\phi) \right\} 
+ \beta \left[ \Delta \phi + \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \right] \}
\]

(15)

where \( \delta_\varepsilon(z) \) is a smooth Dirac function that is the derivative of \( H_\varepsilon \):

\[
\delta_\varepsilon(z) = H'_\varepsilon(z) = \frac{\varepsilon}{(\varepsilon^2 + z^2)\pi}
\]

(16)

## 4 | EXPERIMENTAL RESULTS

In this section, we implemented some experiments to evaluate the proposed Retinex-corrected saliency embedded level set image segmentation method. The experiments are implemented using the MATLAB R2018a on a personal computer, with CPU 2.59 GHz Intel Core-i7, 16 GB RAM, and Windows 10 operating system.

### 4.1 | Effects of Retinex-correction saliency on image segmentation

Saliency information is embedded into the proposed level set segmentation model, which is intensity inhomogeneity corrected by Retinex method in advance (shown in Figure 3). With these operations, this model is expected to acquire better segmentation results. In order to demonstrate this, the following experiment is designed.

By setting different parameters of Equation (9) and different inputs, we can obtain different level set active contour models. In particular, when \( \eta = 1, \kappa = 0, \lambda_1 \neq 0, \lambda_2 \neq 0, \lambda_3 = 0, \lambda_4 = 0 \), it becomes traditional CV model; when \( \eta = 0, \kappa = 1, \) it becomes traditional DRLSE model. The specific parameters are carefully chosen and are given in Table 2. The second column lists the inputs of each model, where “img” refers to the original image, “sm” refers to the saliency map obtained by ref. [27] and “Rsm” refers to the Retinex-corrected saliency map. The \( \Delta t \) of all models are set to 0.1 and a checkerboard shape is defined to initialize all level set functions:

\[
\phi_0(x,y) = \sin \left( \frac{\pi \times H \times x}{1500} \right) \sin \left( \frac{\pi \times W' \times y}{1500} \right)
\]

(17)

where \( H \) and \( W' \) are respectively the height and width of input image.

Figure 4 shows the segmentation results obtained from using the CV model and the proposed SE model, RSE model and RSED model. Figure 4a are the input images from MSRA1000...
Figure 4b–e are the segmentation results of these models respectively. The first image contains a tomato illuminated by sidelight, which has high intensity inhomogeneity. In this case, the CV model cannot handle very well high intensity variation; the SE (saliency-embedded) model also shows poor performance in segmenting tomato, because the saliency map without Retinex-correction is intensity inhomogeneous (as shown in Figure 3, image 3); the RSE (Retinex-corrected saliency embedded) model segments the whole object and the RSED (RSE model combined with DRLSE model) shows more robust segmentation. The second image is an image with textured background and intensity inhomogeneity. With Retinex-corrected saliency information embedded, the proposed RSED model successfully and precisely segments the distinct object. Other cases also demonstrate the effectiveness of the proposed method.

The Dice similarity coefficient (Dice) and the Hausdorff distance (HD) are used to evaluate the performance of the prosed method in this section. The Dice is one the region-based coefficients based upon the measure of spatial overlap; and the HD is one of distance-based coefficients based upon the measure of the distance between the segmentation contour and the true boundary, which is used when the delineation of the boundary is critical [48, 49].

Let \( S_g \) and \( S_t \) stand for the ground segmentation truth and the segmented region by the compared methods, respectively. The Dice and HD are defined as follows:

\[
\text{Dice}(S_g, S_t) = \frac{2|S_g \cap S_t|}{|S_g| + |S_t|} \quad (18)
\]

\[
\text{HD}(S_g, S_t) = \max \left( b(S_g, S_t), b(S_t, S_g) \right) \quad (19)
\]

where \( b(S_g, S_t) \) is called the directed Hausdorff distance and given by \( b(S_g, S_t) = \max_{i \in S_g} \min_{j \in S_t} \| i - j \| \) and \( \| \cdot \| \) is the chosen norm.

