

## Research Article

# Nonlocal Variational Model for Saliency Detection

Meng Li,<sup>1</sup> Yi Zhan,<sup>2</sup> and Lidan Zhang<sup>3</sup>

<sup>1</sup> Department of Mathematics & KLDAIP, Chongqing University of Arts and Sciences, Yongchuan, Chongqing 402160, China

<sup>2</sup> College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China

<sup>3</sup> College of Mathematics and Physics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Correspondence should be addressed to Yi Zhan; zhanyi@ctbu.edu.cn

Received 4 June 2013; Accepted 25 August 2013

Academic Editor: Baocang Ding

Copyright © 2013 Meng Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a nonlocal variational model for saliency detection from still images, from which various features for visual attention can be detected by minimizing the energy functional. The associated Euler-Lagrange equation is a nonlocal  $p$ -Laplacian type diffusion equation with two reaction terms, and it is a nonlinear diffusion. The main advantage of our method is that it provides flexible and intuitive control over the detecting procedure by the temporal evolution of the Euler-Lagrange equation. Experimental results on various images show that our model can better make background details diminish eventually while luxuriant subtle details in foreground are preserved very well.

## 1. Introduction

Saliency is an important and basic visual feature for describing image content. It can be particular location, objects, or pixels which stand out relative to their neighbors and thus capture peoples' attention. The saliency detection technologies, which exploit the most important areas for natural scenes, are very useful in image and video processing such as image retrieval [1], video compression [2], and video analysis [3]. However, saliency detection is still a difficult task because it somewhat requires a semantic understanding of the image. Furthermore, the difficulty arises from the fact that most of the natural images contain variant texture and color information. So far, a large number of good algorithms and methodologies have been developed for this task. Saliency detection methods can be roughly categorized as biologically based [4, 5], purely computational [6–12], or those that combine the two ideas [13–15].

Itti et al. [4] devise their method based on the biologically plausible architecture proposed by Koch and Ullman [16], in which multiple low-level visual features, such as intensity, color, orientation, texture, and motion are extracted from images at multiple scales and are used for saliency computing. They determine center-surround contrast using a difference of Gaussians approach. Inspired by Itti's method,

Frintrop et al. [5] present a method in which they compute center-surround differences with square filters.

Different from the biological methods, the pure computational models [6–12] are not explicitly based on biological vision principles. Ma and Zhang [6] and Achanta et al. [7, 8] measure saliency using center-surround feature distances. Hou and Zhang [9] devise a saliency detection model based on a concept defined as spectral residual (SR). Liu et al. [10] obtain the saliency map of images from the technology of machine learning. The model in [11] achieves the saliency maps by inverse Fourier transform on a constant amplitude spectrum and the original phase spectrum of images. Feng et al. [12] define the multiscale contrast features as a linear combination of contrasts in the Gaussian image pyramid.

The third category of methods are partly based on biological models and partly on computational ones, that is, the combination of the two ideas. For instance, Harel et al. [13] create feature maps adopting Itti's method but perform their normalization by a graph-based approach. In [14], Bruce and Tsotsos present a saliency computation method within the visual cortex which is based on the premise that localized saliency computation serves to maximize information sampled from one's environment. Fang et al. [15] propose a saliency detection model based on human visual sensitivity and the amplitude spectrum of quaternion Fourier transform.

These methods [4–15] build up elegant maps based on biological theories and/or computational framework. However, some key characteristics in the object are still neglected in these models. For example, the saliency maps generated by the methods [4–6, 9, 13] have low resolution. Moreover, the outputs of [4, 5, 13] have ill-defined boundaries, and the methods [6, 9] produce higher saliency values at object edges instead of the whole object. The methods [7, 8, 15] capture the saliency maps of the same size as the input image. Though methods [7, 8] achieve higher precision than methods [4–6, 9, 13], the information in the background cannot be well suppressed. Additionally, the method [15] seems difficult to extract subtle details (e.g., the texture in saliency) which are very important for visual perception and are primary visual cue for pattern recognition. Moreover, the study of human attention mechanism is not mature yet. Therefore, if we are concerned with high level application such as image retrieval and browsing, we should exploit some mechanism producing accurate saliency.

In this paper, we focus on the problem of saliency detection in the variational framework. The main advantage of variational methods for image processes is that they can be easily formulated under an energy minimization framework and allow the inclusion of constraints to ensure image regularity while preserving important features. Over the past decades, many researchers have devoted their work to the development of variational models and proposed many good algorithms to solve important topics in image analysis and computer vision, including anisotropic diffusion for image denoising [17],  $p$ -Laplacian evolution for image analysis [18], nonlocal  $p$ -Laplacian evolution for image interpolation [19], active contour model for image segmentation [20], and complex Ginzburg-Landau equation for codimension-two objects detection [21] and image inpainting [22], respectively. But, to our knowledge, there exist very few saliency detection methods which take benefits of variational framework.

