IMAGE SEGMENTATION TECHNOLOGY BASED ON PARTIAL DIFFERENTIAL EQUATION AND CLUSTER ANALYSIS

by

CHAO ZHANG

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the
Applied Mathematics Program
of Delaware State University

DOVER, DELAWARE
December 2017

This thesis is approved by the following members of the Final Oral Review Committee:

Dr. Fengshan Liu, Committee Chairperson, Department of Mathematical Sciences, Delaware State University
Dr. Jinjie Liu, Committee Member, Department of Mathematical Sciences, Delaware State University
Dr. Xiquan Shi, Committee Member, Department of Mathematical Sciences, Delaware State University
Dr. Mukti Rana, Committee Member, Department of Physics and Pre-Engineering, Delaware State University
Dr. Chandra Kambhamettu, External Committee Member, Department of Computer Sciences, University of Delaware

DEDICATION

This dissertation is dedicated to my beloved wife, Ying Yang.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Fengshan Liu, who provided me with valuable encouragement and guidance in every stage of the writing of this thesis. Without his enlightening instruction, impressive kindness and patience, I could not have completed my thesis.

I would also like to extend my sincere thanks to Dr. Yucui Guo for all his kindness and help.

I would also like to thank other members of my thesis committee, Dr. Jinjie Liu, Dr. Xiquan Shi, Dr. Mukti Rana and Dr. Chandra Kambhamettu for taking the time to read this thesis and for giving helpful suggestions.

I wish to thank College of Mathematics, Natural Sciences and Technology and Delaware State University for providing me with the opportunity, resources and the scholarships that made this research possible.

My sincere appreciation also goes to the faculty and staff of Department of Mathematical Sciences.

I wish to acknowledge the US Department of Defense through the grant "Center for Advanced Algorithms" (W911NF-11-2-0046) for their support.

Last but not least, my thanks would go to my beloved parents for their loving considerations and great confidence in me through all these years. I also owe my sincere gratitude to my wife who gave me her help and time during the writing of this thesis.

Image Segmentation Technology Based on Partial Differential Equation and Cluster Analysis

Chao Zhang

Faculty Advisor: Fengshan Liu

ABSTRACT

Image segmentation is to extract objects that are interesting to users that is the most fundamental and important part in image processing. After decades of development and change, image segmentation methods based on various theories have been put forward. The active contour model based on the variational method and the level set method has become an important method of image segmentation. It demonstrates the superiority of partial differential equation image processing method. It mainly uses the idea of dynamic evolution, and has important significance in the study of image segmentation technique. For the study of the image segmentation based on partial differential equation, the numerical calculation good stability in the processing of discretizing the partial differential equations, and achieves high quality image restoration and accurate segmentation of images. In recent years, the active contour model has become a popular research method, and is widely used in edge detection, medical image segmentation and object tracking.

This paper introduces the background, research status, the current development the purpose and the significance of image segmentation based on partial differential equations, the classical active contour model and the related mathematical theory. By studying the problems that exist in the image segmentation on the active contour model we established a new model of medical image segmentation based on information entropy. At the same time, we also studied the method of image segmentation based on clustering analysis (SFCM) and analyzed its own advantages and disadvantages. Furthermore, we established a fuzzy c-means clustering image segmentation model based on gray space. The main work and achievements of this dissertation are as follows:

1. We studied various methods and models of image theory of partial differential equations based on the segmentation, analyzed the theoretical basis and the mathematical principle of these methods, and compared their respective advantages and disadvantages. On this basis, we established two new models, and the numerical results show that the new models have prominent advantages.

2. When there is weak edge, strong noise, or uneven brightness in the image, the traditional active contour model can not achieve the correct segmentation of the target boundaries, especially for medical magnetic resonance images and ultrasonic images. To solve this kind of problems, we propose an active contour model CER based on CV model and RSF model combined with information entropy. By minimizing the energy functional and taking into account the internal and external energy of the target boundaries, the segmentation of the target boundary in the uneven intensity image is realized. The experimental results show that the algorithm can extract the object with weak edge and uneven brightness. At the same time, the RSF modelâĂŹs robustness to noise is enhanced, and its sensitivity to initial contours is also improved.

3. Image segmentation method based on clustering analysis is one of the popular segmentation methods in recent years. The main feature of the traditional FCM clustering algorithm is that it is an unsupervised segmentation method with fast computation speed. However, it is sensitive to outliers. Therefore, we study an improved SFCM algorithm, whose biggest advantage is that it combines the spatial information of the image and reduces the sensitivity of the traditional FCM algorithm to noise. Based on this advantage, we propose to preprocess the image by SFCM, and to take the edge information of the image after clustering as the initial contour of the CER model. This method not only improves the CER model's robustness to noise, but also solves the initial contour selection problem. In addition, the rationality of the initial contour selection is improved and the computational efficiency is improved to a certain extent. A large number of experimental data show that this method can not only accurately segments the strong noise images, but also improves the initial contour selection problem and reduces the computation time, thus enhances the robustness of the CER model.

Keywords: image segmentation; medical image processing; partial differential equations; level set method; information entropy; spatial fuzzy c-means

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMSS	Affine Morphological Scale Space
CER	CV & Entropy & RSF
CV	Chan-Vese
DSC	Dice Similarity Coefficient
FCM	Fuzzy C-Means
FDM	Finite Difference Method
GAC	Geodesic Active Contour
HCM	Hard C-Means
LBF	Local Binary Fitting
MRI	Magnetic Resonance Imaging
MS	Mumford-Shah
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
PS	Piecewise Smooth
RSF	Region-scalable Fitting
SCER	SFCM & CV & Entropy & RSF
SFCM	Spatial Fuzzy C-Means

Chapter 1
INTRODUCTION

1.1 Image and image processing

Image is a reproduction or imitation of the form of picture or graph. It is the human visual perception of the material reproduction. The definition of image is very extensive. Image can be obtained by optical device, such as camera, mirror, telescope and microscope. And image can also be created by using computer and other equipment. Image can be recorded and be kept on paper and film, and materils which is sensitive to optical signal. With the rapid development of the computer industry, more and more images are stored as digital form in the computer. Therefore, in some condition "image" refers to digital image. There are many kinds of images. According to the color type, image can be divided into grayscale images and color images. According to the movement type, images can be classified into dynamic images and static images. According to the type of dimension, images can be classified into two-dimensional images and three-dimensional images. In this dissertation, we focus on still grayscale images especially medical images.

Digital image processing technology is originated in the 1920s. It includes Image acquirement, image compression, image-compositing, image enhancement, Image Restoration, image segmentation and image recognition. Image segmentation is an important research content in image processing. Image segmentation can be considered as subdividing the obtained images into meaningful and non-coincident sub-regions. In image processing, often people interested in certain parts of images. Usually these parts are divided into foreground or background. In fact, the foreground or background is the image area with a specific property in the image. To achieve the recognition and the analysis of the target, it is necessary to separate the target by image segmentation. Successful segmentation is very helpful for the following higher level image application. Therefore, image segmentation is a very important part in entire image processing [23, 76]. In general, the accuracy of segmentation depends on whether the required areas are accurately segmented or not. For example, in the automated inspection of electronic assemblies, interest lies in analyzing images of products with the objective of determining the presence or absence of specific anomalies, such as missing

components or broken connection paths.

Today, the theory of digital image processing has been further developed. Meanwhile, applications of digital image processing is also a powerful tool for many fields, all aspects of economy and human life.

Now we present some definitions of digital images and image segmentations [28]. Let $\Omega \subset R^2$ be a set. then the image $I(x, y)$ can be treated as a mapping that defined at set Ω,

$$I : (x, y) \in \Omega \to V.$$

Let range V be a set of all grey value in the image. It contains all the gray values from the darkest (pure black) to the brightest (white). An image corresponds to a specific mapping rule I. Depends on different mapping, we will get different images. In general, the mapping is continuous. Since the computer can only accept and process discrete data, we usually need to discretize the image in space and in grayscale. If Ω is discretized into M rows and N columns and greyscale is discretized into 256 levels, then the entire image contains $M \times N$ pixels of equal size. The coordinates are $(x_i, y_i), 1 \leq i \leq M, 1 \leq j \leq N$. Meanwhile, each pixel corresponds to a unique gray value $I(x_i, y_i) \in [0, 1, 2, \cdots, 255]$. Now a discrete image can be defined as:

$$I : (x_i, y_i) \in \Omega \to [0, 1, 2, \cdots, 255].$$

The discrete image is called digital image. The various operations performed on digital images are called digital image processing. Image segmentation is to divide image domain Ω into K disjoint subsets $\Omega_1, \Omega_2, \cdots, \Omega_K$. Meanwhile, the divided subsets should meet the following five conditions:

1. Integrity: $U_{i=1}^{K} \Omega_i = \Omega$;
2. Independence: $\forall i \neq j, \Omega_i \cap \Omega_j = \emptyset$;
3. Homogeneity: $P(\Omega_i) = TRUE, i = 1, 2, \cdots, K$;
4. Heterogeneity: $\forall i \neq j, P(\Omega_i \cup \Omega_j) = FALSE$;
5. Connectivity: $\Omega_i (i = 1, 2, \cdots, K)$ is a connected region,

where $P(\cdot)$ is a function that measure pixels consistency. The value of $P(\cdot)$ is TRUE or FALSE. The early image segmentation method is based on the lower level knowledge of the image, such

as images grayscale (Similarity and Continuity), edge and so on. The most typically segmentation methods includes thresholding approach [59], region growing method [1], and edge detection method [10]. With the increasing of the types of images to be segmentation and the improvement of actual needs, more and more high level knowledge has been used in image segmentation. A clustering-based [6, 31, 51] segmentation method was developed. Clustering-based segmentation method is to use certain similarity measure, and to divide the data into multiple subclasses with similar properties, so that the similarity within the subclasses is greater than the similarity between classes. The method is based on the graph theory [8, 63]. The image is mapped to the weighted undirected graph, a pixel is treated as a node. The best image segmentation result is obtained by using the minimum cut standard. And the method is also based on partial differential equations [11, 12, 15, 32, 39, 41]. Those segmentation methods are classical methods for solving complex image segmentation problems. The modern image processing method mainly includes three categories [58]. The first category is signal analysis methods. Signal analysis mainly includes spatial transformation based image processing methods. It is applied to image denoising, edge detection and segmentation. The second category is Stochastic modeling method. The imaging process of digital images is randomness, so the two-dimensional image domain can be treated as a random field. Stochastic modeling method based on the random field theory is to model the image domain and to describe the conditional distribution of image pixels and their neighborhood pixels, and thus to describe the statistical features of the image. Stochastic modeling method is a good method to describe the periodic image texture characteristics. However, the model parameter establishment process is complex and the stability is poor. The third method is the variational principle and partial differential equation method. Image segmentation methods based on PDE is one of the most popular image segmentation methods. It is also the main research object of this dissertation. The method of partial differential equation is mainly mathematical methods used in image processing in spatial domain. Using spatial domain pixel gray value of the first or the second order partial differential equations of characterization in the image area boundary edge character. The partial differential equations have anisotropic diffusion properties. They have different diffusion ability

in different image feature areas. Therefore, it is possible to reconstruct the smooth characteristic region and preserving the edge feature by equation iterations.

1.2 Research background, significance and development of image segmentation based on partial differential equation

Image segmentation is the most important part of image processing. Partial differential equation method for image segmentation is rapidly developed in recent years. The application of partial differential equations in image processing is originated in the 1980s. The partial differential equation method has been developed rapidly and generated a lot of research results. Due to the fact that partial differential equation, as an important branch of mathematics, has formed a strong theoretical system and good numerical methods. The partial differential equation method is developed by the improved Gaussian smoothing image, which can balance the elimination of noise and feature retention.

Partial differential equations applied to image segmentation is based on the following theories. The first one is the theory of multiresolution analysis. In 1983 Witkin [73] and in 1984 Koenderink [35] introduced a rigorous theory of multiresolution analysis in image processing. The theory of multiresolution analysis has become a basis of partial differential equations. The second one is the Euler-Lagrange equation that is derived from the variational model. Through the analysis of the problem and the establishment of the corresponding "energy" model, the image processing problem is transformed into a differentiable functional optimization problem. The corresponding Euler-Lagrange equation is obtained by calculating the extrema of the model by variational model. Then the stable solution is the solution of the image processing problem. The third one is the curve evolution theory and the Level-set method. With this method, the image is considered as a set of horizontal curves or as the surface in high dimensional space the image processing is achieved and by controlling the evolution of the horizontal lines or the surfaces.

Comparing with the traditional image segmentation method, the image segmentation method based on partial differential equations has the following characteristics [19]:

1. The partial differential equation model is a continuous model of direct analysis of images. Usually, the image is filtered by continuous derivatives, and the discrete filtering is expressed as a continuous differential operator. Using discrete grids, the image of the local processing and analysis is easier to achieve;

2. The method based on partial differential equations has a good mathematical basis. Meanwhile, it can provide good stability for image segmentation. In numerical calculations, a more accurate result can be obtained by using the level set method. And it can also solve the problem of the topology change caused by the merging and the separation of the target;

3. The partial differential equation method has stronger local adaptability. The partial differential equation is built on a continuous image model, which makes the image of a pixel value change in the current time only depending on the pixels of a neighborhood.

As a widely applied mathematics, partial differential equation has become a hot topic in its related fields. So far, it already has a lot of research results. The idea of applying partial differential equations to image processing can be traced back to the research of 1965 Gabor [25] and 1977 Jain [30]. However, the substantial leap is thanks to Koenderink [35]and Within [73]. Based on the theory of multi-scale analysis, Koenderink and within proposed the idea that the convolution of the Gaussian function of the signal and the different scale is equivalent to the thermal diffusion equation with the signal as the initial value in 1984 and 1985 respectively. In 1993, L. Alvarez et al. [4] created a rational system through the process of deriving the AMSS (Affine Morphological Scale Space) equation, which marked the formal formation of the subject of image processing based on partial differential equations.

In 1987, Kass et al. proposed the Snake model (Active Contour Models or Parametric model) [32]. It is based on the perspective of dynamics to study the evolution of the curve. The energy function of the contour curve of the region is established by the internal force of the image and the external force constraint generated by the image information. The internal forces and the external forces under the joint action of the continuous movement finally converge the edge of the target. The

boundary of the interested region of the image is obtained, and the precision reaches the sub-pixel level. This model has the following advantages: (1) No matter how the quality of the image is, we can always get a smooth and closed target boundary. (2) It is computationally efficient and suitable for modeling. And it can also extract any shape of the deformation profile. It is suitable for dealing with individual differences in significant structural complex images. With the development of this model, the active contour model is becoming one of the most successful image segmentation method. Later, a lot of different characteristics model has been derived, such as surface flow, Deformable Model, and deformation surface and so on.

In 1988, Osher and Romeny proposed the Level Set method [54, 57]. They improved the Active contour model. The method was successfully applied to the fields of Fluid dynamics, Computer Graphics and so on. In 1993, Caselles et al. [11], and in 1995, Malladi et al. [50] introduced the level set method into the active contour model. They use the Level Set to represent Snakes. The curve was implicitly expressed as a level set of higher dimensions of the surface (level set function). The method is different from the contour model. Level set method does not track the movement of the curve at different times. It is just fixed in the coordinate system to update the level set function at different times to simulate the evolution of the curve. The method effectively solved the problem of the change of topology mechanisms. It also further improved the theory of geometric contour model [7, 47, 48, 66]. In 1997, Caselles et al. proposed the Geodesic Active Contour model [12]. The profile curve of the Geodesic Active Contour model can only move in a single direction, and the segmentation results are heavily dependent on the initial position. Once the contour curve exceeds the image boundary, it will be difficult to return to the correct position. Finally, the segmentation fails. To solve the problem of Geodesic Active Contour model, Paragios combined with Zhu's [78] regional competition idea and proposed the Geodesic Active Area method [55]. The model has nothing to do with the initial position and is entirely dependent on the prior statistics for the image area and the boundary. However, when the image is divided into multiple regions, it is necessary to complete the coupling of the multiple deformation models, and the calculation process is complicated. In 2004, Chan et al. analyzed that the uniqueness of the existence solution

conditions of the Total Variation model [14], which is established in L^1 space. In the same year, Chan et al. proposed an algorithm and obtained a global minimum [13]. This algorithm can transform nonconvex energy functional in image processing into convex energy functional. The authors used the standard convex functional minimum solution to find the global minimum.

In the recent twenty years, the geometric activity profile model attracts a wide range of attention its advantages. Thus, a lot of classic segmentation models [5, 12, 15, 39, 41, 52, 71] were proposed. According to the energy function defined by the edge map, the Active Contour model can be divided into three categories [36]: the Active Contour model based on edge, the Active Contour model based on area and the mixed Active Contour model. The edge-based active contour model is a Geodesic Active Contour. The curve evolution is stopped by the edge indicator function of the image gradient. It has the advantage of detecting multiple different areas at the same time. However, this model has the disadvantages of being sensitive to noise and it is also needed to manually set the initial contour curve. This model has limited applications. The region-based active contour model utilizes the global information of the image area, which can reduce the problem of sensitivity to noise. But the calculation amount is huge and the calculation costs a lot of time. The mixed active contour model combines the edge and area information. For different initialization, we get different segmentation results. In this article, we introduce the details of those classical models in 1.3. Those models have their own advantages and disadvantages. Therefore, it is necessary to select the corresponding model according to the specific image and the actual needs.

The acquirement of medical images is different from the acquirement of ordinary optical images. Most medical images have the characteristics of ambiguity and uneven grayness [16]: (1) Medical image has fuzziness of grey scale. The noise and the edge of the ultrasound image have large grayscale difference. Especially because of the nonuniformity of imaging organ or tissue structure and the interference of acoustic signals, unique spots exist in ultrasound images. (2) Local body effect. Often the boundary of an image, contains both boundary and the object. Due to some of the diseased tissue intrusion into the surrounding tissue, the edge cannot be clearly defined. (3) Uncertainty. In general, the structures that do not have at a normal tissue or region present in the

case of a lesion, such as the spurs on the surface of the bone. And the appearance creates difficulties for building the model. Aimed at the weakness of medical image, we found that in the process of segmentation, we need to eliminate the noise in the image to get the target area. Therefore, the elimination of the noise of medical images has a special significance for segmentation. In this paper, we mainly study medical image segmentation using Partial Differential Equation method, because the Partial Differential Equation method has a good balance between eliminating noise and keeping the characteristics of target characteristics.

In the recent years, the Partial Differential Equation method has widely used in different kinds of image processing area, such as image denoising, image magnification, image inpainting and segmentation. Not only because the model established by partial differential equations is more intuitive and easy to understand, but also because partial differential equations are easy to be integrated with the classical segmentation theory. So, it is easy to create new model.

With the communications during international conferences and massive publications on the famous international journals, the image segmentation methods based on partial differential equation have been developed quickly. At present, the United States has invested a lot of manpower and material resources to study the subject. Major research centers are established at Brown University, UCLA, Florida University, and other places. At these research centers, researchers not only work on theory, but also directly work on specific image processing projects. In addition, the study of image segmentation technology based on partial differential equation, promotes interdisciplinary research, such as biomedicine and computer vision. The methods of using partial differential equations, not only solve many problems in image segmentation, but also promote the development of partial differential equation theory. Therefore, the research of this subject has great application value and very important scientific significance.

1.3 Segmentations based on partial differential equation

Many of the physical laws in nature usually can be described by mathematical language. Partial differential equations reflect the restrictive relation between the unknown variables and the derivatives with respect to the time variable and the derivative with respect to the spatial variables.

The segmentation method based on partial differential equation has the following advantages:

1. Partial differential equation method is a method that directly analyzing on image. It is implemented on the continuous condition. Thus the resulting image edge can be obtained with higher precision;

2. It has an important effect on the shape analysis and the recognition of an object to use a smooth closed curve to represent the edges of an object;

3. The partial differential equation method can make the fusion of various image processing methods more effectively;

4. The partial differential equation method is convenient for establishing of various flexible mathematical models. The mature theory of numerical analysis and partial differential equations can be used to get good image quality and stable results.

Partial differential equation (PDE) image segmentation is a nonlinear image segmentation method which was developed in 1980s. Its basic idea is to make the continuous mathematical model of the image to follow with a partial differential equation, and the corresponding segmentation result is obtained when the stable solution of the equation is obtained. The most important advantage of this technique is that the curve is continuous and smooth in the process of evolution. The classical model of image segmentation based on PDE is the active contour model which is based on the deformable model. The basic idea of Deformable model is that the energy functional of the region contour curve is established by using the internal force and external force constraint of the image edge. The Euler equation is obtained using variational method to minimize the energy function. This equation is generally a nonlinear partial differential equation. Its stable solution corresponds to the segmentation result of the image. More precisely, the basic principle of the deformable model is that the curve moves and changes under the action of the driving forces including the internal and the external forces, in which the internal force is based on the geometric characteristics of the curve and keeps the smoothness of the curve; external force drives curve to move to the

boundary of the object through prior information of the target and the image information. Under the action of the internal force and the external force, the curve moves to the boundary of the target under the smoothness condition. The active contour model makes deformation of the curve through minimizing the energy. In the active contour model, an energy function of the curve is defined firstly, and then the given energy function is minimized to make the curve evolving to the boundary of the target. In the recent ten years, the image segmentation technology that based on partial differential equations has been greatly developed. And it is successfully applied to the medical and other areas.

1.3.1 Segmentation models based on edges

1. Snake model

The earliest variational mode applied to image segmentation is the Snake model introduced by Kass et al. [32]. The idea of Snake model is derived from the physical Deformable Model. It is assumed that the edge of the object is elastic and can be deformed under the combined action of the internal force and the external force. It has milestone significance of image segmentation field. The model is to get the following energy function by constructing a parametric energy curve $C(p)$, $p \in [0, 1]$:

$$E(C(p)) = \alpha \int_0^1 |C'(p)|^2 dp + \beta \int_0^1 |C''(p)|^2 dp - \lambda \int_0^1 |\nabla I(C(p))| dp. \tag{1.1}$$

The basic idea of Snake model is that the curve $C(p)$ is evolved along the fastest decreasing direction of the function. When it reaches the target boundary, it will stop at the local minimum point of the function.

2. Geodesic Active Contour (GAC) [12]

The Geographic Activity Contour Model is a generalization of the Snake model. Caselles et al assumed that the rigid coefficient $\beta = 0$ in Equation (1.1) and the length in Euclidean space is replaced by the length of authority in Riemannian space [33, 34].

An active contour model which is independent of the curve parameter is proposed. It is named as the Geodesic Active Contour model [12]. The essence of this model is to use the surge of curvature and the force generated by the monotone decreasing edge indicator function $g(\nabla I)$ to make the evolution curve C to stop and to stay on the edge of the target object, where

$$g : [0, +\infty] \to R^+, g(0) = 1, \lim_{s\to\infty} g(s) = 0.$$

In image segmentation, we usually choose function $g(r) = 1/(1 + r^2)$. By selecting the appropriate weighting factors α and λ, the energy functional of this model is equal to solve the Geodesic Curves in Riemannian Spaces [33, 34],

$$\min \int_0^1 g(|\nabla I(C(p))|)|C'(p)|dp. \tag{1.2}$$

Let $ds = |C'(p)|dp$, then the energy functional becomes to $\int_0^1 g(|\nabla I(C(p))|)ds$ for Equation (1.2), where $C(p)$ is the arc length of a curve. Then the above equation can be interpreted as the weighted arc length in the Euclidean space, and the weight is the Riemannian distance $g(|\nabla I(C(p))|)$. Riemannian distance is about the image information, while the Euclidean distance is about the distance between the two points in space.

Using variational method and gradient descent flow, we get the energy functional (1.2) corresponding curve evolution equation:

$$\frac{\partial C}{\partial t} = -g(I)kN - (\nabla g \cdot N)N, \tag{1.3}$$

where k is the curvature, and N is the outward unit normal vector.

From above, Geodesic Active Contour model has a serious limitation. Since the curve with a curvature less than zero will cause the curve move to outward, when the target to be segmented contains a deeper recessed boundary, the curve evolution of the GAC model may cease in a local minimum, not at the edge of the target. To solve this problem, we add a Ballon Force in Equation (1.3) [18]. Thus the direction of the evolution curve always points to the inside of the curve, and is

controlled by $g(|\nabla I|)$. The improved model is shown in below:

$$\frac{\partial C}{\partial t} = g(I)(k+m)N - (\nabla g \cdot N)N, \tag{1.4}$$

where m is a fixed constant.

Meanwhile, V.Caselles et al. proposed a similar equation to (1.4). But it is also based on the idea of the evolution of the average curvature of the movement, and uses level set function. It deals with changes in the topology freely. The model can be presented as the following:

$$\begin{cases} \frac{\partial \phi}{\partial t} = g(|\nabla G_\sigma| * I)\left(div(\frac{\nabla \phi}{|\nabla \phi|}\right) + \nu)|\nabla \phi| \\ \phi(0,x,y) = \phi_0(x,y), \end{cases} \tag{1.5}$$

where $v \geq 0$ is a coefficient, and G_σ is the σ standard deviation of Gaussian kernel function.

As the earliest Geometric Active Contour Model, the basic idea of the GAC model image segmentation is to couple the image data and the curve of the deformation speed, and to make the curve evolution to eventually stop at the target edge position. To maintain the validity of the numerical algorithm and the stability of the level set function in the evolution process, we choose an initial and periodically reinitialize the Signed Distance Function. Since the initial signed distance function needs to be defined by an initial curve in the image area, we choose a suitable initial profile, usually required to enclose the target area or to use multiple initial contours. Therefore, GAC model has the following disadvantage, such as implement complex, great amount calculation and initialization [27].