The larger the Dice or the less the HD means the better the segmentation result achieved. The comparisons of the Dice and HD values of images in Figure 4 is shown in Figure 5. It shows that, the saliency information (SE model) can help improving the performance of segmentation, and the segmentation results of the RSE model are better than those of the SE model, and the RSED model achieves the best segmentation.

The embedded saliency information makes the foreground objects easy to be segmented and the Retinex-correction enhances the intensity homogeneity of both foreground and background. Combined with the DRLSE model, the segmentation results of the proposed RSED method have more smooth edges.

### 4.2 Comparisons on natural images

In this section, we perform experiments on natural images from three public image datasets: MSRA1000 [27], PASCAL-S [50] and DUT-OMRON [51]. The MSRA1000 contains 1000 natural images and the image resolution is [400, 300] or [300, 400]. The images in this dataset are mainly contains one significant object and the backgrounds are usually simple. It is widely used in salient object detection and segmentation community. The PASCAL-S dataset contains 850 images, and the images contains multiple complex objects and cluttered backgrounds. It is a challenging dataset for object detection and segmentation. The DUT-OMRON dataset is composed of 5168 images with complex backgrounds and the one dimension of each image is 400 while the other dimension is less than 400. It mainly contains one significant object but the backgrounds are complex. The DUT-OMRON dataset is also a challenging dataset for object detection and segmentation. Some images and their ground truths of the three datasets are shown in Figure 6.
The proposed model is compared with several other level set segmentation models: CV model [6], LBF model [11], DRLSE model [16], LPF model [15] and SDREL model [8]. For better comparison, we also chose two state-of-the-art salient region detection method iC the RC model [29] and the DSC model [34]. The parameters of this model are set as follows: \( \eta = 0.1, \kappa = 0.9, \varepsilon = 1.5, \mu = 5, \nu = 1.5, \lambda_1 = 0.01, \lambda_2 = 0.01, \lambda_3 = 0.06, \lambda_4 = 0.06, \lambda = 5, \alpha = 1.5, \beta = 0.2 \) and \( \Delta t = 0.1 \). The principles of choosing parameter values are: (a) The parameter value of penalization of the area inside contour is equal to that of penalization of the area outside contour, thus \( \lambda_1 = \lambda_2 \) and \( \lambda_3 = \lambda_4 \); (b) The saliency information contributes more than the original image, thus \( \lambda_3 \) and \( \lambda_4 \) are larger than \( \lambda_1 \) and \( \lambda_2 \); (c) The parameters \( \mu \) and \( \nu \) are chosen according to the CV model, and the parameters \( \lambda, \alpha \) and \( \beta \) are chosen according to the DRLSE model; (d) The parameters \( \eta \) and \( \kappa \) are chosen based on experimental experience. The initial contour of the DRLSE models is set as rectangles 10 pixels away from the object/objects. The CV model, LBF model, LPF model, SDREL model and the proposed model use a checkerboard shape defined in Equation (17) as initial contour.

4.2.1 Results on MSRA1000

Visual comparisons on the MSRA1000 of different level set models are shown in Figure 7, and most images of this dataset are simple ones with less complex backgrounds.

The proposed Retinex-corrected-saliency embedded level set segmentation shows robustness results especially when shadow and reflection of objects occurred in the tomato, sign and boat examples. The Retinex operation to saliency extraction can reduce the effects of illumination and correct the intensity inhomogeneity of objects. Then the embedded saliency information makes the objects more distinct and easier to be segment. The CV model and SDREL model suffer effects of intensity inhomogeneity, and DRLSE model is sensitive to noise and easy to stuck around false boundaries. For images containing one main object such as the dog, fish toy, flower and fish examples, the proposed model shows more precise results than the other models. In the flower example, the DRLSE model and LBF model stuck in local minimum. The SDREL model has better result than CV model while the result of CV model contains much background texture. For images contain multiple objects such as the bird, plane and sign, footmen and boat examples, the proposed model shows more reliable results than others. Take the boat image as an example, the CV model and SDREL model cannot segment foreground properly and the LBF model cannot eliminate the reflection of boat in water. Results of the LPF model is robust with intensity inhomogeneity but it is sensitive to noisy backgrounds. The generation of the RC and DSC model are grey-scale saliency maps, the larger the value, the higher the saliency. The proposed RSED model performs better than the other model on this dataset.

