Accepted Manuscript

Saliency Driven Region-Edge-based Top Down Level Set Evolution Reveals the Asynchronous Focus in Image Segmentation

Xu-Hao Zhi , Hong-Bin Shen

PII: S0031-3203(18)30100-6
DOI: 10.1016/j.patcog.2018.03.010
Reference: PR 6488

To appear in: *Pattern Recognition*

Received date: 9 February 2017
Revised date: 4 December 2017
Accepted date: 11 March 2018

Please cite this article as: Xu-Hao Zhi , Hong-Bin Shen , Saliency Driven Region-Edge-based Top Down Level Set Evolution Reveals the Asynchronous Focus in Image Segmentation, *Pattern Recognition* (2018), doi: 10.1016/j.patcog.2018.03.010

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights

- A new level set energy function has been designed by incorporating both the region and edge information.
- Saliency knowledge has been modeled into the level set evolution.
- A hierarchical evolution approach has revealed the asynchronous focus in image segmentation.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Saliency Driven Region-Edge-based Top Down Level Set Evolution Reveals the Asynchronous Focus in Image Segmentation

Xu-Hao Zhi and Hong-Bin Shen*

Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China

* Address correspondence to H.B. Shen at hbshen@sjtu.edu.cn

Tel: +86-21-34205320

Fax: +86-21-34204022

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Saliency Driven Region-Edge-based Top Down Level Set Evolution Reveals the Asynchronous Focus in Image Segmentation

Xu-Hao Zhi and Hong-Bin Shen

Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China

* Address correspondence to H.B. Shen at hbshen@sjtu.edu.cn

Tel: +86-21-34205320

Fax: +86-21-34204022

ACCEPTED MANUSCRIPT

ABSTRACT

Level set method (LSM) is popular in image segmentation due to its intrinsic features

for handling complex shapes and topological changes. Existing LSM-based segmentation models can be generally grouped into region- and edge-based models. The former often have problems to deal with images whose objects have similar color intensity to that of the background when the region descriptor is insufficient. The latter usually suffer to boundary leakage problem when the images' edges are weak. To overcome these problems, we present a novel hierarchical level set evolution protocol (SDREL), wherein we propose to use both saliency map and color intensity as region external energy to motivate an initial evolution of level set function (LSF), followed by the LSF and further smoothed by an internal energy (regulation term) to recognize a more precise boundary positioning. Our results show that the newly introduced saliency map term improves extracting objects from complex background and the asynchronous evolution of a single LSF results in a better segmentation. The new hierarchical SDREL model has been evaluated extensively and the results indicate that it has the merits of flexible initialization, robust evolution, and fast convergence. SDREL is available at: www.csbio.sjtu.edu.cn/bioinf/SDREL/.

Keywords: Image segmentation; Level set evolution; Saliency map; Edge; SDREL

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. Introduction

Image segmentation is a process of dividing an image into meaningful areas, which aims to simplify information extraction from the image for further content analysis or understanding. In the past decades, a variety of algorithms for image segmentation have been introduced, such as thresholding [1-3], clustering [4, 5], region growing [6], and curve evolution [7, 8], etc. Curve evolution is an effective method to perform image segmentation tasks, which has two formulation models: parametric active contour model (e.g. Snakes [8]) and geometric active contour model (e.g. the level set model). Curve evolution first defines an energy function to measure energy of a curve, which is determined by the curve's geometrical shape and distance to object. Then according to mathematical rule of topological changes, in the process of minimizing energy function, the curve approaches to object and finally converges to object's edge. Essentially, curve evolution formulates image segmentation task into energy function's optimization, which can be solved by mathematical tools like Euler-Lagrange equation and PDE [9]. Snakes is a dynamic parametric contour with an energy function, which evolves during the process of energy-minimizing, driven by external constraint forces and internal energy. The resulting contour is a closed curve with the minimum energy, localizing the nearby edge accurately.

The level set method (LSM) was firstly proposed by Osher and Sethian [10] in 1987 to track moving interface, since then it has been widely used in computer vision and image processing. For example, the LSMs have been implemented for

shape modeling [11], image restoration [12], object tracking [13], image classifications [14], image segmentation [15, 16], cell segmentation [17] and cell migration analysis [18], and etc. The LSM is an effective approach to handle curve evolution, it represents the planar closed contour as the zero level set of a level set function (LSF) in a higher dimension. For example, in Fig.1, a plane curve $C(t)$ is represented as the zero level set of LSF $\varphi(x,y,t)$, and the splitting of C from t_1 to t_2 can be regarded as numerical change of $\varphi(x,y,t)$. In this way, contour parameterization is replaced by geometric representation, and topological change of curves can happen naturally, such as splitting and merging. That's one of the important reasons that LSM becomes an effective framework to implement curve evolution in image segmentation tasks.

Assume a plane curve as $C(p,t)=\left(x(p,t),y(p,t)\right)$, where p is a parameter in $[0,1]$, and t is time in $[0,\infty)$. Then the curve evolution can be represented as:

$$\frac{\partial C(p,t)}{\partial t}=V(C)N \tag{1}$$

where $V(C)$ is a speed function, which determines the evolution speed of every point on curve C, and N is the inward normal vector to curve C. In the level set-based method, the plane closed curve can be implicitly expressed as the zero level set of a level set function φ, which is time independent. The curve can be expressed as $\varphi(C,t)=0$, whose total differential is:

$$\frac{\partial \varphi}{\partial t}+\nabla\varphi\frac{\partial C}{\partial t}=0 \tag{2}$$

ACCEPTED MANUSCRIPT

Assuming LSF is negative inside zero level set and positive outside, then the inward normal vector $N = -\nabla\varphi / |\nabla\varphi|$. Eq. (2) can then be rewritten as:

$$\frac{\partial\varphi}{\partial t} = V|\nabla\varphi| \tag{3}$$

Which is known as the level set evolution function.

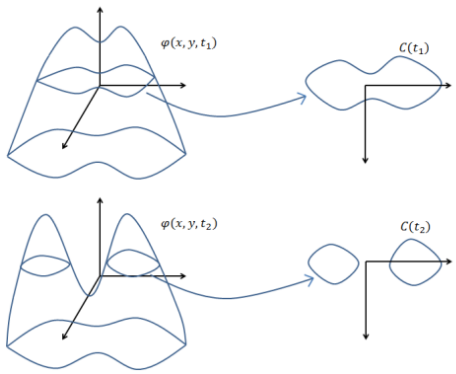


Fig.1 An example of evolution of level set function, where φ is level set function and C is the zero level set of φ.

2. Related Work and Motivation

The existing image segmentation oriented level set models can be generally categorized into two major groups: region-based models and edge-based models. In region-based level set models, the curve evolution does not rely on fitting local edge information, but rather on fitting mathematical models to image attributes like color, intensity or texture [19, 20]. And the cost function has less local minima for realistic images, thus region-based level set methods are less sensitive to noise and initialization of LSF [21]. The region-based level set models can be further divided into global fitting model and local fitting model. The former relies on fitting of global color and intensity information, and the latter depends on fitting of local intensity information [22, 23]. A classic global fitting model is

known as the piecewise smooth (PS) formulation proposed by Mumford and Shah [22], which approximates the image I by a piecewise smooth function u as the optimal solution of the minimization problem:

$$\arg\min_{u,C}(\mu Length(C)+\lambda\int_{\Omega}\left(I(x)-u(x)\right)^2 dx+\int_{\Omega\setminus C}\left|\nabla u(x)\right|^2 dx) \tag{4}$$

where C stands for edge set curve and u is allowed to be discontinuous. The first term keeps curve C to be regular, and the second term ensures u to be close to I, and the third term makes sure u is differentiable on $\Omega\setminus C$.

Inspired by Mumford-Shah model, Chan and Vese proposed an active contour model [7], and the level set formulation is expressed as follows:

$$\begin{aligned} F(c_1,c_2,\varphi) &= \mu\int_{\Omega}\nabla\left|H\left(\varphi(x,y)\right)\right|dxdy+\nu\int_{\Omega}H\left(\varphi(x,y)\right)dxdy \\ &+\lambda_1\int_{\Omega}\left|I(x,y)-c_1\right|^2 H\left(\varphi(x,y)\right)dxdy \\ &+\lambda_2\int_{\Omega}\left|I(x,y)-c_2\right|^2\left(1-H\left(\varphi(x,y)\right)\right)dxdy \end{aligned} \tag{5}$$

where the first term is a penalization of the length of contour, and the second term is a penalization of the area inside contour, c_1 is an average intensity inside contour and c_2 is an average intensity outside contour. The last two terms calculate intensity variance inside and outside the contour respectively.

A classic local fitting model is proposed by Li et al. in [23], which is well known as LBF (local binary fitting). The model uses a region-scalable fitting energy to fit the local intensity of images. The energy function is as follows:

ACCEPTED MANUSCRIPT

$$E(\varphi, f_1, f_2) = \sum_{i=1}^{2} \lambda_i \int (\int K_\sigma(x-y) \left| I(y) - f_i(x) \right|^2 M_i(\varphi(y)) dy) dx + \nu \int \left| \nabla H(\varphi(x)) \right| dx \tag{6}$$

where K_σ is a Gaussian kernel with a scale parameter σ, $f_1(x)$ and $f_2(x)$ are two functions that approximate local image intensities inside and outside the zero level set, $M_1(\varphi) = H(\varphi)$ and $M_2(\varphi) = 1 - H(\varphi)$, $H(\varphi)$ is Heaviside function. The core idea behind LBF is that it considers the region energy fitting locally.

