**Partial Differential Equation Acceleration for level set image segmentation**

Abstract

In this project we further applied a novel framework, which called PDE accelerations, to the problem of Chan-Vese active contours. The original Chan-Vese model convergences slowly and has a bad performance on noisy images. The feature of the level set surface is increased by employing a new model called Fuzzy region competition to replace the initial signed distance surface. The gradient descent method, which is utilized by general PDE problems to find the optimal solution including aforementioned models, has been replaced by our PDE acceleration scheme. A simple explicit finite difference scheme solution of the PDE acceleration is used in our project and in this case the acceleration is achieved by the enhancement in the CFL condition from Δt ~ Δx2 (for diffusion) to Δt ~ Δx (for wave equations). To exploit the potential of PDE acceleration, the hard constrain of surface in Fuzzy region competition is substantiated by soft penalization. We carefully explore our new algorithm on Chan-Vese level set segmentation with both total variation regularization and Beltrami regularization and experiment results are compared with another top-of –art algorithm called Chambolle-Pock primal dual.

Main text

Introduction

Binary segmentation has long been a hot topic in computer vision. The general goal is to detect the object we interested in and separate it from the background. To achieve this, a contour is evolved to overlap with the boundary of the object. There have been lots of methods proposed to solve this problem.

In the early stage classical segmentation models, such as Snake model[] and other active contours[], an edge detector is employed to help the curve find boundary based on the gradient of the image. However, they are not capable of handling complex images. For example, they are susceptible to images with complicated or ambiguous edges and they do not have a good performance on noisy images because in such cases the gradient of image is not a reliable indicator of the boundary. Therefore there is a great need to increase the generality of the active contour and to find more precise boundary.

Recently, region-based segmentation models, such as ChanVese[] or Mumfold-Shah[] are proposed to solve these problems and have been widely used in segmentation tasks. Region-based models segment the image based on the pixel values, assign the pixels with similar value to the same region and keep the regions as uniform as possible. The gradient flow of region based model is not related to the gradient of the image, therefore it could be applied to image segmentation with smooth boundary and discontinued boundary. It could also detect the boundary with meaningful gradient and meaningless gradient. But there still some limitation of these models. The convergence speed is too slow and there still space for improvement of performance on noisy image. Fuzzy region competition model[] is developed to improve the speed of convergence by using a new kind of level set, which constrains the level set between 0 and 1.

However, all the aforementioned models are classical PDE model and the gradient descent method is utilized to find the solution of the PDE. Most of the PDE image segmentation problems have the form

Where L is a convex functional of variable *u* and the gradient descent of this functional is:

It is not efficient to use gradient descent to solve the problem because it is under the constrain of stability condition(CFL condition) . More efficient methods are proposed, such as Chambolle-Pock primal-dual, to avoid this convergence speed limitation.

Accelerated gradient search algorithms have been widely used in machine learning. The gradient descent convergences slowly because on a large scale, the local gradient direction is not precise, resulting in large move in poor direction and large fixation in contrary direction. Accelerated method converges considerably faster and is more robust in searching local direction since pass information is incorporated into the current descent. In Polyak’s heavy ball method[], which is one of the oldest acceleration models, the update of the variable is:

where is the momentum term, which corrects the current gradient direction by incorporating previous gradient direction.

Nesterov[] momentum is a recently acceleration method that is very prevailing among machine learning community.

This was furthered developed by Sundaramoorthi and Yezzi to adapt it to Partial Differential Equation settings, which have been applied to many PDE problems such as active contour, diffusion, optical flow. In this project, we applied this framework to the level set image segmentation problem of with both TV regularization and Beltrami regularization. The hard clip in Fuzzy region competition is replaced by ramp penalization in TV regularization and quadratic penalization in Beltrami regularization to fully explore the potential of momentum method .The comparison is made between our model and other top-of-art model such as Chambolle-Pock.

Chan-Vese level set method

1. The Chan-Vese Model

1.1 Definition of energy function

Chan-Vese model[1] require the definition of an energy functional. The purpose of energy functional is to create a model for the energy of the image that the contour C will attempt to minimize. At each time step, the goal of the active contour is to progress to a lower energy. The Chan-Vese model segment an image into two regions based on pixel values, therefore the goal of energy functional is to minimize the energy in both side of the contour.