Inspired by the nonlocal  $p$ -Laplacian [19, 23] and the complex Ginzburg-Landau model [21, 22], we propose a nonlocal  $p$ -Laplacian regularized variational model for saliency detection. Our work is a pure computational model for saliency extraction from still images. The proposed energy functional is described by a diffusion-based regularization, phase transition, and a reaction term for the fidelity. In the energy functional, the nonlocal  $p$ -Laplacian is introduced to penalize the intermediate values of image intensity, and the phase transition makes the background vanish while preserving visually prominent features. Our approach offers the following technical features. First, we formulate saliency detection as a phase transition over an image domain and then a variational framework for saliency selection is developed. Various visual features can be detected by minimizing the energy functional in the variational framework. Second, a dynamical formulation follows naturally from the definition of the energy functional. The associated Euler-Lagrange equation is a nonlocal  $p$ -Laplacian type diffusion equation with a nonlinear reaction term for saliency extraction and a linear reaction term for the fidelity. It achieves the control of the information flow from original images to saliency maps. So the process of saliency extraction is a nonlinear diffusion.

This makes our method quite different from the existing models for saliency detection. Third, our method employs the nonlocal  $p$ -Laplacian regularization which restricts the features of the resulting image. Compared to the classical  $p$ -Laplacian regularization, the direction of edge curves indicated by nonlocal  $p$ -Laplacian is more accurate than the direction indicated by gradient in  $p$ -Laplacian equation [19]. Due to the accuracy of our model, our saliency maps can be seen as salient objects directly. Experimental results on various images show that our model can better make background details diminish eventually, while luxuriant subtle details in foreground are preserved very well.

The remainder of this paper is organized as follows. In Section 2, we review the Ginzburg-Landau model and the nonlocal evolution equation. The proposed model is introduced in Section 3. Section 4 presents numerical method, followed by experiments and results in Section 5. This paper is summarized in Section 6.

## 2. Background

**2.1. The Ginzburg-Landau Models.** The Ginzburg-Landau equation was originally developed by Ginzburg and Landau [24] to phenomenologically describe phase transitions in superconductors near their critical temperature. The equation has proven to be useful in many areas of physics or chemistry [25]. A lot of mathematical theories about this matter can be found in the literature [26]. Moreover, Ginzburg-Landau equations have already been used for image processing [21, 22, 27, 28]. Most of them rely on the simplified energy as

$$E_\varepsilon(u) = \frac{1}{2} \int_{\Omega} \left( |\nabla u|^2 + \frac{1}{2\varepsilon^2} (1 - |u|^2)^2 \right) dx \quad (1)$$

or on the associated flow governed by the following evolution equation:

$$\frac{\partial u}{\partial t} = \Delta u + \frac{1}{\varepsilon^2} (1 - |u|^2) u, \quad (2)$$

where  $\varepsilon$  is a small nonzero constant and  $u$  is a complex-value function indicating local state of the material: if  $|u| \approx 1$ , the material is in a superconducting phase; if  $|u| \approx 0$ , it is in its normal phase. A rigorous mathematical theory on the Ginzburg-Landau functional shows that there exists a phase transition between the above two states [21]. Minimization of the functional (1) develops homogeneous areas which are separated by phase transition regions. In image processing, homogeneous areas correspond to domains of constant grey value intensities and phase transitions to features.

**2.2. Nonlocal Evolution Equations.** Recently, nonlocal evolution equations have been widely used to model diffusion processes in many areas [19]. Let us briefly introduce some references of nonlocal problem considered along this work.

A nonlocal evolution equation corresponding to the Laplacian equation is presented as follows:

$$\begin{aligned} u_t(x, t) &= J * u - u(x, t) \\ &= \int_{\mathbb{R}^N} J(x - y) (u(y, t) - u(x, t)) dy. \end{aligned} \quad (3)$$

The kernel  $J$  is a nonnegative, bounded continuous radial function with  $\text{supp}(J) \subset B(0, d)$  (compact support set). Equation (3) is called a nonlocal diffusion equation since the diffusion of the density at a point  $x$  and time  $t$  depends not only on  $u(x, t)$  but also on all the values of  $u$  in a neighborhood of  $x$  through the convolution term  $J * u$ . This equation shares many properties with the classical heat equation  $u_t = \Delta u$ . This nonlocal evolution can be thought of as nonlocal isotropic diffusion.

