

A Diffusion Approach to Seeded Image Segmentation

Juyong Zhang Jianmin Zheng Jianfei Cai
Nanyang Technological University, Singapore
S070051, asjmzheng, asjfcai@ntu.edu.sg

Abstract

Seeded image segmentation is a popular type of supervised image segmentation in computer vision and image processing. Previous methods of seeded image segmentation treat the image as a weighted graph and minimize an energy function on the graph to produce a segmentation. In this paper, we propose to conduct the seeded image segmentation according to the result of a heat diffusion process in which the seeded pixels are considered to be the heat sources and the heat diffuses on the image starting from the sources. After the diffusion reaches a stable state, the image is segmented based on the pixel temperatures. It is also shown that our proposed framework includes the RandomWalk algorithm for image segmentation as a special case which diffuses only along the two coordinate axes. To better control diffusion, we propose to incorporate the attributes (such as the geometric structure) of the image into the diffusion process, yielding an anisotropic diffusion method for image segmentation. The experiments show that the proposed anisotropic diffusion method usually produces better segmentation results. In particular, when the method is tested using the groundtruth dataset of Microsoft Research Cambridge (MSRC), an error rate of 4.42% can be achieved, which is lower than the reported error rates of other state-of-the-art algorithms.

1. Introduction

Image segmentation is an important process in computer vision and image processing, which divides an image into a number of disjoint regions such that the pixels have high similarity in each region and high contrast between regions. While unsupervised segmentation groups elements of an image automatically according to some criteria, supervised algorithms incorporate user intervention into the process to influence the segmentation, which have recently become popular. One type of supervised segmentation methods is seeded segmentation, in which the user provides labelling of some pixels (called seeds) as belonging to the foreground or the background and the algorithm completes the labelling

for the remaining pixels.

There have been some seeded image segmentation methods published [6, 4, 5, 10, 15, 2, 3, 17, 14]. They basically treat an image as a weighted graph with nodes corresponding to pixels in the image and edges being placed between neighboring pixels, and minimize a certain energy function on this graph to produce a segmentation. Different energy functions give different behaviors of the corresponding algorithms [15].

In this paper, we propose to conduct the seeded image segmentation based on the result of a heat diffusion process. Our work is motivated by the observation that the image segment boundary is often embedded in a crowd of “spurious” edges due to noise or texture, and the statement that the solution of the heat diffusion equation with the initial signal is equivalent to the convolution of the initial signal with Gaussian filter at each scale [1]. In our diffusion-based segmentation framework, the seeded pixels are treated as the heat sources and the heat diffuses on the image starting from the sources. After the state goes stable, the image can be segmented according to the temperature at each pixel. The segmentation result depends on how the heat diffuses. The diffusion process is controlled by the diffusion direction and velocity. Our goal is to partition the image along semantically meaningful edges and thus our basic idea is that the heat should diffuse more in the direction parallel to the edge and less in the perpendicular one. With this in mind, we incorporate the local geometric structure of the image into the diffusion process and thus suggest an anisotropic diffusion for segmentation. As a consequence, the local structure information guides the diffusion to give more accurate and reliable segmentation.

One major novelty of this work is the brand new formulation of the seeded image segmentation problem based on diffusion with Dirichlet conditions, where the segmentation relies on the pixel temperatures at the steady state. Another major novelty lies in the proposed anisotropic diffusion method that well controls the diffusion process. Particularly, we modify the heat equation so that the diffusion respects both the local geometric structure of the image and the color difference between neighboring pixels. This

modification significantly improves the segmentation performance.

The rest of the paper is organized as follows: Section 2 briefly describes some related work. Section 3 explains the definition and computation of the structure tensor for a pixel in an image, which will be used in Section 4. In Section 4, we propose a diffusion based image segmentation framework, show that the RandomWalk algorithm [10] is a special case of our framework, and further present an anisotropic diffusion method for seeded image segmentation, which uses the geometric structure of the image to guide diffusion. The numerical scheme is suggested in Section 5 and the experiments are conducted in Section 6. Finally, we conclude the paper in Section 7.

2. Related work

2.1. Seeded image segmentation

Three popular seeded image segmentation algorithms are GraphCut [6, 5], RandomWalk [10] and the Geodesic segmentation [2]. They are based on energy functionals which are minimized via discrete optimization. It is shown in [15, 3] that these three methods minimize a similar energy function while under different L^q norm ($q = 1, 2, \infty$).

The GraphCut method [6] treats the foreground/background as source/sink, and uses a “max-flow/min-cut” algorithm to find a set of edges that separates the source and the sink with the minimum total weights. The cut across the found edges is returned as the segmentation boundary.

The RandomWalk method [10] computes for each unseeded pixel the probability that a random walker starting at that pixel first reaches the foreground or background seeds, and then classifies each pixel into the corresponding group according to the maximal probability. In this paper, we will show that RandomWalk can be viewed as a special case under our diffusion-based segmentation framework, in which the diffusion is conducted only along the directions of the two coordinate axes. Our proposed anisotropic diffusion method improves the RandomWalk method by allowing diffusion along arbitrary directions which respect the local geometric structure of the image.