The energy functional could be defined as follows:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

In this equation (1), is the image, and are constant values which represent the average pixel intensity inside and outside the contour respectively. ,, , and are weight of each energy term. The first two terms pertain to the intensity found in the two regions the contour seeks to create. Whatever is inside the contour should be an object of interest and whatever is outside of the contour should be the rest of the image. The first two terms are called fidelity, which seek to minimize the energy on both sides of the contour and keep the regions as uniform as possible. The third and fourth term are called regularity, which seek to penalize the length of the contour, which will lead the contour to attempt to enclose with as little length as needed. The final term seeks to penalize the area of the contour, which will attempt to force it only enclose the smallest area possible. The parameters in front of each term vary the influence each energy term has on the overall energy.

* 1. Level Set formulation

Instead of directly evolving and parameterizing the contour, we can redefine the problem in the level set formalism. A surface *φ(x,y)* is defined in the space *Ω: ¶Rmxn* , which is the same scale of the image and the contour is represented by the zero level set of surface *φ*.

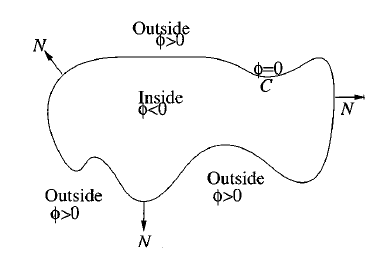


Figure1. Surface and contour

The problem is converted to searching for the solution in terms of surface *φ*. So the energy function need to be rewritten in terms of surface *φ*. To find the minimum value of the functional (), we have to use the Euler -Lagrange equation to obtain the gradient flow. The computation is much simpler if the integral boundary is fixed. Therefore some auxiliary functions are defined to convert the integral region of four integral terms into the image space *Ωmxn*.

The Heaviside function is defined as , which means the value of this function is one inside the contour and zero outside the contour. The delta function is defined as , which is the derivative of Heaviside function in terms of level set.

We express each energy terms by level set :

So the energy function can be rewritten as

|  |  |
| --- | --- |
|  |  |

* 1. Derive gradient descent

The gradient descent method is used to find the minimum of the energy function in terms of *φ*.

* 1. Numerical implementation

For numerical implementation, the Heaviside function could be written in the form as . The delta function is written as .

The above algorithm is expressed in a continuous form, both in time and space. The model has to be discretized with finite difference implicit scheme for implementation purpose. For notation: let Δ*x* be the space between pixels(assume Δ*x* = Δ*y*), Δ*t* represents the time step, *(i,j)* is the space location, n is the iteration time. The finite difference in space could be written as:

The forward difference in time is used in the update of level set :

* 1. Stability condition

In the original ChanVese method, the reinitialization process is utilized to help keep the surface as a signed distance function. Stability condition is not discussed and there is not strict limitation for time step because the reinitialization help prevent the surface from blowing up. However, in our PDE acceleration scheme reinitialization could not be used because it will interrupt the accumulation of the descent. Therefore the time step in our scheme has to be constrained by strict stability condition. Van Neumann analysis is used to determine the CFL condition for these terms.

Using a forward difference in time to approximate the time derivative on the left hand side we obtain:

Taking a Discrete Fourier Transform (DFT) on both sides of the homogeneous part to obtain:

where is called gradient amplifier, defined by .

To prevent blowing up , and the gradient amplifier is real and non-negative. In such cases the stability constraint takes the form of the following CFL condition

To obtain the maximum gradient amplifier, we seek for the maximum gradient amplifier for fidelity term and regularity term separately then we can get the maximum for the whole gradient amplifier.

1.5.1 Stability condition for fidelity term

The gradient flow of the fidelity term is

().

Expand the bracket to get

We have to rewrite it in the form of so that we could obtain the maximum gradient amplifier . Apply Taylor expansion for delta function at the point , where is a very small value so we can obtain , where Therefore .