Later on, many local fitting models have been proposed, such as LIF (local image fitting) [24], LGDF (local Gaussian distribution fitting) [25], LSD (local signed difference energy) [26], and the MRF (Markov Random Filed) embedded level set model [27]. These local fitting models' evolutions are driven by the local image intensity, and can achieve good segmentation results for those images with intensity inhomogeneity.

One of the common problems for the existing level set methods is the periodical re-initialization problem. Due to the complexity of real-world applications, singularities or sharp often occur during the evolution of LSF φ, which will result in inaccuracy of further computation. To avoid this problem, LSF φ has to be re-initialized as a signed distance function periodically, which is complicated and time consuming.

To solve that problem, Li et.al proposed an edge-based model with a regularization term [28, 29], which penalizes the deviation of LSF from signed distance function. The energy function is expressed as follows:

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

$$E(\varphi) = \lambda \int_{\Omega} g\delta(\varphi)|\nabla\varphi| dxdy + v\int_{\Omega} gH(\varphi) dxdy + \mu \int_{\Omega} \frac{1}{2}\left(|\nabla\varphi - 1|^2\right) dxdy \quad (7)$$

where g is an edge indicator, $H()$ is Heaviside function, and $\delta()$ is Dirac function. The first term is a penalization of the length of contour, the second term is a penalization of the area inside contour, and the third term is for regularization to solve the LSF's periodic re-initialization problem.

The C-V model is a typical region-based level set model, which is insensitive to noise and initialization but may not work very well for segmentation, especially for those images whose objects have similar color intensity to that of the background. Li's model is an advanced edge-based level set model, which has a regularization term to get rid of re-initialization, but may suffer to boundary leakage problem when the images' edges are weak.

The visual system of human has an attentional mechanism that can direct vision to saliency region, which is distinguished from its neighbor. In image processing or computer vision system, visual selective attention can split image into compact and separate foreground from background. In image segmentation tasks, saliency can be taken into account to yield promising results. For instance, Bai and Wang proposed the Saliency-SVM model, which takes the saliency information into account and formulates the image segmentation as a binary classification problem [30]. Qin et al. combined the region saliency with affinity propagation clustering algorithm, and used random walks method for segmentation [31].

ACCEPTED MANUSCRIPT

To develop a better LSM-based image segmentation approach, inspired by aforementioned merits and drawbacks in the existing level set models, we propose a novel Saliency Driven Region-Edge based Level set model, called SDREL. A new energy function is designed in SDREL, wherein the region-based information and gradient information are systematically taken into account. One of typical features in the proposed SDREL is that the saliency knowledge of image is effectively embedded into the model, which is sensitive to human visual system. Specifically, in the proposed SDREL model, there are two major parts. The first is the external energy, which is constituted by region-based information (saliency map and color intensity) and edge-based information (gradient map). The second is a regularization term to penalize contour length and discrepancy between level set function and signed distance function. Accordingly, we designed a new top-down two-stage evolution protocol in SDREL. In the first stage, the evolution is dominated by the external energy, and in the second stage, the evolution heavily depends on the internal energy. Our results on both synthetic and real images demonstrate that the new system is efficient, and can give more precise segmentation results.

The main contributions of the new SDREL approach are: (1) This model has designed a new energy term of incorporating the saliency map into the level set energy function, which has the significant potential to improve the level set-based image segmentations. (2) Sensitive to the initialization conditions is a common problem for many level set-based segmentation methods, in our SDREL model,

we have designed a new two-stage evolution protocol, which are dominated by the external and internal energy respectively. This new top-down evolution approach has made our model insensitive to the initializations, resulting in a more robust and fast level set evolution.

3. New SDREL Model

As discussed above, in the level set models, a contour is embedded as the zero level set of a LSF. Let φ be a LSF defined on domain Ω. We define Ω_0 ($\varphi = 0$) as a zero level set, Ω_{in} ($\varphi < 0$) as a domain inside Ω_0, and Ω_{out} ($\varphi > 0$) as a domain outside Ω_0. The energy function $E(\varphi)$ is defined by:

$$E(\varphi) = \varepsilon_{img}(\varphi, g) + \varepsilon_{reg}(\varphi, g) \tag{8}$$

where ε_{img} is an external energy, which is determined by the image attribute, and ε_{reg} is a regulation term describing the internal energy, used as a constraint of the level set evolution. And g is an edge detector, defined by:

$$g = \frac{1}{1 + \left|\nabla G_\sigma * I\right|^2} \tag{9}$$

where I is an input image, ∇ is gradient operator, and G_σ is a Gaussian kernel with standard deviation σ with a default value of 0.8, and the kernel size is set to 3. The convolution is used to reduce the influence of noise.

A. Design of external energy ε_{img}

In our model, the external energy ε_{img} includes color intensity variance and saliency map, both are determined by the input image. Gradient map, as the edge information, is incorporated into external energy as weight. The process of external energy minimization can be regarded as pixels clustering, in which pixels

ACCEPTED MANUSCRIPT

with similar color intensity and saliency value are assigned into Ω_{in} or Ω_{out}. We define the external energy as:

$$\varepsilon_{img}(\varphi, g) = \varepsilon_c(\varphi, g) + \varepsilon_s(\varphi, g) \tag{10}$$

where $\varepsilon_c(\varphi, g)$ is defined by color intensity variance and edge detector, and $\varepsilon_s(\varphi, g)$ is the saliency map with edge detector.

$$\varepsilon_c(\varphi, g) = \lambda_1 \int_{\Omega} g(I - m_1)^2 H(\varphi) dxdy + \lambda_2 \int_{\Omega} g(I - m_2)^2 (1 - H(\varphi)) dxdy \tag{11}$$

where $H()$ is a Heaviside function, φ is level set function, g is an edge detector, I is image intensity, and $m1$ and $m2$ are mean values defined by:

$$m_1 = \frac{\int_{\Omega} I(x, y) H(\varphi) dxdy}{\int_{\Omega} H(\varphi) dxdy} \tag{12}$$

$$m_2 = \frac{\int_{\Omega} I(x, y)(1 - H(\varphi)) dxdy}{\int_{\Omega} (1 - H(\varphi)) dxdy} \tag{13}$$

The two terms of $\varepsilon_c(\varphi, g)$ respectively define the color intensity variance of Ω_{in} and Ω_{out}.

The saliency term $\varepsilon_s(\varphi, g)$ in level set energy function (Eq.(10)) is defined by:

$$\varepsilon_s(\varphi, g) = \alpha_1 \int_{\Omega} g(S - s_1)^2 H(\varphi) dxdy + \alpha_2 \int_{\Omega} g(S - s_2)^2 (1 - H(\varphi)) dxdy \tag{14}$$

where S is the saliency map, g is an edge detector, $H()$ is Heaviside function.

Saliency originates from the visual distinctness and surprise [32, 33]. It is often up to image's attributes, such as color, edge and texture (see Fig.2 as an

ACCEPTED MANUSCRIPT

example). In this paper, the saliency map is formulated as [34]:

$$S(x,y)=\left|I_u - I_G(x,y)\right| \tag{15}$$

where I_u is the mean pixel value of input image, I_G is the image blurred with Gaussian filter, which is defined by:

$$I_G(x,y)=G_\tau * I(x,y) \tag{16}$$

where G_τ is a Gaussian kernel with the standard deviation τ (default value is 0.5), and the kernel size is set to 3.

In Eq.(14), s_1 and s_2 are the mean values of saliency in Ω_{in} and Ω_{out} respectively, which are defined as:

$$s_1=\frac{\int_\Omega S(x,y)H(\varphi)dxdy}{\int_\Omega H(\varphi)dxdy} \tag{17}$$

$$s_2=\frac{\int_\Omega S(x,y)\left(1-H(\varphi)\right)dxdy}{\int_\Omega \left(1-H(\varphi)\right)dxdy} \tag{18}$$

(a) (b)

Fig.2. An illustration of the Saliency map in the image. (a) The original image, and (b) the saliency map of (a) calculated by Eq.(15).

B. Design of internal energy ε_{reg}

The internal energy is to regularize level set function and prevent the

appearance of singularity and sharp shape during the evolution. In this paper, it is defined as:

$$\varepsilon_{reg} = \beta L(\varphi) + \mu P(\varphi) \tag{19}$$

where the first term $L(\varphi)$ is penalization of the length of contour, and the second term $P(\varphi)$ is defined to characterize the discrepancy between level set function and signed distance function, and we have:

$$L(\varphi) = \int_{\Omega} g\delta(\varphi)|\nabla\varphi| dxdy \tag{20}$$

$$P(\varphi) = \int_{\Omega} \frac{1}{2}\left(1 - |\nabla\varphi|\right)^2 dxdy \tag{21}$$

with the term $P(\varphi)$, during the evolution, LSF can keep close to signed distance function. And it prevents the appearance of singularities and sharp shapes.