For the  $p$ -Laplacian equation  $u_t = \text{div}(|\nabla u|^{p-2} \nabla u)$ , a nonlocal counterpart was studied mathematically in the literature [23]

$$\begin{aligned} u_t(x, t) &= \int_{\Omega} J(x - y) |u(y, t) - u(x, t)|^{p-2} (u(y, t) - u(x, t)) dy \end{aligned} \quad (4)$$

with Neumann boundary condition. It was proved that the solution of (4) converges to the solution of the classical  $p$ -Laplace if  $p > 1$  and to the total variation flow when  $p = 1$  with Neumann boundary conditions when the convolution kernel  $J$  is rescaled in a suitable way [23]. The energy functional corresponding to (4) is

$$E_p(u) = \frac{1}{p} \iint_{\Omega} J(x - y) |u(y, t) - u(x, t)|^p dy dx. \quad (5)$$

### 3. The Proposed Model for Saliency Detection

In this section, we propose a variational model (nonlocal  $p$ -Laplacian regularized variational model) whose (local) minima extract salient objects from image background.

**3.1. Nonlocal  $p$ -Laplacian Regularized Variational Model.** Let  $\Omega \subset \mathbb{R}^2$  be an image domain. For a given image  $I : \Omega \rightarrow \mathbb{R}$ , and we construct a complex-value image  $u_0$  from image  $I$  as following. We first rescale the intensity image  $I(x)$  into interval  $[-1, 1]$  by the formula  $v_0 = (-1)^k((2I(x)/255) - 1)$  ( $k = 0$  or  $k = 1$ ) and assume  $\omega_0 = \sqrt{1 - v_0^2}$ ; then  $I(x)$  is identified with real part  $v_0$  of the complex image  $u_0 = v_0 + \omega_0 i$ , so that  $|u_0| = 1$  for all  $x \in \Omega$ . In order to extract salient objects from a still image, we propose the following energy functional:

$$E(u) = E_p(u) + \psi(u) + \frac{\lambda}{2} \int_{\Omega} |u - u_0|^2 dx \quad (6)$$

with

$$\psi(u) = \frac{1}{2\varepsilon^2} \int_{\Omega} (1 - |u|)^2 dx, \quad (7)$$

where  $p > 2$ ,  $\lambda > 0$ ,  $\varepsilon$  is a small constant,  $u$  is a complex-valued function, and  $E_p(u)$  is defined by energy functional (5). Note that the functional  $\psi(u)$  is slightly different from the second term of the Ginzburg-Landau model (1).

In the following, we will explain the proposed energy functional defined as (6).

- (I) The functional  $E_p(u)$  in (6) serves the purpose of penalizing the spatial inhomogeneity of  $u(x)$ . As we know, certain penalties on intermediate densities are equivalent to restrictions on the microstructural configuration [29]. So the nonlocal  $p$ -Laplacian acts as a regularizer to restrict the feature of the resulting images, physically.
- (II) The potential  $\psi(u)$  in (6) has clearly a minimum at  $|u| = 1$ . Thus the minimization of the functional (6) develops homogeneous areas separated by phase transition regions, which makes  $|u| \approx 1$  almost everywhere after enough diffusion except for the regions of the visually prominent features.
- (III) The third term is a fidelity term which forces  $u(x)$  to be a close approximation of the original function  $u_0$ .

**3.2. Behavioral Analysis of Our Model.** In calculus of variations, a standard method to minimize the functional  $E(u)$  is to find steady state solution of the gradient descent flow equation as

$$\frac{\partial u}{\partial t} = -\frac{\partial E(u)}{\partial u}, \quad (8)$$

where  $\partial E(u)/\partial u$  is Gâteaux derivative of the functional  $E(u)$ . Equation (8) is an evolution equation of a time-dependent function with a spatial variable  $(x, y)$  in the domain  $\Omega$  and an artificial time variable  $t \geq 0$ , and the evolution starts with a given initial function  $u(x, 0) = u_0(x)$ . So a dynamical formulation follows naturally from definition of the energy functional (6) as

$$\frac{\partial u}{\partial t} = P_p^J(u) + \frac{1}{\varepsilon^2} |u|^{-1} (1 - |u|) u + \lambda (u - u_0) \quad (9)$$

with the initial condition  $u(x, 0) = u_0(x)$  and the Neumann boundary condition  $\partial u / \partial \vec{n} = 0$  on  $\partial\Omega$  (where  $\vec{n}$  is the outward unit normal to  $\partial\Omega$ ), where

$$P_p^J(u) = \int_{\Omega} J(x - y) |u(y) - u(x)|^{p-2} (u(y) - u(x)) dy. \quad (10)$$

The kernel  $J : \Omega \rightarrow \mathbb{R}$  in (10) is a nonnegative, bounded continuous radial function with  $\text{supp}(J) \subset B(0, d)$  and satisfies the following properties:

- (1)  $J(-z) = J(z)$ ,
- (2)  $J(z_1) \geq J(z_2)$ , if  $|z_1| < |z_2|$ , and  $\lim_{|z| \rightarrow \infty} J(z) = 0$ ,
- (3)  $\int_{\Omega} J(z) = 1$ .