So

We separate the z(x,t) into two parts and , then we search for the maximum value of two parts independently and obtain the maximum value of their product.

For , we analysis the monotonicity of three variables then we obtain when reaches its maximum value

To obtain , we calculate the derivative of , which is , and set it as zero to get the extreme value point.

.

When , we can obtain .

Multiply the maximum value of those two parts we obtain the maximum value of gradient amplifier for the fidelity term:

1.5.2 Stability condition for regularity term

The gradient flow of the regularity term is

It is not easy to derive the gradient amplifier if the gradient flow is in this form, therefore we rewrite it as

Where denotes the normal unite gradient vector of .

As we are calculating the gradient amplifier of the gradient flow, so it doesn’t matter what is the gradient direction. Thus we assume that its gradient direction align with the y-axis, .

Take the discrete Fourier Transform of it

Therefore we obtain the gradient amplifier

As . The maximum of its value is:

Then we sum the maximum value of the gradient amplifier for fidelity tern and regularity term to obtain the maximum value of the total gradient amplifier.

As there is always a point on the surface where , so . In this case, the CFL condition is *dt = 0*. As a result, the surface could not evolve. The problem is from the total variation regularization term.

If we want to use TV regularization for segmentation, we can use a minimum acceptable to plug in the gradient amplifier . A quantization level Q could be utilized to represent digital u and the approximation bounds for

Then we substitute the level set value between two pixels with the quantization interval Q. The quantization level will cause distortions between pixels, but it is not the same as the violation of CFL condition which will cause blowing up. Instead, the instability degree will be constrained by the quantization level. However, in that case, performance of the algorithm declines because the update of level does not strictly consistent with the gradient flow of the energy function.

We then plug

into the gradient amplifier to obtain:

* 1. Beltrami regularization

Beltrami regularization term is defined as . The new energy function is

The new gradient flow is:

To get the gradient amplifier of new gradient flow, we calculate the fidelity term and regularity term independently and then combine them as what we have done in the last part. As the fidelity term remains unchanged, we only need to calculate the regularity term.

As when , .

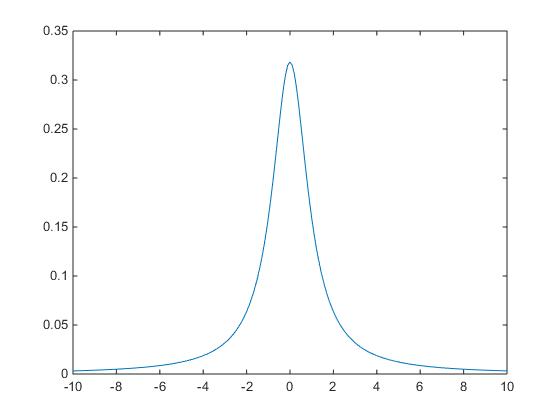
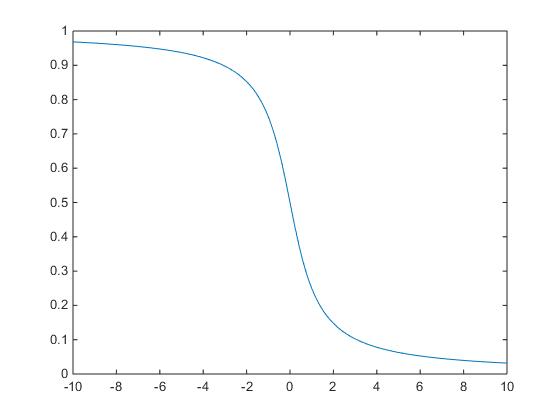
Plugging this into the gradient flow yield:

Take the discrete Fourier Transform of it

Remind that , therefore the upper bound for gradient amplifier

1. Fuzzy region competition

The convergence speed of traditional Chan-Vese model is considerable slow because of the delta function. In the gradient flow[], all the terms are multiplied with delta function. Recall that delta function is the derivative of the Heaviside function.



As shown in the plot, the delta function has a high peak value in the vicinity of 0 level set, which corresponds to the position of the curve. Therefore, for each iteration, the evolvement of level set surface will only happen in the vicinity of the curve. The convergence speed is barricaded by the property of delta function.