The Geodesic algorithm [2] classifies the unseeded pixels according to their geodesic distances to the “foreground” and “background”. It intrinsically belongs to the L^∞ -norm approach [3].

As have been shown in [15], the GraphCut method is sensitive to seed quantity and has the problem of “small cut” behavior because it tries to minimize the total edge weights in the cut, the RandomWalk tends to give the “average” cut result, and the L^∞ -norm approach is strongly influenced by the position of the seeds.

2.2. Diffusion for image processing

Partial differential equation (PDE) based methods have been widely used in many image processing tasks such as restoration, multiscale representation, inpainting, smoothing and edge detection [9, 11]. In particular, the diffusion equation has been successfully used for image smoothing, restoration and regularization [11, 8, 1, 13, 18, 16]. The diffusion equation is an important partial differential equation that describes the distribution of heat in a given region over time. For a function $\mathbf{U}(t; \mathbf{p})$ with the time variable t and the location variable \mathbf{p} , the heat diffusion problem can be written as

$$\frac{\partial \mathbf{U}(t; \mathbf{p})}{\partial t} = \text{div}(\mathbf{D}(t; \mathbf{p}) \nabla \mathbf{U}(t; \mathbf{p})), \quad (1)$$

s.t. $\mathbf{U}(0; \mathbf{p}) = \mathbf{U}_0(\mathbf{p})$

where $\mathbf{D}(t; \mathbf{p})$ is the diffusion conductance or the diffusivity at location \mathbf{p} at time t and $\mathbf{U}_0(\mathbf{p})$ is the function at the initial state. It has been known that this diffusion process is similar to the convolution with a Gaussian kernel of variance t [1]. Much attention has been paid to apply this diffusion equation to various image processing problems. However, little work has been done to use the equation for seeded image segmentation. In this paper we explore the use of this diffusion equation in seeded image segmentation. While in most previous work the function $\mathbf{U}(t; \mathbf{p})$ stands for the color of the image and the image processing problem is formulated as an initial value problem of a PDE, we treat $\mathbf{U}(t; \mathbf{p})$ as a scalar quantity called temperature, which we assign to each pixel of the image and use to perform the segmentation. We focus on the steady state of the diffusion equation, and therefore our formulation is more like a boundary value problem than an initial value one.

3. Differential geometry of images

To analyze the local geometric structure of an image, the classic differential geometry theory [7] provides an elegant method if we consider that the image is obtained by discretizing a differentiable surface. For simplicity, we assume that an image is defined by a differentiable function $\mathbf{I}(\mathbf{p}) : \Omega \rightarrow R^n$ where $\Omega \subset R^2$ is the domain of the image (a $2D$ rectangular area), $n \in N^+$ is the number of image channels, and $\mathbf{p} = (x, y)$ represents a point on the domain. We examine the difference of image colors at two points \mathbf{p}_1 and \mathbf{p}_2 , which is given by $\mathbf{I}(\mathbf{p}_1) - \mathbf{I}(\mathbf{p}_2)$. When the distance between \mathbf{p}_1 and \mathbf{p}_2 tends to zero, the difference can be written

$$d\mathbf{I} = \frac{\partial \mathbf{I}}{\partial x} dx + \frac{\partial \mathbf{I}}{\partial y} dy = \mathbf{I}_x dx + \mathbf{I}_y dy.$$

Thus, its squared magnitude is

$$\begin{aligned} \|d\mathbf{I}\|^2 &= d\mathbf{I}^T d\mathbf{I} = \mathbf{I}_x^T \mathbf{I}_x dx^2 + 2\mathbf{I}_x^T \mathbf{I}_y dx dy + \mathbf{I}_y^T \mathbf{I}_y dy^2 \\ &= \begin{bmatrix} dx & dy \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \end{aligned}$$

where $g_{11} = \mathbf{I}_x^T \mathbf{I}_x$, $g_{12} = g_{21} = \mathbf{I}_x^T \mathbf{I}_y$ and $g_{22} = \mathbf{I}_y^T \mathbf{I}_y$. This quadratic expression is actually the first fundamental form of $\mathbf{I}(\mathbf{p})$, which measures the changes of the image. Specifically, for a unit vector $\theta = (\theta_1, \theta_2)$,

$$\|d\mathbf{I}\|^2(\theta) = \begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad (2)$$

is a measure of the rate of change of the image in direction θ .