Equation (9) is a nonlocal  $p$ -Laplacian type diffusion equation with nonlinear reaction terms. Here we will explain

further the nonlocal  $p$ -Laplacian equation. The nonlocal  $p$ -Laplacian  $P_p^J(u)$  in (9) acts as a regularizer to restrict features of the output images. First, the regularizer  $P_p^J(u)$  shares many properties of the classical  $p$ -Laplacian regularization. In the case of saliency detection, we can achieve a reasonable balance between penalizing irregularities (often due to noise) and reserving intrinsic image features by the regularizers with different values of  $p$ . Second, the regularizer  $P_p^J(u)$  improves the classical  $p$ -Laplacian regularization based on local gradient because the nonlocal diffusion at a point  $x$  and time  $t$  depends on all the values of  $u$  in a larger neighborhood of  $x$ . The evolution process at artificial time  $t$  given by (9) is viewed as an anisotropic energy dissipation process. The direction of anisotropic diffusion is indicated by  $|u(y, t) - u(x, t)|^{p-2}$  in a larger neighborhood. It approximates to the direction of edge curve more accurately than the direction indicated by gradient.

We conclude this subsection by discussing dynamical behavior of the formula (9). The temporal evolution of the dynamical formulation (9) makes the energy in (6) decrease monotonically in time. We may make a supposition that the regions with less activity in temporal evolution have rich information and are most likely to attract human attention. Therefore, the irrelevant information will be suppressed gradually, and the visual features can be preserved to the last. This achieves the control of information flow from original images to saliency maps.

#### 4. Numerical Algorithms

In this section, we briefly present the numerical algorithm and procedure to solve the evolution equation (9). In this paper,  $u$  is a complex-valued function. Let  $u = (v, \omega)$ , and we can get the following Euler-Lagrange equations with (9):

$$\begin{aligned} \frac{\partial v}{\partial t} &= \int_{\Omega} J(x-y) \left( (v(y) - v(x))^2 + (\omega(y) - \omega(x))^2 \right)^{(p-2)/2} \\ &\quad \times (v(y) - v(x)) dy \\ &\quad + \frac{1}{\varepsilon^2} (v^2 + \omega^2)^{-1/2} \left( 1 - (v^2 + \omega^2)^{1/2} \right) v + \lambda (v - v_0), \\ \frac{\partial \omega}{\partial t} &= \int_{\Omega} J(x-y) \left( (v(y) - v(x))^2 + (\omega(y) - \omega(x))^2 \right)^{(p-2)/2} \\ &\quad \times (\omega(y) - \omega(x)) dy \\ &\quad + \frac{1}{\varepsilon^2} (v^2 + \omega^2)^{-1/2} \left( 1 - (v^2 + \omega^2)^{1/2} \right) \omega + \lambda (\omega - \omega_0) \end{aligned} \quad (11)$$

with the initial condition  $v(x, 0) = v_0(x)$  and  $\omega(x, 0) = \omega_0(x)$ .

Equation (9) can be implemented via a simple explicit finite difference scheme. Let  $h$  and  $\Delta t$  be space and time steps, respectively, and let  $(i, j) = (ih, jh)$  be the grid points. Let

$u_{i,j}^n = u(i, j, n\Delta t)$  with  $n \geq 0$ . Then we discretize time variable using explicit Euler method for (9) and we have

$$\begin{aligned} &\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} \\ &= \sum_{(k,l) \in \Omega} \left[ J((k,l) - (i,j)) |u_{k,l}^n - u_{i,j}^n|^{p-2} (u_{k,l}^n - u_{i,j}^n) \right] \\ &\quad + \frac{1}{\varepsilon^2} u_{i,j}^n |u_{i,j}^n|^{-1} \left( 1 - |u_{i,j}^n| \right) - \lambda (u_{i,j}^n - (u_0)_{i,j}). \end{aligned} \quad (12)$$