3. No-need reinitialization model

Aimed at the reinitialization problem in the GAC model, Li et al. proposed a level set model that does not need reinitialization [41]. It is named as the Distance regularized level set model [38]. This model has improved the shortcomings of the GAC model. The model arises from the perspective of functional, with the addition of an internal energy functional $P(\phi)$. This functional is called no-need reinitialization item or distance regularization item, which is used to measure the

deviation between the level set function and the Signed distance function. Thus, this model solves the periodic initialization problem of the level set function. The energy functional is:

$$\begin{aligned} E(\phi) &= \mu P(\phi) + \lambda L_g(\phi) + \nu A_g(\phi) \\ &= \frac{1}{2}\mu \int_\Omega (|\nabla\phi| - 1)^2 dxdy + \lambda \int_\Omega g\delta(\phi)|\nabla\phi| dxdy + \nu \int_\Omega gH(-\phi) dxdy, \end{aligned} \tag{1.6}$$

where, $\mu, \lambda > 0$, ν is constant, g is edge indicator function, $L_g(\phi)$ and $A_g(\phi)$ are the weight lengths of the zero level set ($\phi = 0$) and the weighted area, respectively. They are collectively called the external energy. Their role is to force the level set to move toward to the target boundary. In addition, $H(\cdot)$ and $\delta(\cdot)$ are one-dimension Heaviside function and the one-dimension Dirac function respectively.

$$H(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}, \ \delta(z) = \frac{d}{dz}H(z).$$

Use the variational principle [5] and the steepest decent method, we can get the minimized energy functional $E(\phi)$ which is corresponding to the gradient descent flow equation:

$$\frac{\partial\phi}{\partial t} = \mu\left(\triangle\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right) + \lambda\delta(\phi)\, div\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + \nu g\delta(\phi). \tag{1.7}$$

To solve Equation (1.7), we use the regularized functions $\delta_\varepsilon(\phi)$ to approximate functions $\delta(\phi)$:

$$\delta_\varepsilon(\phi) = \begin{cases} \frac{1}{2\varepsilon}\left[1 + \cos\left(\frac{\pi x}{\varepsilon}\right)\right], & |x| \leq \varepsilon \\ 0, & |x| > \varepsilon. \end{cases} \tag{1.8}$$

Compared with the traditional level set method, the no need reinitialization level set model has the following advantages [72]:

1. It is not necessary to periodically initialize the level set function in the evolutionary process, which saves a significant amount of time. Meanwhile, it avoids a series of problems caused by repeated initialization of the level set functions, such as the complexity of algorithm.

2. Due to the introduction of the distance regularization $P(\phi)$, in the evolution process, the level

set function can always maintain the stability form the numerical calculation. In addition, due to the use of the simple finite difference scheme and the significantly larger time step, the speed of evolution of the level set was greatly improved.

3. The initialization level set function does not have to be a Signed Distance Function. It can be more flexible defined as a simple piecewise constant function. Comparing with the signed distance function, this initialization method is widely used in practical problems because of its simplicity, effectiveness and flexibility [38, 41]. But the model still has the disadvantage in defining the initial contour. It is means that it still cannot avoid the problems caused by the contour initialization, such as how and where to define an initialization contour define the initialization contour.

1.3.2 Segmentation models based on area

Compared with the segmentation models based on edge, the area based segmentation model defines the energy functional by using internal and external or global information of the active contour and it does not use the edge detection operator. Therefore, it is more suitable for image segmentation for the increased-noise images.

1. Mumford-Shah (MS) model

In the 1980s, Mumford and Shah proposed an effective active contour model which uses curve evolution method for solving image segmentation or target detection problems. This model is named as the Mumford-Shah (MS) model. The model created an important field of image segmentation which is based on edge partial differential equations. MS model uses a specific law of a closed curve or surface deformation to define an energy function of a curve or surface:

$$E^{MS}(u, C) = \mu \int_{\Omega \setminus C} |\nabla u|^2 dxdy + \int_{\Omega} (u - I)^2 dxdy + \nu |C|, \tag{1.9}$$

where $I : \Omega \subset R^2 \to R$ is the image with noise, Ω is the area of image I, $\mu, \nu > 0$ are the fixed weighting parameters, C is a unknown curve, $|C|$ is the length of the curve, and u is the Piecewise

Smooth function which is used to approximate the grayscale distribution of image I, and to allow the discontinuity when acrossing the curve C. The first item is the regularization term which keeps the function u not changing dramatically in each sub-region, and the function u as smooth as possible. The second term is used to measure the degree of similarity between the function u and the image I, add to keep the function u as close to the image I as possible. The third item makes the curve C short and smooth.

The MS model transforms the image segmentation problem into the optimization problem. When minimizing the energy functional $E^{MS}(u, C)$, we get the piecewise smooth function u and its boundary C at the same time, where u is the image without noise and the curve C is the boundary of the image segmentation area. Although the MS model has a good theoretical basis from mathematics, but this model contains an unknown discontinuous set, which makes the functional minimization and the numerical solution of the model very difficult. Therefore, the implementation of the MS model becomes more complex, computationally intensive, and difficult to apply to practical image processing.

2. Chan-Vese (CV) model

From the above MS model, we know that the model has some defects in the practical applications. Since the image segmentation model depends on the gradient as a stop condition. Aimed at this problem, Chan and Vese [15] proposed the Chan-Vese model which uses the variational level set method without using the gradient as the stopping condition. CV model effectively improves the defects of the MS model, which greatly promotes the application of MS model with a great influence [67, 71].

The basic idea of the CV model is to search a closed curve C. This closed curve divides the image area into $in(C)$ and $out(C)$, i.e. the target and the background. The closed curve C maximizes the difference average gray value of the image in $in(C)$ and the image in $out(C)$. This closed curve can be seen as the contour of the target area which is the edge of the image. Given an image $I : \Omega \subset R^2 \to R$ for segmentation. C is an evolving curve. The CV model does the

segmentation by minimizing the following energy functional:

$$\begin{aligned} E^{CV}(c_1, c_2, C) =& \nu \cdot Length(C) + \mu \cdot Area\,(inside(C)) \\ &+ \lambda_1 \int_{inside(C)} |I - c_1|^2 dxdy + \lambda_2 \int_{outside(C)} |I - c_2|^2 dxdy, \end{aligned} \tag{1.10}$$

where $\nu, \mu \geq 0, \lambda_1, \lambda_2 > 0$ are weight coefficients, in general, $\nu = 0$, c_1 and c_2 are defined as the average gray values of the image in area of $in(C)$ and $out(C)$ respectively.

Assume $\phi(x, y, t)$ is the level set function, $\delta(x)$ is the function of Dirac, $H(x)$ is the function of Heaviside. For efficient numerical calculation, the regular Heaviside function $H_\varepsilon(z)$ is usually used to approximate $H(z)$:

$$H_\varepsilon(z) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{z}{\varepsilon}\right)\right),$$

where ε is a minimum value selected for numerical calculation. Then the energy functional (1.10) can be expressed in level set as the following:

$$\begin{aligned} E_\varepsilon^{CV}(c_1, c_2, \phi) =& \nu \int_\Omega \delta_\varepsilon(x)|\nabla\phi| dxdy + \mu \int_\Omega H_\varepsilon(\phi) dxdy \\ &+ \lambda_1 \int_\Omega |I - c_1|^2 H_\varepsilon(\phi) dxdy + \lambda_2 \int_\Omega |I - c_2|^2 (1 - H_\varepsilon(\phi)) dxdy. \end{aligned} \tag{1.11}$$

By using the variational principle as the steepest descent flow, we derive the Euler-Lagrange equation which controls the evolution of level set:

$$\frac{\partial\phi}{\partial t} = \nu\delta_\varepsilon(\phi) div\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \delta_\varepsilon(\phi)\left(\lambda_2\left(1 - c_2\right)^2 - \lambda_1\left(1 - c_1\right)^2\right). \tag{1.12}$$

The expression of the average gray value of c_1 and c_2 is shown in below:

$$c_1 = \frac{\int_\Omega I(x,y) H_\varepsilon\left(\phi(x,y)\right) dxdy}{\int_\Omega H_\varepsilon\left(\phi(x,y)\right) dxdy}, c_2 = \frac{\int_\Omega I(x,y)\left(1 - H_\varepsilon\left(\phi(x,y)\right)\right) dxdy}{\int_\Omega \left(1 - H_\varepsilon\left(\phi(x,y)\right)\right) dxdy}. \tag{1.13}$$

Chan and Vese's Chan-Vese model is a special case of the Mumford-Shah functional. Since the existence of the minimum value of the MS functional has already been proved, the minimum value of the energy functional $E^{CV}(c_1, c_2, C)$ must also exist. The main advantage of Chan-Vese model is that the initial curve does not need to be completely located at inside or outside of homogeneous region. Meanwhile, CV model does not use gradient information, so it has good global optimization characteristics. And it also reduced the impact of noise on the image. However, the CV model only uses global area information to optimize c_1 and c_2. It does not take care of the image grey values changed at the target and background area.

Therefore, the CV model cannot segment uneven grayscale images. It can be used for the images which contains two different mean regions of objects and backgrounds. Thus the scope of its application is limited. In addition, to ensure the stability of the numerical solution of the level set [22], the level set evolution process requires periodic initialization signed distance function which costs a lot of computation time.

3. Region-scalable Fitting (RSF) model

To solve the problem of segment with uneven grayscale images, Li proposed a local region level set model which is based on variational framework. It was named as RSF model [40] or LBF model [39]. This model firstly defined two Region-scalable Fitting Energy $f_1(x)$ and $f_2(x)$, not using c_1 and c_2, as the local averages of the gray values of the image on the two sides of the contour line. It can be obtained by minimizing the energy functional. For any point $x \in \Omega$, the energy functional of RSF is as follows:

$$\begin{aligned} E^{RSF}(f_1, f_2, \phi) =& \lambda_1 \int_{in(C)} K_\sigma(x-y)|I(y) - f_1(x)|^2 dy \\ &+ \lambda_2 \int_{out(C)} K_\sigma(x-y)|I(y) - f_2(x)|^2 dy, \end{aligned} \tag{1.14}$$

where $K(x)$ is a kernel function that satisfies the local nature. It is usually selected as a Gaussian kernel function. Therefore, $f_1(x)$ and $f_2(x)$ are mainly determined by the gray value near x with local characteristics. The RSF model transforms two value global fitting energy functional of the

CV model into the local binary fitting energy with the Gaussian function as the kernel function. The RSF model has effectively solved the problem that CV model cannot deal with gray scale image. However, RSF model is very sensitive to the selection of the initial contour. We need to choose a suitable initial contour manually. Because of the improvement of the RSF model and the existence of CV model is the focus of this dissertation we will introduce the details of the RSF model in Chapter 3.

1.3.3 The existing problem of image segmentation method based on partial differential equation

Although in recent years, the image segmentation methods based on partial differential equations have been developed rapidly. Even though we already have a lot of research results, there are still some issues need to be solved. The problems are shown in below:

1. Evolution speed problem

As described earlier in this dissertation, most models such as GAV, MS and CV have evolutionary speed problems. This is mainly because of that most theoretical results of level set are based on the smoothness of the surface. However, in practical applications, the singular value surface often appears in the evolution process. In order to ensure the smoothness conditions and the stableness and effectiveness of the level set evolution, it is necessary to periodically repeat the reinitialization of the level set function. This process costs a lot of calculation time. To solve this problem, researchers proposed many new methods. But these methods are limited to applications, since they are just the improvement to the specific needs of each mode. Thus these improved acceleration algorithms is lack of extensive adaptability. Therefore, it is necessary to find an algorithm that does not depend on certain specific environment, but enhances the scope of adaptation, and effectively improves the speed of segmentation.

2. Contour initialization problem

Similar to the sensitiveness of the RSF model in selecting the initial contour, the essence of most models is to find the numerical solution of partial differential equations with initial conditions and

boundary conditions. The segmentation results are often dependent on the given initial conditions. Most of the existing models need to manually define the initial contour. Therefore, it is created the problems of how to initialize and where to initialize the contours. Traditional level set segmentation model GAV needs to define the initialization level set as a signed distance function. No-initialization model needs to define a piecewise constant function. Although they need to define different types of functions, they need to define the initial contour in advance. Therefore, solving the problem of contour initialization is a hot topic in the recent years.

3. Segmentation of complicated background images

Most of the existing models can only handle simple images. It means that there is a clear boundary between the target and the background. Because of the imaging system and many uncertainty factors in real world, real images are often more complex. For example, medical images contain a lot of noise, have weak edge in the image and intensity inhomogeneity. Most of the existing research involved less in complex image. Therefore, applying the partial differential equations to image filtering, image restoring and image deniosing becomes a major research topic in the recent year. We hope that these preprocessing can be applied to image segmentation to effectively solve the problem of images with complex background.

1.4 Main work and the structure arrangement of this paper

1.4.1 Main work of this dissertation

Image segmentation is an important part of image processing, and is also a difficult part. The key issues of segmentation are manifested in two aspects: (1) Image segmentation problem involves much uncertainty and the solution is often not unique. Image contain a wealth of information, such as edge, color, and texture. But it is difficult to use a unified segmentation method. Therefore, although the partial differential equation method has many advantages, there is still no uniform segmentation algorithm. Meanwhile, a lot of research is still in its infancy, particularly in the areas related to the sensitivity of contour initialization, the robustness to strong noise and the evolutionary

speed. In this dissertation, we aim to solve these problems by combining the local entropy and the clustering in the research. The main work is shown as below:

1.) By combining with the local entropy, we propose CER model for improving the famous RSF (LBF) model and CV model. The CER model enhanced the robustness of the original model to contour initialization and to strong noise, and in the meanwhile increased the velocity of evolution.

The RSF model is a recently-proposed variational segmentation model that effectively segments the grayscale image, such as magnetic resonance images of the brains. For any point $x \in \Omega$, the energy functional is defined as:

$$\begin{aligned} E^{RSF}(f_1, f_2, \phi) =& \lambda_1 \int_{in(C)} K_\sigma(x-y)|I(y) - f_1(x)|^2 dy \\ &+ \lambda_2 \int_{out(C)} K_\sigma(x-y)|I(y) - f_2(x)|^2 dy. \end{aligned}$$

In RSF model, $f_1(x)$ and $f_2(x)$ are the fitting energy of the image at point x. It is a local energy associated with variance. Therefore, RSF model is based on the local information of the image. It is precisely due to the introduction of this local energy that the RSF model has a good segmentation result for gray uneven images. However, the model is more sensitive to the contour initialization with poor noise immunity.

The CV model is a classical geometric active contour model that is based on the region. Its energy functional is shown as below:

$$\begin{aligned} E^{CV}(c_1, c_2, C) =& \nu \cdot Length(C) + \mu \cdot Area(inside(C)) + \lambda_1 \int_{inside(C)} |I - c_1|^2 dxdy \\ &+ \lambda_2 \int_{outside(C)} |I - c_2|^2 dxdy. \end{aligned}$$

For CV model, c_1 and c_2 are the average gray values of the image in the internal and the external contour curves, respectively. They are all global variables. Therefore, CV model is a segmentation model based on global information. Although the CV model has advantage in anti-noising and simplicity for calculation, it cannot be used to segment the image of uneven gray.

Based on the above analysis, this dissertation considers combining the above two models, utilizing their respective advantages and disadvantages, to achieve complementarity. It's energy functional is shown as below:

$$E^{NLGF}(c, f_1, f_2) = \omega E^{GIF} + (1-\omega)E^{NLIF} + \nu L(\phi) + \mu P(\phi).$$

Meanwhile, we use the local entropy $E_I(x)$ to define the below weighted energy functional:

$$E(c, f_1, f_2) = \int_{\Omega} E_I(x) E_x(c, f_1(x), f_2(x)) dx.$$

This dissertation uses the finite difference algorithm. The proposed CER model is simulated by MAELAB programming. The result shows that the CER model has a very good result. The CER model improved the noise sensitivity and the poor noise immunity of the RSF model. Meanwhile, the CER model also resolved the problem of the CV model, in failing the segmentation of images with uneven gray. CER improved the speed of curve evolution.

2.) The traditional FCM clustering algorithm is a fast image segmentation method with no monitoring. However, it cannot accurately segment the image with strong noise or image with weak edge. Therefore, in this dissertation, we adopted SFCM clustering algorithm that combines image gray features and the association of the neighbor pixel. This method takes into account the image space information, so it has strong robustness to noise. This dissertation proposed to use SFCM to pre-process the image and at the same time extract the image edge information of the segmentation result. We use the extracted edge information as the initial model of the CER model, which improves the problem in choosing the initial model. The accurate choice of the initial model also increases the evolution speed of the CER model, thereby enhances the calculation efficiency. More importantly, thanks to the SFCM clustering algorithm's strong robustness to noise, the SCER model enjoys strong robustness than the CER model, thus segments the image more accurately.

1.4.2 Structure arrangement

The main contributions of the dissertation are shown as below:

Chapter 1: Introduction. Firstly, we introduce the basic concept of image and image processing, background, significance, and development of Partial Differential Equation Image Segmentation. Secondly, we focus on some classical Partial Differential Equation image segmentation models. Thirdly, we introduce our main research work.

Chapter 2: This chapter introduces mathematical principle of image segmentation method based on partial differential equations, including partial differential equations (PDE) and the basic knowledge of the variational method; plane curve and the characteristics of mathematical representation and curve evolution and the variational level set method and so on.

Chapter 3: Since the Chan-Vese and RSF models have the problems of sensitivity to the initialization contours, poor noise immunity and inability to segment the grayscale images, we proposed the CER model using a local entropy weight.

Chapter 4: Based on the SFCM method, we improved the CER model and proposed SCER model.

Chapter 5: Conclusion. This chapter summarizes our main work and our research results, and proposed the future development of the application partial differential equation to image segmentation.

Chapter 2
FUNDAMENTALS OF MATHEMATICS

It has been pointed out in the first chapter that the main work of this dissertation is to use partial differential equations for image segmentation. This chapter introduces the mathematical principle of this work, including the partial differential equation (PDE) and the basic knowledge of the variational method, plane curve, mathematical representations of the characteristics and curve evolution, the variational level set method and so on.

2.1 Introduction to partial differential equations and variational methods

The basic idea of image processing method based on partial differential equations is that in the continuous image model, we make the image the follow a designated partial differential equation, and solving the partial differential equation is to obtain the processing results. Therefore, the first step of the image processing method based on partial differential equations is to establish a partial differential equation that meets the requirements, namely, to establish a mathematical model. The commonly used modeling methods are: (1) Establishing the "energy" functional by variational method, through which the Euler-Lagrange equation [29] is the needed partial differential equation; (2) Carrying out analogy among the image changes expected to achieve with a certain physical process (such as, analogy among smoothing the image and impurity diffusion) to establish the corresponding partial differential equation. This dissertation mainly adopts the first method to establish models. We start with partial differential equations and variational methods to discuss the mathematical basis of this paper.

2.1.1 Partial differential equation and its definite solution condition

1. Partial differential equation [58]

We know that an equation involving derivatives or differentials with unknown functions is called differential equation. When the unknown function is a univariate function, the differential

equation is called an ordinary differential equation (ODE). When the unknown function is a multivariate function, the differential equation is called a partial differential equation (PDE). The partial differential equation of a space-time differentiable function $u(x, y, z, t)$ is one mathematical relation between u and its partial derivatives,

$$F\left(x, y, z, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}, \frac{\partial u}{\partial t}, \frac{\partial^2 u}{\partial x_2}, \frac{\partial^2 u}{\partial y_2}, \frac{\partial^2 u}{\partial z_2}, \frac{\partial^2}{\partial t_2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y \partial z}, \frac{\partial^2 u}{\partial z \partial x}, \ldots\right) = 0, \tag{2.1.1}$$

where F represents the functional relation, (x, y, z) is called the space independent variable, t is called the time independent variable, and u is the unknown function of (x, y, z) and t. The highest order of partial derivatives in the equation is called the order of the partial differential equation, for example, first order partial differential equations:

$$F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0; \tag{2.1.2}$$

second order partial differential equation:

$$F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x_2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y_2}\right) = 0, \tag{2.1.3}$$

and so on. When F is linear, the equation is called linear equation. The general form of the second order linear partial differential equations of the two independent variables is

$$\begin{aligned} A(x, y)u_{xx} + 2B(x, y)u_{xy} + C(x, y)u_{yy} + D(x, y)u_x \\ + E(x, y)u_y + F(x, y)u = f(x, y). \end{aligned} \tag{2.1.4}$$

Linear partial differential equations contain only the first power of unknown function and its derivatives, and the partial differential equations that do not satisfy this condition are collectively referred to as nonlinear partial differential equations.

The classical linear partial differential equations include:

(1) Wave Equation

$$\frac{\partial^2 u}{\partial t^2} - a^2 \nabla^2 u = 0, \tag{2.1.5}$$

where ∇^2 is called Laplace, and in the 3D rectangular coordinate system $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Wave equations mainly describe the propagation of waves, such as vibration of strings, vibrations of thin films, acoustics and propagation of electromagnetic waves, etc.

(2) Heat Conduction Equation

$$\frac{\partial u}{\partial t} - k \nabla^2 u = 0, \tag{2.1.6}$$

where k is the coefficient of thermal conductivity. The equation describes heat conduction, quantum flow, and the growth and diffusion processes in biology; it also describes vortices, diffusion and so on.

(3) Laplace Equation

$$\nabla^2 u = 0. \tag{2.1.7}$$

The equation is used to describe the potential, gravitational field, translational motion of the elastic film, fluid velocity field and the temperature distribution of the steady heat conduction in the passive electrostatic field.

(4) Poisson Equation: the non-homogeneous form of Laplace equation is called Poisson equation, which represents the phenomenon of Laplace equation in active or leaky conditions, and the general form of Poisson equation is as follows

$$\nabla^2 u = f(x, y). \tag{2.1.8}$$

(5) Helmholtz Equation:

$$\nabla^2 u + \lambda u = 0, \tag{2.1.9}$$

where λ is the constant and this equation represents sound radiation field in acoustics.

(6) Telegraph Equation: the telegraph equation is shown as follows:

$$\frac{\partial^2 u}{\partial t^2} + a\frac{\partial u}{\partial t} + bu = \frac{\partial^2 u}{\partial x^2}, \tag{2.1.10}$$

where a and b are constants that describe the law of propagation of electrical signals in the cable as well as the propagation of blood pressure waves in the arteries.

The second-order linear partial differential equation (2.1.4) has mathematical classification besides the above physical division. In equation (2.1.4), we regard $\Delta(x,y) = B^2(x,y) - A(x,y)C(x,y)$ as a discriminant. If $\Delta(x,y) > 0$, in $D \in R^2$ then we say (2.1.4) within D is a hyperbolic equation; if $\Delta(x,y) \leq 0$, in $D \in R^2$ we say that (2.1.4) in D is parabolic a equation; if $\Delta(x,y) < 0$, in $D \in R^2$ then we say that (2.1.4) in D is an elliptic equation.

In most cases, the general solutions of partial differential equations contains arbitrary functions and have certain uncertainties. Thus they are not very useful. Therefore some definite solution conditions must be added to obtain a particular solution. The definite solution conditions are divided into the initial condition and/or boundary condition. The partial differential equation and these definite solution conditions together form the definite solution problem.

The initial condition, also known as Cauchy condition, is the initial displacement and initial velocity condition

$$u(x,y,z,t)|_{t=0} = \varphi(x,y,z) \ \ and \ \ u_t(x,y,z,t)|_{t=0} = \psi(x,y,z), \tag{2.1.11}$$

where φ and ψ are known functions. Boundary conditions can be categorized as the following three conditions:

The first kind of boundary condition, also called di Reeh Lai (Dirichlet), directly gives the distribution of the physical quantity u on the boundary,

$$u(M,t)|_{M\in\partial\Omega} = f_1(M,t). \tag{2.1.12}$$

The second kind of boundary condition, also known as Neumann condition, gives the distribution of the gradient of the physical quantity u on the boundary

$$\left.\frac{\partial u}{\partial n}\right|_{M\in\partial\Omega} = f_2(M,t), \tag{2.1.13}$$

where $\vec{n}$ is the outer normal direction on the boundary $\partial\Omega$.

The third kind of boundary condition, also called Robin condition, gives the linear relation between the physical quantity u and its normal derivative on the boundary

$$\left.\left(u+\sigma\frac{\partial u}{\partial n}\right)\right|_{M\in\partial\Omega} = f_3(M,t) \tag{2.1.14}$$

where σ is constant.

In fact, the three types of boundary conditions can be written uniformly

$$\left.\left(\alpha u+\beta\frac{\partial u}{\partial n}\right)\right|_{M\in\partial\Omega} = f(M,t). \tag{2.1.15}$$

The function f in equation (2.1.15) and the functions $f_i(i=1,2,3)$ in equation (2.1.12), (2.1.13) and (2.1.14) are known functions

2. Nonlinear partial differential equations

A nonlinear partial differential equation that can not be written as equation (2.1.4) form, i.e., at least one the powers of the unknown equations or its partial derivative is not one, is called a nonlinear partial differential equation. The general form is shown as in equation (2.1.1) or (2.1.2).

Typical nonlinear partial differential equations include KdV equation:

$$u_t + uu_x - \mu u_{xxx} = 0; \tag{2.1.16}$$

Sine-Gordon equation

$$u_{xx} - u_{tt} = \sin u; \tag{2.1.17}$$

And nonlinear Schrodinger equation

$$iu_t + u_{xx} + \beta u|u|^2 = 0 \tag{2.1.18}$$

and so on

2.1.2 Calculus of variations–Euler-Lagrange equation

1. Functional and functional extreme value

Functional is the generalization of the concept of function. We will give an example to explain the problem.