In the fuzzy region competition model[3], the surface is constrain within [0,1] and the contour is 0.5 level set. We use *u* to denote the new surface. Chan-Vese model, the initial surface is a signed distanced function, which means the slope of the surface is 1 and its shape is like a cone. There is no limitation of the surface value. Whereas in the fuzzy region competition model, the surface is strictly constrain with in [0,1]. If during an update the value of a point on the surface becomes larger than 1, then it will be hard clipped to 1. Same for 0. There is no need to utilize Heaviside function construct the link between surface value and category of that point(inside, outside or on the curve), so the delta function is on longer an impedance of the convergence speed. The new energy function is:

Beltrami regularization:

TV regularization:

Obtain the gradient flow with Euler-Lagrange Equation

Beltrami regularization

TV regularization

In the gradient flow() and (), the fidelity is homogenous term, therefore it has no contribution to the stability limitation. The gradient amplifier is only determined by regularity term.

Beltrami regularization:

TV regularization:

In this new level set we, the update of the surface is constrained within [0,1]

1. PDE acceleration scheme

3.1 General PDE acceleration framework

We define the new energy functional which integrates kinetic energy:

The new action integral comprises two parts. is the potential energy that we have already defined in the last part and represents the kinetic energy. denotes the weight of potential energy and is another time-dependent coefficient. denotes the distribution of mass. To obtain PDE accelerated motion equation for the action integral in the Lagrange sense, we take a variation on the action integral:

Obtain the derivative of action integral

Integral by part for we get

Remind that , therefore

We denote and we assume that and are constant as value one in space and time. We have

where *a* is called damping coefficient. Compared with gradient descent

* 1. Explicit Euler discretization for accelerated PDE

We now start to focus on the discretization for accelerated PDE. We use Euler discretization framework here, which means central difference in time and forward difference in space.

Rearrange the expression we have

Apply Discrete Time Fourier transform to [] and substitute , where ξis the update amplifier. Then we obtain

We denote the coefficient for the second order term as A, coefficient for the first order term as B, constant as C. To maintain the stability, we have to keep .

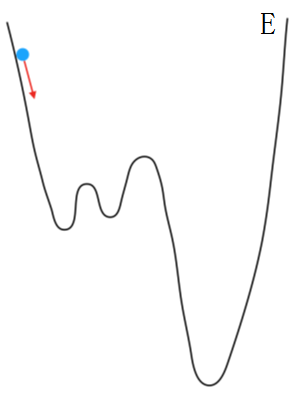
According to the Root Amplitude Lemma, we have to satisfy

For the first condition, as *a* and are larger than zero, is always satisfied. Now we start to consider the second condition. It is . Then we analyze the absolute value. For , we get , which is always satisfied. For , we get . We plug in the maximum value of to obtain the extreme case. Remind that , so we obtain the CFL condition

For the update scheme we have

Rewrite the expression in the form of increment .

From here we clearly see a traditional momentum style structure and how previous descent information integrates into the current descent.

3.3 Physical interpretation of PDE acceleration

The searching for the minimum value of a potential functional can be viewed as a physical ball rolling down a valley guided by gravity. The gravitational potential energy of the ball is the potential functional and its kinetic energy is .

In gradient descent, the velocity of that ball is proportional to the slope of current position, which is the gradient . In contrast, the speed is accumulated in kinetic energy in the PDE acceleration scheme.

There are two benefits using acceleration optimization. First, previous momentum is integrated into current momentum. So even when the current gradient is small, previous momentum could energize the ball to move forward. Second, momentum will prevent the optimization process from being trapped in local minimum, which is a very intractable problem. If the current gradient is zero, the momentum will still move the ball forward to look for potentially better minimizer.

1. Accelerated PDE for Fuzzy region competition

4.1 Level set update scheme

Substitute the maximum gradient amplifier, damping coefficient and gradient flow of both Beltrami regularization and TV regularization into accelerated framework we have

Beltrami:

Total Variation:

Then we have to find the optimal damping coefficient for accelerated PDE. Damping coefficient plays a critical role in the performance of momentum. Large damping coefficient will lead to decline of the convergence speed, while small damping coefficient will result in instability.