Let \mathbf{G} be the 2×2 matrix $[g_{ij}]$. It is a symmetric and semi-positive-definite matrix. \mathbf{G} is called the structure tensor [19] because it indicates the local geometry of the image. In fact, the extrema of $\|d\mathbf{I}\|^2(\theta)$ of (2) are obtained in the directions of the eigenvectors of \mathbf{G} and the values attained are just the corresponding eigenvalues [13]. Denote by λ_+ and λ_- the maximal and minimal eigenvalues and by θ_+ and θ_- the eigenvectors corresponding to λ_+ and λ_- . Then θ_+ and θ_- are two orthogonal vectors giving the direction of maximal and minimal changes at a given point in the image and λ_+ and λ_- tell the corresponding change rates. In particular,

- when $\lambda_+ \gg \lambda_-$, there are a lot of variations. The point may be located on an edge. θ_+ and θ_- are respectively the directions across the edge and along the edge.
- when $\lambda_+ \simeq \lambda_- \simeq 0$, there are very few variations around the point. The region is almost flat and does not contain any edge or corner.
- when $\lambda_+ \simeq \lambda_- \gg 0$, the point is located on a saddle point of the surface, which is possibly a corner structure in the image.

Note that for scalar images (i.e., $n = 1$), the eigenvectors $\theta_{+/-}$ and the eigenvalues $\lambda_{+/-}$ of \mathbf{G} are

$$\theta_+ = \frac{\nabla \mathbf{I}}{\|\nabla \mathbf{I}\|}, \quad \theta_- = \frac{\nabla \mathbf{I}^\perp}{\|\nabla \mathbf{I}\|}$$

and

$$\lambda_+ = \|\nabla \mathbf{I}\|^2, \quad \lambda_- = 0.$$

That is, the direction of the maximal change is the gradient direction, the maximal eigenvalue is the squared magnitude of the gradient, and the minimal eigenvalue is zero.

4. Diffusion based segmentation framework

We first would like to explain our principle by looking at a simple physical problem which may give more intuition. The physical model is an object composed of two different materials as shown in Figure 1. The green part is wood and the blue part is copper. We place two heat sources on the object: *Hot* with the constant temperature of 100 and *Cold* with the constant temperature of 0. We can imagine that there is an obvious temperature difference between the wood and copper after the temperature goes stable on the object. The *barrier* can be found by checking the temperature against 50. Actually, this process can be modeled by the heat diffusion equation with spatially varying diffusivity.

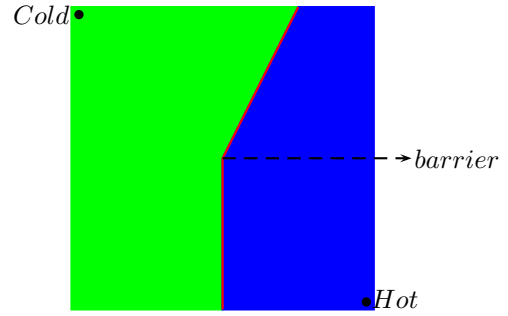


Figure 1. A physical model for the seeded segmentation algorithm.

Now we are ready to describe our diffusion based segmentation framework. Suppose we are given an image \mathbf{I} defined over domain Ω . Denote by F the set of foreground seeds and by B the set of background seeds. We assign a scalar value called temperature to each pixel of the image such as the pixels of F have a constant temperature of 1 and the pixels of B have a constant temperature of 0. Denote by $u(t; \mathbf{p})$ the temperature of pixel \mathbf{p} at time t . We formulate the diffusion process for seeded image segmentation as

$$\begin{aligned} \frac{\partial u(t; \mathbf{p})}{\partial t} &= \text{div}(\mathbf{D}(\mathbf{p}) \mathbf{W}(\mathbf{p}) \nabla u(t; \mathbf{p})), \\ \text{s.t. } u(t; \mathbf{p}) &\equiv 1 \quad \text{for } \mathbf{p} \in F \\ u(t; \mathbf{p}) &\equiv 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (3)$$

where $\mathbf{D}(\mathbf{p})$ is a 2×2 symmetric matrix called the diffusion tensor and $\mathbf{W}(\mathbf{p})$ is a 2×2 diagonal matrix called the inhomogeneous metric tensor. After the diffusion process goes steady, the segmentation is done by examining the temperature at each pixel. That is, pixels with temperature $u \geq 1/2$ are labelled as the foreground and pixels with temperature $u < 1/2$ are labelled as the background. Similar to the classical RandomWalk algorithm, the general model (3) can also be used for multiple objects segmentation.

Note that the proposed equation (3) is a diffusion equation with Dirichlet boundary conditions, which is slightly

different from the one used in regularization or smoothing (1). In addition, the inhomogeneous metric tensor $\mathbf{W}(\mathbf{p})$ that changes in response to color changes is introduced to amend the gradient [15]. The diffusion tensor $\mathbf{D}(\mathbf{p})$ and the inhomogeneous metric tensor $\mathbf{W}(\mathbf{p})$ serve different purposes: $\mathbf{D}(\mathbf{p})$ describes the diffusivity, controlling the diffusion directions and velocities, and $\mathbf{W}(\mathbf{p})$ respects the color variance between neighboring pixels.