C. Gradient flow for energy minimization

To solve the optimization problem of the energy function, the evolution of LSF should satisfy the Euler-Lagrange equation $\partial E / \partial \varphi = 0$. A standard method to find solution of Euler-Lagrange equation is the gradient flow defined by:

$$\frac{\partial \varphi}{\partial t} = -\frac{\partial E}{\partial \varphi} \tag{22}$$

where $\partial E / \partial \varphi$ is the derivative of energy E with respect to φ. It is an evolution equation of function φ, which is time-dependent with spatial variable (x, y) in the domain Ω. The derivative of ε_{img} (Eq. (10)) with respect to φ is:

$$\frac{\varepsilon_{img}}{\partial \varphi} = \lambda_1 g\left(I - m_1\right)^2 - \lambda_2 g\left(I - m_2\right)^2 + \alpha_1 g\left(S - s_1\right)^2 - \alpha_2 g\left(S - s_2\right)^2 \tag{23}$$

ACCEPTED MANUSCRIPT

The derivative of ε_{reg} (Eq.(19)) with respect to φ is:

$$\frac{\partial \varepsilon_{reg}}{\partial \varphi} = -\beta\delta(\varphi) div\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right) - \mu\Delta\varphi \tag{24}$$

where $div()$ is the divergence operator and Δ is a Laplacian operator. Therefore, with Eqs.(23) and (24), the evolution of φ (Eq.(8)) with respect to t is:

$$\begin{aligned}\frac{\partial\varphi}{\partial t} = -\frac{\partial E}{\partial\varphi} = &-\lambda_1 g\left(I-m_1\right)^2 + \lambda_2 g\left(I-m_2\right)^2 - \alpha_1 g\left(S-s_1\right)^2 \\ &+ \alpha_2 g\left(S-s_2\right)^2 + \beta\delta(\varphi) div\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right) + \mu\Delta\varphi\end{aligned} \tag{25}$$

The direction of LSF evolution is in the opposite direction of the derivative of energy function E, i.e., $-\partial E/\partial\varphi$, and it is the steep descent direction of $E(\varphi)$. Thus, the method to find level set evolution is also called steepest decent method.

D. Convergence analysis

The convergence proof of Eq.(25) is as follows. Firstly, we have three Lemmas and an inference [35].

Lemma 1: If f is a continuous function on R^n, $d \neq 0$, then the algorithm mapping M is closed at (x,d).

$$M(x,y) = \left\{ y | y = x + \bar{\lambda} d, \bar{\lambda} \text{ satisfies } f\left(x+\bar{\lambda}d\right) = \min_{0\le\lambda<\infty} f\left(x+\lambda d\right)\right\} \tag{26}$$

Lemma 2: Assume $B: X \to Y$ and $C: Y \to Z$ are mappings from point to set, and B is closed at point x, C is closed at $B(x)$. Assume that $x^{(k)} \to x$ and $y^{(k)} \in B\left(x^{(k)}\right)$, then a convergent subsequence $\left\{y^{(k_j)}\right\} \subset \left\{y^{(k)}\right\}$ exists. Then the synthetic mapping $A = CB$ is closed at x.

ACCEPTED MANUSCRIPT

From Lemma 2, we have the following inference:

Inference 1: Assume $B: X \to Y$ is a mapping from point to point, and $C: Y \to Z$ is a mapping from point to set. If B is continuous at x and C is closed at $B(x)$, then the synthetic mapping $A = CB$ is closed at x.

Lemma 3: Assume A is an algorithm defined on X, Ω is the solution set. With a given start point $x^{(1)}$, do the following iteration:

(1) If $x^{(k)} \in \Omega$, stop the iteration, else $x^{(k+1)} \in \mathrm{A}\left(x^{(k)}\right)$.

(2) Use $k+1$ instead k and repeat (1).

Through the iteration, we can get a sequence $\left\{x^{(k)}\right\}$. If these conditions are satisfied: $\left\{x^{(k)}\right\}$ is contained in a compact set of X, there is a continuous descent function of Ω and A, the mapping A is closed in the complementary set of Ω. Then any convergent subsequence of $\left\{x^{(k)}\right\}$ has the limit belonging to Ω.

Theory 1: Assume $f(x)$ is a continuous differentiable function and the solution set $\Omega = \left\{\bar{x} \mid \nabla f(\bar{x}) = 0\right\}$. If the sequence $\left\{x^{(k)}\right\}$ generated by steepest decent method is contained in a compact set, then $\left\{x^{(k)}\right\}$ is convergent.

Proof: please see appendix for detailed mathematical proof.

E. Extension to color image

SDREL can be easily extended to segment the color images. The energy function is still the same as Eq.(8), except some changes in external energy ε_{img}.

To extend energy function of Eq.(10) for features of color, we rewrite Eq.(14) as:

$$S(x, y) = \left\| I_u - I_G(x, \mathrm{y}) \right\| \tag{27}$$

ACCEPTED MANUSCRIPT

And the external energy ε_{img} can be extended to:

$$\begin{aligned}\varepsilon_{img} &= \lambda_1 \int_{\Omega} g \left\| I - m_1 \right\|^2 H(\varphi)\, dxdy + \lambda_2 \int_{\Omega} g \left\| I - m_2 \right\|^2 (1 - H(\varphi))\, dxdy \\ &+ \alpha_1 \int_{\Omega} g \left\| S - s_1 \right\|^2 H(\varphi)\, dxdy + \alpha_2 \int_{\Omega} g \left\| S - s_2 \right\|^2 (1 - H(\varphi))\, dxdy\end{aligned} \tag{28}$$

Here, the range of I is *d*-dimensional, and m_1, m_2, s_1 and s_2 are vectors of size *d*, and $\|\bullet\|$ is L2 norm. By comparison with the original energy function of Eq.(10), the values of $(I - m_1)$, $(I - m_2)$, $(S - s_1)$ and $(S - s_2)$ are replaced with the Euclidean vector norms. The energy minimization is the same as the original process.

4. Implementation of SDREL

A. Numerical scheme

In practice, the Dirac function $\delta(x)$ is smoothed as $\delta_\varepsilon(x)$, which is defined by:

$$\delta_\varepsilon(x) = \begin{cases} 0, & |x| > \varepsilon \\ \dfrac{1}{2\varepsilon}\left[1 + \cos\left(\dfrac{\pi x}{\varepsilon}\right)\right], & |x| \le \varepsilon \end{cases} \tag{29}$$

In application, we use $\delta_\varepsilon(x)$ as Dirac function with $\varepsilon = 1.5$.

B. Initialization of LSF

In this proposed SDREL model, the region-based information, consisting of color intensity variance and saliency map, is a global feature determined by the input image. Therefore, the initialization of LSF can be very flexible. Different from conventional initialization of LSF by drawing box inside or outside the object, we use a checkerboard shape to initialize the level set function. The definition of checkerboard is as follows:

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

$$\varphi_0(x,y) = c_0 * sin\left(\frac{\pi * R * x}{1500}\right)\sin\left(\frac{\pi * C * y}{1500}\right) \tag{30}$$

where R is the height of input image, and C is the width of input image, c_0 is a positive number defining the amplitude of initial LSF, and in our application, we set $c_0 = 1$. With this checkerboard shape initialization, Ω_{in} and Ω_{out} of Eq.(11) and Eq.(14) can be evenly distributed into different positions of an image, and when the image contains multiple objects, this can also make sure most objects are contained or partly contained in Ω_{in}, so that the region-based information of objects can be used more sufficiently and the evolution of LSF can have faster convergence.

C. Top-down two-stage evolution protocol

In our level set model, LSF is driven by input image's attribute (external energy) and its own attribute (internal energy). These two types of energy play different roles in the evolution of LSF, wherein external energy helps extract object from background and LSF is smoothed by internal energy. With the initialization of Eq.(30), LSF is assigned as a checkerboard at the very beginning. So external energy plays the key role in the beginning of evolution to extract object and drive contour toward object's edge. After the extraction of objects, changes of LSF become flat and the contour positioning needs further precision, when internal energy plays a key role.

To make sufficient use of external energy and internal energy in different stages of evolution, we implement the evolution of LSF as a two-stage process, where external energy is a dominant factor of the first stage while internal energy

plays a dominant role in the second stage. The iteration step of LSF is used as the criterion to judge the current stage:

$$\left|\frac{\partial \varphi}{\partial t}\right| < T \tag{31}$$

where T is the input threshold.

In the first stage, the evolution mostly depends on external energy ε_{img}. In this stage of evolution, pixels are clustered into Ω_{in} or Ω_{out} so that objects are extracted from background. During the stage one, the edge detector g is combined with region-based information to position boundary. However, after the stage one, some object's edges are not accurately located during the evolution, and some singularities and irregular shape of LSF may occur. Therefore, in the second stage, evolution of LSF mainly depends on the internal energy, which consists of penalization of contour length $L(\varphi)$ and regulation term $P(\varphi)$, which plays a key role to smooth the level set function, eliminate singularities, and enhance the accuracy of edge location. This top-down two-stage evolution can be summarized as the following Fig.3:

Step 1) Input image, and initialization parameters, including weight λ_1, λ_2, α_1, α_2, β, μ, Δt and threshold T_1 and T_2.