We look into the Branchistochrone problem in mechanics. As shown in Figure 2.1, A and B are two points neither in the same vertical line nor in the same horizon line. The goal is to look for such a curve between A and B, when a particle under the action of gravity moves without friction along this curve from A to B, such that the time required T is minimum. Since there are countless lines that connect A and B, this is a problem about extreme value. Here is a mathematical expression for this problem. By kinematic knowledge, we know that the velocity of a particle is $\frac{ds}{dt} = \sqrt{2gy}$. So the time needed to slip from A to B is

$$T = \int_{t(A)}^{t(B)} dt = \int_A^B \frac{ds}{\sqrt{2gy}} = \int_A^B \frac{1+y'^2}{\sqrt{2gy}} dx, \quad namely \quad T[y(x)] = \int_A^B \frac{1+y'^2}{\sqrt{2gy}} dx. \tag{2.1.19}$$

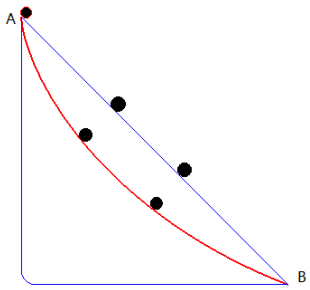


Figure 2.1: Branchistochrone.

We call the above T as the functional of $u(x)$, and all such functions $u(x)$ as a desirable function class, called the definition domain of functional $T[u(x)]$. To make it simply, functional is the function of a function (not the meaning of a composite function). Let C be a set of functions, and E be a set of real or complex numbers. If each element of C has an corresponding element J in E, then we say J is the functional of u(x), and

$$J = J[u(x)]. \tag{2.1.20}$$

It must be noted that functional is different from the normally stated function. The factor determining the value of the usual function is the value of the independent variable, while the factor determining the functional value is the shape of the function. For example, the change of functional T in the above example is caused by the change of the function itself (that is, the different curves from A to B). Its value depends neither on a value nor on a value , but rather on the functional relationship in the entire C set.

Functional usually appears in integral form, such as the formula for the Branchistochrone problem equation (2.1.19) described above. In general, the simplest and typical functional can be described as

$$J[u(x)] = \int_a^b F(x, u, u')dx, \tag{2.1.21}$$

where $F(x, u, u')$ is called the kernel of a functional.

2. Variations and Calculus of Variations

After the concept of functional is introduced, the above Branchistochrone becomes a minimalization problem of the functional $T[y(x)]$. The extreme value problem of functions is very common in physics. The image processing based on partial differential equation is solved by partial differential equations deduced from energy extreme value theorem. The method of finding the extremum of function is called variational method. Therefore we introduce the variational of function.

Assume there is continuous function $u(x)$, and we change it into $u(x) + t\eta(x)$, where t is a small parameter. We say $t\eta(x)$ is the variation of $u(x)$, namely

$$\delta u = t\eta(x), \tag{2.1.22}$$

then $u'(x)$ will change accordingly $\lim\limits_{\Delta x \to 0} \frac{\Delta(u+t\eta)}{\Delta x} = u'(x) + t\eta'(x)$. Thus we have

$$\delta u' = t\eta'(x) = \frac{d}{dx}(\delta u). \tag{2.1.23}$$

This shows that for a given function, the order of variation and differentiation can be interchanged.

Assume in equation (2.1.21), F is second order continuous function to x, u, u'. Therefore, when $u(x)$ has variation δu, then the change of J is shown as below

$$\begin{aligned}
\Delta J &= J[u(x) + t\eta(x)] - J[u(x)] \\
&= \int_a^b [F(x, u + t\eta, u' + t\eta') - F(x, u, u')]dx \\
&= \int_a^b \left[\frac{\partial F}{\partial u} t\eta + \frac{\partial F}{\partial u'} t\eta' + higher\ power\ terms\ of\ t\right] dx.
\end{aligned}$$

We call the linear principal part of the upper right (that is, omitting the higher order infinitesimal part) as the first variation of the functional, namely

$$\delta J = \int_a^b \left(\frac{\partial F}{\partial u} \delta u + \frac{\partial F}{\partial u'} \delta u' \right) dx. \tag{2.1.24}$$

3. Necessary conditions for the extreme value of functional–Euler-Lagrange function

Assume the extreme value of $J[u(x)]$ has a solution

$$u = u(x). \tag{2.1.25}$$

The differential equations satisfied by this solution are derived. Assume this solution has variation $t\eta(x)$, then $J[u(x) + t\eta(x)]$ can be regarded as the function of parameter t : $\Phi(t) = J[u(x) + t\eta(x)]$. When $t = 0$, $u(x) + t\eta(x) = u(x)$ is corresponding to equation (2.1.25), namely the extreme value of $J[u(x) + t\eta(x)]$. Thus, the extreme problem of the original functional is reduced to the extreme problem of a general function. The necessary condition of extreme value from function is $\frac{d\Phi}{dt}\big|_{t=0} = 0$. We have $\frac{\partial J[u(x)+t\eta(x)]}{\partial t}\big|_{t=0} = 0$. Substituting the formula equation (2.1.21), that is

$$\int_a^b \left[\frac{\partial}{\partial t} F(x, u + t\eta, u' + t\eta') \right]_{t=0} dx = 0 \;\Rightarrow\; \int_a^b \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right) dx = 0$$

multiply the two sides with t and we have

$$\int_a^b \left(\frac{\partial F}{\partial u} \delta u + \frac{\partial F}{\partial u'} \delta u' \right) dx = 0. \tag{2.1.26}$$

Comparing with equation (2.1.24), the solution equation (2.1.25) of the function equation (2.1.21) must satisfy $\delta J = 0$. This is the necessary condition for the extreme value of the functional. That is, the extreme function $u(x)$ of functional J must be the functional class $u(x)$ that satisfies the variational function $\delta J = 0$. Therefore, the extreme value problem of functional is called variational problem. Therefore, we regard the extreme value problem of function as a problem of variation. Note that both δu and $\delta u'$ exist in equation (2.1.26). Appling the partial integration method to the

second term, we have only δu under the integral sign,

$$\int_a^b \frac{\partial F}{\partial u'}\delta u' dx = \int_a^b \frac{\partial F}{\partial u'}\frac{d}{dx}(\delta u)dx = \frac{\partial F}{\partial u'}\delta u\Big|_a^b - \int_a^b \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\delta u dx.$$

In the simple variation, we keep $\delta u|_{x=a} = 0$ and $\delta u|_{x=b} = 0$, that is, the function value of the endpoint is fixed. So the first item on the upper right is zero, and the equation (2.1.26) becomes

$$\int_a^b \left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right]\delta u dx = 0. \tag{2.1.27}$$

The above equation is valid with the given (a, b) and any δu. Therefore

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) = 0. \tag{2.1.28}$$

That is, the functional equation (2.1.21) has the necessary condition for extreme values, and can be expressed as equation equation (2.1.28). This equation (2.1.28) is called the Euler-Lagrange equation of the extreme value problem of functional equation (2.1.21). Therefore, the extremum problem of functional equation (2.1.21) can be solved by solving the corresponding Euler-Lagrange equation.

Now we will discus the solution of Euler-Lagrange (2.1.27). The equation (2.1.28) is expanded to

$$\frac{\partial F}{\partial u} - \frac{\partial^2 F}{\partial u''\partial x} - \frac{\partial^2 F}{\partial u''\partial u}u' - \frac{\partial^2 F}{\partial u'\partial u'}u'' = 0.$$

When $\frac{\partial^2 F}{\partial u'\partial u'} \neq 0$, the above equation is a second order ordinary differential equation about $u(x)$. But it is not necessarily a linear differential equation to have only constant coefficients. So it is not easy to obtain its general solution. In some cases, however, it can be simplified.

(1) When the kernel function $F(x, u, u')$ in equation (2.1.21) does not contain u explicitly, namely $F = F(x, u')$, so obviously we have

$$\frac{\partial F}{\partial u'} = C. \tag{2.1.29}$$

(2) When the kernel function $F(x, u, u')$ in equation (2.1.21) does not explicitly contain x, namely $F = F(u, u')$, because

$$\frac{d}{dx}(F - u'F_{u'}) = F_x + F_u u' + F_{u'}u'' - \left(u''F_{u'} + u'\frac{dF_{u'}}{dx}\right) \overset{F_x=0}{=\!=} u'\left(F_u - \frac{dF_{u'}}{dx}\right) \overset{(2.1.28)}{=\!=} 0$$

thus we have

$$F - u'\frac{\partial F}{\partial u'} = C, \ or \ u'\frac{\partial F}{\partial u'} - F = C, \tag{2.1.30}$$

where C is the integral constant. Here are two examples.

Example 2.1.1. Solving the Branchistochrone problem, is the variational problem,

$$\delta \int_A^B \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} dx = 0.$$

Solution:

Since $F = \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}$ does not explicitly contain x, then from equation (2.1.30), Euler-Lagrange-Lagrange equation is

$$y'\frac{\partial}{\partial y}\sqrt{\frac{1 + y'^2}{y}} - \sqrt{\frac{1 + y'^2}{y}} = \frac{y'^2}{\sqrt{y(1 + y'^2)}} - \sqrt{\frac{1 + y'^2}{y}} = c \ \Rightarrow \ \frac{1}{y(1 + y'^2)} = c^2.$$

Let $\frac{1}{c^2} = c_1$, after separating variables, we get $\frac{\sqrt{y}dy}{\sqrt{c_1 - y}} = dx$ (c_1 is arbitrary constant). Then let $y = c_1 \sin^2 \frac{\theta}{2}$. Substituting it into the above equation, we have

$$dx = c_1 \sin^2 \frac{\theta}{2} d\theta = \frac{c_1}{2}(1 - \cos\theta)d\theta \ \Rightarrow \ \begin{cases} x = \frac{c_1}{2}(\theta - \sin\theta) + c_2 \\ y = \frac{c_1}{2}(1 - \cos\theta). \end{cases}$$

This is the parametric equation of the cycloid, and the integral constant can be determined by the positions A and B.

Example 2.1.2. Solving the shortest part of the curve between two points on the connection plane.

Solution:

This is about variation. Assume (x_0, y_0) and (x_1, y_1) are two fixed points on a plane, then the length of the curve connecting the two points can be expressed as

$$s[y(x)] = \int_{(x_0,y_0)}^{(x_1,y_1)} ds = \int_{(x_0,y_0)}^{(x_1,y_1)} \sqrt{(dx)^2 + (dy)^2} = \int_{(x_0,y_0)}^{(x_1,y_1)} \sqrt{1 + y'^2} dx.$$

Therefore, to get the minimum of the desired curve requires the extreme value of this functional. Here

$$F = \sqrt{1 + y'^2}, \ \ \frac{\partial F}{\partial y} = 0, \ \ \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}}.$$

So its Euler-Lagrange is

$$\frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}} = 0 \ \Rightarrow \ \frac{y'}{\sqrt{1 + y'^2}} = C \ \Rightarrow \ y'^2 = C^2(1 + y'^2) \ \Rightarrow \ y(x) = C_1 x + C_2.$$

It is a geometric axiom that the straight line between two points is the shortest line.

4. Euler-Lagrange Equation of complex functional

The Euler-Lagrange equation of the most simple functional equation (2.1.21) is equation (2.1.28) and the Euler-Lagrange equation of the relatively complex functional can be derived by referring to the above method. But the form can be even more complex.

(1) Euler-Lagrange that depends on the functional of multivariate function.

Functional that depends on the binary function $u(x, y)$

$$J(u) = \iint_G F(x, y, u, u_x, u_y)dxdy, \tag{2.1.31}$$

where F is the given function. On the boundary ∂G of G, we have

$$u(x, y)|_{\partial G} = u_0(x, y) \tag{2.1.32}$$

What we need to do is under the condition equation (2.1.32), to get $u(x, y)$ and the obtain the extreme value of the functional $J(u)$ in equation (2.1.31).

Assume the function $u(x, y)$ makes the functional $J(u)$ get extreme value and that this solution has variation $\alpha\eta(x, y)$, in which α is the minimum parameter, that is to get comparative function in a neighborhood of $u(x, y)$.

Let

$$u^*(x, y) = u(x, y) + \alpha\eta(x, y) \ \ and \ \ \eta(x, y)|_{\partial G} = 0. \tag{2.1.33}$$

Then the J value corresponding to $u^*(x, y)$ is

$$\begin{aligned} J(u^*) &= \iint_G F(x, y, u^*, u_x^*, u_y^*)dxdy \\ &= \iint_G F(x, y, u + \alpha\eta, u_x + \alpha\eta_x, u_y + \alpha\eta_y)dxdy \equiv J(\alpha). \end{aligned} \tag{2.1.34}$$

Since $J(u)$ gets extreme value in $u(x, y)$, namely when $\alpha = 0$, then $J(\alpha)$ gets extreme value. Therefore from the necessary condition of the extreme value, we get

$$\begin{aligned} \left.\frac{dJ(\alpha)}{d\alpha}\right|_{\alpha=0} = \iint_G (F_u\eta + F_{u_x}\eta_x + F_{u_y}\eta_y)dxdy &= \iint_G \left(F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y}\right)\eta(x, y)dxdy \\ &+ \iint_G \left(\frac{\partial(F_{u_x}\eta)}{\partial x} + \frac{\partial(F_{u_y}\eta)}{\partial y}\right)dxdy. \end{aligned} \tag{2.1.35}$$

The second integral can be formula into integral on boundary ∂G by plan Green's Formula,

$$\iint_G \left(\frac{\partial(F_{u_x}\eta)}{\partial x} + \frac{\partial(F_{u_y}\eta)}{\partial y} \right) dxdy = \int_{\partial G} \eta(F_{u_y}dy - F_{u_x}dx). \tag{2.1.36}$$

From equation (2.1.33), we know this integral is 0. Substituting it into equation (2.1.35), we have

$$\iint_G \left(F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y} \right) \eta(x,y)dxdy = 0. \tag{2.1.37}$$

Since both G and $\eta(x,y)$ are arbitrary, so satisfied with the following

$$F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y} = 0. \tag{2.1.38}$$

This formula is the necessary condition of the extreme value of the functional (2.1.31) of the binary function (Euler-Lagrange). It is obviously a partial differential equation.

The same, the variational problem of the functional $J(u)$ which is depends on ternary function $u(x,y,z)$ is $\delta \iiint_v F(x,y,z,u,u_x,u_y,u_z)dxdydz = 0$. The corresponding Euler-Lagrange is the partial differential equation

$$\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial F}{\partial u_z}\right) = 0. \tag{2.1.39}$$

For the case of n independent variables, the functional is

$$J(u) = \iint_G F(x_1, x_2, \ldots, x_n, u, u_{x_1}, u_{x_2}, \ldots, u_{x_n})dx_1dx_2\cdots dx_n. \tag{2.1.40}$$

Its first order variation is

$$\delta J(u) = \iint_G \left(\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_{x_1}}\delta u_{x_1} \ldots + \frac{\partial F}{\partial u_{x_n}}\delta u_{x_n} \right) dx_1dx_2\cdots dx_n$$

$$= \iint_G \left\{ \left[\frac{\partial F}{\partial u} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial u_{x_i}} \right) \right] \delta u + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial u_{x_i}} \delta u \right) \right\} dx_1 dx_2 \cdots dx_n. \tag{2.1.41}$$

Using the Green's Formula in n-dimensional space

$$\iint_G \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial u_{x_i}} \delta u \right) dx_1 dx_2 \cdots dx_n = \sum_{i=1}^{n} \int_{\partial G} \frac{\partial F}{\partial u_{x_i}} \cos\theta \delta u dS.$$

If $\delta u|_{\partial G} = 0$, then

$$\delta J(u) = \iint_G \left[\frac{\partial F}{\partial u} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial u_{x_i}} \right) \right] \delta u dx_1 dx_2 \cdots dx_n. \tag{2.1.42}$$

Through the arbitrariness of G and δu, we get the Euler-Lagrange equation of (2.1.42),

$$\frac{\partial F}{\partial u} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial u_{x_i}} \right) = 0. \tag{2.1.43}$$

Example 2.1.3. Write the Euler-Lagrange equation of functional $J[u(x,y)] = \iint_D [u_x^2 + u_y^2 + 2uf(x,y)]dxdy$.

Solution:

For functional

$$J[u(x,y)] = \iint_D [u_x^2 + u_y^2 + 2uf(x,y)]dxdy \;\Rightarrow\; F = u_x^2 + u_y^2 + 2uf(x,y),$$

from equation (2.1.37), we obtain

$$F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 2f(x,y) - 2u_{xx} - 2u_{yy} = 0.$$

So the Euler-Lagrange equation is $u_{xx} + u_{yy} = f(x,y)$. This is a two dimensional Poisson equation.

When the mathematical model is established, solving the resulting partial differential equations becomes the most important problem. The inherent discontinuity of the image function, the nonlinearity of the partial differential equation obtained by the mathematical model and the large amount of image data all together make it difficult to solve the equation. So, the partial differential equation in image processing, numerical implementation and the establishment of the model are all challenging tasks. This chapter focuses on the mathematical foundation of mathematical modeling and numerical methods, and mainly considers the numerical stability, accuracy and efficiency in numerical implementations.

2.1.3 Numerical solution of partial differential equation

For the image segmentation, the partial differential equations obtained from the energy functional are often nonlinear which complex forms. It is difficult to obtain analytic solutions, and thus numerical solutions are often needed. The common numerical methods for solving partial differential equations include the finite difference method, the finite element method and the boundary element method. The finite difference method is the main application in our thesis. Brief introduction is as follows.

Finite difference method (FDM) is a numerical method for solving differential equations by approximation them with difference equations, in which finite differences are used to approximate the derivatives. For example, forward difference is used to approximate the partial derivatives ($\frac{\partial u}{\partial t}$) with respect to time, namely

$$\left.\frac{\partial u}{\partial t}\right|_i^n \approx \frac{u_i^{n+1} - u_i^n}{\Delta t} = D_t^{(+)} u|_i^n. \tag{2.1.44}$$

The first derivative $\frac{\partial u}{\partial x}$ with respect to space x has the following choices, such as

$$\begin{cases} \frac{\partial u}{\partial x}|_i^n \approx \frac{u_{i+1}^n - u_i^n}{\Delta x} = D_x^{(+)} u|_i^n \\ \frac{\partial u}{\partial x}|_i^n \approx \frac{u_i^n - u_{i-1}^n}{\Delta x} = D_x^{(-)} u|_i^n \\ \frac{\partial u}{\partial x}|_i^n \approx \frac{u_{i+1}^n - u_{i-1}^n}{\Delta x} = D_x^{(0)} u|_i^n. \end{cases} \tag{2.1.45}$$

They are called the forward difference, the backward difference and the central difference, respectively. Using the Taylor expansion, we obtain

$$u(x+\Delta x) = u(x) + \frac{\partial u}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(\Delta x)^2 + \ldots,$$

and

$$\frac{\partial u}{\partial x} = \frac{u(x+\Delta x) - u(x)}{\Delta x} - \frac{1}{2}\frac{\partial^2 u}{\partial x^2}\Delta x - \ldots = D_x^{(+)} + O(\Delta x).$$

It is known that the forward difference and the backward difference are of first order accuracy. From the Taylor's expansion, we have

$$u(x+\Delta x) = u(x) + \frac{\partial u}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(\Delta x)^2 + \ldots,$$

and

$$u(x-\Delta x) = u(x) - \frac{\partial u}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(\Delta x)^2 - \ldots.$$

We get

$$u(x+\Delta x) = \frac{u(x+\Delta x) - u(x-\Delta x)}{2\Delta x} + O(\Delta x) - \ldots.$$

We can be know that central difference is of second order accuracy. When partial differential equations contain second order derivatives, they can also be approximated by finite difference. The first derivative of the central differential of the two half points can be obtained first

$$\left.\frac{\partial u}{\partial x}\right|_{i+\frac{1}{2}}^{n} \approx \frac{u_{i+1}^n - u_i^n}{\Delta x} \quad and \quad \left.\frac{\partial u}{\partial x}\right|_{i-\frac{1}{2}}^{n} \approx \frac{u_i^n - u_{i-1}^n}{\Delta x}.$$

And then using these two first order difference to do a central difference for the second derivative,

we have

$$\left[\frac{\partial^2 u}{\partial x^2}\right]\Big|_i^n \approx \left(\frac{\partial u}{\partial x}\Big|_{i+\frac{1}{2}}^n - \frac{\partial u}{\partial x}\Big|_{i-\frac{1}{2}}^n\right)\Big/\Delta x = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} = D_{xx}^{(0)}|_i^n. \tag{2.1.46}$$

In the two-dimensional case, we also need the approximate expression of the second order partial derivatives $\frac{\partial^2 u}{\partial x \partial y}$, which can be obtained by the least number of central difference of the half points, namely

$$\left[\frac{\partial^2 u}{\partial x \partial y}\right]\Big|_{i,j}^n = \frac{\partial u}{\partial y}\Big|_{i,j+\frac{1}{2}}^n - \frac{\partial u}{\partial y}\Big|_{i,j-\frac{1}{2}}^n \approx \left(\frac{u_{i+1,j+\frac{1}{2}} - u_{i-1,j+\frac{1}{2}}}{2\Delta y} - \frac{u_{i+1,j-\frac{1}{2}} - u_{i-1,j-\frac{1}{2}}}{2\Delta y}\right)\Big/\Delta x.$$

The half points can be approximated as

$$u_{i\pm1,j+\frac{1}{2}} \approx \frac{1}{2}(u_{i\pm1,j+1} + u_{i\pm1,j}) \; and \; u_{i\pm1,j-\frac{1}{2}} \approx \frac{1}{2}(u_{i\pm1,j-1} + u_{i\pm1,j}).$$

Thus

$$\left[\frac{\partial^2 u}{\partial x \partial y}\right]\Big|_{i,j}^n = D_{xy}^{(0)} u|_{i,j}^n = \frac{u|_{i+1,j+1}^n + u|_{i-1,j-1}^n - u|_{i+1,j-1}^n - u|_{i-1,j+1}^n}{4\Delta x \Delta y}. \tag{2.1.47}$$

2.1.4 Gradient descent flow method

From the above discussion we can see that the extreme value problem of the energy functional becomes a problem of solving the corresponding Euler-Lagrange equation. In general, the Euler-Lagrange equation is nonlinear PDE, and the discretization will lead to a nonlinear simultaneous algebraic equations. Numerical calculation is difficult. By introducing a "time" auxiliary variable, the static nonlinear PDE problem can be transformed into a dynamic PDE problem. When the evolution reaches the steady state, we obtain the solution of the Euler-Lagrange equation of the variational problem. This is the gradient descent flow method that will be discussed in the next section.

1. Gradient descent flow

Assume that the solution we look for changes over time, that is, it can be expressed as $u(\cdot,t)$, and this change with time always reduces the energy $J(u(\cdot,t))$. How should $u(\cdot,t)$ change to meet this requirement? Again we take one-dimensional variation as an example, and the variation $\delta u(\cdot)$ in the equation (2.1.27) is caused by the amount of change in $u(\cdot,t)$ from t to $t+\Delta t$. Namely $\delta u = \frac{\partial u}{\partial t}\Delta t$. Then equation (2.1.27) can be written as $J(\cdot,t+\Delta t) = J(\cdot,t) + \Delta t \int_a^b \frac{\partial u}{\partial t}\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right]dx$. So, only by making

$$\frac{\partial u}{\partial t} = -\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right] = \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) - \frac{\partial F}{\partial u}, \tag{2.1.48}$$

we can make $J(y(\cdot,t))$ gradually reduce because

$$\Delta J = J(\cdot,t+\Delta t) - J(\cdot,t) = -\Delta t \int_a^b \left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right]^2 dx \leq 0.$$

So we call equation (2.1.48) as the gradient descent flow corresponding to the variational problem equation (2.1.21).

In this way, we use an properly selected initial function u_0 and carry out iteration calculation based on equation (2.1.48) until u reaches a stable solution, and then $\frac{\partial u}{\partial t} = 0 \Rightarrow \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) - \frac{\partial F}{\partial u} = 0$. It can be see that the stable solution of the descending gradient is the soliton of the Euler-Lagrange equation equation (2.1.28).

Similarity for two-dimensional variational problems, a gradient descent flow can be obtained $\frac{\partial u}{\partial t} = \frac{d}{dx}\left(\frac{\partial F}{\partial u_x}\right) + \frac{d}{dy}\left(\frac{\partial F}{\partial u_y}\right) - \frac{\partial F}{\partial u}$. It is worth noting that only when $J(u)$ is convex, it has a unique minimum value, so that the gradient descent flow can be obtained with the unique solution independent of the initial conditions. When $J(u)$ is non-convex, the gradient descent flow may obtain different local minima due to different initial conditions instead of the global minimum.

Example 2.1.4. When we discuss an alternating projection method, which uses the edge information to characterize signals, the basic idea is that the reconstructed two dyadic wavelet coefficients satisfy the value of $|g(x)|$ should be as small as possible and at the same time $g(x)$ should

be as smooth as possible under the condition that at two endpoints $g(x_0) = a, g(x_1) = b$ (a, b are known). This leads to a typical one-dimensional variational problem with fixed endpoints: $J[g(x)] = \int_{x_0}^{x_1} \left[g^2 + \left(\frac{\partial g}{\partial x} \right)^2 \right] dx$. From $F[g, g'] = g^2 + \left(\frac{\partial g}{\partial x} \right)^2$, we have Euler-Lagrange equation $g - \frac{d^2 g}{dx^2} = 0$. This is a second order linear ordinary differential equation with constant coefficients, and its general solution is $g = \alpha e^x + \beta e^{-x}$, where α and β can be ascertained from boundary condition $g(x_0) = a$, $g(x_1) = b$. However, it is very rare for the Euler-Lagrange equation to obtain the analytical solution directly as in this case. The following example illustrates the necessity of using the gradient descent flow to obtain numerical solutions.