For our segmentation purpose, we are more interested in the situation when the diffusion has reached the steady state. Without ambiguity, we omit the time t in our notation and just write $u(\mathbf{p})$ after the diffusion becomes stable. Thus, we seek the solution to the following problem:

$$\begin{aligned} \operatorname{div}(\mathbf{D}(\mathbf{p})\mathbf{W}(\mathbf{p})\nabla u(\mathbf{p})) &= 0 \\ \text{s.t. } u(\mathbf{p}) &= 1 \quad \text{for } \mathbf{p} \in F \\ u(\mathbf{p}) &= 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (4)$$

Given the temperatures at the source pixels, the diffusion process is closely related to the choice of $\mathbf{D}(\mathbf{p})$ and $\mathbf{W}(\mathbf{p})$ as well. Different conduction directions and velocities lead to different stable states.

4.1. Isotropic diffusion

Let us first examine a simple situation, in which $\mathbf{D}(\mathbf{p})$ is chosen to be a 2×2 identity matrix. Thus, (4) becomes $\operatorname{div}(\mathbf{W}(\mathbf{p})\nabla u(\mathbf{p})) = 0$. This is the Euler-Lagrange equation of the following minimization problem:

$$\begin{aligned} \min E(u) &= \int_{\Omega} \|\mathbf{W}^{\frac{1}{2}}\nabla u\|^2 d\Omega \\ \text{s.t. } u(\mathbf{p}) &= 1 \quad \text{for } \mathbf{p} \in F \\ u(\mathbf{p}) &= 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (5)$$

By appropriately discretizing, the problem (5) can be changed to a discrete version:

$$\begin{aligned} \min E(u) &= \sum_{\|\mathbf{q}-\mathbf{p}\|_1=1} w_{\mathbf{p}\mathbf{q}}(u(\mathbf{p}) - u(\mathbf{q}))^2 \\ \text{s.t. } u(\mathbf{p}) &= 1 \quad \text{for } \mathbf{p} \in F \\ u(\mathbf{p}) &= 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (6)$$

If the weights are chosen to be

$$w_{\mathbf{p}\mathbf{q}} = \exp(-\beta\|\mathbf{I}(\mathbf{p}) - \mathbf{I}(\mathbf{q})\|^2), \quad (7)$$

the above minimization problem just describes the RandomWalk algorithm [10, 15]. The temperature of an unseeded pixel \mathbf{p} at the stable state can be explicitly obtained: $u(\mathbf{p}) = \frac{1}{d_{\mathbf{p}}} \sum_{\|\mathbf{q}-\mathbf{p}\|_1=1} w_{\mathbf{p}\mathbf{q}}u(\mathbf{q})$, where $d_{\mathbf{p}}$ is the degree of pixel \mathbf{p} . Apparently, from the viewpoint of the isotropic diffusion, we can see that the RandomWalk algorithm implicitly admits the the following two assumptions:

- The more similar the colors between adjacent pixels are, the more likely the pixels belong to the same region.
- The heat diffuses only along the x- and y-axes.

However, they are not always true in practice.

4.2. Anisotropic diffusion

Note that our diffusion model (3) is very flexible and allows us to locally design the tensors to affect the diffusion. While the inhomogeneous metric tensor can be set up based on the popular weight choice (7) used in the RandomWalk algorithm, the diffusion tensor should be carefully designed to well control the diffusion process. It is natural to expect that the diffusion should obey the geometric structure of the image, e.g. performing a local conduction more in the direction along the edge and less in the perpendicular one. Diffusion only along the x- and y-axes is in general not a good choice. Therefore, we need to measure the local geometry of an image first. The structure tensor we discussed in Section 3 provides us such desired information. We now proceed to derive the diffusion tensor based on the the structure tensor.

For a pixel, suppose that we have already computed its structure tensor \mathbf{G} with the maximal and minimal eigenvalues λ_+ and λ_- and the corresponding unit eigenvectors θ_+ and θ_- . By some matrix conversion, we can rewrite \mathbf{G} as

$$\mathbf{G} = \lambda_- \theta_- \theta_-^T + \lambda_+ \theta_+ \theta_+^T.$$

Although θ_+ and θ_- are two good orthogonal directions for guiding the diffusion, \mathbf{G} is not suitable to be used as the diffusion tensor. This is because the eigenvalues of the diffusion tensor reflect the diffusion velocities along the directions of the corresponding eigenvectors. Directly employing the structure tensor \mathbf{G} as the diffusion tensor will lead to fast diffusion across the edge and slow diffusion along the edge, which is opposite to our intention.

Therefore, we construct the diffusion tensor \mathbf{D} as

$$\mathbf{D} = \lambda_1 \theta_- \theta_-^T + \lambda_2 \theta_+ \theta_+^T \quad (8)$$

where

$$\lambda_1 = \frac{1}{1 + \lambda_-}, \quad \lambda_2 = \frac{1}{1 + \lambda_+}. \quad (9)$$

This diffusion tensor \mathbf{D} is a 2×2 symmetric and positive-definite matrix with two positive eigenvalues λ_1, λ_2 and two corresponding eigenvectors θ_-, θ_+ . It can be visualized with an ellipse, oriented by vectors θ_-, θ_+ and elongated by λ_1 and λ_2 , as illustrated in Figure 2. It satisfies our requirement and gives good diffusion rates and directions at each pixel. In fact,

- when the pixel locates on an edge, $\lambda_1 \gg \lambda_2$ and thus the diffusion velocity along the edge is larger than the one perpendicular to the edge. This means the diffusion is hard to cross the edge;
- at a flat region, $\lambda_1 \simeq \lambda_2 \gg 0$ which gives large diffusion velocities along the two directions, making the temperatures in this region almost the same;
- at a corner, the heat almost stops flowing.