The iteration formulas are given by

$$\begin{aligned} &\frac{v_{i,j}^{n+1} - v_{i,j}^n}{\Delta t} \\ &= \sum_{(k,l) \in \Omega} \left[ J((k,l) - (i,j)) \left( (v_{k,l}^n - v_{i,j}^n)^2 + (\omega_{k,l}^n - \omega_{i,j}^n)^2 \right)^{(p-2)/2} (v_{k,l}^n - v_{i,j}^n) \right] \\ &\quad + \frac{1}{\varepsilon^2} v_{i,j}^n \left( (v_{i,j}^n)^2 + (\omega_{i,j}^n)^2 \right)^{-1/2} \left( 1 - \left( (v_{i,j}^n)^2 + (\omega_{i,j}^n)^2 \right)^{1/2} \right) \\ &\quad - \lambda (v_{i,j}^n - (v_0)_{i,j}), \\ &\frac{\omega_{i,j}^{n+1} - \omega_{i,j}^n}{\Delta t} \\ &= \sum_{(k,l) \in \Omega} \left[ J((k,l) - (i,j)) \left( (v_{k,l}^{n+1} - v_{i,j}^{n+1})^2 + (\omega_{k,l}^n - \omega_{i,j}^n)^2 \right)^{(p-2)/2} (\omega_{k,l}^n - \omega_{i,j}^n) \right] \\ &\quad + \frac{1}{\varepsilon^2} \omega_{i,j}^n \left( (v_{i,j}^{n+1})^2 + (\omega_{i,j}^n)^2 \right)^{-1/2} \left( 1 - \left( (v_{i,j}^{n+1})^2 + (\omega_{i,j}^n)^2 \right)^{1/2} \right) \\ &\quad - \lambda (\omega_{i,j}^n - (\omega_0)_{i,j}). \end{aligned} \quad (13)$$

*Remark 1.* In all numerical experiments, we choose the following kernel function:

$$J(x) := \begin{cases} C \exp\left(\frac{1}{|x|^2 - d^2}\right) & \text{if } |x| < d \\ 0 & \text{if } |x| \geq d. \end{cases} \quad (14)$$

The constant  $C$  is selected such that  $\int_{\Omega} J(x) = 1$ .

*Remark 2.* For color image, let  $r$ ,  $g$ , and  $b$  be red, green, and blue channels of the input image, respectively.  $I$ , the intensity channel, is defined to be used in our model, where

$$I = \frac{r + g + b}{3}. \quad (15)$$

#### 5. Experiments and Results

The proposed nonlocal  $p$ -Laplacian regularized variational model has been applied to detect saliency in varied image scenes. In our experiments, the function  $v_0$  is initialized to  $v_0 = (-1)^k ((2I(x)/255) - 1)$  ( $k = 0$  or  $k = 1$ ) and  $\omega_0 = \sqrt{1 - v_0^2}$ . If the saliency in image is brighter than the background,  $k = 1$ . If the saliency is darker than the