Example 2.1.5. If $F = \rho(|\nabla u|) = \rho(u_x^2 + u_y^2)^{\frac{1}{2}}$ in equation (2.1.31), then the corresponding Euler-Lagrange equation is (see equation (2.1.38))

$$\frac{d}{dx}\left(\frac{\partial F}{\partial u_x}\right) + \frac{d}{dy}\left(\frac{\partial F}{\partial u_y}\right) = 0 \;\Rightarrow\; \frac{d}{dx}\left(\rho'|\nabla u|\frac{u_x}{|\nabla u|}\right) + \frac{d}{dy}\left(\rho'|\nabla u|\frac{u_y}{|\nabla u|}\right) = 0$$

$$\Rightarrow\; div\left(\rho'|\nabla u|\frac{\nabla u}{|\nabla u|}\right) = 0. \tag{2.1.49}$$

In this way, we obtain a static nonlinear second order PDE. Finding numerical solutions of the direct equation (2.1.49) leads to the solution of a nonlinear simultaneous algebraic equations. It is relatively simple to solve the problem using the gradient descent flow method. Taking the gradient descent flow $\frac{\partial u}{\partial t} = div\left(\rho'|\nabla u|\frac{\nabla u}{|\nabla u|}\right)$, we transform the same variational problem into a initial boundary value problem. So it is possible to realize numerical calculation conveniently by explicit scheme.

Now we discuss two specific cases of significance in this case: 1) If $\rho(r) = r$, then the gradient descent flow is simplified as $\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{|\nabla u|}\right)$. This equation is closely related to the "total variation" image restoration model. 2) If $\rho(r) = r^2$, then $\rho'(r)2r$. We neglect the constant factor, and then the gradient descent flow can be reduced to $\frac{\partial u}{\partial t} = div\nabla u = \nabla^2 u$. So in this case, the gradient descent is linear heat flow.

2.1.5 The extremum principle of partial differential equations

In the partial differential equation (PDE) theory, the maximum principle is a very important basic concept, which can be generally expressed as: for certain types of PDF, its solution's maximum (small) value can only appear on the (time and space) defined boundary. For example, a continuous function $u(x)$ defined in $[a, b]$ intervals, if it satisfies the below in the open interval (a, b)

$$u'' + g(x)u' > 0, \quad x \in (a, b), \tag{2.1.50}$$

where $g(x)$ is any bounded function, and then the maximum value of u is only possible at the endpoints a and b. If at an internal point c the function reaches the maximum value, then $u'(c) = 0$, and $u''(c) < 0$, $a < c < b$. But this obviously contradicts equation (2.1.50), so equation (2.1.50) follows the extremum principle. We apply the above inference method to Laplace equation $\nabla^2 u = 0$. If solution u reach the maximum in the strict sense at any point c inside range Ω, then $\nabla u(c) = 0$, $\nabla^2(c) < 0$, which contradicts the requirement of $\nabla^2 u(c) = 0$. Thus we reach the conclusion: if the solution u of a Laplace equation reaches the maximum (minimum) value at any point c in range Ω, then $u = C$ (constant) in Ω.

The principle of extreme value can often be used to prove the uniqueness of solutions of partial differential equations. For example, let us consider the Poisson equation

$$\begin{cases} \nabla^2 u = f, & x \in \Omega \\ f = g, & x \in \partial\Omega. \end{cases} \tag{2.1.51}$$

Let u_1 and u_2 be two solutions, and $u := u_1 - u_2$. Then u satisfies

$$\begin{cases} \nabla^2 u = 0, & x \in \Omega \\ u = 0, & x \in \partial\Omega. \end{cases}$$

Since u is continuous on $\Omega \cup \partial\Omega$, then according to the above theory, $u \equiv 0$, that is $u_1 = u_2$. This proves the uniqueness of the solution of Poisson equation with the first kind of boundary condition.

The extreme value principle also has important applications in the study of other basic properties of partial differential equations. For example, in the study of two closed curve according to a partial differential equation. If in the initial time, one curve is strictly contained inside the another curve, then in order to prove the inclusion relation always established in the process of evolution, we only need to prove that the distance between the two curves meets the maximum principle. Similar methods can be used to detect the two images according to whether order is kept between the two images during the evolution following a certain PDE. For example, if the initial image has $u_1 < u_2$, then whether this relationship maintains during the evolution.

2.1.6 Variational and partial differential equations in image processing

Partial differential equations were introduced into image processing in 1980s and have witnessed sound development in the 90s. The partial differential equations in the image domain are represented as: $F(x, y, I, \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial^2 I}{\partial x^2}, \frac{\partial^2 I}{\partial x \partial y}, \frac{\partial^2 I}{\partial y^2}, \ldots) = 0$, where, x, y represent two-dimensional coordinates of the image, and I represents the gray values of pixels. From the above analysis we can see that the partial differential equation has the following advantages:

1. Linear superposition property

If F is a linear operator, then superposition principle is established, that is if $F(I_1) = 0$, $F(I_2) = 0$, $I = \alpha I_1 + \beta I_2$, α, β are constant, then

$$F(I) = \alpha F(I_1) + \beta F(I_2) = 0.$$

2. Uniqueness of solutions of partial differential equation model

By setting the initial conditions and diffusion coefficients of partial differential equations, the existence, uniqueness and stability of the model solution in the image domain can be guaranteed.

3. Local feature preserving performance

Partial differential equations can be used to process images according to local features, and preserve geometric features such as region boundaries. Partial differential equations can model the geometric

information of the gradient, curvature, tangent direction and normal direction of the image directly, so the visual effect of the image after processing is good. For different image processing tasks, scientists proposed a large number of different partial differential equations models. For example: Different energy functionals were applied to image de-noising, enhancement, segmentation, amplification, repair and compression. Such models include tensor diffusion partial differential equation model, diffusion equation models based on the theory of manifold, Mumford-Shah free boundary model, parabolic hyperbolic model, based on Euler (Euler-Lagrange) energy functional high order partial the differential equation model, and gradient vector and gray information diffusion model.

In the field of image processing, partial differential equations have achieved remarkable results, mainly in the areas of PDE image de-noising, PDE image zooming, PDE image segmentation and PDE image restoration.

In this dissertation, we mainly study the PDE segmentation, and adopt the high order partial differential equation model based on the Euler-Lagrange energy functional.

In the following sections, we will present the basics of partial differential equations and variational methods, and the relationship to image segmentation techniques.

2.2 Mathematical representation of plane curves and their characteristics

2.2.1 Differential properties of plane curves

A mapping from one-dimensional real domain to two dimensional real domain $C(p) : [a, b] \in R \to R^2$ defines plane curve [68], in which the p is the parameter of the curve, that is a point on the curve is obtained for any $p \in [a, b]$,

$$C(p) = (x(p), y(p)). \tag{2.2.1}$$

Therefore, $C(p)$ actually can be regarded as a vector, and its derivative of P is also a vector $C_p = (x_p, y_p)$. Its direction is the tangent direction of the curve, so it is called the tangent vector and its length is $|C_p| = \left((x_p)^2 + (y_p)^2\right)^{\frac{1}{2}}$. It can be understood that the position of the point on the curve increases with the increase of p, so C_p is also called the velocity vector.

Now we consider the distance traveled from the starting point $p_0 = a$ to the point p, i.e. the length of the arc:

$$s(p) = \int_a^p |C_p(\tau)| d\tau. \tag{2.2.2}$$

So we have

$$\frac{ds}{dp} = |C_p|.$$

When we use arc length as the parameter of the curve, the leftside is 1 and the rightside is $|C_s|$. Thus $|C_s| = 1$. That is to say, if the arc length s is the parameter, the tangent vector is always the unit length. Let $T = C_s$. Then

$$|T| \equiv 1. \tag{2.2.3}$$

At this point, the arc length between any two points on the curve can be expressed as

$$L(p_1, p_2) = \int_{s(p_1)}^{s(p_2)} ds = \int_{p_1}^{p_2} \sqrt{x_p^2 + y_p^2} dp. \tag{2.2.4}$$

Another fundamental concept of plane curves is curvature. Since it is a unit vector, its inner product is 1, that is $\langle C_s, C_s \rangle = \|C_s\|^2 = 1$. Taking the derivative to both sides, we obtain $\langle C_s, C_{ss} \rangle = 0$. It is obvious that the vector C_{ss} and the unit tangent vector C_s are orthogonal. We define the unit vector that constitutes the right coordinate system with T as the normal vector N. Then we can see that C_{ss} and T are collinear, and

$$C_{ss} = \kappa N. \tag{2.2.5}$$

The scale factor κ in the equation is called curvature. When C_{ss} and N are parallel, $\kappa > 0$ and conversely, when C_{ss} and N antiparallel, $\kappa < 0$. However, the direction of N is determined by T in

the right-hand coordinate system, while the direction of T is the direction of the parameter going, depending on the starting point of the curve. Therefore, for an open curve, plus or minus of the curvature depends on the position of the starting point. For a closed curve, the starting point and the end point is the same. We prescribe the direction of the curve counterclockwise and let N point to the inside of the closed curve. So when the curve is concave out, the curvature is positive; the other way around, if the curve is concave in, the curvature is negative.

According to the equation (2.2.3), the equation (2.2.5) can also be expressed as

$$T_s = \kappa N. \tag{2.2.6}$$

This means that the geometric meaning of curvature $\kappa(s)$ is the change rate of the tangent vector T with the arc length. This point can be seen more clearly from Figure 2.2. Assume that the tangent vectors and normal vectors on a point s of a curve can be represented respectively as $T(s) = (\cos\theta, \sin\theta)$ and $N(x) = (-\sin\theta, \cos\theta)$, where θ represents the inclination angle between the T and the X axis. When the curve moves to $s + \Delta s$, we have

$$\frac{T(s+\Delta s) - T(s)}{\Delta s} \cong \left(-\sin\theta\frac{\Delta\theta}{\Delta s}, \cos\theta\frac{\Delta\theta}{\Delta s}\right) = \frac{\Delta\theta}{\Delta s}N(s).$$

Let $\Delta s \to 0$, and compare with equation (2.2.6). We have

$$\kappa = \frac{d\theta}{ds}. \tag{2.2.7}$$

This shows that the curvature is the angular velocity of the tangential vector (see figure 2.2(a)), as well as the normal vector of the angular velocity of rotation (see figure (2.2(b)). Furthermore when the curve moves counter clockwise from $N(s)$ to $N(s + \Delta s)$, the curvature is positive, otherwise curvature is negative. In addition, we can further have

$$\frac{N(s+\Delta s) - N(s)}{\Delta s} \cong \left(-\cos\theta\frac{\Delta\theta}{\Delta s}, -\sin\theta\frac{\Delta\theta}{\Delta s}\right) = -\frac{\Delta\theta}{\Delta s}T(s). \tag{2.2.8}$$

So

$$N(s) = -\kappa T.$$

Equation (2.2.6) and equation (2.2.8) are called the Frenet equations.

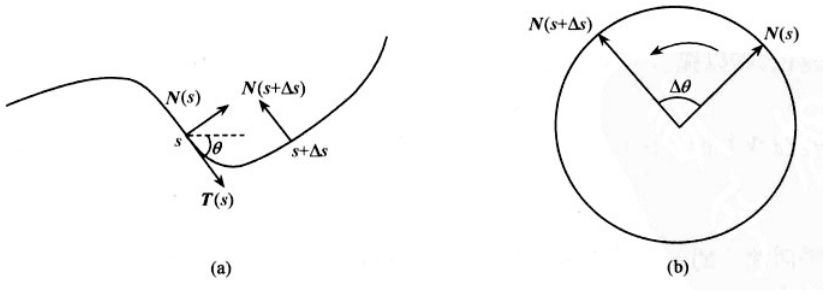


Figure 2.2: Interpretation of geometric meaning of curvature.

Now letus consider a circle with a radius of a. From $s = a\theta \Rightarrow \frac{d\theta}{ds} = \frac{1}{a} \Rightarrow \kappa = \frac{1}{a}$. It can be seen that a circle is a closed curve whose curvature is constant, and the curvature is the reciprocal of its radius.

For an arbitrary curve, $C : y = f(x)$, and a point M on the C. We find another point D, on the normal of C direction with a distance of $1/\kappa(s)$ from M. We make a circle whose center is on M with a radius of $1/\kappa(s)$. Then the circle is tangent to the curve at the point M, and it has a same curvature with curve C. Such a circle is called an Osculating Circle of the curve C at its point M. It can be said that curvature is the reciprocal of the radius of the close curve (see Figure 2.3).

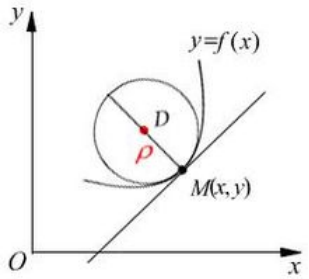


Figure 2.3: Osculating Circle.

In addition, from

$$C_p = (x_p, y_p)\frac{dp}{ds} = \frac{(x_p, y_p)}{\sqrt{(x_p^2 + y_p^2)}},$$

we have

$$\frac{dC_p}{dp} = \frac{x_p y_{pp} - x_{pp} y_p}{\sqrt{(x_p^2 + y_p^2)^3}}(-y_p, x_p),$$

and

$$C_{ss} = \frac{dC_p}{dp}\frac{dp}{ds} = \frac{x_p y_{pp} - x_{pp} y_p}{\sqrt{(x_p^2 + y_p^2)^3}}\frac{(-y_p, x_p)}{\sqrt{(x_p^2 + y_p^2)}} = \frac{x_p y_{pp} - x_{pp} y_p}{(x_p^2 + y_p^2)^{3/2}} N.$$

Then another expression of curvature is obtained

$$\kappa = \frac{x_p y_{pp} - x_{pp} y_p}{(x_p^2 + y_p^2)^{3/2}}. \tag{2.2.9}$$

Next, we derive another important curvature expression. Let unit normal vector $N = (n_1, n_2) = (-\sin\theta, \cos\theta)$. Then

$$\frac{\partial n_1}{\partial x} = \frac{\partial n_1}{\partial \theta}\frac{\partial \theta}{\partial s}\frac{\partial s}{\partial x} = -\cos\theta \cdot \kappa \cdot \frac{\partial s}{\partial x}, \quad \frac{\partial n_2}{\partial y} = -\sin\theta \cdot \kappa \cdot \frac{\partial s}{\partial y}. \tag{2.2.10}$$

Since $ds = dx\cos\theta + dy\sin\theta$, comparing with $ds = \frac{\partial s}{\partial x}dx + \frac{\partial s}{\partial y}dy$, we have

$$\frac{\partial s}{\partial x} = \cos\theta, \quad \frac{\partial s}{\partial y} = \sin\theta.$$

From equation (2.2.10), obtain $\frac{\partial n_1}{\partial x} = -\kappa\cos^2\theta$, $\frac{\partial n_2}{\partial y} = -\kappa\sin^2\theta$. Thus

$$\kappa = -\left(\frac{\partial n_1}{\partial x} + \frac{\partial n_2}{\partial y}\right) = -div(N). \tag{2.2.11}$$

It can be proved that the arc length and the curvature are two Euclidean invariants, and the curvature and the arc length of the plane curve remain unchanged from rotation and translation motion.

2.2.2 A level set representation method for plane closed curves

For plane closed curves, besides using equation (2.2.1) and $C(a) = C(b)$ as the explicit equation, we can also use the implicit equation: $C = \{(x, y), u(x, y) = c\}$, where $u(x, y)$ is a binary function, that is to say, C satisfies the point set of $u(x, y) = c$, called a level (line) set of $u(x, y)$. At this point, we call the $u(x, y)$ as an embed function of C. Figure 2.4 gives a sketch of the level set. When the constant $c = 0$, it is called the zero level set.

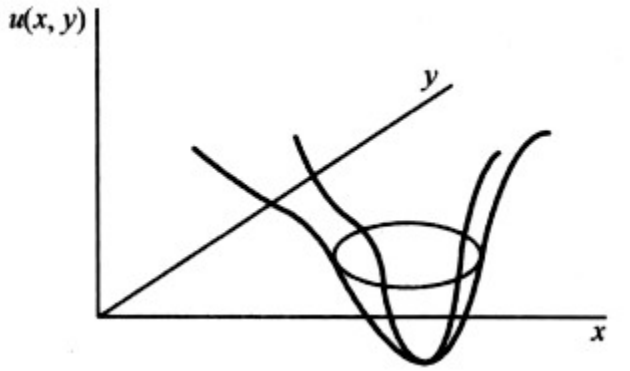


Figure 2.4: Embedded functions and level sets.

If a direction derivative is taken to the function $u(x, y)$ in the tangent direction of the level set at a point p, since $u(x, y)$ remains unchanged along the level set, then $\frac{du}{dT} = \frac{\partial u}{\partial x}\cos\theta + \frac{\partial u}{\partial y}\sin\theta = 0$, where θ represents the inclination angle between the vector T and the axis x. The descendant gradient of u

$$\nabla u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \tag{2.2.12}$$

is perpendicular to the tangent vector $T = (\cos\theta, \sin\theta)$ of the level set, that is, parallels to the

normal vector of the level set. On the other hand, according to the definition (2.2.12), the direction of the gradient always points to the direction in which the value of the function $u(x, y)$ increases. Therefore, the unit normal vectors of level sets can be represented as

$$N = \pm \frac{\nabla u}{|\nabla u|}. \tag{2.2.13}$$

If $u(x, y)$ is positive inside the zero level set, and negative outside the zero level set, we let the sign in equation (2.2.13) be plus, and otherwise minus. In this way, N always points to the interior of the closed curve, which is consistent with the definitions of the closed curve law vector N in the last section.

In the next section when discussing the level set method for curve evolution, the zero level set of the embedded function $u(x, y)$ is always specified. For the curve we are concerned with, it is stipulated that $u(x, y) < 0$ inside the zero level set, and $u(x, y) > 0$ outside the zero level set. Therefore, under this agreement, equation (2.2.13) takes negative. Substituting equation (2.2.13) into equation (2.2.10), the curvature of the embedded function $u(x, y)$ level set can be obtained

$$\kappa = \frac{u_{xx}u_y^2 - 2u_x u_y u_{xy} + u_x^2 u_{yy}}{(u_x^2 + u_y^2)^{3/2}}. \tag{2.2.14}$$

It is worth noting that equation (2.2.14) can not only be used to compute the curvature of zero level set $u(x, y)$, but also to compute the curvature $u(x, y)$ of all levels set.

2.2.3 Global properties of plane curves

For plane closed curve (namely $C(a) = C(b)$), we have the below global properties:

1. A circle is a unique closed curve with constant curvature;

2. Each closed curve has at least four vertices. The so-called vertices means that the first derivatives of the curvature at these points are zeros. Obviously, in this definition, every point on the circle is a vertex;

3. The total curvature of a closed curve is an integral multiple of 2π, i.e. $\int \kappa(s)ds = 2\pi k$;

4. Perimeter inequality. Let A and L be the area and perimeter of the closed curve, respectively. Then we have $L \geq 2\sqrt{\pi A}$. If and only if the closed curve is round, the equal sign in the upper form is valid.

2.3 Curve evolution and variational level set method

2.3.1 A general equation for geometric evolution of curves

First we consider a closed curve sequence $C(p,t)$, $t \geq 0$ that follows the following PDE evolution:

$$\frac{\partial C(p,t)}{\partial t} \equiv V = \alpha(p,t)T + \beta(p,t)N, \;\; C(p,0) = C_0(p), \tag{2.3.1}$$

where α and β are the tangential velocity and the normal velocity respectively. The value of y of a curve can be expressed as a function of x, $y = r(x)$. That is, the curve C is represented by x as a parameter, $C(x) = (x, r(x))$. So, the tangential vector is $C_x = (1, r_x)$. The unit tangent vector and the normal vector are $T = \frac{(1,r_x)}{\sqrt{1+r_x^2}}$, and $N = \frac{(-r,1)}{\sqrt{1+r_x^2}}$. Therefore, when the curve C is evolved according to the equation (2.3.1), x and y at any point will move in accordance with the following equations

$$\begin{aligned}
\frac{dy}{dt} &= \alpha\frac{r_x}{\sqrt{1+r_x^2}} + \beta\frac{1}{\sqrt{1+r_x^2}}, \\
\frac{dx}{dt} &= \alpha\frac{1}{\sqrt{1+r_x^2}} + \beta\frac{-r_x}{\sqrt{1+r_x^2}}.
\end{aligned}$$

Given that

$$\begin{aligned}
&\frac{dy}{dt} = r_x\frac{dx}{dt} + r_t \\
&\Rightarrow \; r_t = \frac{dy}{dt} - r_x\frac{dx}{dt} = \alpha\frac{r_x}{\sqrt{1+r_x^2}} + \beta\frac{1}{\sqrt{1+r_x^2}} - \alpha\frac{r_x}{\sqrt{1+r_x^2}} + \beta\frac{r_x^2}{\sqrt{1+r_x^2}} = \beta\sqrt{1+r_x^2},
\end{aligned}$$

this formula shows that the change of the geometric shape of the curve is only related to the normal β of V, not to the tangential α of the velocity of motion. Therefore, when we discuss the evolution

of geometric curves, we only need to consider the normal velocity (Figure 2.5). In this way, the general equation of curve evolution can be simplified as

$$\frac{\partial C}{\partial t} = \beta N. \tag{2.3.2}$$

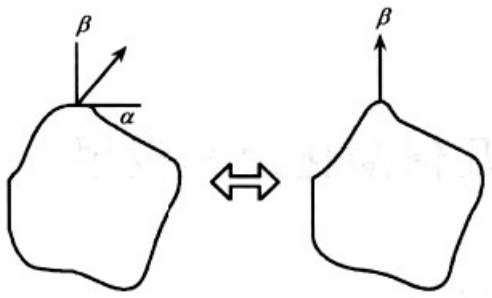


Figure 2.5: Curve evolution is independent of tangential velocity in geometry.

2.3.2 Level set method for curve evolution

1. Basic concept

It has been pointed in section 2.2.2 that a closed plane curve can adopt implicit expression, which is defined as a level set of a two-dimensional function $u(x, y)$, i.e. $C = \{(x, y), u(x, y) = c\}$. Then, if C has any kind of changes, then we can say function $u(x, y)$ has corresponding changes. More precisely, the closed curve that changes with time can be expressed as a level set of a 2D function $u(x, y)$ that changes with time, namely $C(t) = \{(x, y).u(x, y, t) = c\}$. So, how does the embedding function $u(x, y, t)$ evolve when the curve $C(t)$ evolves as in equation (2.3.1)? Taking the derivative $\frac{du}{dt}$ for the function u in the above equation, using the chain rule of the composite function, we have $\frac{du}{dt} = \frac{\partial u}{\partial t} + \nabla u \cdot \frac{\partial (x,y)}{\partial t} = 0$. Using equation (2.3.1), we obtain

$$\frac{\partial u}{\partial t} = -\nabla u \cdot \frac{\partial (x, y)}{\partial t} = -\nabla u \cdot V = -|\nabla u| \frac{\nabla u}{|\nabla u|} \cdot V = |\nabla u| N \cdot V = \beta |\nabla u|, \tag{2.3.3}$$

where $\beta = V \cdot N$ is the normal component of the motion velocity. equation (2.3.3) is the basic equation of curve evolution level set method.

In above derivation, we used equation (2.2.13) with negative sign, which means that it has been assumed that $u(x, y) > c$, for (x, y) outside the closed curve C; $u(x, y) < c$, for (x, y) inside the closed curve C; $u(x, y) = c$, for (x, y) on the closed curve C. In addition, it is noted that the derivation of equation (2.3.3) is independent of the value of constant C. So $c = 0$ is often taken for convenience, that is, the curve we are concerned with is the zero level set of the embedded function.

We can see that the evolution of equation (2.3.2) under the initial condition C_0 of the closed curve C, according to is equivalent to the evolution of the embedding function $u(x, y)$ according to the equation (2.3.3) under the given initial value u_0 condition (the zero level of u_0 corresponds to C_0). That is to say the current curve $C(t)$ can be determined at any time t by taking the level set of $u(x, y, t) = 0$.

2. Variational level set method

When the curve evolution is applied to image processing problems, an curve motion equation often comes from an energy functional for minimized a closed curve C. For example the very famous geodesic active contour model for image segmentation is to minimize the following functional,

$$E(C) = \oint g(|\nabla I[C(s)]|)ds. \tag{2.3.4}$$

Now we will describe this theory and prove [68] that the descent gradient of equation (2.3.4) is

$$\frac{\partial C}{\partial t} = [g(c)\kappa - \nabla g \cdot N]N. \tag{2.3.5}$$

(1) Selection of energy functional

The Fermat theorem in optics tells us that when light travels in an inhomogeneous medium, the light will not be a straight line; rather its path is determined by the shortest path of the light. If the refractive index of the medium is $n(x, y, z)$, then the route of light propagation from A to B

should satisfy a optical path such that

$$L_R = \int_A^B n(s)ds \tag{2.3.6}$$

reaches minimum, where ds in the formula represents the Eulidean arc length. That is to say that light is always transmitted through the local minimum of n. The principle of Fermat can explain the phenomena of optical phenomena such as refraction and reflection in geometrical optics. The idea of this principle is consistent with the idea that the local minima on active contour should be obtained by the local minimization of $g(\nabla I)$ in image segmentation snake model [32]. Therefore it is put forward in [12] to determine the active contour by minimizing the energy functional:

$$L_R(C) = \int_0^{L(C)} g(|\nabla I[C(s)]|)ds, \tag{2.3.7}$$

where $L(C)$ represents the arc length of the closed curve C and $L_R(C)$ is the weighted arc length. This equation can also be written as

$$E(C) = \oint g(|\nabla I[C(s)]|)ds. \tag{2.3.8}$$

(2) Prove equation (2.3.5).