So far we have described how to appropriately set up the diffusion tensor, thus providing an anisotropic diffusion method for seeded image segmentation.

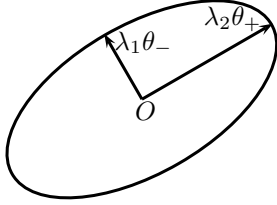


Figure 2. Illustration of tensor \mathbf{D} .

Finally we give an example to illustrate how \mathbf{G} and \mathbf{D} work on images. The example image is chosen from the top of the Lena hat as shown in Figure 3(a). There is an apparent edge along the diagonal. The structure tensor \mathbf{G} in Figure 3(b) has large variance across the edge and small changes along the edge. In addition, the flat area has smaller variance compared with the edge and corner areas. The field of the diffusion tensor \mathbf{D} is shown in Figure 3(c), from which we can see that the diffusion rate along the direction perpendicular to the edge is very small.

4.3. Discussion

As pointed out in [10, 15], the RandomWalk algorithm can be seen as a Dirichlet integral minimization problem with boundary conditions. Similarly, the problem (4) can be viewed as the energy minimization problem with an appropriate tensor \mathbf{T} :

$$\begin{aligned} \min E(u) &= \int_{\Omega} \|\mathbf{T}\nabla u\|^2 d\Omega \\ \text{s.t. } u(\mathbf{p}) &= 1 \quad \text{for } \mathbf{p} \in F \\ u(\mathbf{p}) &= 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (10)$$

This tensor-driven Dirichlet integral expression implies that our model has many good properties. In particular, energy function (10) has a unique minimum value since it is convex. The solution to the Dirichlet integral is a harmonic function, which gives the maximum/minimum principle, uniqueness and mean value properties. The maximum/minimum principle guarantees any unseeded pixel's

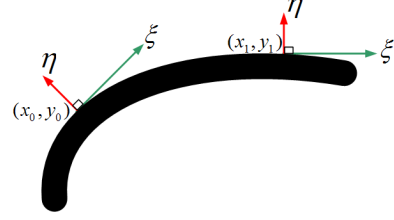


Figure 4. At each pixel, heat diffuses along orthogonal axes ξ and η with different flow rates.

temperature to be in the range of $[0, 1]$. The mean value property ensures that the pixels belonging to the same segment are connected.

Our anisotropic diffusion may also be seen as the simultaneous juxtaposition of two oriented 1D heat flows, leading to 1D Gaussian smoothing processes along orthogonal directions $\xi \perp \eta$, with different weights c_1 and c_2 [13]:

$$\frac{\partial u}{\partial t} = c_1 u_{\xi\xi} + c_2 u_{\eta\eta} \quad (11)$$

The diffusion weights c_1, c_2 and directions ξ, η are induced from the spectral elements λ_{\pm} and θ_{\pm} of \mathbf{G} , in order to perform edge-preserving diffusion, mainly along the direction θ_- orthogonal to the image discontinuities. As shown in Figure 4, each pixel has two diffusion directions. The one along the edge has larger diffusion rate.

5. Numerical scheme

We have formulated our segmentation problem based on a heat diffusion process and we are especially interested in the steady state of the diffusion, which is described by (4). In this section, we present a simple numerical scheme based on \mathbf{D} and \mathbf{W} to approximate the solution. The numerical scheme has the advantage that the maximum principle is preserved due to the fact that the local filtering is done only with normalized kernels.

Using an approach similar to the one given in [16], we can derive an approximate solution to (3), which is:

$$u_t = u_0 * K^{(\mathbf{D}\mathbf{W}, t)}, \quad (12)$$

where $*$ stands for the convolution operator and $K^{(\mathbf{D}\mathbf{W}, t)}$ is an oriented Gaussian kernel defined by:

$$K^{(\mathbf{D}\mathbf{W}, t)}(\mathbf{x}) = \frac{1}{4\pi t} \exp\left(-\frac{\mathbf{x}^T (\mathbf{D}\mathbf{W})^{-1} \mathbf{x}}{4t}\right). \quad (13)$$

Since we only need the temperature at the stable state, there is no need to compute the convolution repeatedly. Considering the boundary conditions and the fact that the temperature goes stable when the convolution process will not change the temperature anymore, we can directly get the temperature at the stable state. After calculating the diffusion tensor \mathbf{D} and inhomogeneous metric tensor \mathbf{W} at each