In this section, we employ \( F \)-Measure to evaluate the segmentation performance of these models, which is defined as:

\[
F_\beta = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}}
\]

where precision = \( \frac{|A \cap B|}{|A|} \), recall = \( \frac{|A \cap B|}{|B|} \), \( A \) is the detected object and \( B \) represents the ground truth. When \( \beta^2 = 1 \) (precision and recall are equally important), we call it \( F_1 \)-Measure, which is mathematically equivalent to Dice (defined in Equation (18)); when \( \beta^2 > 1 \), it means recall is more important than precision; when \( \beta^2 < 1 \), it means precision is more important than recall.

With simple binarization operation, we can obtain the \( F_1 \)-Measures of each image shown in Figure 7, which are presented in Figure 8. In most cases, the proposed RSED model achieves higher \( F_1 \)-Measure than that of the other models. In image 5, all the models have lower \( F_1 \)-Measure, because the ground truth only contains the bird in the centre of image while the active contour models segment the branch as well. In image 10, the proposed model has a considerable performance improvement. The average precision, average recall, average \( F_1 \)-Measure (\( AF_1 \)), average \( F_2 \)-Measure (\( AF_2 \)) and average \( F_0 \)-Measure (\( AF_0 \)) of these 10 images are given in Figure 9. The DRLSE model achieves the highest average recall, because it is sensitive to the initial contour and the background and easy to stop evolution when the level set function arrives at the strong edges of objects or background. The saliency detection models usually segment the saliency map using the mean-shift method, Ostu method or
GrabCut method. Therefore, the RC and DSC model cannot achieve higher F1-Measure than the proposed model with simple binarization operation.

With different $\xi^2$ in Equation (20), we get different $F_{\xi}$ values, but the $F_{\xi}$'s varying tendencies of all those 10 images are the same. Based on the previous experimental analysis, we use the $F_{0.5}$-Measure ($\xi^2$ is set 0.5 to emphasize precision) to compare the RSED model with the other models on these three datasets in this work. The comparison of average $F_{0.5}$-Measure on MSRA1000 is shown in Figure 12a. It has shown that the proposed model has a higher $F_{0.5}$-Measure than the other compared models.

4.2.2 Results on PASCAL-S

The reason to include PASCAL-S dataset is to assess performance of models over scenes with multiple objects with structurally complex scenes. The visual comparisons of PASCAL-S is shown in Figure 10 and the comparisons of average $F_{0.5}$-Measure on this dataset is shown in Figure 12b.

Both the LBF model and the LPF model are local region-based models, so their segmentation results have great similarities, and the LPF model are more effective dealing with intensity inhomogeneity. The saliency detection method RC model does not achieve good performance on this dataset. Compared with ground truth, the proposed model tends to generate segmentation results with clear boundaries of objects more than salient regions (especially in the third case).

4.2.3 Results on DUT-OMRON

The reason to include DUT-OMRON dataset is to assess performance of models on a large scale dataset. The visual comparisons of DUT-OMRON is shown in Figure 11 and the comparisons of average $F_{0.5}$-Measure on this dataset is shown in Figure 12c. The proposed model can generate segmentation results with clear boundaries and shows better robustness on cases with noise and intensity inhomogeneity.
4.3 Comparisons on synthetic, noisy and medical images

In this section, the proposed model is compared with several other level set segmentation models on synthetic, noisy and medical images. The parameters of this model are set as follows: $\eta = 0.1$, $\kappa = 0.9$, $\varepsilon = 1.5$, $\mu = 5$, $\nu = 1.5$, $\lambda_1 = 0.01$, $\lambda_2 = 0.01$, $\lambda_3 = 0.06$, $\lambda_4 = 0.06$, $\lambda = 5$, $\alpha = 1.5$, $\beta = 0.2$ and $\Delta t = 0.1$.