Using the parametric representation of a curve, the closed curve C is expressed as $C(p) : [0, 1] \to R^2$, $C(0) = C(1)$. Given that $ds = |C_p|dp$, equation (2.3.8) can be rewritten into

$$E(C(p)) = \int_0^1 g(C(p))|C_p(p)|dp. \tag{2.3.9}$$

In order to deduce the gradient descent flow of this formula, the auxiliary variable t is introduced into C, and the formula is further rewritten into

$$E(C(p,t)) = \int_0^1 g(C(p,t))|C_p(p,t)|dp. \tag{2.3.10}$$

So we have

$$\frac{dE}{dt} = \int_0^1 \left\{ \frac{d}{dt} g(C(p,t)) \right\} |C_p(p,t)| dp + \int_0^1 g(C(p,t)) \left\{ \frac{d}{dt} |C_p(p,t)| \right\} dp$$
$$= \int_0^1 \{\nabla g(C(p,t)) \cdot C_t(p,t)\} |C_p(p,t)| dp + \int_0^1 g(C(p,t)) \{C_{pt}(p,t) \cdot T\} dp. \tag{2.3.11}$$

In the formula, T represents the unit tangent vector of C. The partial integration method is applied to the second item of equation (2.3.11) and we take into account that C is a closed curve, then we have

$$\int_0^1 g(C(p,t)) \{C_{pt}(p,t) \cdot T\} dp = -\int_0^1 \{g(C(p,t))T\}_p C_t(p,t) dp$$
$$= -\int_0^1 \{\nabla g(C(p,t)) \cdot C_p(p,t)\} [T \cdot C_t(p,t)] dp - \int_0^1 \{g(C(p,t))\} T_p \cdot C_t(p,t) dp.$$

Substituting it into equation (2.3.11) and using s to replace the parameters p in the above equation, and $|C_s(s,t)| \equiv 1$, we have

$$\frac{dE}{dt} = \int_0^{L(C)} \{\{\nabla g(C(s,t)) \cdot C_t(s,t)\} - \{\nabla g(C(s,t)) \cdot T\}[T \cdot C_t(s,t)]\} ds$$
$$- \int_0^{L(C)} \{g(C(s,t))\} T_p \cdot C_t(p,t) ds$$
$$= \int_0^{L(C)} \{\nabla g(C) - [\nabla g(C) \cdot T]T - g(C)T_s\} \cdot C_t ds.$$

Because $\nabla g(C) - [\nabla g(C) \cdot T]T = [\nabla g(C) \cdot N]N$, $T_s = \kappa N$ and $T_s = \kappa N$, we have

$$\frac{dE}{dt} = \int_0^{L(C)} \{\nabla g(C) \cdot N - g(C)\kappa\} N \cdot C_t ds.$$

In order to make $\frac{dE}{dt} < 0$, C_t should satisfy $C_t = g(C)\kappa N - \nabla g(C) \cdot N$. Now the equation (2.3.5) is approved.

(3) Variational level set representation

If the level set method discussed above is used, then the corresponding PDE for the embedded function is

$$\frac{\partial u}{\partial t} = [g\kappa - \nabla g \cdot N]|\nabla u| = |\nabla u| div\left(g\frac{\nabla u}{|\nabla u|}\right). \tag{2.3.12}$$

A new level set method, called variational level set method, is proposed in the literature [77] for the curve evolution problem, derived from the energy functional minimization of curves.

First, we use the special function defined below (called Heaviside functions),

$$H(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0. \end{cases}$$

The loop integral formula equation (2.3.4) along the C can be rewritten as an area division $\oint_C g(C)ds = \iint_\Omega g(x,y)|\nabla H(u)|dxdy$. Because $\nabla H(u) = \delta(u)\nabla u,\ \delta(z) = \frac{d(H(z))}{dz}$, equation (2.3.4) can be rewritten as the functional of imbed function u

$$J(u) = \iint_\Omega g(x,y)\delta(u)|\nabla u|dxdy. \tag{2.3.4a}$$

Using the variational method, the gradient descent flow of the upper type can be obtained

$$\frac{\partial u}{\partial t} = \delta(u) div\left(g\frac{\nabla u}{|\nabla u|}\right). \tag{2.3.13}$$

In order to make it a computable PDE, δ in the formula needs to be approximated by regularization δ_ε, that is equation (2.3.13) can be rewritten as

$$\frac{\partial u}{\partial t} = \delta_\varepsilon(u) div\left(g\frac{\nabla u}{|\nabla u|}\right),\ \ \delta_\varepsilon(z) = \frac{dH_\varepsilon(z)}{dz}. \tag{2.3.14}$$

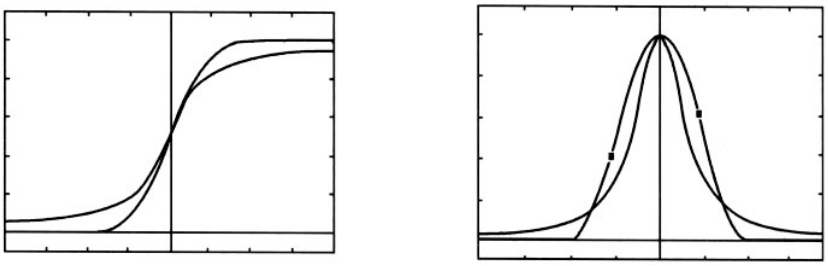
Here, $H_\varepsilon(z)$ is called the regularized Heaviside function. In principle, it can be any function that satisfies the following conditions: $H_\varepsilon(z) \xrightarrow{z\to 0} H(z)$.

For example, the following functions can be used as regularized Heaviside functions:

$$H_\varepsilon^{(1)}(z) = \begin{cases} 1, & z > \varepsilon \\ 0, & z < -\varepsilon \\ \frac{1}{2}\left(1 + \frac{z}{\varepsilon} + \frac{1}{\pi}\sin\frac{\pi z}{\varepsilon}\right), & other \end{cases}, \quad H_\varepsilon^{(2)}(z) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\frac{z}{\varepsilon}\right). \tag{2.3.15}$$

They are all odd symmetric functions, in which the arguments are used to control the speed of the function from 0 to 1. Figure 2.6(a) is an image of these two regularized Heaviside. The image of their derivative $\frac{dH_\varepsilon^{(1)}}{dz}$ and $\frac{dH_\varepsilon^{(2)}}{dz}$ are shown as in Figure 2.6. They can be regarded as $\varepsilon_z(u)$ function in equation (2.3.14).

PDEs (2.3.12) and (2.3.14) look like the same. The difference is just $|\nabla u|$ and $\varepsilon_z(u)$. However, mathematically, the two PDEs have an essential difference. Equation (2.3.12) belongs to the hyperbolic type, while equation (2.3.14) belongs to the parabolic type. The stability of the later one is higher than the former one. Thus, a larger time step is needed in numerical implementations and often without the need for re-initialization of the embedded function.



(a) (b)

Figure 2.6: Regularized Heaviside graphs, and $\frac{dH_\varepsilon^{(1)}}{dz}$ and $\frac{dH_\varepsilon^{(2)}}{dz}$ graphs

However, this does not mean that the variational level set method can replace the level set method, because the premise of using the variational level set method to solve the problem of curve evolution is: the problem is derived from minimizingthe "energy" functional $J(C)$ of the C curves. This is

done by introducing the embedded function u and using the Heaviside function to transform $J(C)$ into $J(U)$. The PDE of u is obtained by calculus of variations. The level set method is: first, the variational method is used to minimize the energy functional of C, and the equation of motion of C is obtained. Then the embedded function is introduced to obtain the PDE of u. The problem is that not all the evolution of curves and surfaces is derived from minimization of the energy functional. There are many kinds of interface evolution problems in fluid mechanics, material science and many other fields. In these cases, the evolution equation (PDE) of the interface is derived directly from the laws of physics. In these cases, level set methods may be applied. Therefore, the level set method is more suitable than the variational level set method.

3. Improved variational level set method

In order to completely avoid re-initialization, an improved variational level set method is proposed in literature [41] by adding an item in the "energy" functional related to the embedded function u, i.e. $P(u) = \iint\limits_{\Omega} \frac{1}{2}(|\nabla u| - 1)^2 dxdy$. Its gradient descent flow is $\frac{\partial u}{\partial t} = \left[\nabla^2 u - div\left(\frac{\nabla^2 u}{|\nabla u|}\right)\right]$. Obviously, minimizing $P(u)$ means that the requirement of $|\nabla u| = 1$, that is, the embedded function u should be kept as a distance function as far as possible in the evolution process. The above equation can be rewritten as

$$\frac{\partial u}{\partial t} = div\left[\left(1 - \frac{1}{|\nabla u|}\right)\nabla u\right]. \tag{2.3.16}$$

This is a nonlinear heat conduction equation with a conductivity rate of $\alpha = 1 - \frac{1}{|\nabla u|}$. Obviously when $|\nabla u| > 1$ and $\alpha > 0$, the heat is conducted to the outside, and $|\nabla u|$ is reduced; The other way around, when $|\nabla u| > 1$ and $\alpha < 0$, $|\nabla u|$ is increased. As a result, any deviation from the local $|\nabla u| = 1$ will be corrected in the subsequent evolution. Thus, re-initialization is not necessary at all.

For example, after adding $P(u)$ to the functional of equation (2.3.4a), we have

$$J(u) = \mu \iint\limits_{\Omega} \frac{1}{2}(|\nabla u| - 1)^2 dxdy + \iint\limits_{\Omega} g(x, y)\delta(u)|\nabla u| dxdy. \tag{2.3.4b}$$

The Corresponding gradient descent flow is

$$\frac{\partial u}{\partial t} = \mu\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right) + \delta_\varepsilon(u) div\left(g\frac{\nabla u}{|\nabla u|}\right)$$

where μ is a chosen constant.

The advantage of the improved variational level set method is that it not only completely avoids the problem of re-initialization, but also simplifies the work of initializing the embedded function $u_0(x, y)$. Since the model itself has the ability to automatically approximate the embedded function u as a distance function, initialization does not necessarily strictly require that $u_0(x, y)$ be a distance function C_0. Based on this consideration, literature [32] proposed the below initialization scheme

$$u_0 = \begin{cases} -\alpha, & (x, y) \in int\ ernal\ of\ C_0 \\ 0, & (x, y) \in C_0 \\ \alpha, & (x, y) \in external\ of\ C_0, \end{cases}$$

where $\alpha \approx 2\varepsilon$, and ε is the parameter of regularized Heaviside function (see equation (2.3.15)).

Chapter 3
THE MEDICAL IMAGE SEGMENTATION MODEL BASED ON ENTROPY

In 2001, Chan and Vese assumed that the image consists of two homogeneous regions of the target and the background and proposed the famous CV model [15]. As described in Section 1.3.2, CV model is a Geometric Activity Contour Model Based on Region. The model has the advantages of low computational complexity, strong anti-noise and insensitivity to the initial contour. Since the CV model depends on the grayscale homogeneity of the to-be segmented region. Therefore, the segmentation speed is very slow for the image with intensity inhomogeneity and sometimes the image cannot even be segmented. In addition, CV model evolution requires the periodical re-initialization of the level set function to ensure the stability of the computation [56]. It costs a lot of time.

In order to improve segmentation processing speed, many excellent algorithms are proposed [9, 13, 26]. Good segmentation results have been achieved. But for those intensity inhomogeneity images, these algorithms still cannot get a good result. Aimed at this problem, Vese proposed a complicated Piecewise Smooth (PS) model [67]. The PS model uses two smooth functions to approximate the grayscale of the to-be segmented image, which solves the problem of that the CV model cannot segment the uneven grayscale image. However, the PS model calculation is more complex and the evolution time is too long. In order to improve the existing problems of the CV model, Li proposed the famous RSF model [40]. The RSF model transforms the two-value global fitting energy functional of the CV model into the local binary fitting energy with the Gaussian function as the kernel function. The RSF model has effectively solves the problem that CV model cannot deal with uneven grayscale image. However, RSF model is very sensitive to the selection of the initial contour. When the initial contour selection is not appropriate, it is easy to fall into the local minimum. Meanwhile, RSF model is sensitive to noise. In 2009, Wang et al proposed the LGIF (Local and Global Intensity Fitting) model [69]. For the first time, they integrated the global and local information of the image into an energy functional. The LGIF model used the CV model

is not sensitive to the initial position and noise and the RSF model can deal with gray scale image. Based on the LGIF model, we proposed CER model by using information entropy in this paper. Numerical experiments show that CER model improved the problems that the CV model cannot segment uneven grayscale image and that RSF model is sensitive to initial contours and has poor noise immunity. Meanwhile, CER model also improved the speed of curve evolution.

3.1 Chan-Vese model and Region-scalable Fitting model

3.1.1 CV model [15]

The CV model can be defined as that the image set I defined in Ω is divided into two parts by the active contour C, and the two parts are recorded as $inside(C)$ and $outside(C)$. We use c_1 and c_2 to represent the average value of the internal gray level of the curve and the average value of the external gray level of the curve respectively. The fitting energy functional is defined as follows:

$$\begin{aligned} E^{CV}(c_1, c_2, C) =& \nu \cdot Length(C) + \mu \cdot Area(inside(C)) \\ &+ \lambda_1 \int_{inside(C)} |I - c_1|^2 dxdy + \lambda_2 \int_{outside(C)} |I - c_2|^2 dxdy, \end{aligned} \tag{3.1}$$

where $Length(C)$ is the length of the contour curve which plays a role of regularization, $Area(inside(C))$ is the internal area of the contour curve, $\mu, \nu \geq 0$ are two parameters, λ_1, λ_2 are two weight coefficients, and $\lambda_1, \lambda_2 > 0$. The first two terms on the right hand side of the formula are called the "fitting" terms which are mainly used to control the fitting error for the contour curve fit edge. The other two terms are called the "smooth" terms for getting a smooth curve in the process of evolution. The final position of the contour curve can be obtained by minimizing the energy functional.

Let the level set function u be represented by the following formula:

$$\begin{cases} u(x, y) = 0, & (x, y) \in C \\ u(x, y) > 0, & (x, y) \in inside(C) \\ u(x, y) < 0, & (x, y) \in outside(C). \end{cases}$$

In order to apply the level set method, the Heaviside function is introduced

$$H(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

The Dirac function is

$$\delta(x) = \frac{d}{dx} H(x).$$

In this method, the Heaviside function is used to divide the region of evolution, and the Dirac function is used to define the value of evolution around the zero level set function. Here,

$$\begin{aligned} Length(C) = \int_0^{L(C)} ds = \int_\Omega \delta(u)|\nabla u| dxdy, \\ Area(inside(C)) = \int_\Omega H(u) dxdy, \\ \int_{inside(C)} |I(x,y) - c_1|^2 dxdy = \int_\Omega |I(x,y) - c_1|^2 H(u) dxdy, \\ \int_{outside(C)} |I(x,y) - c_2|^2 dxdy = \int_\Omega |I(x,y) - c_2|^2 (1 - H(u)) dxdy. \end{aligned}$$

The regularized Heaviside function is often used in the numerical calculation as

$$H_\varepsilon(x) = \frac{1}{2}\left[1 + \frac{1}{\pi} \arctan\left(\frac{x}{\varepsilon}\right)\right].$$

For the constant ε, the corresponding Dirac function is

$$\delta_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + x^2}.$$

The goal is to expand the capture range and to prevent the occurrence of singular cases. Thus, the

energy functional of the level set function is obtained as

$$E_{\varepsilon}^{CV}(c_1, c_2, u) = \nu \int_{\Omega} \delta_{\varepsilon}(x)|\nabla u| dxdy + \mu \int_{\Omega} H_{\varepsilon}(u) dxdy + \lambda_1 \int_{\Omega} |I - c_1|^2 H_{\varepsilon}(u) dxdy + \lambda_2 \int_{\Omega} |I - c_2|^2 (1 - H_{\varepsilon}(u)) dxdy. \tag{3.2}$$

At this point, the level set of the image $I(x, y)$ is

$$I(x, y) = c_1 H_{\varepsilon}(u) + c_2(1 - H_{\varepsilon}(u)).$$

Fixing u and minimizing the energy functional $E_{\varepsilon}^{CV}(c_1, c_2, u)$, one has

$$c_1 = \frac{\int_{\Omega} I(x, y) H_{\varepsilon}(u(x, y)) dxdy}{\int_{\Omega} H_{\varepsilon}(u(x, y)) dxdy}, c_2 = \frac{\int_{\Omega} I(x, y)(1 - H_{\varepsilon}(u(x, y))) dxdy}{\int_{\Omega} (1 - H_{\varepsilon}(u(x, y))) dxdy}. \tag{3.3}$$

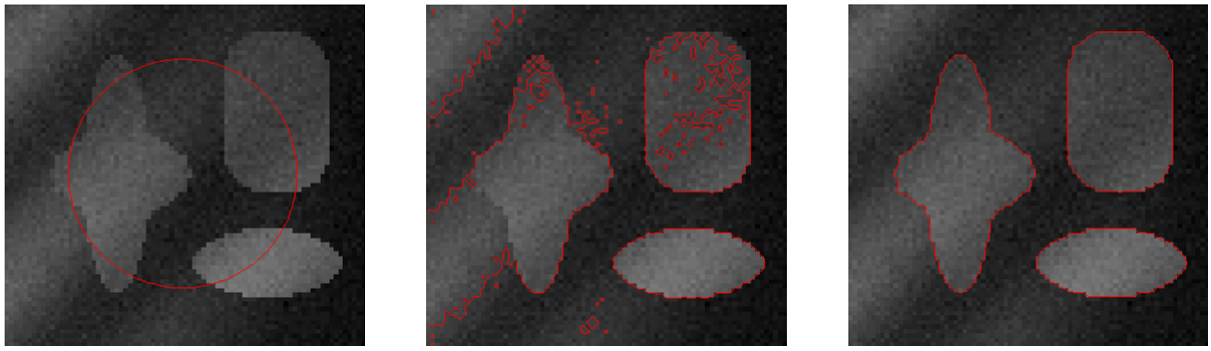
In the CV model, we usually set $\nu = 0$. The energy functional $E_{\varepsilon}^{CV}(c_1, c_2, u)$, is minimized using the variation method and the gradient descent flow technique and Euler-Lagrange equation is obtained as

$$\frac{\partial u}{\partial t} = \nu \delta_{\varepsilon}(u) div(\frac{\nabla u}{|\nabla u|}) + \delta_{\varepsilon}(u)(\lambda_2(1 - c_2)^2 - \lambda_1(1 - c_1)^2). \tag{3.4}$$

The literature [44] points out that the Dirac function is narrowly defined, which limits the globality of the edge of the image, and it cannot be detected stably since the target edge away from the closed curve. CV model is only using global area information to optimize c_1 and c_2. It does not take care of the grey value changed in target and the background area of the image. Therefore, CV model cannot segment the image with uneven greyscale. Equation (3.4) contains a curvature term, and it is really complex to achieve. We must use a specific numerical method to maintain numerical stability. In addition, to maintain numerical stability in CV model, the level set function $u(x, y, t)$ needs to be initialized as a symbol distance function, and it need to be reinitialized after each update. Meanwhile, reinitialization may cause the position of zero level set shift. It also increases

the amount of computation and consumes a significant amount of time. The iterative equation is shown in below:

$$H(x) = \begin{cases} \psi_t = sign(u(t))(1 - |\nabla\psi|) \\ \psi(0,\cdot) = u(t,\cdot). \end{cases}$$



(a) Initial contour (b) CV, 200 iterations (c) CER, 200 iterations

Figure 3.1: Results of both CV model and our model for image with low contrast.

Figure 3.1 showes the segmentation result for low contrast (edge blur) images by using CV [42] model and our CER model in the same initial conditions respectively. The results show that the CV model cannot get the correct segmentation results for gray-scale images. Sometimes, it cannot even complete the segmentation. But our CER model can quickly and accurately get the segmentation results.

3.1.2 RSF model

Aimed at the problems of the CV model, Li et al proposed the famous RSF model which is based on variable region fitting energy functional [40]. It is also called LBF (Local Binary Fitting) model [39]. The RSF model uses two local fitting functions $f_1(x)$ and $f_2(x)$ to approximate the average gray-value of the local area on both sides of the target.

For any point $x \in \Omega$ in the image area and contour curve C, and the fitting energy is

$$E_x^{RSF}(f_1, f_2, C) = \lambda_1 \int_{inside(C)} K_\sigma(x-y)|I(y) - f_1(x)|^2 dy + \lambda_2 \int_{outside(C)} K_\sigma(x-y)|I(y) - f_2(x)|^2 dy. \tag{3.5}$$

where λ_1 and λ_2 are two positive constant numbers (usually fixed to constant 1), $f_1(x)$ and $f_2(x)$ are gray fitting values of the image at the point of x. Grey value $I(y)$ is limited to a local area centered on the point x, and the size of this area is controlled by the kernel function K. Thus, equation (3.5) is also referred, in the literature [40], as a regional scale variable fitting (RSF) energy of contour C at point x. $K : R^2 \to R$ is a kernel function [2, 3]. It is satisfied the local property.

1. $K(u) = K(|u|)$;
2. If $|u| < |v|$, then $K(u) \geq K(v)$, meanwhile $\lim_{|u|\to\infty} K(u) = 0$;
3. $\int_{R^2} K(u)du = 1$.

Because the kernel function $K(u)$ has locality (property 2), so $f_1(x)$ and $f_2(x)$ are mainly determined by the nearby gray values. It has "local" characteristics. Obviously, in the RSF model, it is very important to choose the right kernel function. There are many kernel functions that satisfy these local properties. The Gaussian kernel function is used in the original literature [39, 40],

$$K_\sigma(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{|u|^2}{2\sigma^2}}, u \geq 0,$$

where, $\sigma > 0$ is a scale parameter.

Let u be the level set function of the contour curve C. Using the level set method, one has

$$E_x^{RSF}(f_1, f_2, C) = \lambda_1 \int_{inside(C)} K_\sigma(x-y)|I(y) - f_1(x)|^2 H(u) dy + \lambda_2 \int_{outside(C)} K_\sigma(x-y)|I(y) - f_2(x)|^2 (1 - H(u)) dy. \tag{3.6}$$

The total fitting energy function is

$$E^{RSF} = \int_{\Omega} E_x^{RSF} dx.$$

In order to avoid periodic initialization of the level set, the RSF model combines the methods that do not need reinitialization [41]. Adding the set regularization $\int_{\Omega} \frac{1}{2}(|\nabla u| - 1)^2 d\sigma$ term, the total energy functional of the RSF model is

$$\begin{aligned} E_x^{RSF}(f_1, f_2, u) =& \lambda_1 \int_{inside(C)} K_\sigma(x-y)|I(y) - f_1(x)|^2 H(u) dy dx \\ &+ \lambda_2 \int_{outside(C)} K_\sigma(x-y)|I(y) - f_2(x)|^2 (1 - H(u)) dy dx \\ &+ \nu \int_{\Omega} |\nabla H(u(x))| dx + \mu \int_{\Omega} \frac{1}{2}(|\nabla u(x)| - 1)^2 dx. \end{aligned} \tag{3.7}$$

Firstly fixing u in above formula, after minimizing the energy functional $E_x^{RSF}(f_1, f_2, u)$ with respect to $f_1(x)$ and $f_2(x)$, we have

$$f_1(x) = \frac{\int_{\Omega} K_\sigma(x-y) I(y) H\left(u(y)\right) dy}{\int_{\Omega} K_\sigma(x-y) H\left(u(y)\right) dy}, f_2(x) = \frac{\int_{\Omega} K_\sigma(x-y) I(y) \left(1 - H\left(u(y)\right)\right) dy}{\int_{\Omega} K_\sigma(x-y) \left(1 - H\left(u(y)\right)\right) dy}.$$

In fact, the fitting function $f_1(x)$ and $f_2(x)$ are the weighted averages of the image grey values $I(y)$ in area $inside(C)$ and $outide(C)$. The weight is $K_\sigma(x-y)$. Due to the local character of the Gaussian kernel K_σ, the effect of the gray value $I(y)$ on $f_1(x)$ and $f_2(x)$ decreases as the distance between y and x increases. When $|x-y| > 3\sigma$, the effect of $I(y)$ is almost reduced to zero. Therefore, the values of the fitting functions $f_1(x)$ and $f_2(x)$ are determined primarily by the gray value of point y in $\{y : |x-y| \leq 3\sigma\}$.