Step 2) Initialize LSF as $\varphi_0(x,y) = c_0 * \sin\left(\frac{\pi * R * x}{1500}\right)\sin\left(\frac{\pi * C * y}{1500}\right)$, let $c0 = 1$.

Step 3) Calculate edge detector $g = \frac{1}{1+\left|\nabla G_\sigma * I\right|^2}$.

Step 4) Calculate saliency map, get $S(x,y) = \left|I_u - I_G(x,y)\right|$.

Step 5) Calculate mean color intensity value $m1$, $m2$ and mean saliency values s_1 and s_2.

Step 6) Start stage-one level set evolution, the evolution equation is:

$$\frac{\partial \varphi}{\partial t} = -\lambda_1 g\left(I - m_1\right)^2 + \lambda_2 g\left(I - m_2\right)^2 - \alpha_1 g\left(S - s_1\right)^2$$

$$+\alpha_2 g\left(S - s_2\right)^2 + \beta\delta\left(\varphi\right) div\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right) + \mu\Delta\varphi.$$

If $\left|\frac{\partial \varphi}{\partial t}\right| < \mathrm{T}_1$, let $\lambda_1 = \lambda_2 = \alpha_1 = \alpha_2 = 0$, and jump to Step 7).

Else $\varphi_{n+1} = \varphi_n + \frac{\partial \varphi}{\partial t}\Delta t$, jump to Step 5).

Step 7) Start stage-two level set evolution, the evolution equation is:

$$\frac{\partial \varphi}{\partial t} = \beta\delta\left(\varphi\right) div\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right) + \mu\Delta\varphi.$$

If $\left|\frac{\partial \varphi}{\partial t}\right| < \mathrm{T}_2$, stop evolution.

Fig.3 The flowchart of the proposed SDREL approach.

As shown by an example of the energy graph in Fig.4, the evolution can be intuitively divided into two parts: in the first five iterations, energy sharply deceases. In the last fifteen iterations, energy goes flat and convergent. As shown in this figure, after the first stage of evolution, the object is roughly extracted from background, but there are still some singularities, and some background regions are contained inside the contour. In the second stage, with the penalization of length $L(\varphi)$ and regulation term $P(\varphi)$, level set function is smoothed and the segmentation result becomes more precise.

ACCEPTED MANUSCRIPT

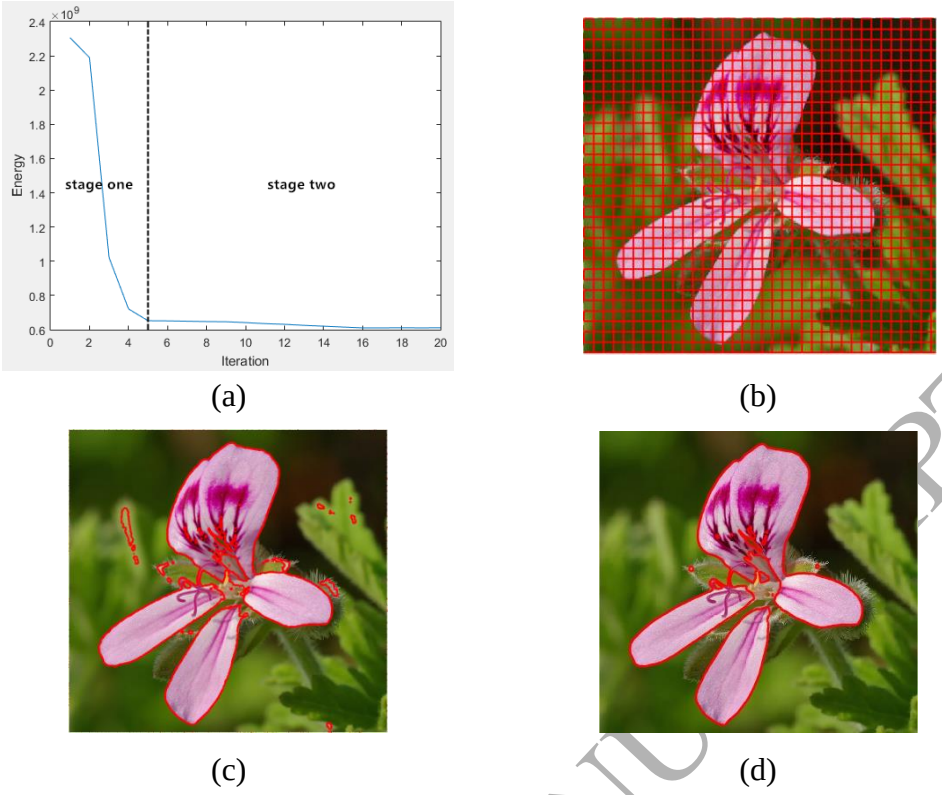


Fig.4 An example showing the two-stage implementation of SDREL. (a) Energy graph of LSF. (b) The original image with initial checkerboard contour. (c) Evolution result after the 1st stage. (d) The final result of contour after the 2nd stage evolution.

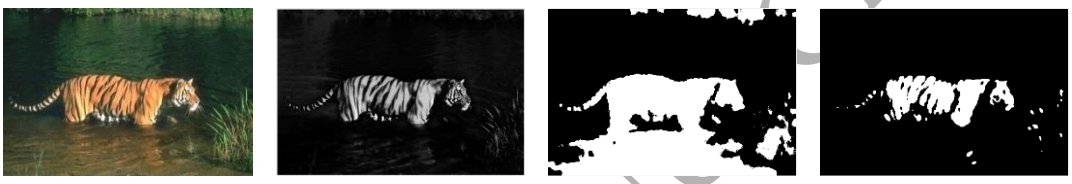
5. Experimental Results

In this section, we test SDREL on both synthetic and real images. There are initial parameters in our model and time step Δt for implementation. In all the experimental tests, the parameters are fixed as: $\lambda 1 = 0.01$, $\lambda 2 = 0.01$, $\alpha 1 = 0.06$, $\alpha 2 = 0.06$, $\beta = 5$, $\mu = 0.2$, $\Delta t = 1$, $T_1 = N * 0.005$, and $T_2 = N * 0.001$, where N is the number of pixels of the input image.

A. Effects of the newly introduced saliency knowledge in SDREL

Different from other level set models, saliency map is newly incorporated into our model, which is expected to help distinguish object from background.

With this saliency information, segmentation result is expected to be more reliable. Fig.5 gives an example comparing the energy function with/without the saliency term. We can see that the segmentation result generated by the level set model without saliency map is not very precise (Fig.5c), where some water regions that have similar color to tiger are recognized as part of the foreground. Fig.5d is the segmentation result with the saliency term, where the fallout ratio is much lower when comparing to Fig.5c.



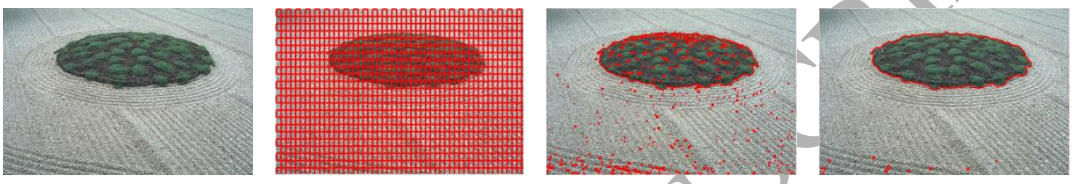
(a) (b) (c) (d)

Fig.5 Segmentation results with and without saliency map. (a) The original image, (b) the saliency map of (a), (c) the segmentation result generated from level set model without saliency term, and (d) segmentation result generated from level set model with saliency map.

B. Effects of the two-stage asynchronous energy evolution

Considering that external energy and internal energy play different roles in the evolution of LSF, we implement a top-down two-stage evolution protocol. The protocol separates the evolution into two parts, where external energy plays a key role to extract object from the background in the first stage and LSF is further smoothed in the second stage by internal energy. Fig.6 compares the effects of the two-stage evolution approach with the evolution at a single stage, where the external and internal energies are equally important in all evolution scales

(denoted as synchronous evolution). Fig.6c and Fig.6d are the segmented results from the synchronous and asynchronous (two-stage) evolutions, where we can intuitively see that through the asynchronous process, final segmentation results are more reliable and precise. Our results show an interesting phenomena that a more accurate image segmentation is from rough to precise, where the energy term focuses are different.