During the experiments we have discovered that, the segmentation results of DRLSE model can obtain better performance when the initial contour entirely contains objects or between objects. Thus the initial contour of the DRLSE models is set as rectangles 10 pixels away from the object/objects. The CV model, LBF model, LPF model SDREL model and the proposed RSED model use a checkerboard shape defined in Equation (17) as initial contour. Figure 13 shows the comparison of segmentation results of these models.

For a synthetic image contains one single object, all these models show ideal results of the image segmentation, in particular, the last four methods have more smooth boundaries than the first two methods.

For synthetic image contains multiple objects, the DRLSE method generates results with imprecise boundary where the separated objects are detected in adhesion, the reason is that the boundaries of objects which are too close become fuzzy and broaden after smooth filtering. The SDREL method gets irregular boundary because of its gradient terms are also sensitive under this condition. The segmentation boundaries of the CV, LBF and LPF models are not as smooth as that of the proposed model.

For a synthetic image with heavy salt and pepper noise, the segmentation results of the CV method and the SDREL
method cannot eliminate most noise of the image, and the segmentation results of the DRLSE method, LBF method and LPF method are stuck in local minimum due to the severe noise. The proposed RSED method demonstrates robustness performance under this situation, whose result has precise boundary and eliminates most noise.

For a medical image, the local fitting-based active contour models - the LBF method and LPF method generate the most ideal segmentation results, and the proposed RSED method also get more precise segmentation results than the other three methods.

5 | CONCLUSIONS

In this paper, we propose a hybrid level set active contour model for image segmentation with Retinex-corrected saliency information embedded (RSED). The proposed method aims to solve some serious problems of natural image segmentation, such as the intensity inhomogeneity and complicated backgrounds. With the Retinex-corrected saliency information, on one hand, the Retinex correction can suppress the intensity inhomogeneity on both saliency detection and subsequent segmentation. On the other hand, the corrected saliency information can help the foreground objects more distinct to be segmented. Combined with the edge information, the boundaries of the segmentation results can be more precise and smooth. Experiments have demonstrated its efficiency and effectiveness. However, there are still some improvements of this method, for example, an automatic weights calculation of the level set model. We will explore these problems in our future work.

ACKNOWLEDGMENT

This work was supported by National Key R&D Program of China (No. 2018YFB1305300), Shandong Provincial Key Research and Development Program (Major Scientific and Technological Innovation Project) (Nos. 2019JZZY010130, 2018CXGC0907), China Natural Science Foundation Committee (Nos. 6173244, 61703240), and the Taishan Scholar Project of Shandong, China (No. ts20190924).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in reference number [27], reference number [50] and reference number [51].

ORCID

Dongmei Liu https://orcid.org/0000-0002-8061-1797
Faliang Chang https://orcid.org/0000-0003-1276-2267
Huaxiang Zhang https://orcid.org/0000-0001-6259-7533
Li Liu https://orcid.org/0000-0002-9121-5124