Then fixing the fitting factors $f_1(x)$ and $f_2(x)$, after minimizing the energy functional E^{RSF} with respect to u, the evolution equation of the energy functional E^{RSF} is obtained using a gradient

descent flow:

$$\frac{\partial u}{\partial t} = -\delta_\varepsilon(u)(\lambda_1 e_1 - \lambda_2 e_2) + \mu\left[\nabla^2 u - div\left(\frac{\nabla u}{|\nabla u|}\right)\right] + \nu\delta_\varepsilon(u) div\left(\frac{\nabla u}{|\nabla u|}\right), \tag{3.8}$$

where

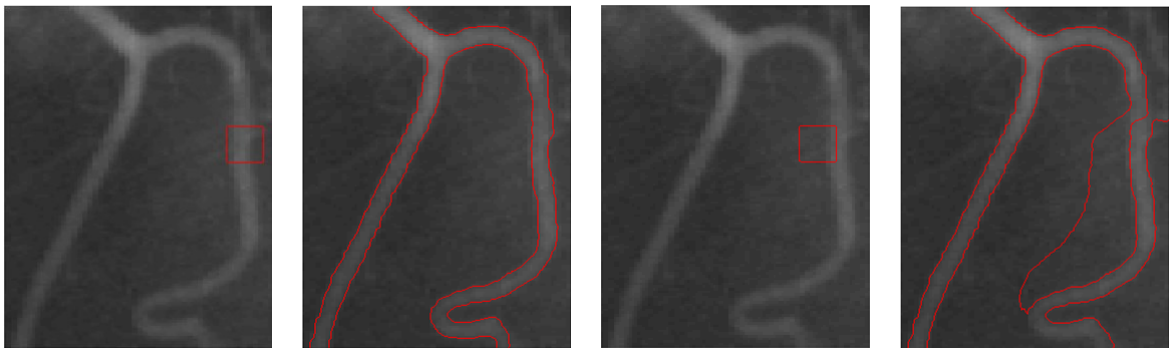
$$e_1 = \int_\Omega K_\sigma(x-y)|I(x) - f_1(y)|^2 dy, \qquad e_2 = \int_\Omega K_\sigma(x-y)|I(x) - f_2(y)|^2 dy.$$

The RSF model is the energy functional, which is defined by the local binary fitting energy based on Gaussian kernel function. The functions $f_1(x)$ and $f_2(x)$ are a local energy which are related to variance. Therefore, the RSF model can better segment the image with intensity inhomogeneity. In addition, to avoid periodic initialization of the level set function, the RSF model has combined models with no reinitialization [41]. It greatly accelerates the evolution rate.

Although the RSF model can better segment the image with intensity inhomogeneity, the RSF model is still very sensitive to the initial position of active contour. When the initial contour selection is not appropriate, it is easy to fall into the local minimum value. Then it will produce the wrong segmentation result. This phenomenon is essentially caused by the local characteristics of the RSF model. The fitting functions $f_1(x)$ and $f_2(x)$ in the model are fitted to the local grayscale of the image. Therefore, the different fitting values obtained at different locations results in that the RSF model being sensitive to the initial contour position.

Figure 3.2 is a X-ray vessel image. When the initial contour is chosen as the red rectangle in Figure 3.2(a), the RSF [43] model has the right segment result of the red curve shown in Figure 3.2(b). Meanwhile, we also have the correct segmentation curve (red curve). It is the local minimum value of the energy functional $E_x^{RSF}(f_1, f_2, u)$. The corresponding fitting functions $f_1(x)$ and $f_2(x)$ for each point on the red curve are the best grayscale approximations of the two regions where the neighbourhood is divided by the curve, except that the two fitting functions are very close. However, when we move the initial contour slightly to the left, we get the red rectangle shown in Figure 3.2(c). The segment result by using RSF model is shown in the red curve in Figure 3.2(d)

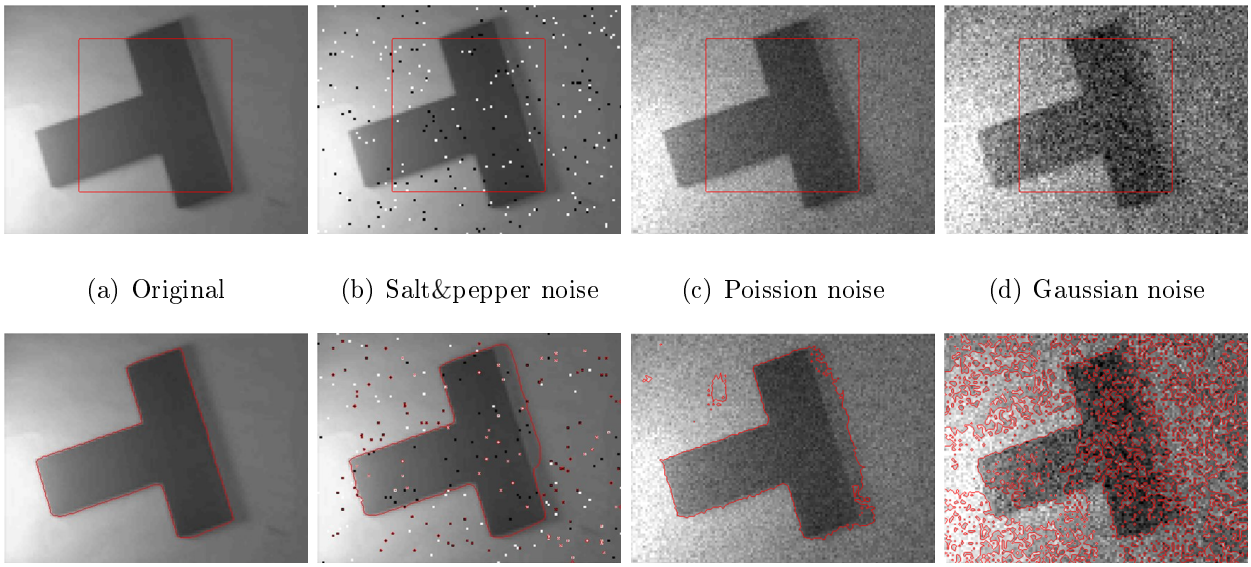
that some of the evolved contours are stuck in the background area and we cannot get the correct target boundary. However, if we change the coordinate $(47:60, 86:99)$ of the initial contour in Figure 3.2(a) to $(47:60, 77; 90)$ in Figure 3.2(c), we get a different result. It is also explained the sensitivity of the RSF model to the initial contour position.



(a) Initial contour (b) 220 iterations (c) Initial contour (d) 500 iterations

Figure 3.2: Sensitivity of the RSF model to the initial contour locations.

RSF model is very sensitive to the noise. When the noise intensity in the image is high, it is difficult to get the correct segmentation result, as shown in Figure 3.3 of the four real T-type images. Figure 3.3(a) is the original uneven grayscale image. Figure 3.3(b), (c) and (d) are the images that obtained by adding salt & pepper, poission with noise density of 0.02 and the Gaussian noise with mean value of 0 and variance of 0.01. The initial contour is selected as the rectangle with the same size and same location (as shown in the red line). From the segmentation result we can see that for Figure 3.3(a) without noise, the RSF model can obtain the correct result as shown in Figure 3.3(e); For Figure 3.3(b) with weaker noise, the RSF can obtain the approximate correct segmentation result shown in Figure 3.3(f); However, for Figure 3.3(c) and Figure 3.3(d) with strong noises, the contours evolving stays at the non-target boundary because of the noise interference or even the segmentation cannot be completed, as shown in the red curves in Figure 3.3(g) and Figure 3.3(h). There is not a clear conclusion of the reason for the problem caused by RSF model shown in above. One of the possible reasons is that energy $E_x^{RSF}(f_1, f_2, C)$ only considers the spatial information of the image pixel without considering its grayscale variation information.



(a) Original (b) Salt&pepper noise (c) Poission noise (d) Gaussian noise

(e) 300 iterations (f) 300 iterations (g) 300 iterations (h) 300 iterations

Figure 3.3: Sensitivity of the RSF model to noise.

Based on the advantages and disadvantages analysis of the CV model and the RSF model above, we combined with information entropy and created a new local fitting energy functional to replace the ones in the CV and the RSF models. Finally, we proposed a new CER model to improve the CV model in which it cannot segment the image with uneven gray scale and the RSF model in which it is sensitive to the initial position and strong noise.

3.2 The CER model based on information entropy

3.2.1 The definition of information entropy

Entropy is the degree of chaos in the system. It has important applications in the fields of cybernetics, number theory, astrophysics, life sciences and so on. There are also more specific definitions in different disciplines. Entropy is an important parameter in various fields. Entropy is proposed by Rudolf Clausius. It is first applied in thermodynamics. Later Shannon proposed the concept of information entropy by generalizing Boltzmann's formula [61]. The information entropy is described the uncertainty of the source. It is the average amount of information for all targets in the source. The amount of information is the central concept of information theory, the entropy

as a random event of uncertainty or the amount of information measurement. The information entropy is considered from the statistical characteristics of the whole source, and it is the totality of the information source from the mean. The greater the uncertainty of the variable, the greater the entropy. The introduction of entropy has established the scientific theoretical foundation of modern information theory, which greatly promoted the development of information theory. In 1972, Freden first introduced the concept of entropy into image processing [24]. It has become a powerful tool in image segmentation right now. The image entropy is expressed as the number of bits of the image gray set, and it is described the average information of the image information source. Based on Shannon's information theory, if p_i is a given distribution of a given image $I = (I_i)$, then image entropy is:

$$E_I = -\sum_i p_i log_2 p_i, \tag{3.9}$$

where the distribution function p_i of the image is the probability that a certain gray scale appears in the image. The distribution function p_i can be defined in a variety of ways, it is depending on the method used to segment the image. For example, in the edge extraction, in order to consider the gray scale of the image, Shiozaki [64] defined the p_i as the grey scale distribution $p_i = I_i / \sum_j I_j$ in the image, where I_i is the greyscale of pixel i. In the image thresholding method [60], distribution function p_i can be approximately obtained from a given histogram. In this section, our distribution function is obtained through the image histogram.

The image segmentation method based on entropy is to minimize the loss of image information. Therefore, we can use image entropy to repair background. Meanwhile, the image entropy reflects the degree of dispersion of the image gray scale. When the gray distribution of the image is relatively even, the value of entropy is large. However, the entropy is small when the image gray distribution has bigger discreteness. Entropy is the result of the interaction of all the pixels in the image. It is not sensitive to single point noise. Therefore, it has a certain anti-noise and filtering capabilities. While the image entropy is a value, the local entropy is a matrix. Compared to the convolution operation on the matrix, the image entropy saves a lot of time.

3.2.2 The effect of adding information entropy to RSF model

As already pointed out before, the CV model is a model based on the global property of the image. The main problem of the CV model is that it cannot segment the uneven grey scale image. The main reason is that the CV model does not consider the grey value changed in the target and the background area of the image. However, the RSF model (LBF model) is a model based on local information of the image. It can effectively segment the uneven grey scale image. Such as MRI image. But the localized property of the model makes the model easy to fall into the local extremum, which leads to the model being very sensitive to contour initialization. Meanwhile, RSF model does not have a strong robustness to noise. The main reason for these problems is that the gray value change of the image is not considered when the local fitting energy $E_x^{RSF}(f_1, f_2, C)$ is defined. In order to overcome the above shortcomings of the RSF model, we introduce the concept of information entropy (see equation (3.9)) in this paper, the energy functional in the RSF model is rewritten as follows:

$$E^{ZRSF}(C, f_1, f_2) = \int_\Omega E_r(x) E_x(C, f_1(x), f_2(x)) dx. \tag{3.10}$$

A new model is formed by combining the global energy and the local energy. In equation (3.10), $E_r(x) = E(x, B(x, r))$ is the local information entropy of point $x \in \Omega$, where $B(x, r) = \{y : |x-y| \leq 0\}$, $r > 0$ is the circle centered on point x with radius r.

In addition, the smoothness of the image contour C is ensured by controlling its length $|C|$, So combining with the above formula, the new energy functional can be written as the following:

$$\begin{aligned} E^{ZRSF1}(C, f_1, f_2) &= E^{ZRSF}(C, f_1, f_2) + \nu|C| \\ &= \sum_{i=1}^{2} \lambda_i \int_\Omega E_r(x) \left(\int_\Omega K_\sigma(x-y) |I(y) - f_i(x)|^2 dy \right) dx + \nu|C|. \end{aligned} \tag{3.11}$$

Using Heaviside function, the energy function can be expressed as follows:

$$
\begin{aligned}
E^{ZRSF1}(u, f_1, f_2) &= E^{ZRSF}(u, f_1, f_2) + \nu \int_{\Omega} |\nabla H(u(x))| dx \\
&= \sum_{i=1}^{2} \lambda_i \int_{\Omega} E_r(x) \left(\int_{\Omega} K_\sigma(x-y) |I(y) - f_i(x)|^2 M_i(u(y)) dy \right) dx \\
&+ \nu \int_{\Omega} |\nabla H(u(x))| dx,
\end{aligned} \tag{3.12}
$$

where $M_1 = H(u)$, $M_2 = (1 - H(u))$, $\int_{\Omega} |\nabla H(u(x))| dx$ is the length of the contour C (i.e. the zero level set of u), and $\delta(\cdot)$ is the derivative of $H(\cdot)$.

Meanwhile, in order to avoid the periodic initialization of the level set function, we consider an internal energy term for the level set function in the energy functional $E^{ZRSF1}(u, f_1, f_2)$ of equation (3.12):

$$
P(u) = \int_{\Omega} \frac{1}{2} |\nabla u(x) - 1|^2 \, dx. \tag{3.13}
$$

When $|\nabla u(x)| = 1$, $P(u)$ is minimized. That is the level set function is required to remain as a sign distance function as much as possible during the evolution. Thus the level set of energy functional based on information entropy is

$$
\begin{aligned}
E^{ZRSF1}(u, f_1, f_2) &= E^{ZRSF}(u, f_1, f_2) + \nu \int_{\Omega} \delta(u(x)) |\nabla u(x)| dx + \mu \int_{\Omega} \frac{1}{2} |\nabla u(x) - 1|^2 \, dx \\
&= \sum_{i=1}^{2} \lambda_i \int_{\Omega} E_r(x) \left(\int_{\Omega} K_\sigma(x-y) |I(y) - f_i(x)|^2 M_i(u(y)) dy \right) dx \\
&+ \nu \int_{\Omega} \delta(u(x)) |\nabla u(x)| dx + \mu \int_{\Omega} \frac{1}{2} |\nabla u(x) - 1|^2 \, dx,
\end{aligned} \tag{3.14}
$$

where $\mu > 0$ is weight parameter.

3.2.3 CER model

The weighted ZRSF model based on local entropy is proposed above to enhance the robustness of the original model to the initial contour position and strong noise. In fact, In the image with

uneven grayness, due to the heterogeneity of gray scale, the gray scale difference between different regions is often different. Therefore, in order to better deal with the intensity inhomogeneity phenomenon, considering the advantages and the disadvantages of the CV model and the RSF model [69], we create a new local fitting energy functional to replace the ones in the CV and the LBF models. The new energy functional is defined as:

$$E^{CER}(C, c_1, c_2, f_1, f_2) = \omega E^{CV}(C, c_1, c_2) + (1-\omega)E^{ZRSF}(C, f_1, f_2), \tag{3.15}$$

where the range of ω is $0 \leq \omega \leq 1$. constants c_1 and c_2 are the average gray values in the inside and the outside of evolution curve of the original image area, respectively, and f_1 and f_2 are the fitting value of the image at the point x. Using Heaviside function $H(u)$, the level set of energy function is:

$$\begin{aligned} E^{CER}(u, c_1, c_2, f_1, f_2) &= \omega E^{CV}(u, c_1, c_2) + (1-\omega)E^{ZRSF}(u, f_1, f_2) \\ &+ \nu \int_\Omega \delta(u(x))|\nabla u(x)|dx + \mu \int_\Omega \frac{1}{2}\left|\nabla u(x) - 1\right|^2 dx \\ &= \omega \left[\lambda_1 \iint_{insides(C)} |I(x,y) - c_1|^2 dxdy + \lambda_2 \iint_{outside(C)} |I(x,y) - c_2|^2 dxdy \right] \\ &+ (1-\omega)\sum_{i=1}^{2} \lambda_i \int_\Omega E_r(x) \left(\int_\Omega K_\sigma(x-y)|I(y) - f_i(x)|^2 M_i(u(y))dy \right) dx \\ &+ \nu \int_\Omega \delta(u(x))|\nabla u(x)|dx + \mu \int_\Omega \frac{1}{2}|\nabla u(x) - 1|^2 dx. \end{aligned} \tag{3.16}$$

3.2.4 Euler-Lagrange equation in CER model

We will use the variational idea to transfer the energy functional minimization problem to finding the steady–state solution of the corresponding Euler - Lagrange Equation.

Based on the information entropy in equation (3.14), the energy functional is about u, f_1 and f_2 variables, and the functions f_1 and f_2 are defined by the inner and outer regions determined by u. Therefore, when the energy function is minimized, u, f_1 and f_2 will be fixed respectively.

First, in the case of fixed u, the following theorem holds for the functions f_1 and f_2.

Theorem 3.1, for the fixed u, the necessary condition for the energy function $E^{ZRSF}(u, f_1, f_e)$ to reach the minimum is that f_1 and f_2 satisfy the following: For a.e. $x \in \Omega$, $f_i(x)\int_\Omega K_\sigma(x-y)M_i(u(y))dy = \int_\Omega K_\sigma(x-y)I(y)M_i(u(y))dy,\ i = 1, 2$. Thus

$$f_i = \frac{\int_\Omega K_\sigma(x-y)I(y)M_i(u(y))dy}{\int_\Omega K_\sigma(x-y)M_i(u(y))dy},\ \ i = 1, 2. \tag{3.17}$$

Proof: For fixed u and f_2, taking the derivative of f_1 for the energy functional $E^{ZRSF}(u, f_1, f_2)$, we have,

$$\frac{\partial E^{ZRSF}}{\partial f_1} = -2\int_\Omega K_\sigma(x-y)[I(y) - f_1(x)]M_1(u(y))dy.$$

The necessary condition for obtaining the minimum value of the function is $\frac{\partial E^{ZRSF}}{\partial f_1} = 0$, therefore,

$$\begin{gathered}\int_\Omega K_\sigma(x-y)[I(y) - f_1(x)]M_1(u(y))dy = 0 \\ \Rightarrow f_1(x)\int_\Omega K_\sigma(x-y)M_1(u(y))dy = \int_\Omega K_\sigma(x-y)I(y)M_1(u(y))dy \\ \Rightarrow f_1(x) = \frac{\int_\Omega K_\sigma(x-y)I(y)M_1(u(y))dy}{\int_\Omega K_\sigma(x-y)M_1(u(y))dy}.\end{gathered}$$

The same, the expression for a function f_2 can be proved similarly.

Now in the case of fixed f_1 and f_2, the model (3.14) can be transformed into a form represented by only the level set function u,

$$\begin{aligned}E^{ZRSF2}(u) &= E^{ZRSF}(u) + \nu L(u) + \mu P(u) \\ &= \sum_{i=1}^{2}\lambda_i\int_\Omega E_r(x)\left(\int_{\Omega_i} K_\sigma(x-y)|I(y) - f_i(x)|^2 M_i(u(y))dy\right)dx \\ &+ \nu\int_\Omega \delta(u(x))|\nabla u(x)|dx + \mu\int_\Omega \frac{1}{2}|\nabla u(x) - 1|^2 dx.\end{aligned} \tag{3.18}$$

In order to solve the minimization problem (3.18), we first deduce the Euler-Lagrange equation

corresponding to the function, as follows:

Theorem 3.2, assume $W^{1,2}(\Omega)$, the necessary conditions for the energy functional (3.18) to reach the minimum is

$$\begin{cases} \delta_\varepsilon(u)(\lambda_1 e_1 - \lambda_2 e_2) - \nu\delta_\varepsilon(u) div\left(\frac{\nabla u}{|\nabla u|}\right) - \mu\left[\nabla^2 u - div\left(\frac{\nabla u}{|\nabla u|}\right)\right] = 0, & u \in \Omega \\ \frac{\partial u}{\partial n} = 0, & u \in \partial\Omega \end{cases} \tag{3.19}$$

where

$$e_i = e_i(x) = \int_\Omega K_\sigma(x-y)E_r(y)|I(x) - f_i(y)|^2 dy, \quad i = 1, 2. \tag{3.20}$$

The second equation is called the Neumann boundary condition, and n is the outer normal vector of the boundary $\partial\Omega$.

Proof: The necessary condition for obtaining the extremum from the function is known, the necessary condition for $E^{ZRSF2}(u)$ minimization is that the corresponding functional's Gâteaux derivative is zero. For any test function $h \in C^\infty(\Omega)$,

$$\frac{dE^{ZRSF2}(u)}{du} = \lim_{\tau\to 0}\frac{E^{ZRSF2}(u+\tau h) - E^{ZRSF2}(u)}{\tau} = 0.$$

Considering the linear property of the derivative of the Gâteaux, we can rewrite the Gâteaux derivative of the energy functional $E^{ZRSF2}(u)$ to the following:

$$\frac{dE^{ZRSF2}(u)}{du} = \frac{dE^{ZRSF}(u)}{du} + \nu\frac{dL(u)}{du} + \mu\frac{dP(u)}{du}. \tag{3.21}$$

Now we derive the Gâteaux derivative of the energy functional E^{ZRSF2}, $L(u)$, $P(u)$, respectively. Firstly, we derive the Gâteaux derivative of the energy functional $E^{ZRSF}(u)$ in the h direction at u. So we first exchange the variables x and y in the functional, that is $x = y$, $y = x$. Then we

have

$$E^{ZRSF}(u) = \sum_{i=1}^{2} \lambda_i \int_{\Omega} E_r(y) \left(\int_{\Omega_i} K_\sigma(y-x)|I(x) - f_i(y)|^2 M_i(u(x))dx \right) dy.$$

Changing the order of the integral and pay attention to that $K_\sigma(-u) = K_\sigma(u)$, we obtain

$$\begin{aligned} E^{ZRSF}(u) &= \sum_{i=1}^{2} \lambda_i \int_{\Omega} M_i(u(x)) \left(\int_{\Omega_i} E_r(y) K_\sigma(y-x)|I(x) - f_i(y)|^2 dy \right) dx \\ &= \sum_{i=1}^{2} \lambda_i \int_{\Omega} M_i(u(x)) e_i(x) dx. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \frac{dE^{ZRSF}(u)}{du} &= \lim_{\tau \to 0} \frac{E^{ZRSF}(u+\tau h) - E^{ZRSF}(u)}{\tau} \\ &= \lim_{\tau \to 0} \sum_{i=1}^{2} \lambda_i \int_{\Omega} \frac{M_i(u+\tau h)(x) - M_i(u(x))}{\tau h} h e_i(x) dx \\ &= \int_{\Omega} \delta_\varepsilon(u)[e_1(x) - e_2(x)] h dx. \end{aligned} \tag{3.22}$$

And then derive the Gâteaux derivative of the energy functional $L(u)$. The detail is show in the following.

$$\begin{aligned} \frac{dL(u)}{du} &= \lim_{\tau \to 0} \frac{L(u+\tau h) - L(u)}{\tau} \\ &= \lim_{\tau \to 0} \int_{\Omega} \frac{\delta_\varepsilon(u+\tau h)|\nabla(u+\tau h)| - \delta_\varepsilon(u)|\nabla u|}{\tau} dx \\ &= \lim_{\tau \to 0} \left\{ \frac{[\delta_\varepsilon(u+\tau h) - \delta_\varepsilon(u)]|\nabla u + \tau \nabla h|}{\tau h} h + \frac{\delta_\varepsilon(u)|\nabla u + \tau \nabla h| - \nabla u}{\tau \nabla h} \nabla h \right\} dx \\ &= \int_{\Omega} \delta'_\varepsilon(u) h |\nabla u| dx + \int_{\Omega} \delta_\varepsilon \frac{\nabla u}{|\nabla u|} \nabla h dx \\ &= \int_{\Omega} \delta'_\varepsilon(u) h |\nabla u| dx + \int_{\partial \Omega} \delta_\varepsilon(u) \frac{\nabla u}{|\nabla u|} h \cdot n dS - \int_{\Omega} div \left(\delta_\varepsilon(u) \frac{\nabla u}{|\nabla u|} \right) h dx. \end{aligned}$$

Since second term of the above equation is zero by the zero Neumann boundary condition, so

that the above equation can be simplified as

$$\begin{aligned}\frac{dL(u)}{du} &= \int_\Omega \delta'_\varepsilon(u)h|\nabla u|dx - \int_\Omega div\left(\delta_\varepsilon(u)\frac{\nabla u}{|\nabla u|}\right)hdx \\ &= \int_\Omega \delta'_\varepsilon(u)h|\nabla u|dx - \int_\Omega \delta'_\varepsilon(u)\nabla u\frac{\nabla u}{|\nabla u|}hdx - \int_\Omega \delta_\varepsilon(u)div\left(\frac{\nabla u}{|\nabla u|}\right)hdx \\ &= \int_\Omega -\delta_\varepsilon(u)div\left(\frac{\nabla u}{|\nabla u|}\right)hdx.\end{aligned} \tag{3.23}$$

Finally, we derive the Gâteaux derivative of the energy functional $P(u)$.

$$\begin{aligned}\frac{dP(u)}{du} &= \lim_{\tau\to 0}\frac{P(u+\tau h)-P(u)}{\tau} \\ &= \lim_{\tau\to 0}\frac{1}{2}\int_\Omega \frac{(|\nabla(u+\tau h)|-1)^2-(|\nabla u|-1)^2}{\tau}dx \\ &= \lim_{\tau\to 0}\frac{1}{2}\int_\Omega \left\{\frac{(|\nabla(u+\tau h)|-|\nabla u|)(|\nabla(u+\tau h)|+|\nabla u|-2)}{\tau}\right\}dx \\ &= \lim_{\tau\to 0}\frac{1}{2}\int_\Omega (|\nabla(u+\tau h)|+|\nabla u|-2)\frac{(|\nabla(u+\tau h)|-|\nabla u|)}{\tau\nabla h}\nabla h dx \\ &= \int_\Omega (|\nabla u|-1)\frac{\nabla u}{|\nabla u|}\nabla h dx = \int_{\partial\Omega}(|\nabla u|-1)\frac{\nabla u}{|\nabla u|}h\cdot ndS - \int_\Omega\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right)hdx.\end{aligned}$$

Since the first term of the above equation is zero by the zero Neumann boundary condition, so that the above equation can be simplified as

$$\frac{dP(u)}{du} = -\int_\Omega\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right)hdx. \tag{3.24}$$

Combining the equation (3.22), equation (3.23) and equation (3.24), the Gâteaux derivative of $E^{ZRSF2}(u)$ can eventually be expressed as

$$\begin{aligned}\frac{dE^{ZRSF2}(u)}{du} &= \frac{dE^{ZRSF}(u)}{du} + \nu\frac{dL(u)}{du} + \mu\frac{dP(u)}{du} \\ &= \int_\Omega\left\{\delta_\varepsilon(u)[e_1(x)-e_2(x)] - \nu\delta_\varepsilon(u)div\left(\frac{\nabla u}{|\nabla u|}\right) - \mu\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right)\right\}hdx.\end{aligned} \tag{3.25}$$

To make all $h \in C^{\infty}(\Omega)$, the value of the above formula is zero. Thus

$$\delta_{\varepsilon}(u)[e_1(x) - e_2(x)] - \nu\delta_{\varepsilon}(u)div\left(\frac{\nabla u}{|\nabla u|}\right) - \mu\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right) = 0.$$

Theorem 3.2 is proved.