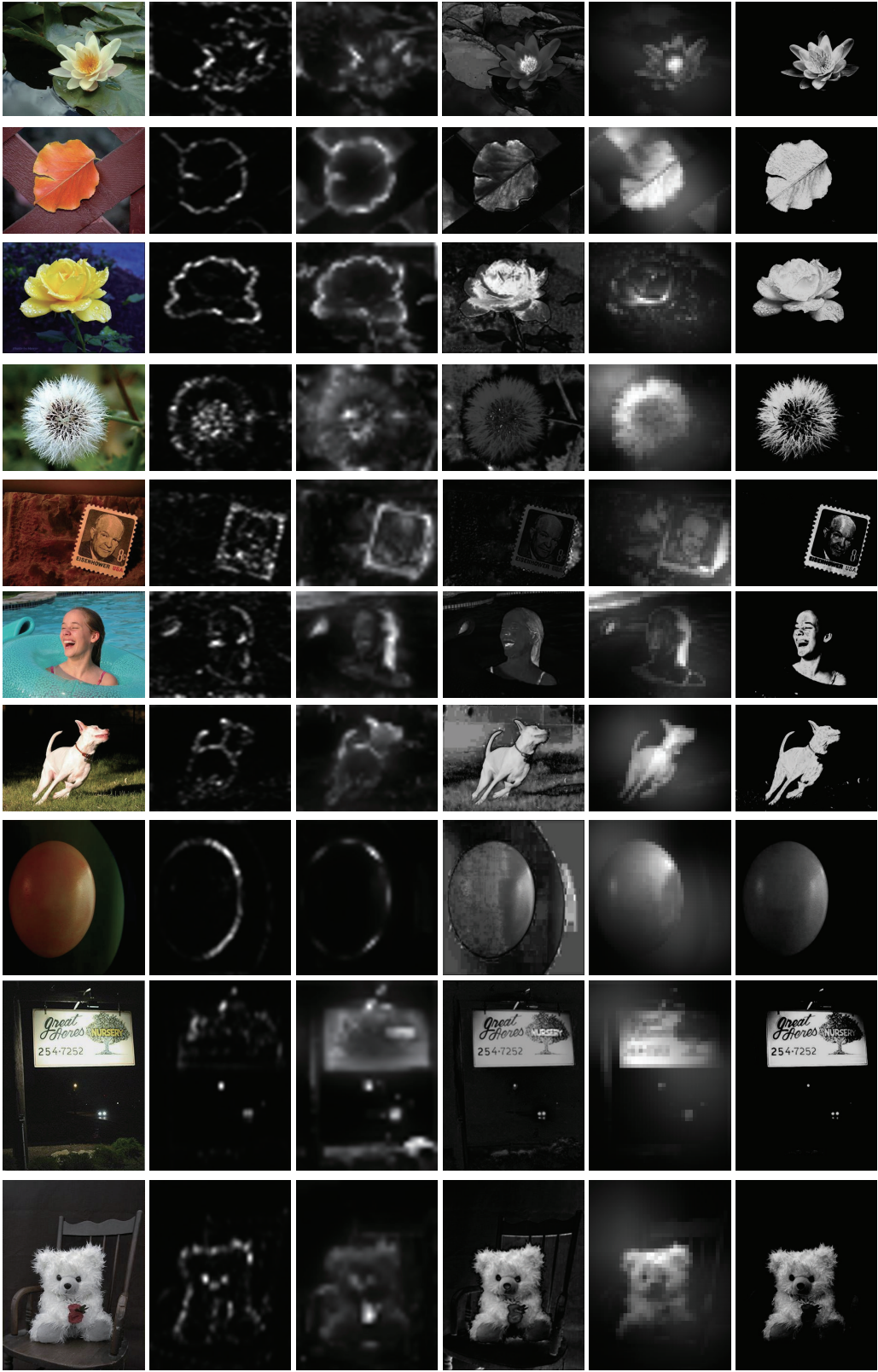


FIGURE 1: Visual comparison of saliency maps for images under complex background. Column 1: original images. Column 2: the SR model [9]. Column 3: the GB model [13]. Column 4: the IG model [8]. Column 5: the AS model [15]. Column 6: our model.

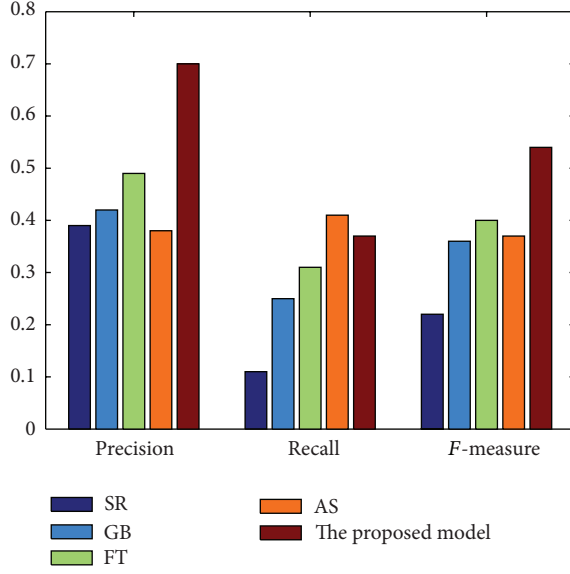


FIGURE 2: Overall mean scores of precision, recall, and F-measure from five different algorithms for 1000 images.

background,  $k = 0$ . We always use the following parameter setting for all the experiments:  $p = 3$ ,  $\varepsilon = 0.5$ ,  $\lambda = 0.1$ , and time step  $\Delta t = 0.03$ . Our saliency maps are displayed by  $\tilde{I} = (255 \cdot (v + 1))/2$ , and the initial state  $\tilde{I}(x, 0) = \tilde{I}_0(x)$  corresponds to the intensity image  $I(x)$ . The reason for using  $\tilde{I}$  as the saliency maps is that  $\tilde{I}$  is the rescaling of  $v$ , which is anisotropic diffusion from initial data  $v_0$ , by the evolution equations (13).