REFERENCES

1. Wu, K., Yu, Y.: Automatic object extraction from images using deep neural networks and the level-set method. IET Image Proc. 12(7), 1131–1141 (2018)
2. Smulders, A. et al.: Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. 22(12), 1349–1380 (2000)
3. Zhou, S. et al.: Active contour model based on local and global intensity information for medical image segmentation. Neurocomputing 186, 107–118 (2016)
4. Kim, S. et al.: Image segmentation using higher-order correlation clustering. IEEE Trans. Pattern Anal. Mach. Intell. 36(9), 1761–1774 (2014)
5. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13(6), 583–598 (1991)
6. Chan, T., Vese, L.: Active contours without edges. IEEE Trans. Image Process. 10(2), 266–277 (2001)
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int. J. Comput. Vision 1(4), 321–331 (1988)
8. Zhi, X-H., Shen, H-B.: Saliency driven region-edge-based top down level set evolution reveals the asynchronous focus in image segmentation. Pattern Recognit. 80, 241–255 (2018)
9. Osher, S., Sethian, J.: Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. J. Comput. Phys. 79(1), 12–49 (1988)
10. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. 42(5), 577–685 (1989)
11. Li, C., Kao, C., Gore, J.: Minimization of region-scalable fitting energy for image segmentation. IEEE Trans. Image Process 17(10), 1940–1949 (2008)
12. Zhang, K., Song, H., Zhang, L.: Active contours driven by local image fitting energy. Pattern Recognit. 43(4), 1199–1206 (2010)
13. Yang, X., Gao, X., Tao, D., et al.: An efficient MRF embedded level set method for image segmentation. IEEE Trans. Image Process. 24(1), 9–21 (2015)
14. Zhang, K. et al.: A level set approach to image segmentation with intensity inhomogeneity. IEEE Trans. Cybern. 46(2), 546–557 (2016)
15. Ding, K., Xiao, L., Weng, G.: Active contours driven by local pre-fitting energy for fast image segmentation. Pattern Recognit. Lett. 104, 29–36 (2018)
16. Li, C., Xu, C., Gui, C., et al.: Distance regularized level set evolution and its application to image segmentation. IEEE Trans. Image Process 19(12), 3243–3254 (2010)
17. Wang, X. et al.: Enhanced distance regularization for re-initialization free level set evolution with application to image segmentation. Neurocomputing 141, 223–235 (2014)
18. Liu, C., Liu, W., Xing, W.: A weighted edge-based level set method based on multi-local statistical information for noisy image segmentation. J. Visual Commun. Image Represent. 59, 89–107 (2019)
19. Ren, Z. et al.: Region-based saliency detection and its application in object recognition. IEEE Trans. Circuits Syst. Video Technol. 24(5), 769–779 (2014)
20. Guo, C., Zhang, L.: A novel multisresolution spatiotemporal saliency detection model and its applications in image and video compression. IEEE Trans. Image Process 19(1), 185–198 (2010)
21. Yang, F., Xu, Q., Li, B.: Ship detection from optical satellite images based on saliency segmentation and structure-Ibp feature. IEEE Geosci. Remote Sens. Lett. 14(5), 602–606 (2017)
22. Borji, A., Sihite, D., Itti, L.: What/where to look next? Modeling top-down visual attention in complex interactive environments. IEEE Trans. Syst. Man Cybern. Syst. 44(5), 523–538 (2014)
23. Yang, J., Yang, M-H.: Top-down visual saliency via joint crf and dictionary learning. IEEE Trans. Pattern Anal. Mach. Intell. 39(3), 576–588 (2017)
24. Zhou, Y. et al.: Salient object detection via fuzzy theory and object-level enhancement. IEEE Trans. Multimedia 21(1), 74–85 (2019)
25. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20(11), 1254–1259 (1998)
26. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. Adv. Neural Inform. Process. Syst. 19, 545–552 (2007)
27. Achanta, R. et al.: Frequency-tuned salient region detection. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1597–1604. IEEE, Piscataway, NJ (2009)