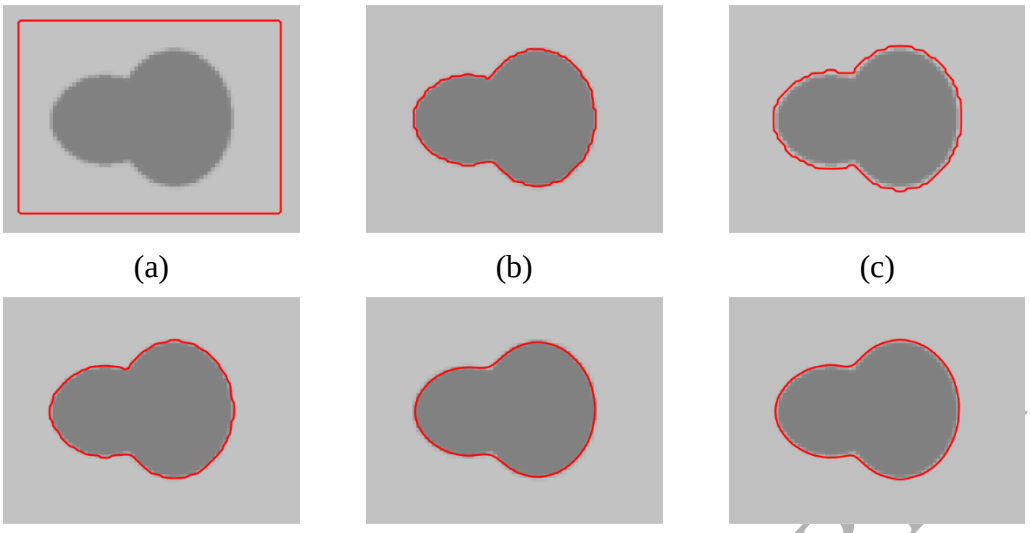
(a) (b) (c) (d)

Fig.6 Segmentation results of synchronous and two-stage implementation. (a) The original image, (b) original image with initial contour, (c) the segmentation result generated from synchronous implementation, and (d) segmentation result generated from two-stage asynchronous implementation.

C. Results on synthetic images

To prove the efficiency of SDREL, we compare it with C-V model [7], Li's model [28], LBF [23] and Zhang's model [16]. Results in Fig.7 demonstrate that region-based level set model, no matter global fitting model like C-V model or local fitting model like LBF and Zhang's model, cannot generate results with precise boundary. For those models considering edge information like Li's model and SDREL, they can get more precise segmentation results.

ACCEPTED MANUSCRIPT

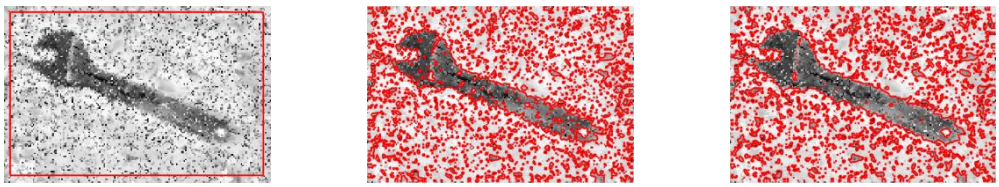


(a) (b) (c)

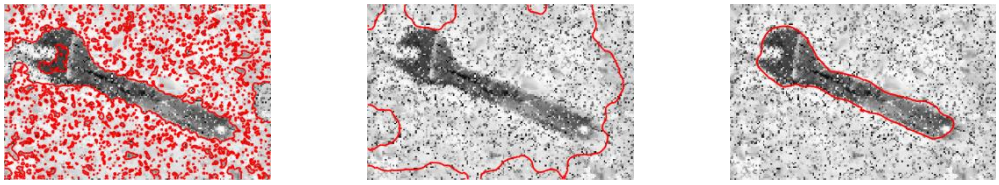
(d) (e) (f)

Fig.7 Segmentation results of a synthetic image. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

To test the robustness of the proposed method in case of dealing with noise, we use a synthetic image with Salt-and-Pepper noise (as shown in Fig.8a). Region-based models are not capable of eliminating noise of the image, the segmentation results (Fig.8b-d) contain lots of noise. And Li's model is stuck in local minima because of the noise. Evolution of our level set SDREL is divided into two stages. In the first stage, the object is extracted from background. In the second stage, level set function is smoothed by regulation term. During the process of stage-two evolution, most noise is eliminated, resulting in a better contour detection and image segmentation results (Fig.8f).



(a) (b) (c)

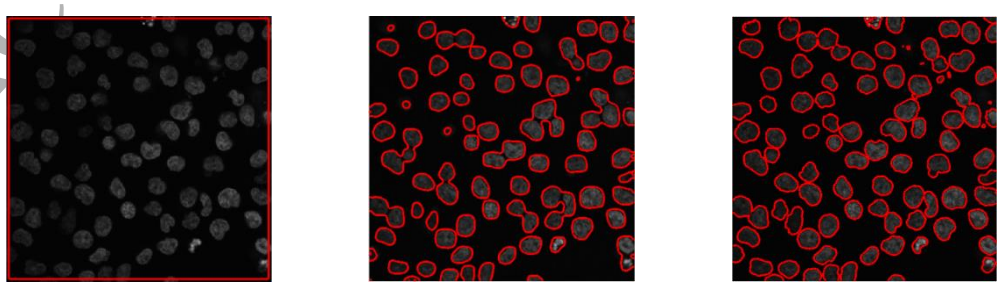


(d) (e) (f)

Fig.8 Segmentation results of a synthetic image. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

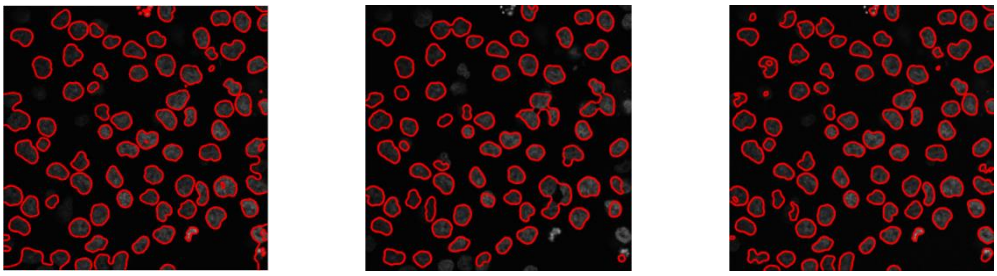
D. Results on real images

SDREL has also been tested on real images. For example, Fig.9a is a microscope gray scale image of cells with initial contour, and the goal of segmentation is to extract all the cells from the dark background. Because the intensity of some cells are similar to the background and the edge of these cells are weak, boundary leakage occurs during the evolution of Li's model, and some cells are undetected (Fig.9e). When using the region-based model for cell detection, Fig.9b-d show that most cells are detected but the edge position is not precise for some targets, where some separated cells are detected in adhesion. In the evolution of our SDREL model, there is less adhesion of cells, as shown in Fig.9f.



(a) (b) (c)

ACCEPTED MANUSCRIPT

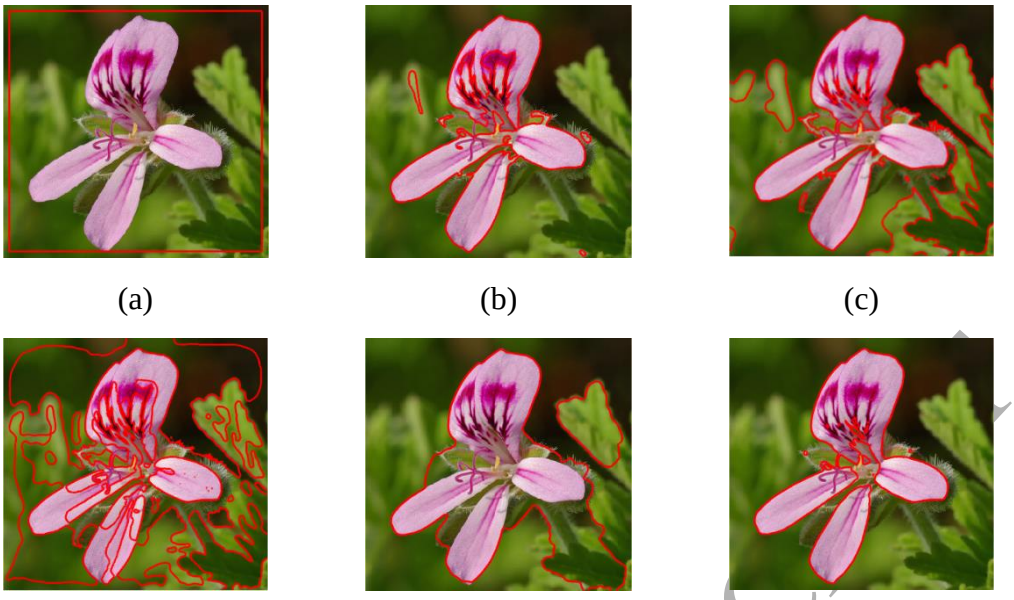


(d) (e) (f)

Fig.9 Segmentation results of a cell image. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

Different from the cell images, natural images have much more information, so they are more difficult to be analyzed and segmented. For region-based models, C-V model mainly consider the global color information; local fitting model like LBF and Zhang's model consider the color intensity in local regions, and the segmentation results of these three models could contain some background area, when the color information of background and foreground are similar. Li's model depends on the external force which drives it to shrink and fit to the edge information. When the external force is too weak, the evolution may stuck on local minima, like Fig.10e, and when the external force is too strong, the boundary leakage problem may occur, like Fig.11e. In practice, it is a tradeoff between local minima and boundary leakage. The results of our SDREL model are more reliable, because saliency map helps distinguish object from the background and the two-stage implementation makes the edge positioning more precise.