Figure 1 demonstrates the effects of the proposed method on various images with objects having luxuriant subtle details and/or complex background. We compare our saliency maps with four state-of-the-art methods. The four saliency detectors are Hou and Zhang [9], Harel et al. [13], Achanta et al. [8], and Fang et al. [15], hereby referred to as SR, GB, IG, and AS, respectively. The codes of AS model are cited from <http://qtfm.sourceforge.net/>. And the results of SR, GB, and IG models are cited from [http://ivrg.epfl.ch/supplementary\\_material/RK\\_CVPR09/index.html](http://ivrg.epfl.ch/supplementary_material/RK_CVPR09/index.html). From Figure 1, we can see that our saliency maps have well-defined borders, highlight whole objects features, and suppress background better than the other methods even in the presence of complex backgrounds. In addition, our saliency maps have higher accuracy than the previous approaches. We can see from Column 6 that the preservation of subtle details in foreground is very good for these test images; for example, subtle details such as the textures in petals, the downy flower of a dandelion, and the hairs of dog are maintained clearly. In methods SR and GB, Col. 2 and Col. 3 show that the net information retained from original image contains very few details and represents a very blurry version of the original image. In method IG, Col. 4 shows that the high frequencies from the original image are retained in saliency maps whereas some details in background are still clear. In method AS, Col. 5 shows that the background information is suppressed better, but some subtle details in

saliency are smoothed out, and the saliency maps suffer from “stair-case” effects for smooth-texture salient objects, for example, the egg in Figure 1. Due to the accuracy of our model, our saliency maps can be seen as salient objects directly. However, in order to segment a salient object, the other methods need to binarize saliency map such that ones (white pixels) correspond to salient object pixels while zeros (black pixels) correspond to the background [8].

In order to perform an objective comparison of the quality of the saliency maps with other methods, we adopt the precision, recall, and F-measure used by Achanta et al. [8] and Fang et al. [15] to evaluate these methods. The quantitative evaluation of this experiment is based on 1000 images which come from the experimental settings of Achanta et al. [8]. This image database includes original images and their corresponding ground-truth saliency maps. The quantitative evaluation for a saliency detection algorithm is to see how much the saliency map from algorithm overlaps with the ground-truth saliency map. And then for a ground-truth saliency map  $G$  and the detected saliency map  $S$ , we have precision =  $\sum_x g_x S_x / \sum_x S_x$  and recall =  $\sum_x g_x S_x / \sum_x g_x$ , with a nonnegative  $\alpha$ :  $F_\alpha = ((1 + \alpha) * \text{precision} * \text{recall}) / (\alpha * \text{precision} + \text{recall})$ . We set  $\alpha = 0.3$  in this experiment as in the literature [15] for fair comparison. The comparison results are shown in Figure 2. It can be clear that the overall performance of our proposed model for 1000 images is better than the others under comparison in terms of all three measures.

## 6. Conclusion

In this paper, we develop a variational model for saliency detection, which bases on the phase transition theory in the fields of mechanics and material sciences. The dynamics of the system, that is, the temporal evolution from the energy functional, yields information of attention. And the process of saliency extraction is interface diffusion. Compared to the existing models for saliency detection, our method provides flexible and intuitive control over the detecting procedure. Experimental results show that the proposed method is effective in extracting important features in terms of human visual perception.

## Acknowledgments

This work was supported by the NSF of China (nos. 61202349 and 61271452), the Natural Science Foundation Project of CQ CSTC (no. cstc2013jcyjA40058), the Education Committee Project Research Foundation of Chongqing (nos. KJ120709 and KJ131209), the Research Foundation of Chongqing University of Arts and Sciences (no. R2012SC20), and the Innovation Foundation of Chongqing (no. KJTD201321).

## References

- [1] H. Fu, Z. Chi, and D. Feng, “Attention-driven image interpretation with application to image retrieval,” *Pattern Recognition*, vol. 39, no. 9, pp. 1604–1621, 2006.
- [2] C. Guo and L. Zhang, “A novel multiresolution spatiotemporal saliency detection model and its applications in image and video