According to [], for a linear Dirichlet problem

where L denotes a uniformly elliptic second order partial differential operator, *b* and *c* are the weights for the second order term and the first order term respectively, f denotes the constant term. The problem is defined in the domain . The optimal damping coefficient is

where is the first the first Dirichlet eigenvalue of *L*. Now we apply this analysis to accelerated PDE motion function of Fuzzy region competition with Beltrami regularization. Rewrite [] in a linear form

We notice that there is no first order term in the expression, therefore The coefficient for the second term and the second order elliptic operator . Remind that in our setting , so the first Dirichlet eigenvalue here is . The optimal damping coefficient of Fuzzy region competition with Beltrami regularization

For TV regularization, there is no optimal damping coefficient. It is because we plug in quantization level Q to substitute the minimum acceptable gradient between points, therefore the TV regularization is no longer a strictly PDE problem. The analysis of

1. Primal dual method

Primal-dual optimization is another successful technique and it has been widely utilized by many algorithms to solve PDE problems. These methods[] are easy and simple to implement and converge considerable faster than gradient descent. Primal-dual TV algorithm was first proposed by Chambolle[] and further developed by [] to make it more efficient. Recently, it has been extended to Beltrami regularization on inverse problems.

The primal-dual algorithm is to find the solution for a saddle point problem, which is established using Legendre-Fenchel transform

where and are convex functional, is a continuous linear operator, is the convex conjugate of , and are the primal and dual variable respectively. For detailed proof please check [].

In this section, we apply both primal-dual TV algorithm and primal-dual Beltrami algorithm to level set segmentation problem.

5.1 Beltrami regularization

Now we derive our Beltrami segmentation model [] guided by primal-dual Beltrami-denoising optimization problem developed by []. By comparison our potential energy functional[] with the equivalence [], we get the following expression

where the linear operator , integral domain , .

We set . The Legendre transform of is

Hence we obtain the convex conjugate of

where . Plug and into we get our primal-dual Beltrami segmentation problem:

Further, we integrate the by part and rewrite the problem as

We now use the projected gradient method to solve the problem

To avoid the denominator become zero, we multiply the second term with .Then projection algorithm is used to yield the update scheme

where and are the update time steps for prime and dual variables respectively. To achieve the constrain on the variable domain we have

Plug into () we finally obtain the primal-dual Beltrami segmentation algorithm

* 1. TV regularization

By comparison TV regularization potential energy functional [] with the equivalence [], we get the following expression

where the linear operator , integral domain ,

We set . The Legendre transform of is

Hence we obtain the convex conjugate of

where . Plug and into we get our primal-dual TV segmentation problem:

Further, we integrate the by part and rewrite the problem as

We now use the projected gradient method to solve the problem

Then projection algorithm is used to yield the update scheme

To achieve the constrain on the variable domain we have

Plug into () we finally obtain the primal-dual Beltrami segmentation algorithm

1. Soft penalization

The acceleration method takes advantage of integrating previous momentum into current momentum. However, in the Fuzzy region competition model we hard clip the level set surface that extend the boundary, which suddenly reduce the momentum and result in the interruption of acceleration process. We assume that hard clip process may have deteriorating effect on the performance of accelerated descent.

We substitute hard-clip with soft penalization to further increase the performance of PDE acceleration on obstacle problem. We consider the *L2* penalization(quadratic)

Where the second term is the penalization for level set above one and the third term is the penalization for level set below zero. and are weights of two penalization terms. and are the subspace of Ω. and Apply the new energy functional to the update scheme() we can obtain

Apply Euler discretization framework() to the update scheme

Then we get the explicit solution

The CFL condition for this scheme is only determined by and is unrelated to the penalty and . And we have proved that the accelerated optimization for obstacle problem is independent from the value of and . Hence we take the limit of and

Substitute the limitations into the update scheme and therefore we could obtain a simpler scheme

Surprisingly, this indicates that the soft penalization is equivalent to hard-clip. In order to further improve the performance of