Then we use the steepest descent method to convert the Euler-Lagrange equation (3.19) into the corresponding gradient descent flow equation. Introducing the "time" auxiliary variable $t \geq 0$, and based on the variational method (see section 2.1.2), the gradient descending flow equation corresponding to the variational problem equation (3.18) is

$$\begin{aligned}\frac{\partial u}{\partial t} &= -\frac{dE^{ZRSF2}(u)}{du} \\ &= -\delta_{\varepsilon}(u)[e_1(x) - e_2(x)] + \nu\delta_{\varepsilon}(u)div\left(\frac{\nabla u}{|\nabla u|}\right) + \mu\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right),\end{aligned} \tag{3.26}$$

where $e_1(x)$ and e_2 are shown as equation (3.20). This is the gradient downflow equation that considers the information entropy of the model (3.14). It is also the level set evolution equation.

Combining with the gradient descent flow equation (3.4) of the CV model and the gradient descent flow equation (3.26) which considers the information entropy model (3.14), the level set evolution equation of the CER model (3.16) in this dissertation is:

$$\frac{\partial u}{\partial t} = -\delta_{\varepsilon}(u)(F_1 + F_2) + \nu\delta_{\varepsilon}(u)div\left(\frac{\nabla u}{|\nabla u|}\right) + \mu\left(\nabla^2 u - div\frac{\nabla u}{|\nabla u|}\right), \tag{3.27}$$

where

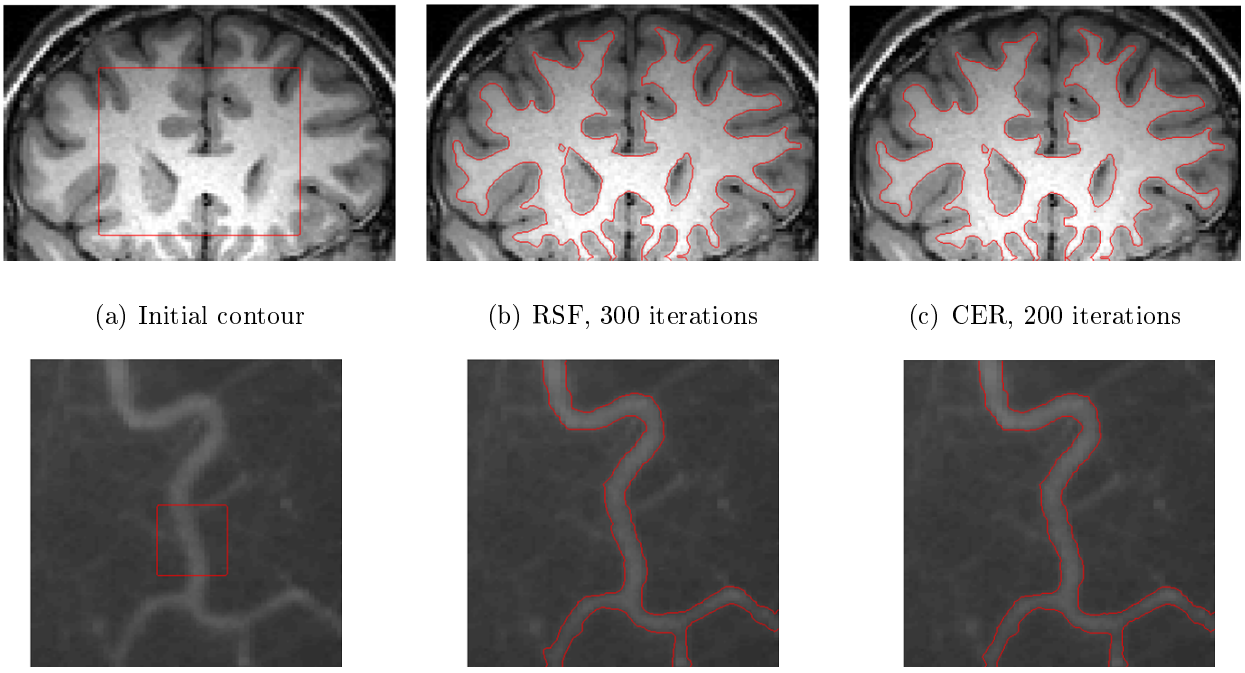
$$F_1 = \omega(\lambda_1(I(x) - c_1)^2 - \lambda_2(I(x) - c_2)^2)$$
$$F_2 = (1-\omega)\left(\lambda_1\int_{\Omega} K_{\sigma}(x-y)E_r(y)(I(y) - f_1(y))^2 dy - \lambda_2\int_{\Omega} K_{\sigma}(x-y)E_r(y)(I(y) - f_2(y))^2 dy\right).$$

It considers the image boundary and the regional influences at the same time.

3.3 Research result

With this model, several groups of images are segmented by using the finite difference method and MATLAB. The experimental results show that the proposed CER method not only segment the image with uneven gray scale better and faster, but also has stronger robustness to the initial contour position and strong noise than the original RSF model.

3.3.1 Segmentation time



(a) Initial contour (b) RSF, 300 iterations (c) CER, 200 iterations

(d) Initial contour (e) RSF, 150 iterations (f) CER, 50 iterations

Figure 3.4: Segmentation results of RSF and CER model for two typical images with intensity inhomogeneity.

Figure 3.4 shows the segmentation results of two typical gray inhomogeneous images which both are used in the literature [24]. Figure 3.4(a) is a human brain Magnetic Resonance Imaging (MRI) (119×78). Due to the nonuniform magnetic field generated by the radio-frequency coil, the gray scale is uneven. From Figure 3.4(a) we can see that the brightness of some places above the white matter is lower than the brightness of some places below the gray matter. Figure 3.4(d)

is the x-ray blood vessel image (111×110). We can see that the brightness of some parts of the background is even higher than the brightness of certain parts of the blood vessels, and some of the margins of the blood vessels are weak. The presence of these grayscale inhomogeneity increases the difficulty of segmentation. In Figure 3.4(a) and Figure 3.4(d) show the two original images and initial contours. The initial contour is selected as the position shown by the red box in the image. The Figure 3.4(b) and Figure 3.4(e) is the segmentation result by using RSF model. Figure 3.4(c) and Figure 3.4(f) are the segmentation results by using new CER model shown by the red curve.

In order to qualitatively evaluate the performance of RSF model and CER model [74], we use the Dice similarity coefficient (DSC) [62], FNR, FPR [20] and RES [49] four regional overlap measures. Assume that S_1 represents a reference foreground area for a given image (such as the real boundaries), and S_2 represents the foreground area obtained by model segmentation. Then the above four measures can be defined as

$$DSC = \frac{2N(S_1 \cap S_2)}{N(S_1) + N(S_2)}, \quad FNR = \frac{N(S_1 \setminus S_2)}{N(S_1)},$$
$$FPR = \frac{N(S_2 \setminus S_1)}{N(S_2)}, \quad RSE = \frac{N(S_2 \setminus S_1) + N(S_1 \setminus S_2)}{N(\Omega)},$$

where $N(\cdot)$ represents the number of pixels in a closed area, and Ω is the area of the image. The more the value of DSC is closer to 1, the closer the value of FNR, FPR, and RSE are to 0, the better the segmentation results we have. The segmentation results of the two images are obtained

Table 3.1: DSC, FNR, FPR, and RSE values for the images in Figure 3.4.

Image	DSC	FRN	FPR	RSE
Figure 3.4(a)	0.9705	0.0572	0	0.0053
Figure 3.4(d)	0.9923	0.0142	0	0.000

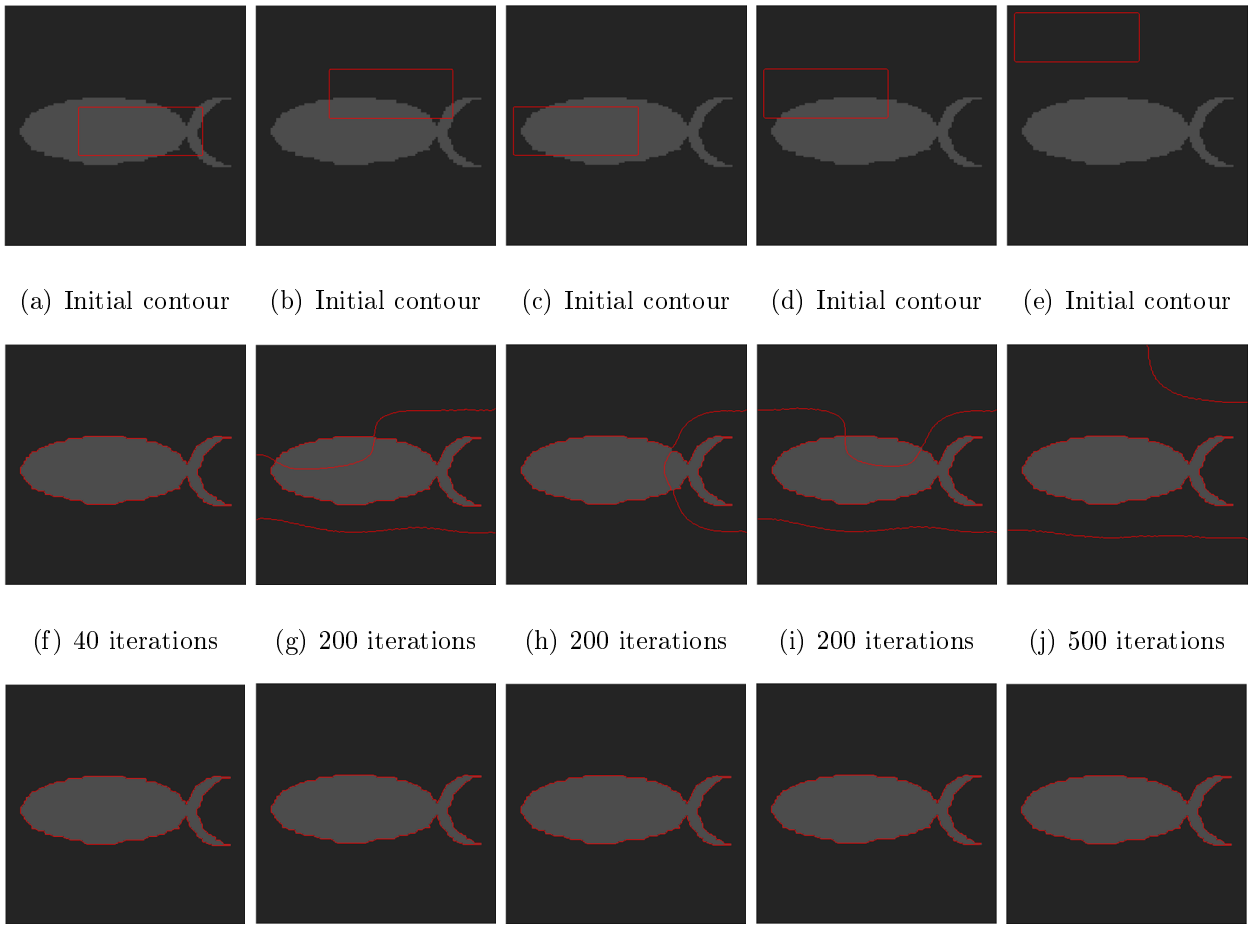
from the CER model and the RSF model respectively using the same parameters of $\lambda_1 = \lambda_2 = 1$, $\mu = 1$, $\Delta t = 0.1$, $h = 1$, $\sigma = 3$, and $\omega = 0.01$ and 0.065 for CER model. It is not difficult to see that the CER model achieves almost the same segmentation results as the RSF model from the figures. Table 3.1 is further confirmed that the CER model has almost the same segmentation result as the

RSF model. However, the CER model used fewer iterations and less segmentation time. The speed of the CER model is approximately 1/2 of the RSF model. It is shown that the CER model has greatly improved the speed of segmentation while ensuring the accuracy.

3.3.2 The comparison of initial contour position

In this section, we will compare the robustness of the RSF model and the CER model to the initial contour position. In each set of experiments, we will select the initial contours of the same size at different locations (As shown in Figure 3.5(a) and the red rectangle of section 3.6(a)). Although the computational speed of the CER model is significantly higher than that of the RSF model, but in order to compare the robustness of the two models, we select the same number of iterations to get the results for comparison. We selected an artificial fish image (Figure 3.5). The graph consists of only two different gray scale levels, representing the target (fish) and the background, respectively; and picture of salpingogram (Figure 3.6). These two images are typical images with grayscale unevenness and low contrast.

Figure 3.5 shows the segmentation results of an artificial fish image using the RSF and CER models at five different initial contour positions. The first line shows the original image and the initial contours of the five different positions (shown in the red rectangle). The second and third line represent the segmentation results of the RSF model and the CER model, respectively. From the above experimental results, it can be seen that the RSF model is sensitive to the initial contour position. It only gets the correct segmentation results in the first initial contour (Figrue 3.5(a)) position. In the other four initial contour position, it cannot get the correct segmentation results. In contrast, the CER model improves the problem that RSF model is sensitive to the initial contour position. CER model achieved satisfactory segmentation results in five different positions. But at the different initial contour position, the evolution speed is different. For example, at the first initial contour position the iteration is 40 times. At the 2-4 initial contour position the iteration is 200 times. At the fifth initial contour position the iteration is 500 times.



(a) Initial contour (b) Initial contour (c) Initial contour (d) Initial contour (e) Initial contour

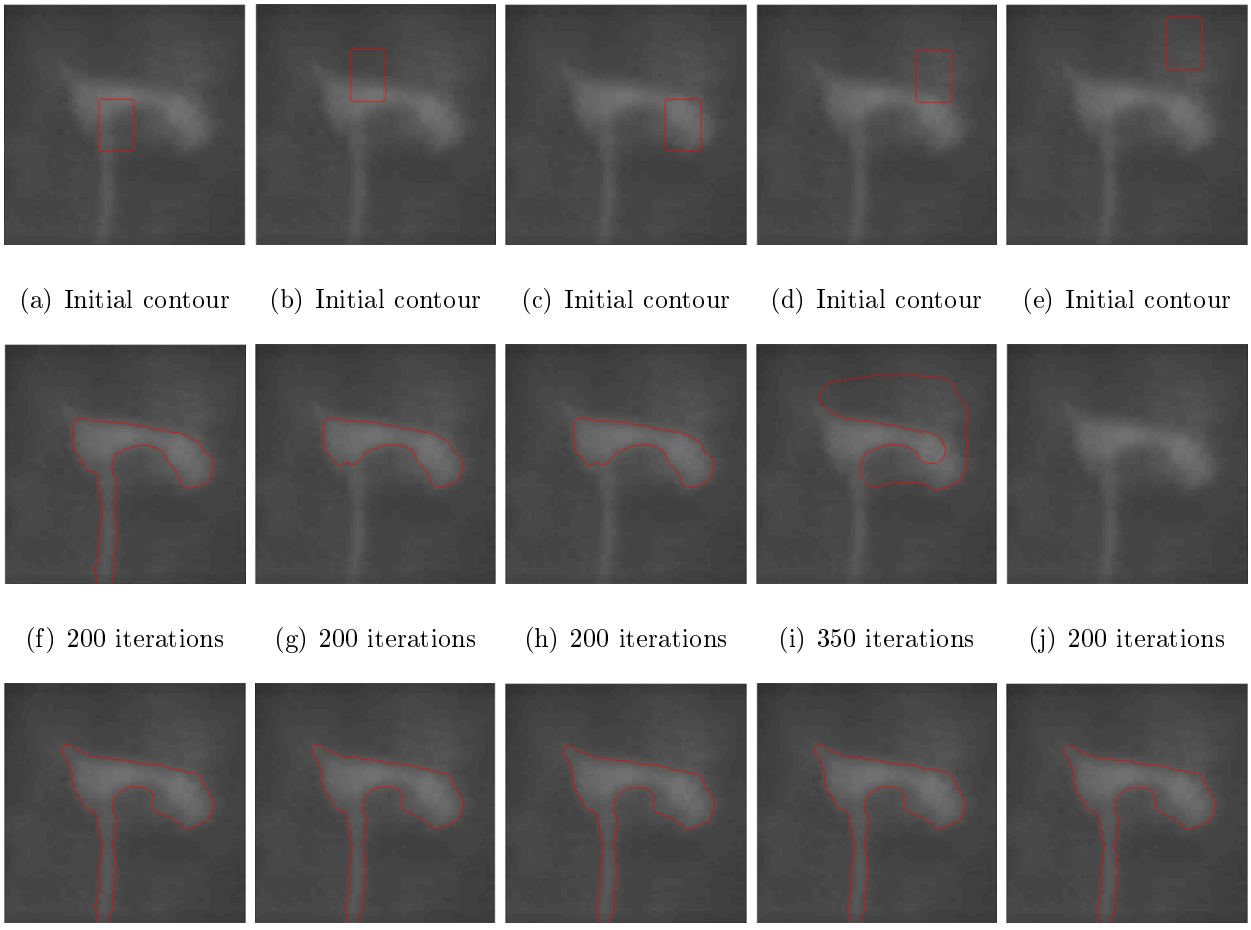
(f) 40 iterations (g) 200 iterations (h) 200 iterations (i) 200 iterations (j) 500 iterations

(k) 40 iterations (l) 200 iterations (m) 200 iterations (n) 200 iterations (o) 500 iterations

Figure 3.5: The segmentation result by using RSF model and CER model for an artificial synthetic image with different initial contours.

Figure 3.6 is the segmentation result by using RSF model and CER model for a real salpingographic image, respectively. The first line is the original image and the initial contour. The second and third line represent the segmentation results of the RSF model and the CER model, respectively. This image is a typical low contrast image, the target is very close to the background, and the target edge is very weak. The same as in Figure 3.5, we selected five initial contours of the same size and at the different position, as shown in the red rectangle in Figure 3.6. The second line shows the segmentation result of the RSF model. It only gets approximate correct segmentation results in the first initial contour position. In the other four initial contour position, it cannot get

the correct segmentation results. Even in the fifth initial contour position, it cannot complete the segmentation. However, CER model were successfully in extracting the target image in the five initial contour positions and in getting the correct segmentation results.



(a) Initial contour (b) Initial contour (c) Initial contour (d) Initial contour (e) Initial contour

(f) 200 iterations (g) 200 iterations (h) 200 iterations (i) 350 iterations (j) 200 iterations

(k) 200 iterations (l) 200 iterations (m) 200 iterations (n) 350 iterations (o) 200 iterations

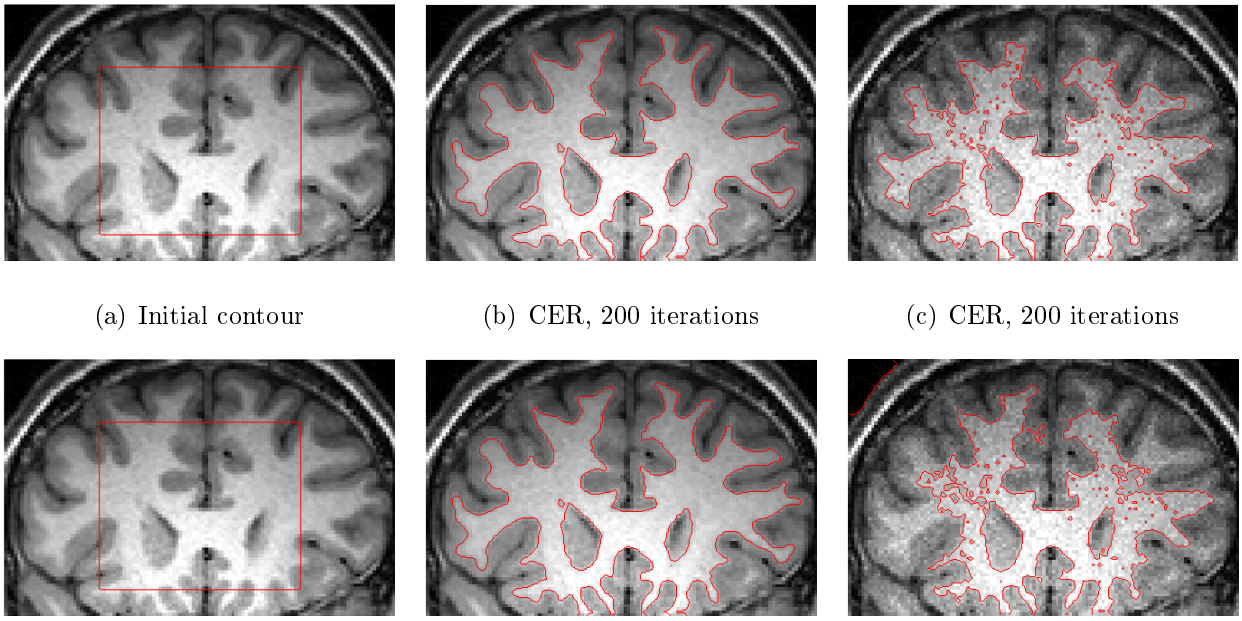
Figure 3.6: The segmentation result by using RSF model and CER model for a real salpingographic image with different initial contour.

From the two sets of experiments above, it is easy to see that the CER model improves the problem that RSF model is sensitive to the initial contour position. Meanwhile, the CER model can obtain a better segmentation result even for the uneven grey scale image or low contrast grayscale image. Because most of the medical images are uneven grayscale or low contrast images, so our

method has a better segmentation effect on this type of medical images. We will make more explanation in the following chapters.

3.3.3 Noise sensitivity degree

In this section, we test the robustness to noise of the CER model. As described in 3.1.2, we know that the RSF model is sensitive to noise. Therefore, we proposed a CER model to improve this problem. The experimental results show that the CER model has strong robustness to noise.



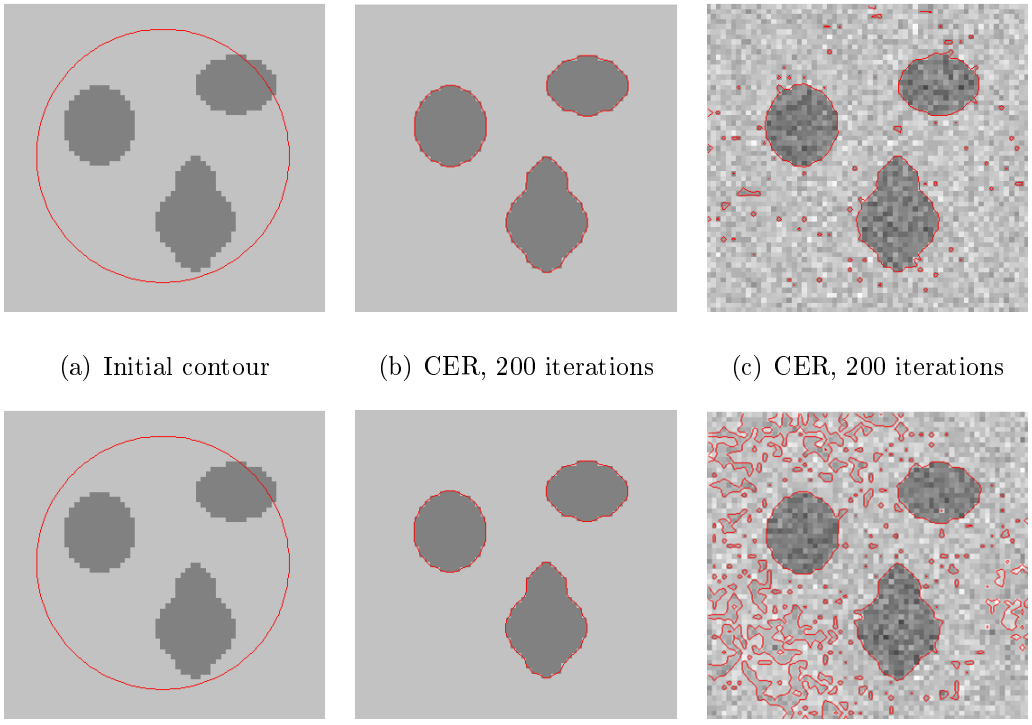
(a) Initial contour (b) CER, 200 iterations (c) CER, 200 iterations

(d) Initial contour (e) RSF, 300 iterations (f) RSF, 300 iterations

Figure 3.7: The segmentation result by using RSF model and CER for a real magnetic resonance images of the brains with noise.