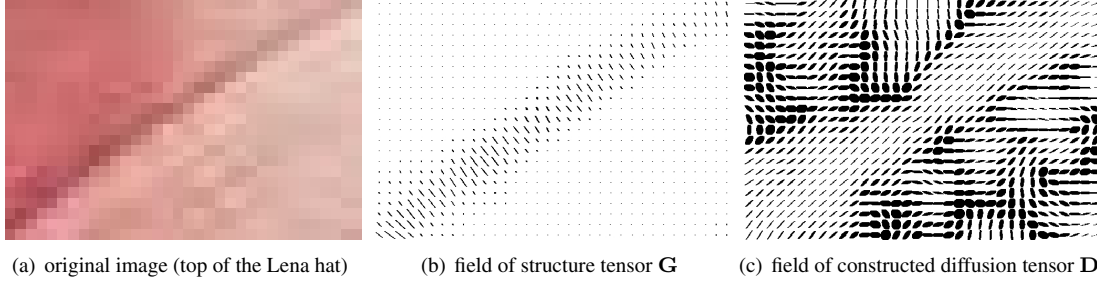


Figure 3. \mathbf{G} reveals the image geometry structure and \mathbf{D} suggests the diffusion directions and amplitudes. Both tensors at each point are illustrated by ellipses.

pixel \mathbf{p} , we then get

$$K(\mathbf{p}, \mathbf{q}) = \exp\left(-\frac{(\mathbf{p} - \mathbf{q})^T (\mathbf{D}\mathbf{W})^{-1} (\mathbf{p} - \mathbf{q})}{\sigma_{\mathbf{p}}}\right), \quad (14)$$

where \mathbf{W} at pixel \mathbf{p} is a diagonal matrix whose element is the square of the forward color difference with periodic boundary condition and its value is normalized to $[0, 1]$ over the entire image, and $\sigma_{\mathbf{p}}$ is a pixel-varying parameter and is set to the inverse of the maximal eigenvalue of $\mathbf{D}\mathbf{W}$ at pixel \mathbf{p} for normalization. Since \mathbf{D} is symmetric and semi-positive definite and \mathbf{W} is a diagonal matrix with nonnegative elements, $\mathbf{D}\mathbf{W}$ is semi-positive definite. This property guarantees that the value of (14) is within interval $(0, 1]$ and matrix $\mathbf{D}\mathbf{W}$ has two nonnegative eigenvalues.

If we choose a $k \times k$ convolution kernel to represent the filter (14), the discrete formulation of our diffusion process for seeded image segmentation (3) becomes

$$\begin{aligned} u(\mathbf{p}) &= \frac{1}{d_{\mathbf{p}}} \sum_{\|\mathbf{q}-\mathbf{p}\|_{\infty} \leq \frac{k-1}{2}} K(\mathbf{p}, \mathbf{q}) u(\mathbf{q}), \\ \text{s.t. } u(\mathbf{p}) &= 1 \quad \text{for } \mathbf{p} \in F \\ u(\mathbf{p}) &= 0 \quad \text{for } \mathbf{p} \in B \end{aligned} \quad (15)$$

where $d_{\mathbf{p}} = \sum_{\|\mathbf{q}-\mathbf{p}\|_{\infty} \leq \frac{k-1}{2}} K(\mathbf{p}, \mathbf{q})$. This is a linear system which can be written in the form of $AU = b$, where the coefficient matrix A is sparse and positive-definite. Thus we use the preconditioned conjugate gradient method to solve the system. In our study, we find setting $k = 3$ is sufficient for most images. In this case, there are at most 9 nonzero elements in each row of the coefficient matrix A while it is 5 for the original random walks [10]. Thus, for all the images tested, our proposed system runs almost as fast as the original random walks, which can provide instantaneous feedback.

6. Experiment results

In this section, we conduct experiments to verify the effectiveness of the proposed anisotropic diffusion approach. We first use a synthetic image to illustrate the limitation

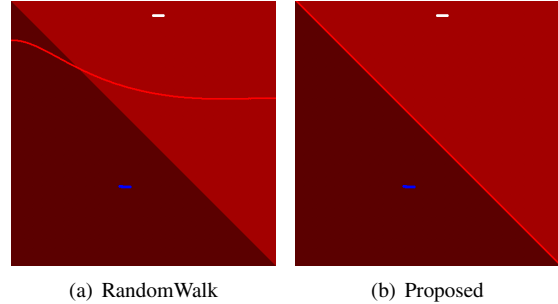


Figure 5. The segmentation results of a synthetic image with white (blue) strokes indicating foreground (background) and red lines indicating the cut boundary. Note that the bottom-left region actually contains the two colors of $(128, 128, 128)$ and $(156, 116, 56)$, and the top-right region actually contains the two colors of $(128, 128, 128)$ and $(100, 140, 200)$.

of isotropic diffusion, i.e. the classic RandomWalk algorithm. In particular, the RGB values of the synthetic image in Fig. 5 are constructed as

$$\begin{aligned} \mathbf{I}(x, y) &= (128, 128, 128), & x + y = 2n \\ \mathbf{I}(x, y) &= (100, 140, 200), & x + y = 2n + 1, x > y \\ \mathbf{I}(x, y) &= (156, 116, 56), & x + y = 2n + 1, x \leq y \end{aligned}$$

As shown in the figure, the RandomWalk algorithm gives the ‘‘average’’ cut between the foreground and the background seeds, which does not follow the apparent diagonal edge at all. This is mainly because the exactly same weight of $w_{\mathbf{p}\mathbf{q}} = \exp(-\beta\|\mathbf{I}(\mathbf{p}) - \mathbf{I}(\mathbf{q})\|^2)$ is used over the entire image. In comparison, the proposed anisotropic diffusion method incorporates the local geometry structure information into the weights, which well controls the heat diffusion and leads to accurate cut.