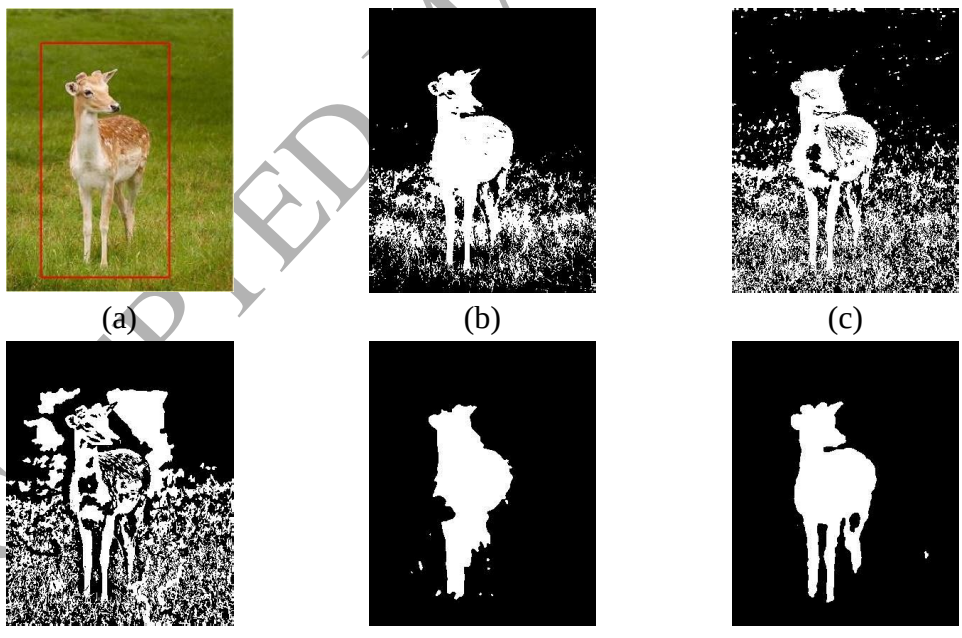
ACCEPTED MANUSCRIPT



(a) (b) (c)

(d) (e) (f)

Fig.10 Segmentation results of a natural image. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

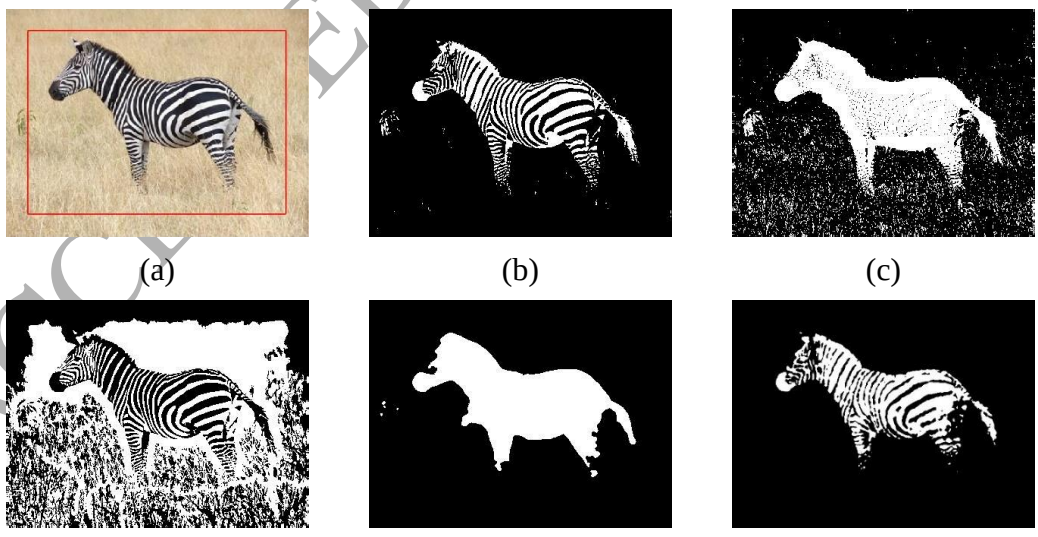


(a) (b) (c)

(d) (e) (f)

Fig.11 Segmentation results of a natural image. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

Figure 12 is an image with textured structure. From Fig.12f, we can see that the segmentation result has space to be further improved from the background, where the stripe structure cannot be completely detected in current model. Region information used in SDREL, including color intensity and saliency map, are both global information, and the local correlations of pixels from textured structure are not considered into our energy function, which can be a future improvement direction. For images with clutter background, it is more challenging to realize an accurate segmentation. For example, in Fig.13, the segmentation results of tested five level set models have space to improve. This is because the background is complex and has very similar color information to the object. To improve the segmentation on image with clutter background, the spatial correlations of pixels need to be considered during the evolution of level set model, and this is still a challenging topic in image segmentation task.



(a) (b) (c)

(d) (e) (f)

Fig.12 Segmentation results of an image with textured structure. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d)

result of LBF, (e) result of Li's model, and (f) result of SDREL.

Fig.13 Segmentation results of an image with clutter background. (a) Original image with initial contour, (b) result of C-V model, (c) result of Zhang's model, (d) result of LBF, (e) result of Li's model, and (f) result of SDREL.

E. Results on large-scale image set

To further test the effect of SDREL on image segmentation, we did ex periments based on MSRA 1000 dataset, which is derived from a public im age dataset [36] by Achanta et.al [34]. MSRA 1000 contains 1000 color im ages and ground truth, which is available at http://ivrlwww.epfl.ch/suppleme ntary_material/RK_CVPR09/index.html.

We compared our SDREL with Li's model, C-V model, LBF, Zhang's model and a state-of-the-art salient region detection method RC (region-based contrast) [32]. We use five metrics to objectively evaluate SDREL, the rand index (RI) [37], the Global Consistency Error (GCE) [38], the Variation of Information (VOI), the F-measure based on precision and recall and the F-measure based on sensitivity

and specificity (SF-Measure).

Table 1. Statistical results of evaluation on MSRA1000 dataset.

	RI	GCE	VOI	F-Measure	SF-Measure
C-V [7][a]	0.693	0.179	1.058	0.567	0.744
Li's [28] [a]	0.707	0.161	0.918	0.615	0.742
LBF [23] [a]	0.535	0.277	1.568	0.361	0.586
Zhang's [16][a]	0.736	0.175	1.006	0.613	0.79
SDREL [a]	0.781	0.139	0.822	0.682	0.818
RC [32][b]	0.857	0.086	0.559	0.733	0.769

[a] Level-set-based image segmentation methods.
[b] Salient region detection method.

Table 1 shows the statistical results of the objective evaluation of SDREL and the other five methods. And it can be seen that among all the tested level set-based models (C-V model, Li' model, LBF, and Zhang's model), SDREL has the highest RI, F-Measure, SF-Measure and lowest GCE, VOI. RC is a salient region detection method, which has the best RI, F-Measure, GCE and VOI, but our SDREL has a higher SF-Measure. Table 2 illustrates the average executive times of the 5 level-set methods, and the results show that SDREL performs faster than the edge-based Li's model and local fitting region-based models (LBF, Zhang's model). SDREL is a little bit slower than C-V model, it is because SDREL needs to generate the saliency map.

Table 2. Average executive times of level-set methods.

	C-V	Li's	LBF	Zhang's	SDREL
Time (s)	1.273	9.835	4.295	16.643	1.658

F. Comparison with CNN-based segmentation method

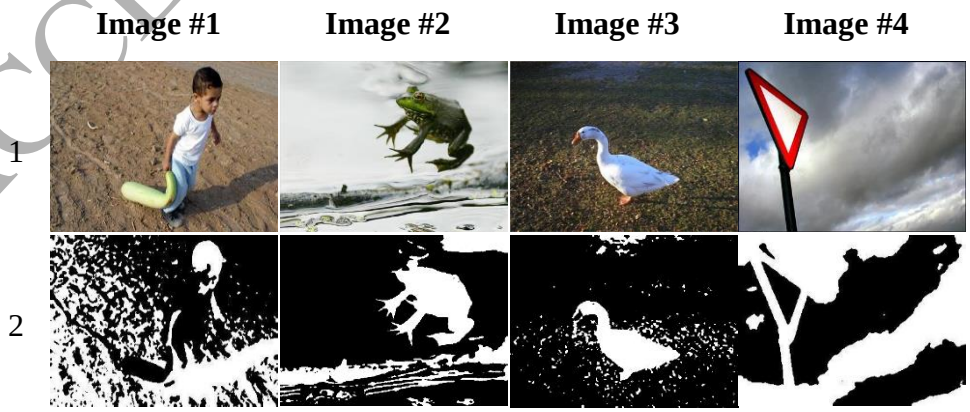
Convolutional networks (ConvNets) have achieved great success in image processing [39]. U-net [40] is a state-of-the-art convolution network for image

ACCEPTED MANUSCRIPT

segmentation, and we compare it with our SDREL on the MSRA1000 dataset. We randomly chose 800 images to train U-net and used the left 200 images to test the performance. Table 3 shows the results of U-net and SDREL on 200 test images from MSRA1000. Our results show that the results of SDREL have higher F-Measure, SF-Measure, RI and lower GCE. The potential reason is that current dataset size is small, 800 training images are not enough for training a good ConvNet. On this relatively small dataset, ConvNets may get overfitting. The other difference between the U-net and SDREL is that the former is a supervised model, while the latter is an unsupervised approach, which does not need the training samples. Some segmentation examples of different methods are shown in Fig.14.