28. Imamoglu, N., Lin, W., Fang, Y.: A saliency detection model using low-level features based on wavelet transform. IEEE Trans. Multimedia 15(1), 96–105 (2013)
29. Cheng, M., Mitra, N., Huang, X.: Global contrast based salient region detection. IEEE Trans. Pattern Anal. Mach. Intell. 37(5), 569–582 (2015)
30. Liu, D., Chang, F., Liu, C.: Salient object detection fusing global and local information based on nonsubsampled contourlet transform. J. Opt. Soc. Am. A 33(8), 1430–1441 (2016)
31. Ishikura, K. et al.: Saliency detection based on multiscale extrema of local perceptual color differences. IEEE Trans. Image Process. 27(2), 703–717 (2018)
32. Li, X. et al.: Saliency transfer: An example-based method for salient object detection. International Joint Conferences on Artificial Intelligence, pp. 3411–3417. AAAI Press, Palo Alto, CA (2016)
33. Li, X. et al.: Contour knowledge transfer for salient object detection. European Conference on Computer Vision, pp. 355–370. Springer, Berlin (2018)
34. Wu, X. et al.: Salient object detection via deformed smoothness constraint. 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 2815–2819. IEEE, Piscataway, NJ (2018)
35. Land, E., Mccann, J.: Lightness and Retinex theory. J. Opt. Soc. Am. 61(1), 1–11 (1971)
36. Kim, K., Bae, J., Kim, J.: Natural HDR image tone mapping based on Retinex. IEEE Trans. Consum. Electron. 57(4), 1807–1814 (2011)
37. Li, M., Liu, J., Yang, W., et al.: Structure-revealing low-light image enhancement via robust Retinex model. IEEE Trans. Image Process 27(6), 2828–2841 (2018)
38. Zosso, D. et al.: Image segmentation with dynamic artifacts detection and bias correction. Inverse Problems Imaging 11(3), 577–600 (2017)
39. Jobson, D., Rahman, Z., Woodell, G.: A multiscale Retinex for bridging the gap between color images and the human observation of scenes. IEEE Trans. Image Process 6(7), 965–976 (1997)
40. Goferman, S., Zelnik-Manor, L., Tal, A.: Context-aware saliency detection. IEEE Trans. Pattern Anal. Mach. Intell. 34(12), 1915–1926 (2012)
41. Li, X. et al.: Multi-scale cascade network for salient object detection. Proceedings of the 25th ACM International Conference on Multimedia, pp. 439–447. ACM Press, New York (2017)
42. Luo, A. et al.: Webly-supervised learning for salient object detection. Pattern Recognit. 103, 107308 (2020)
43. Luo, A. et al.: Cascade graph neural networks for rgb-d salient object detection. European Conference on Computer Vision, pp. 346–364. Springer, Berlin (2020)
44. Yang, F., Li, X., Cheng, H., et al.: Multi-scale bidirectional fcn for object skeleton extraction. AAAI Conference on Artificial Intelligence, pp. 7461–7468. AAAI Press, Palo Alto, CA (2018)
45. Li, C., Xu, C., Gui, C.: Level set evolution without re-initialization: a new variational formulation. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 430–436. IEEE, Piscataway, NJ (2005)
46. Qin, C. et al.: Integration of the saliency-based seed extraction and random walks for image segmentation. Neurocomputing 129, 378–391 (2014)
47. Wang, Y., Xu, X.: An improved level set method to image segmentation based on saliency. J. Inf. Process. Syst. 15(1), 7–21 (2019)
48. Chang, H-H. et al.: Performance measure characterization for evaluating neuroimage segmentation algorithms. NeuroImage 47(1), 122–135 (2009)
49. Taha, A., Hanbury, A.: Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. BMC Med. Imaging 15(1), 29 (2015)
50. Li, Y. et al.: The secrets of salient object segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 280–287. IEEE, Piscataway, NJ (2014)
51. Yang, C. et al.: Saliency detection via graph-based manifold ranking. 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3166–3173. IEEE, Piscataway, NJ (2013)

How to cite this article: Liu D, Chang F, Zhang H, Liu L. Level set method with Retinex-corrected saliency embedded for image segmentation. IET Image Process. 2021;15:1530–1541. https://doi.org/10.1049/ipr2.12123