Fig. 6 and Fig. 7 show the segmentation results on several natural images and medical images, respectively. We compare the proposed anisotropic diffusion approach with not only the RandomWalk algorithm but also the Grabcut algorithm [12], which improves the classical GraphCut algorithm by incorporating better prior models and various types of user inputs. From the results, we can see that GraphCut has the disconnection problem, i.e. its segmentation results contain many isolated regions, as shown in

the first two rows of Fig. 6, and the “small cut” problem as shown in the first row of Fig. 7. As expected, the results of RandomWalk often miss the edges. All the results reported above for the RandomWalk and Grabcut algorithms are the best results obtained through exhaustively searching for their optimal scaling parameter β . We also test our method using the MSRC benchmark dataset, which contains 50 test images with ground truth. As shown in Table 1, the error rate (percentage of mislabeled pixels) of our method is lower than other approaches. These extensive experiments demonstrate that our method has a consistent performance. In general, our approach achieves better segmentation results with more accurate boundary.

Table 1. Error rate comparison using the benchmark dataset with exactly the provided trimaps.

Segmentation model	Error rate
GMMRF [4]	7.9%
GrabCut [12]	5.6%
Random Walker [10]	5.4%
Proposed	4.42%

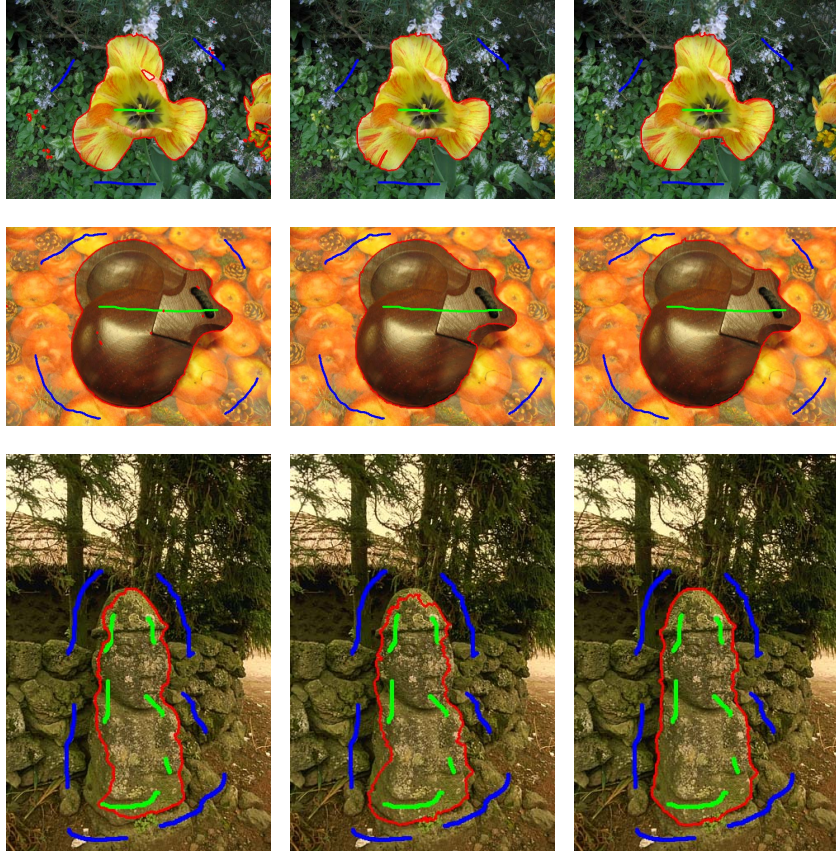
7. Conclusions

In this paper, we have formulated the seeded image segmentation based on the result of a heat diffusion process. The RandomWalk algorithm is showed to be an isotropic diffusion approach, i.e. a special case of our proposed diffusion framework. To overcome the limitation of isotropic diffusion, we further proposed an anisotropic diffusion method for seeded segmentation, which takes the local geometric structure of the image into consideration. The experimental results show that the proposed anisotropic diffusion approach generally has good diffusion direction and velocity than the RandomWalk algorithm and thus provides more accurate segmentation. As for future work, we plan to incorporate the tensor based weights into other seeded image segmentation methods such as GraphCut and the Geodesic algorithm to improve their performance.