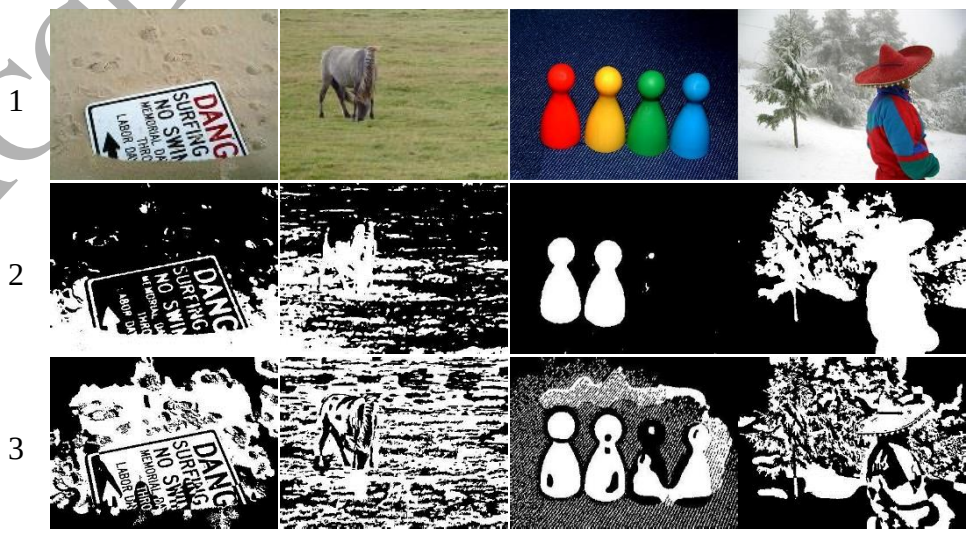
Table 3. Evaluation of segmentation results of U-net and SDREL on 200 test images from MSRA1000.

	RI	GCE	VOI	F-Measure	SF-Measure
U-net	0.712	0.159	0.773	0.643	0.699
SDREL	0.773	0.144	0.833	0.678	0.803



ACCEPTED MANUSCRIPT

Image #5 **Image #6** **Image #7** **Image #8**



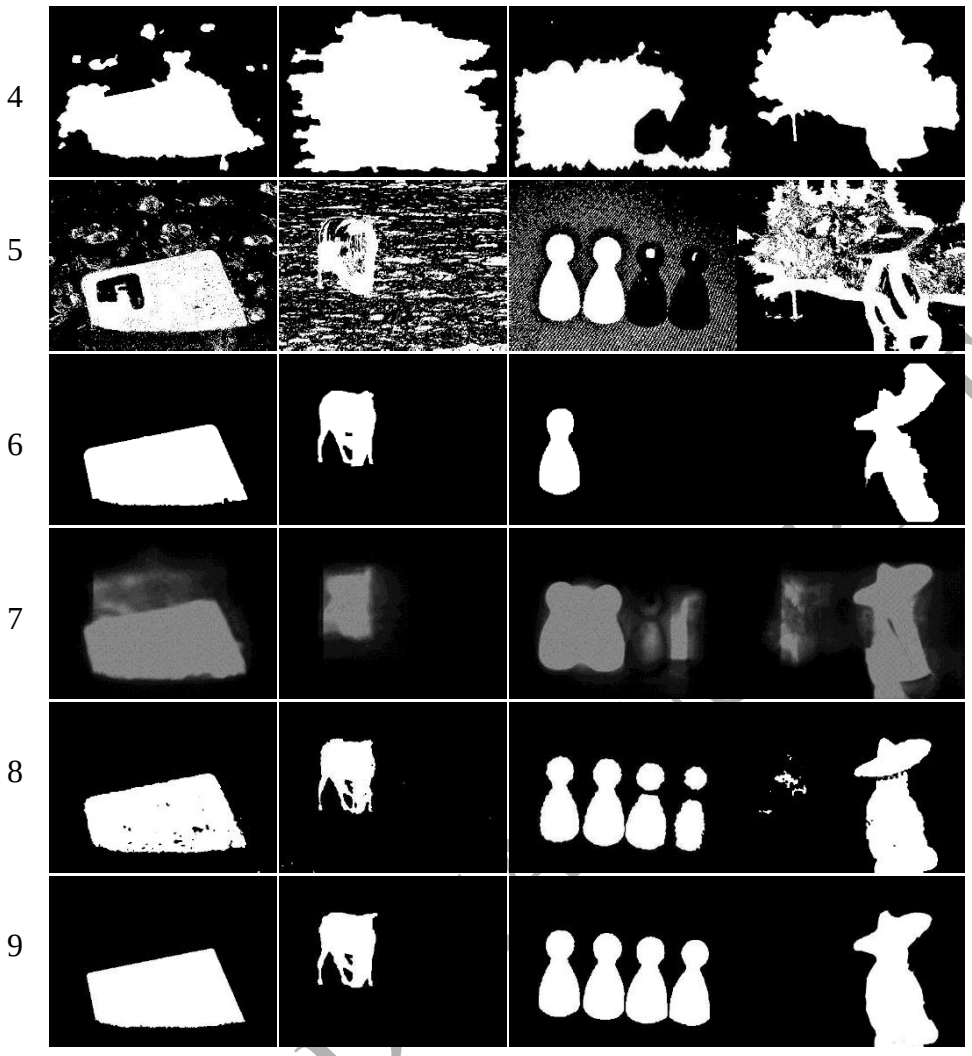


Fig.14 Comparison of segmentation results of 8 images from MSRA1000. Line number 1: test images. Line number 2: results of C-V model. Line number 3: results of LBF. Line number 4: results of Li's model. Line number 5: results of Zhang's model. Line number 6: results of RC. Line number 7: results of U-net. Line number 8: results of SDREL. Line number 9: ground truth.

6. Discussion and Conclusion

The level set models can be roughly divided into two groups, region-based models and edge-based models. Region-based level set models can be further divided into global fitting model and local fitting model. The global fitting model relies on fitting to global color and intensity information, and the local fitting

ACCEPTED MANUSCRIPT

model depends on fitting to local intensity information. In a region-based level set model, certain region descriptor drives the evolution of level set function. And color variance is often used to constitute region descriptor, which is difficult to describe images whose objects have similar intensity to that of the background. Therefore, segmentation results generated by region-based models may contain some region of background. Li's model is a typical edge-based level set model, which relies on edge information for image segmentation, and can usually get precise boundary positioning. However, this type of level set methods is sensitive to initial conditions and often suffers to boundary leakage problems when applied to images with weak object boundaries. In Li's model, a level set regularization term defined by the distance between signed function is introduced in energy function. With this regularization term, during the evolution of level set function, sharp shape and singularities can be avoided, so the re-initialization is no longer needed, which can improve the evolution efficiency.

SDREL considers the saliency map into account and combines region information with edge information into the level set model, which has advantages such as initialization condition flexibility, fast evolution and the ability to extract object from background with precise boundary positioning. Area of the region inside zero level set is a penalization in both C-V model and Li's model, which constraints the direction of evolution to expansion or shrink before the evolution starts. But in our model, this area term is no longer considered into our energy function, which makes the evolution of LSF more flexible. Different from

ACCEPTED MANUSCRIPT

existing level set-based image segmentation models, our SDREL model considers saliency map information into energy function. Saliency map originates from visual distinctness and surprise, which can help extract object from the background. Through experiments, we have found that saliency map and the designed two-stage energy evolution protocol achieve more precise segmentation results. By comparing the segmentation results of SDREL with that of other level set methods on MSRA1000, we show that our SDREL is an efficient approach. In the experiments of SDREL, we have fixed its parameters on all the tested images. Our local tests show that if we further optimize the parameters for each single image, we could get better results.

The region information used in current SDREL model includes color intensity and saliency map, which are both global information, and the pixels are assumed to be spatial independent in the segmentation of image. Local region information, like texture, is not considered into our model. This is the reason that for those images with clutter background or textured images (refer to Fig.12 and Fig.13 as examples), which are characterized by local correlations of pixels, SDREL's performance still has space for further improvements. In our future work, we will consider local correlations of pixels into our model to enhance the segmentation results for those textured images or images with clutter background. In addition, we will also try to find an automatic dynamic way to calculate these weights during the process of evolution according to the real-time segmentation result on specific images.

Appendix A

Proof of Theory 1.

In our model (refer to Eq.(8)), assume the solution set:

$$\Omega = \left\{ \bar{\varphi} | \nabla E\left(\bar{\varphi}\right) = 0 \right\} \tag{A1}$$

The gradient algorithm of our method can be regarded as the synthetic mapping:

$$A = MD \tag{A2}$$

where

$$D\left(\varphi\right) = \left(\varphi, -\nabla E\left(\varphi\right)\right) \tag{A3}$$

is a mapping from R^n to $R^n \times R^n$ and M is a mapping from $R^n \times R^n$ to R^n. With a given point φ, through mapping D, we can get the point φ^k and its negative gradient:

$$\begin{aligned} -\nabla E\left(\varphi^k\right) = & -\lambda_1 g\left(I - m_1\right)^2 + \lambda_2 g\left(I - m_2\right)^2 - \alpha_1 g\left(S - s_1\right)^2 \\ & + \alpha_2 g\left(S - s_2\right)^2 + \beta\delta\left(\varphi^k\right) div\left(g\frac{\nabla\varphi^k}{\left|\nabla\varphi^k\right|}\right) + \mu\Delta\varphi^k \end{aligned} \tag{A4}$$

With a given point φ^k and a direction $d = -\nabla E\left(\varphi^k\right)$, through mapping M, we can get a new point:

$$\varphi^{k+1} = \varphi^k + d\Delta t \tag{A5}$$

where the energy $E\left(\varphi^{k+1}\right)$ is smaller comparing to the former iteration point φ^k.