- compression,” *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 185–198, 2010.
- [3] K. Rapantzikos, N. Tsapatsoulis, Y. Avrithis, and S. Kollias, “Bottom-up spatiotemporal visual attention model for video analysis,” *IET Image Processing*, vol. 1, no. 2, pp. 237–248, 2007.
  - [4] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
  - [5] S. Frintrop, M. Klodt, and E. Rome, “A real-time visual attention system using integral images,” in *Proceedings of the International Conference on computer Vision Systems*, 2007.
  - [6] Y. F. Ma and H. J. Zhang, “Contrast-based image attention analysis by using fuzzy growing,” in *Proceedings of the 11th ACM International Conference on Multimedia (MM ’03)*, pp. 374–381, November 2003.
  - [7] R. Achanta, F. Estrada, P. Wils, and S. Süssstrunk, “Salient region detection and segmentation,” in *Proceedings of the International Conference on Computer Vision Systems*, pp. 66–75, 2008.
  - [8] R. Achanta, S. Hemami, F. Estrada, and S. Süssstrunk, “Frequency-tuned salient region detection,” in *Proceedings of the International on Computer Vision and Pattern Recognition*, pp. 1597–1604, 2009.
  - [9] X. Hou and L. Zhang, “Saliency detection: a spectral residual approach,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’07)*, June 2007.
  - [10] T. Liu, J. Sun, N. N. Zheng, X. Tang, and H. Y. Shum, “Learning to detect a salient object,” in *Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’07)*, June 2007.
  - [11] C. Guo, Q. Ma, and L. Zhang, “Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform,” in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR ’08)*, June 2008.
  - [12] S. Feng, D. Xu, and X. Yang, “Attention-driven salient edge(s) and region(s) extraction with application to CBIR,” *Signal Processing*, vol. 90, no. 1, pp. 1–15, 2010.
  - [13] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” *Advances in Neural Information Processing Systems*, vol. 19, pp. 545–552, 2007.
  - [14] N. D. B. Bruce and J. K. Tsotsos, “Saliency, attention and visual search: an information theoretic approach,” *Journal of Vision*, vol. 9, no. 3, article 5, 2009.
  - [15] Y. Fang, W. Lin, B. S. Lee, C. T. Lau, Z. Chen, and C. W. Lin, “Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum,” *IEEE Transactions on Multimedia*, vol. 14, no. 1, pp. 187–198, 2012.
  - [16] C. Koch and S. Ullman, “Shifts in selective visual attention: towards the underlying neural circuitry,” *Human Neurobiology*, vol. 4, no. 4, pp. 219–227, 1985.
  - [17] H. C. Li, P. Z. Fan, and M. K. Khan, “Context-adaptive anisotropic diffusion for image denoising,” *Electronics Letters*, vol. 48, no. 14, pp. 827–829, 2012.
  - [18] A. Kuijper, “Image analysis using  $p$ -Laplacian and geometrical PDEs,” *Proceedings in Applied Mathematics and Mechanics*, vol. 7, no. 1, pp. 1011201–1011202, 2007.
  - [19] Y. Zhan, “The nonlocal  $p$ -Laplacian evolution for image interpolation,” *Mathematical Problems in Engineering*, vol. 2011, Article ID 837426, 11 pages, 2011.
  - [20] Q. Xin, C. Mu, and M. Li, “The Lee-Seo model with regularization term for bimodal image segmentation,” *Mathematics and Computers in Simulation*, vol. 81, no. 12, pp. 2608–2616, 2011.
  - [21] G. Aubert, J. F. Aujol, and L. Blanc-Féraud, “Detecting codimension-two objects in an image with Ginzburg-Landau models,” *International Journal of Computer Vision*, vol. 65, no. 1–2, pp. 29–42, 2005.
  - [22] H. Grossauer and O. Scherzer, “Using the complex Ginzburg-Landau equation for digital inpainting in 2D and 3D,” *Lecture Notes in Computer Science*, vol. 2695, pp. 225–236, 2003.
  - [23] F. Andreu, J. M. Mazón, J. D. Rossi, and J. Toledo, “A nonlocal  $p$ -Laplacian evolution equation with Neumann boundary conditions,” *Journal de Mathématiques Pures et Appliquées*, vol. 90, no. 2, pp. 201–227, 2008.
  - [24] V. L. Ginzburg and L. D. Landau, “On the theory of superconductivity,” *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, vol. 20, pp. 1064–1082, 1950.
  - [25] M. Ipsen and P. G. Sørensen, “Finite wavelength instabilities in a slow mode coupled complex Ginzburg-Landau equation,” *Physical Review Letters*, vol. 84, no. 11, pp. 2389–2392, 2000.
  - [26] L. Ambrosio and N. Dancer, *Calculus of Variations and Partial Differential Equations: Topics on Geometrical Evolution Problems and Degree Theory*, Springer, Berlin, Germany, 2000.
  - [27] F. Li, C. Shen, and L. Pi, “A new diffusion-based variational model for image denoising and segmentation,” *Journal of Mathematical Imaging and Vision*, vol. 26, no. 1–2, pp. 115–125, 2006.
  - [28] Y. Zhai, D. Zhang, J. Sun, and B. Wu, “A novel variational model for image segmentation,” *Journal of Computational and Applied Mathematics*, vol. 235, no. 8, pp. 2234–2241, 2011.
  - [29] M. P. Bendsoe, “Variable-topology optimization: status and challenges,” in *Proceedings of the European Conference on Computational Mechanics Wunderlich*, W. Wunderlich, Ed., Munich, Germany, 1999.