The major limitation of the proposed anisotropic diffusion approach is that its performance is still affected by the seed locations, although not as sensitive as the RandomWalk algorithm. As shown in the experiments, the strokes need to be placed in a way that the desired boundary is roughly in the middle of background and foreground strokes. If one type of strokes is relatively too far away from the boundary, the diffusion based approach cannot achieve satisfactory performance.

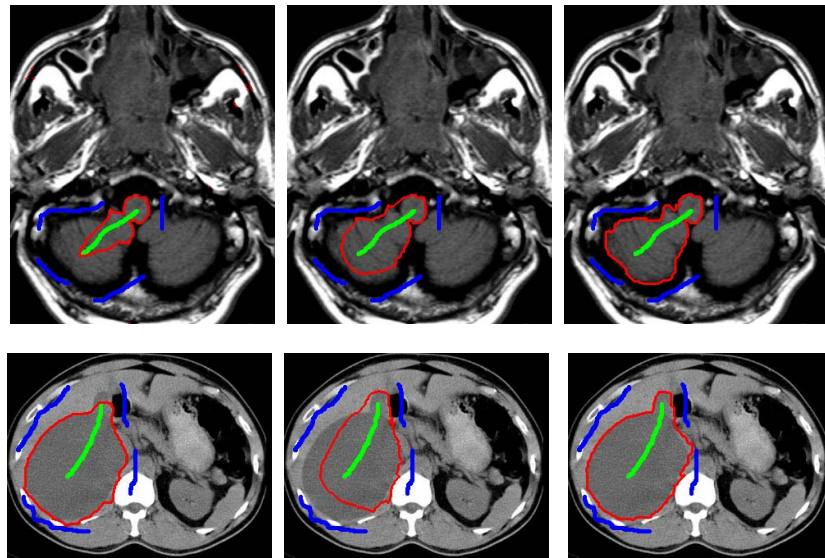
References

- [1] L. Alvarez, P.-L. Lions, and Jean Michel Morel. Image selective smoothing and edge detection by nonlinear diffuse (II). *SIAM J.Numer.Anal.*, 29(3):845–866, June 1992. 1, 2
- [2] X. Bai and G. Sapiro. A Geodesic Framework for Fast Interactive Image and Video Segmentation and Matting. In *Proc. of ICCV 2007*. IEEE Computer Society, IEEE, Oct 2007. 1, 2
- [3] X. Bai and G. Sapiro. Geodesic Matting: A Framework for Fast Interactive Image and Video Segmentation and Matting. *International Journal of Computer Vision*, 82(2):113–132, 2009. 1, 2
- [4] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr. Interactive image segmentation using an adaptive GMMRF model. In *Proc. of ECCV 2004*, pages 428–441. IEEE Computer Society, IEEE, 2004. 1, 7
- [5] Y. Boykov and G. Funka-lea. Graph Cuts and Efficient N-D Image Segmentation. *Int. J. Comput. Vision*, 70(2):109–131, November 2006. 1, 2
- [6] Y. Boykov and M.-P. Jolly. Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. In *Proc. of ICCV 2001*, pages 105–112. IEEE Computer Society, IEEE, July 2001. 1, 2
- [7] M. D. Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976. 2
- [8] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29:182–193, Feb 1992. 2
- [9] T. F. Chan and J. Shen. *Image Processing and Analysis*. SIAM, Philadelphia, 2005. 2
- [10] L. Grady. Random Walks for Image Segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, Nov. 2006. 1, 2, 4, 5, 6, 7
- [11] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990. 2
- [12] C. Rother, V. Kolmogorov, and A. Blake. “GrabCut”: Interactive Foreground Extraction using Iterated Graph Cuts. *ACM Trans. Graph.*, 23(3):309–314, 2004. 6, 7
- [13] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5(11):1582–1586, 1996. 2, 3, 5
- [14] D. Singaraju, L. Grady, and R. Vidal. P-Brush: Continuous Valued MRFs with Normed Pairwise Distributions for Image Segmentation. In *Proc. of CVPR 2009*. IEEE Computer Society, IEEE, June 2009. 1
- [15] A. K. Sinop and L. Grady. A Seeded Image Segmentation Framework Unifying Graph Cuts and Random Walker Which Yields a New Algorithm. In *Proc. of ICCV 2007*. IEEE Computer Society, IEEE, Oct. 2007. 1, 2, 4, 5
- [16] D. Tschumperlé and R. Deriche. Vector-Valued Image Regularization with PDE’s: A Common Framework for Different Applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(4):506–517, April 2005. 2, 5
- [17] S. Vicente, V. Kolmogorov, and C. Rother. Graph cut based image segmentation with connectivity priors. In *CVPR*, 2008. 1
- [18] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, 1998. 2
- [19] S. D. Zeno. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, Jan 1986. 3



(a) GrabCut (b) RandomWalk (c) proposed

Figure 6. The segmentation results of three natural images, “flower”, “music” and “stone”, with green (blue) strokes indicating foreground (background) and red lines indicating the cut boundary.



(a) GrabCut (b) RandomWalk (c) proposed

Figure 7. The segmentation results of two medical images with green (blue) strokes indicating foreground (background) and red lines indicating the cut boundary.