According to Lemma 1, M is a closed mapping when $\nabla E\left(\varphi\right) \neq 0$.Because the energy function $E\left(\varphi\right)$ is continuous and differentiate, the mapping D is continuous.

ACCEPTED MANUSCRIPT

According to inference 1, the mapping A is closed at $\varphi\left(\nabla E(\varphi)\neq 0\right)$. When $\varphi\notin\Omega$, we can get $d=-\nabla E(\varphi)\neq 0$ and $\nabla E(\varphi)^T d<0$, so $E(\varphi)$ is a decent function of A and φ. In the implementation of our algorithm, the evolution of level set function is within limits:

$$\varphi(x,y)=\begin{cases}\varphi(x,y) & if\ \ |\varphi(x,y)|<L\\ -L & if\ \ \varphi(x,y)<-L\\ L & if\ \ \varphi(x,y)>L\end{cases} \tag{A6}$$

where L is a positive number. So we can define a compact set:

$$\Gamma:\left\{\varphi(\varphi_1,\ldots,\varphi_N):\varphi\in R^N,|\varphi_i|\le L\forall i\in[1,N]\right\} \tag{A7}$$

where N is the number of pixels of input image. Obviously, sequence $\left\{\varphi^{(k)}\right\}$ is contained in Γ. According to Lemma 3, the algorithm is convergent.

QED. □

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61671288, 91530321, 61603161, 61725302), and Science and Technology Commission of Shanghai Municipality (No. 16JC1404300, 17JC1403500).

References

[1] H. Cai, Z. Yang, X. Cao, W. Xia, X. Xu, A New Iterative Triclass Thresholding Technique in Image Segmentation, IEEE Trans. Image Process. 23 (2014) 1038.
[2] P. Ghamisi, M.S. Couceiro, F.M.L. Martins, J. Atli Benediktsson, Multilevel Image Segmentation Based on Fractional-Order Darwinian Particle Swarm Optimization, IEEE Trans. Geosci. Remote Sens. 52 (2014) 2382–2394.
[3] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man. Cybern. 9 (1979) 62–66.
[4] D. Comaniciu, P. Meer, Mean Shift: A Robust Approach Toward Feature Space

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 603–619.
[5] S. Kim, S. Nowozin, P. Kohli, D.Y. Chang, Higher-Order Correlation Clustering for Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2015) 1761–1774.
[6] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Trans. Pattern Anal. Mach. Intell. 13 (1991) 583–598.
[7] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. A Publ. IEEE Signal Process. Soc. 10 (2001) 266–277.
[8] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, Int. J. Comput. Vis. 1 (1988) 321–331.
[9] Y. Wu, C. He, A convex variational level set model for image segmentation, Signal Processing. 106 (2015) 123–133.
[10] S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, J. Comput. Phys. 79 (1988) 12–49.
[11] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape Modeling with Front Propagation: A Level Set Approach, IEEE Trans. Pattern Anal. Mach. Intell. 17 (1995) 158–175.
[12] Z. Huan, L. Kong, X. Zuo, A feature-oriented forward-backward diffusion model for intensity image restoration based on level set motion, Int. J. Comput. Math. 86 (2009) 2072–2094.
[13] X. Sun, H. Yao, S. Zhang, A novel supervised level set method for non-rigid object tracking, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2011: pp. 3393–3400.
[14] C. Samson, L. Blanc-Féraud, G. Aubert, J. Zerubia, A level set model for image classification, Int. J. Comput. Vis. 40 (2000) 187–197.
[15] C. Li, R. Huang, Z. Ding, J.C. Gatenby, D.N. Metaxas, J.C. Gore, A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI., IEEE Trans. Image Process. A Publ. IEEE Signal Process. Soc. 20 (2011) 2007–2016.
[16] K. Zhang, L. Zhang, K.M. Lam, D. Zhang, A Level Set Approach to Image Segmentation With Intensity Inhomogeneity, IEEE Trans. Cybern. 46 (2016) 546–557.
[17] A. Gharipour, A.W.-C. Liew, Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting, Pattern Recognit. 58 (2016) 1–11.
[18] O. Dzyubachyk, W.A. van Cappellen, J. Essers, W.J. Niessen, E. Meijering, Advanced Level-Set-Based Cell Tracking in Time-Lapse Fluorescence Microscopy, Med. Imaging IEEE Trans. 29 (2010) 852–867.
[19] D. Cremers, M. Rousson, R. Deriche, A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape, Int. J. Comput. Vis. 72 (2010) 195–215.
[20] H. Min, W. Jia, X.-F. Wang, Y. Zhao, R.-X. Hu, Y.-T. Luo, F. Xue, J.-T. Lu, An intensity-texture model based level set method for image segmentation, Pattern Recognit. 48 (2015) 1547–1562.

ACCEPTED MANUSCRIPT

[21] X.-F. Wang, H. Min, L. Zou, Y.-G. Zhang, A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement, Pattern Recognit. 48 (2015) 189–204.
[22] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems †, Commun. Pure Appl. Math. 42 (1989) 577–685.
[23] C. Li, C.-Y. Kao, J.C. Gore, Z. Ding, Minimization of region-scalable fitting energy for image segmentation, IEEE Trans. Image Process. 17 (2008) 1940–1949.
[24] K. Zhang, H. Song, L. Zhang, Active contours driven by local image fitting energy, Pattern Recognit. 43 (2010) 1199–1206.
[25] L. Wang, J. Macione, Q. Sun, D. Xia, C. Li, Level set segmentation based on local gaussian distribution fitting, in: Asian Conf. Comput. Vis., Springer, 2009: pp. 293–302.
[26] L. Wang, H. Wu, C. Pan, Region-based image segmentation with local signed difference energy, Pattern Recognit. Lett. 34 (2013) 637–645.
[27] X. Yang, X. Gao, D. Tao, X. Li, J. Li, An efficient MRF embedded level set method for image segmentation, IEEE Trans. Image Process. 24 (2015) 9–21.
[28] C. Li, C. Xu, C. Gui, M.D. Fox, Distance Regularized Level Set Evolution and Its Application to Image Segmentation, IEEE Trans. Image Process. 19 (2010) 3243–3254.
[29] C. Li, C. Xu, C. Gui, M.D. Fox, Level Set Evolution without Re-Initialization: A New Variational Formulation, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2005: pp. 430-436.
[30] X. Bai, W. Wang, Saliency-SVM: an automatic approach for image segmentation, Neurocomputing. 136 (2014) 243–255.
[31] C. Qin, G. Zhang, Y. Zhou, W. Tao, Z. Cao, Integration of the saliency-based seed extraction and random walks for image segmentation, Neurocomputing. 129 (2014) 378–391.
[32] M.-M. Cheng, N.J. Mitra, X. Huang, P.H.S. Torr, S.-M. Hu, Global contrast based salient region detection, IEEE Trans. Pattern Anal. Mach. Intell. 37 (2015) 569–582.
[33] C. Gong, D. Tao, W. Liu, S.J. Maybank, M. Fang, K. Fu, J. Yang, Saliency propagation from simple to difficult, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015: pp. 2531–2539.
[34] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2009: pp. 1597–1604.
[35] B. Chen, Optimization theory and algorithm, second ed., Tsinghua University Press, Beijing, 1989.
[36] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, H.Y. Shum, Learning to detect a salient object, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2010) 353–367.
[37] W.M. Rand, Objective criteria for the evaluation of clustering methods, J. Am. Stat. Assoc. 66 (1971) 846–850.
[38] D. Martin, C. Fowlkes, D. Tal, J. Malik, A database of human segmented natural

images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Comput. Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE Int. Conf., IEEE, 2001: pp. 416–423.
[39] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015: pp. 3431–3440.
[40] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Int. Conf. Med. Image Comput. Comput. Interv., Springer, 2015: pp. 234–241.

Xu-Hao Zhi received the B.S. degree in department of automation from Shanghai Jiao Tong University in 2015. He is currently a graduate student in the Institute of Image Processing and Pattern Recognition of Shanghai Jiao Tong University. His current research interests include image processing and tracking.

Hong-Bin Shen received his Ph.D. degree from Shanghai Jiao Tong University China in 2007. He was a postdoctoral research fellow of Harvard Medical School from 2007 to 2008, and a visiting professor of University of Michigan in 2012. Currently, he is a professor of Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. His research interests include pattern recognition and bioinformatics. He has published more than 100 journal papers and constructed 35 bioinformatics severs in these areas, and he is the associate editor of BMC Bioinformatics and serves the editorial members of several international journals.

ACCEPTED MANUSCRIPT