Figure 3.7(a) and Figure 3.7(d) are the original image with the initial contour position (as shown in the red rectangle). Figure 3.7(b) and Figure 3.7(e) are the segmentation results of human brain magnetic resonance imaging (MRI) (119×78) using the CER and the RSF model under the same parameters. It is easy to see that the results are same under both models without noise; Furthermore, the CER model is faster than RSF. Figure 3.7(c) and Figure 3.7(f) are segmentation results with Poisson noise using the two models. We can see from the results that noise has some

effect for both models, but it has more effect for the RSF model.



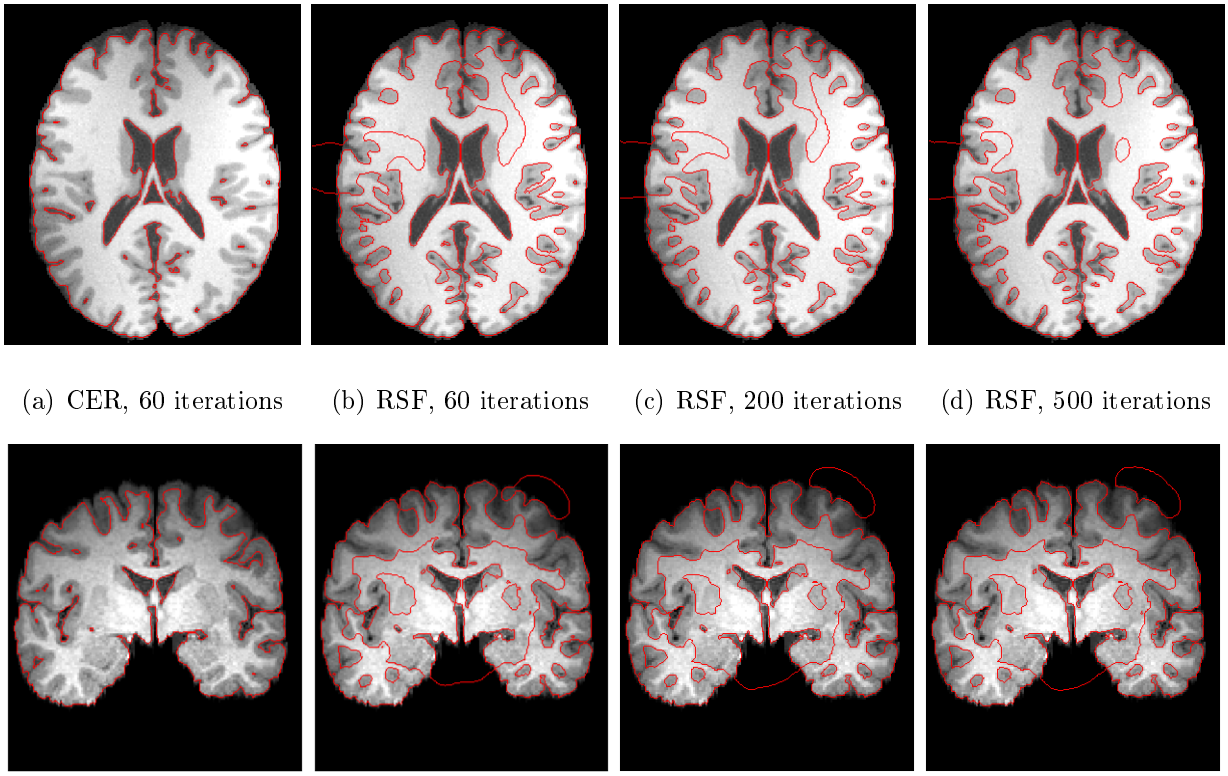
(a) Initial contour (b) CER, 200 iterations (c) CER, 200 iterations

(d) Initial contour (e) RSF, 200 iterations (f) RSF, 200 iterations

Figure 3.8: The segmentation result by using RSF model and CER for a synthetic image with noise.

Figures 3.8(a) and Figure 3.8(d) are the original synthetic images and the initial contours, as shown by the red circular curve. Figure 3.8(b) and Figure 3.8(e) are segmentation results for a simple synthetic image (64×61) using the two models. This image is relatively simple for segmentation because of the target and the background have obvious contrast. Therefore, both two models achieve good results. But CER model is faster than the RSF model. Figure 3.8(c) and 3.8(f) are segmentation results for the image with Gaussian noise (mean value is 0, variance is 0.005). Even though we have segmented some part of the target. It is obvious that CER model obtained less unnecessary segmentation. It means that CER greatly reduces the sensitivity to noise.

3.4 Segmentation for magnetic resonance imaging (MRI)

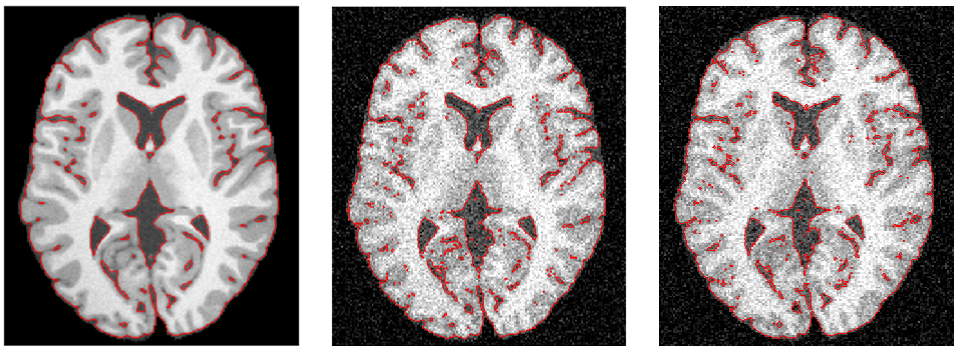


(a) CER, 60 iterations (b) RSF, 60 iterations (c) RSF, 200 iterations (d) RSF, 500 iterations

(e) CER, 80 iterations (f) RSF, 80 iterations (g) RSF, 200 iterations (h) RSF, 500 iterations

Figure 3.9: The segmentation result by using RSF model and CER for a medical image.

Segmentation methods are widely used for dealing with MRI images. Figure 3.9(a) ($189 \times 164 \times 3$) and 3.9(e) (157×141) are two real human brain magnetic resonance images. The segmentation results of the two images are obtained from the CER model and the RSF model with the same parameters of $\lambda_1 = \lambda_2 = 1$, $\mu = 1$, $\Delta t = 0.1$, $h = 1$, $\sigma = 3$, and $\omega = 0.35$ in our CER model. Figure 3.9 (b, c, d) and Figure 3.9 (f, g, h) are segmentation results obtained by using the RSF model with 60, 200 and 500 iterations respectively. The segmentation results show that our CER model has better segmentation ability for complex brain MRI images. Our model ensures the accuracy of segmentation and improves segmentation speed. RSF model cannot get accurate segmentation results on complex images, and therefore it shows that the CER model works well with variance of grayscale level of the image to a certain extent.



(a) Original (b) 1% Gaussian noise (c) 2% Gaussian noise

Figure 3.10: The segmentation result by using CER model (60 iterations) for medical image with noise.

Figure 3.10(a) shows the segmentation result of a human brain MRI ($189 \times 163 \times 3$) for the CER model. Figure 3.10(b) and 3.10(c) show the results with Gaussian noise with standard deviations of 0.01 and 0.02, respectively. All of these three images were accurately segmented. It can be seen from the experimental results that the noise has a certain impact on our model, but the impact is minimal. It also shows that our CER model not only deals with uneven grayscale images, but also has strong robustness to the initial contour position and noise.

The experimental results illustrate further that the CER model has more advantages on medical images with gray inhomogeneity, weak edge images and complex background images. CER has strong robustness to noise.

Chapter 4
THE IMAGE SEGMENTATION MODEL SCER BASED ON GRAY - SPACE FUZZY C - MEANS CLUSTERING

Clustering is the process of organizing physical or abstract objects into groups consisting of similar objects. Cluster is a collection of data objects. The objects in the same cluster have similarity, and the objects in the different clusters are different. Clustering first classifies objects in a multi-dimensional space, and then automatically divides them into clusters according to character of the degree of affinity between the samples. There are usually two ways to describe the degree of affinity between samples. One is to consider each sample as a point at multidimensional space. In multidimensional coordinates, we determine the distances between points, classes, and use the distances between points to describe the degree of affinity between the samples. The other one is to calculate the similarity coefficient between the samples, and to use a similarity coefficient to describe the degree of affinity between the samples.

4.1 Traditional fuzzy c-means clustering algorithm

Fuzzy c-means is a classic gray-based unsupervised image segmentation clustering algorithm. Meanwhile, it is widely used. It optimizes the objective function, then obtains the membership degree of each sample point to each cluster center, and finally achieves the purpose of automatically classifying the data samples [53].

4.1.1 Fuzzy theory

Fuzzy theory consists of fuzzy set theory, fuzzy logic, fuzzy reasoning, fuzzy control and so on. In 1965, Zadeh put forward the classical fuzzy set theory, and created a new science–fuzzy mathematics. Fuzzy set theory is a generalization based on traditional set theory. In the traditional set theory, an element either belongs to this collection or not. However, for a fuzzy set, each element belongs to a set with relative membership, or it can belong to multiple sets at the same time. For example, "young people", "middle-aged" and "old" are three fuzzy concept. Which class should

50-year-old person should belong to? We can think that the 50-year-old belongs to the "young" level is 0, to "middle-aged" is 0.8, and to "old" is 0.2. Therefore, a 50-year-old person belong to both the "middle-aged" and the "old" [45].

The fuzzy set theory reflects the degree of uncertainty that the object belongs to different categories, thus objectively reflecting the real world. In order to describe the fuzzy phenomenon, the discrete points 0 and 1 can be extended to the continuous state interval $[0, 1]$. In this way, the feature function in the general set is extended into the membership function of the fuzzy set. Thus, we can give the definition of fuzzy sets: Give a domain U, then a mapping $\mu_A : x \in U \to [0, 1]$ from U to unit interval $[0, 1]$ is called a fuzzy set on U, denoted as A. The mapping $\mu_A(\cdot)$ or $A(\cdot)$ is called the membership function of the fuzzy set A. $\mu_A(x)$ is called the membership degree of element x to fuzzy set A.

From the above definition, we can see that the fuzzy set A is all described by its membership function. The membership function μ_A maps each element x in U to a value on $[0, 1]$, indicating the degree of the element belongs to A, the greater the value is the better the degree of membership. When the value of $\mu_A(x)$ is only 0 or 1, the fuzzy set A becomes an ordinary set. In general, if the domain is discrete and finite $U = \{x_1, x_2, \ldots, x_n\}$, its fuzzy sets have the following representations:

1. Zadeh's: $A = \sum_{i=1}^{n} \frac{\mu_A(x_i)}{x_i}$. Denominator x_i is the element of the domain, the numerator $\mu_A(x_i)$ is membership degree of the element x_i to A. Sometimes, if the membership is 0, the item can be ignored;

2. Ordered pair method: $A = \{(x_1, \mu_A(x_1)), (x_2, \mu_A(x_2)), \ldots, (x_n, \mu_A(x_n))\}$, the former is the element of the domain, the latter is the membership degree of the element;

3. Vector Method: To specify the elements of a finite field an order of expression, the above ordered pair method can be abbreviated as a vector of membership: $A = (\mu_A(x_1), \mu_A(x_2), \ldots, \mu_A(x_n))$.

4.1.2 Theoretical basis of c-means clustering algorithm

C-means clustering is also called k-means clustering. It has been successfully applied to various fields. The main idea is shown in below: Suppose that the samples in the sample set can be divided into c clusters; Select c initial cluster centers; Assign each sample to a cluster according to the minimum distance principle; Iterate through the calculation of various types of clustering centers and according to the new cluster center to adjust the clustering situation until the iterative convergence.

Assume the sample set is $\{x_1, x_2, \ldots, x_n\}$, clusters are $\omega_1, \omega_2, \ldots, \omega_c$, cluster centers are $m_1, m_2, \ldots, m_c$. Then the objective function (is a function that reflects the order relation of the elements in the set) of this method is:

$$J = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{x_j \in \omega_i} \|x_j - m_i\|^2, \tag{4.1}$$

where $J_i = \sum_{X_j \in \omega_i} \|x_j - m_i\|^2$ is the objective function within cluster ω_i. Thus, the value of J_i depends on the geometric properties of ω_i and the position of m_i.

In general, the class ω_i after classification is defined by a two-dimensional membership matrix U of $c \times n$. If the j data point x_j belongs to class ω_i, then element u_{ij} in U is 1. Otherwise, the element is 0. Once the clustering center m_i is determined, we can derive the minimize u_{ij} in equation (4.1),

$$u_{ij} = \begin{cases} 1, & \forall k \neq i,\ \|x_j - m_i\|^2 \leq \|x_j - m_k\|^2 \\ 0, & otherwise. \end{cases} \tag{4.2}$$

If m_i is the nearest clustering center of x_j, then x_j is belong to cluster ω_i. Since a given data can only belong to a cluster, the membership matrix U has the following properties.

$$\sum_{i=1}^{c} u_{ij} = 1, \quad \forall j = 1, 2, \ldots, n \ and \ \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} = n. \tag{4.3}$$

The algorithm repeats the following steps, to determine the clustering center m_i and the

membership matrix U, the processing is show in below:

1. Initial clustering center m_i, $i = 1, 2, \ldots, c$, the typically way is taking any c points from all the data, set the iteration stop threshold ε and the iteration counter;

2. Use equation (4.2) to determine the membership matrix U, measure the distance of each element to the cluster center m_i,then put it to the nearest cluster ω_i;

3. Use equation (4.1) to calculate objective function. If it is less than the threshold set in step 1, or it is less than a certain threshold relative to the value of the last objective function, then the algorithm stops;

4. According to equation (4.3) recalculate and fix each cluster center which was obtained;

5. Repeat steps 2-4 until the new cluster center is equal to or less than the specified threshold. The algorithm ends.

For the first step initial cluster center, if the number of cluster c cannot be determined by prior knowledge, it can be approximated by graphing method. The c-mean algorithm is used to cluster c from small to large, in different c value to obtain a different J value, and then do $J - c$ curve. If there is an inflexion point on the curve, the number of cluster corresponding to the inflexion point is the optimal number of cluster. If the inflection point is not obvious, then the method fails.

Meanwhile, from the above processing we can see that the algorithm is iterative and does not ensure that it converges to the optimal solution. The property of c-means algorithm is based on the initial position of cluster center. Therefore, in order to get the better result, we need either use some prior knowledge to determine the initial center or use a different initial center for multiple operations. In addition, the above algorithm is only a representative method. For example, we can also initialize an arbitrary membership matrix, and then do the iterative process.

4.1.3 The theoretical basis of fuzzy c-means clustering algorithm

In 1973, Bezdek proposed the Fuzzy C-Means clustering algorithm. As an improvement of the early hard C-means clustering (HCM) method, the algorithm has been widely applied and

developed. The fuzzy c-means algorithm is a clustering algorithm that uses membership degrees to determine the degree of each data point belonging to a cluster.

Assume a sample set $X = \{x_1, x_2, \ldots, x_n\}$ is the whole cluster analysis object. Each element x_i, $i = 1, 2, \ldots, n$ in X can be represented by a finite number of parameter values. And each parameter value represents a property of the element x_i [65]. The clustering process is to analyze the similarity of n elements in the domain, to divide c fuzzy cluster according to the degree of affinity between the elements, and then find the clustering center of each cluster, so that the objective function of non-similarity index is minimized.

The main difference between FCM and HCM is that the FCM is divided by fuzzy, i.e. each given element uses the degree of membership between $[0, 1]$ to determine the degree of belonging to each cluster. So that the membership matrix U allows elements with values between $[0, 1]$, not just 0 and 1. With the normalization of the provisions, the sum of the membership of a data set is always equal to 1 (equation (4.3)).

Then the objective function of FCM can be defined as

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{h} \|x_j - m_i\|^2, \tag{4.4}$$

where n is the number of sample set, c is the given number of cluster, $1 < c < n$, h is the weighted power exponent, m_i is the clustering center of cluster i, and u_{ij} is the degree of sample j belonging to the cluster i.

The FCM algorithm is a clustering result that requires J to reach the minimum value. Therefore, the mathematical model of FCM algorithm is a conditional extreme value problem, i.e. taking the derivative of J with respect to u_{ij} and m_i respectively, making their derivatives to be 0, and substitute into $\sum_{i=1}^{c} u_{ij} = 1$. In order to solve the problem of conditional extreme value, we need to refer to Lagrange multiplier method to convert it to unconditional extreme value problem. Therefore, we

introduce n Lagrange factors, then equation (4.4) becomes

$$J = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{h}\|x_j - m_i\|^2 + \sum_{j=1}^{n}\lambda_j\left(\sum_{i=1}^{c} u_{ij} - 1\right). \tag{4.5}$$

Taking derivative with respect to m_i:

$$\frac{\partial J}{\partial m_i} = \sum_{i=1}^{c}\sum_{j=1}^{n}\frac{\partial u_{ij}^{h}\|x_j - m_i\|^2}{\partial m_i} - \frac{\partial}{\partial m_i}\sum_{j=1}^{n}\lambda_j\left(\sum_{i=1}^{c} u_{ij} - 1\right), \tag{4.6}$$

where

$$\frac{\partial}{\partial m_i}\sum_{j=1}^{n}\lambda_j\left(\sum_{i=1}^{c} u_{ij} - 1\right) = 0.$$

Then

$$\begin{aligned}\frac{\partial J}{\partial m_i} &= \sum_{i=1}^{c}\sum_{j=1}^{n}\frac{\partial u_{ij}^{h}\|x_j - m_i\|^2}{\partial m_i} = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{h}\frac{\partial}{\partial m_i}\|x_j - m_i\|^2 \\ &= \sum_{j=1}^{n} u_{ij}^{h}\frac{\partial}{\partial m_i}\|x_j - m_i\|^2 = -2\sum_{j=1}^{n} u_{ij}^{h}(x_j - m_i) = 0.\end{aligned}$$

Therefore

$$\sum_{j=1}^{n} u_{ij}^{h} m_i - \sum_{j=1}^{n} u_{ij}^{h} x_j = 0.$$

Hence

$$m_i = \frac{\sum\limits_{j=1}^{n} u_{ij}^{h} x_j}{\sum\limits_{j=1}^{n} u_{ij}^{h}},\ i = 1, 2, \ldots, c. \tag{4.7}$$

Taking the derivative with respect to u_{ij}:

$$\begin{aligned}
\frac{\partial J}{\partial u_{ij}} &= \frac{\partial}{\partial u_{ij}} \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{h} \|x_j - m_i\|^2 + \frac{\partial}{\partial u_{ij}} \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{c} u_{ij} - 1 \right) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{c} \frac{\partial}{\partial u_{ij}} u_{ij}^{h} \|x_j - m_i\|^2 + \frac{\partial}{\partial u_{ij}} \sum_{j=1}^{n} \left(\sum_{i=1}^{c} \lambda_j u_{ij} - \lambda_j \right) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{c} h u_{ij}^{h-1} \|x_j - m_i\|^2 + \sum_{j=1}^{n} \sum_{i=1}^{c} \frac{\partial}{\partial u_{ij}} \lambda_j u_{ij} \\
&= \sum_{j=1}^{n} \sum_{i=1}^{c} h u_{ij}^{h-1} \|x_j - m_i\|^2 + \sum_{j=1}^{n} \sum_{i=1}^{c} \lambda_j \\
&= \sum_{j=1}^{n} \sum_{i=1}^{c} h u_{ij}^{h-1} \|x_j - m_i\|^2 + \sum_{j=1}^{n} c\lambda_j \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{c} h u_{ij}^{h-1} \|x_j - m_i\|^2 + c\lambda_j \right) = 0.
\end{aligned} \tag{4.8}$$

Therefore

$$\sum_{i=1}^{c} h u_{ij}^{h-1} \|x_j - m_i\|^2 + \sum_{i=1}^{c} \lambda_j = \sum_{i=1}^{c} (h u_{ij}^{h-1} \|x_j - m_i\|^2 + \lambda_j) = 0.$$

Then

$$h u_{ij}^{h-1} \|x_j - m_i\|^2 + \lambda_j = 0, \ \forall i = 1, 2, \ldots, c.$$

Hence

$$u_{ij} = \left(\frac{-\lambda_j}{h \|x_j - m_i\|^2} \right)^{\frac{1}{h-1}}, \tag{4.9}$$

and

$$\sum_{k=1}^{c} u_{kj} = 1, \ \forall i = 1, 2, \ldots, c.$$

Then we have

$$\sum_{k=1}^{c}\left(\frac{-\lambda_j}{h\|x_j-m_k\|^2}\right)^{\frac{1}{h-1}}=\frac{(-\lambda_j)^{\frac{1}{h-1}}}{\sum\limits_{k=1}^{c}(h\|x_j-m_k\|^2)^{\frac{1}{h-1}}}=1.$$

So

$$(-\lambda_j)^{\frac{1}{h-1}}=\sum_{k=1}^{c}(h\|x_j-m_k\|^2)^{\frac{1}{h-1}}. \tag{4.10}$$

Combining equation (4.10) and equation (4.9), we get the following

$$u_{ij}=\left(\frac{\sum\limits_{k=1}^{c}h\|x_j-m_k\|^2}{h\|x_j-m_i\|^2}\right)^{\frac{1}{h-1}}=\left(\frac{\|x_j-m_i\|^2}{\sum\limits_{k=1}^{c}\|x_j-m_k\|^2}\right)^{\frac{-1}{h-1}}. \tag{4.11}$$

If the number of clusters of known domain X clustering is c and the weight value is m, then equation (4.7) and equation (4.11) can achieve the best fuzzy clustering matrix [75] and clustering center.

The fundamental method of using Alternate optimization of iterative methods to solve these problem is gradient descent. By continuously iterative operating the fuzzy cluster matrix U and the clustering center matrix M, we obtain the correcting the value of clustering center and the membership degree of each element with each cluster center. We finally find the property of the cluster of the sample data set and divide the elements into the largest cluster of its membership. When J has the minimized value, the parameters of the sample set can be expressed as the matrix and the clustering center matrix. The specific iterative optimization process [70] is show in below:

1. Initialization. Given the class number for clustering $c(1 < c < n)$, set the iteration stop threshold ε and the iteration counter;

2. Using the random value at $[0, 1]$ to initialize the membership matrix U, such that it satisfies the constraint of equation (4.3);

3. Using equation (4.7) to calculate c cluster centers m_i;

4. Calculating the objective function based on equation (4.4) . If it is less than the threshold seted in step 1, or it is less than a certain threshold relative to the value of the last objective function, then the algorithm stops;

5. Using equation (4.11) to recalculate the membership matrix U, back to step 2.

Of course, the above algorithm can also initialize the cluster center first, and then do the iterative calculation.

The advantages of the fuzzy c-means clustering algorithm as described in Section 4.1 are the unsupervised classification method, which does not require any human intervention in the clustering process, and the efficiently. It is suitable for automatic segmentation area. However, fuzzy c-means algorithm still has some disadvantage. For example, the fuzzy c-means clustering algorithm requires a prior knowledge of the number c of the clustering class. But for most data, there is not much prior knowledge about the spatial distribution and the structure of the data, and most do not. Therefore, on the one hand the algorithm affects the unsupervised performance of clustering algorithm, on the other hand it has the problem of effectively judging the clustering results, such as the correctness of the cluster and the rationality of the number of clusters and so on. In addition, once a cluster number c is given, the classification of c clustering must be obtained. So even if we cannot cluster all the data in the data set, we can still cluster using of fuzzy c-means clustering algorithm. Therefore, it is difficult to correctly reveal the structural information contained in the data.

4.2 Improvement of fuzzy c-means algorithm combined with spatial feature of gray image

The main purpose of the FCM algorithm is to divide the sample points of the vector space into c subspaces according to distance measure. Because the standard FCM algorithm does not consider the gray scale property of each point and the degree of correlation of its neighborhood pixels in image segmentation, and even in the same cluster the gray scales of the pixels and the noise for two pints with the same membership are very different, so the algorithm's anti-noise function is poor

or we can say it has poor robustness. However, the image in the image processing is inevitably interferenced by different noise. This requires that the noise information be processed together with the real information and that the noise effect is suppressed as much as possible during the segmentation process, which cannot be achieved in the standard FCM clustering process. For this reason, Chuang et al [17] proposed a FCM cluster which is combined with spatial information. It is Spatial Fuzzy c-means. This algorithm is first uses the standard FCM to calculate the membership degree of each points in the image, and then use the sliding window template to determine the degree of similarity between the points in the cluster and the neighborhood point, to reduce the membership degree of the noise point. Finally, the new spatial membership function is constructed by using the pixel clustering distribution statistics. This algorithm can correct the wrong clustering pixels.

An important property of an image is that the neighbourhood pixels have a high correlation. In other words, these adjacent pixels have similar eigenvalues, and they should have the same membership degree. So the probability is very large that they belong to the same cluster. Therefore, spatial relations are very important in clustering, but are not applied in FCM algorithms. In order to use the correlation of neighborhood pixels, a sliding window template is used to define the spatial neighborhood function as

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}, \tag{4.12}$$

where $NB(x_j)$ is a square sliding window centered on pixel x_j. We use 5×5 square sliding window in this dissertation. Similar to the membership function, the spatial function h_{ij} reflects the probability that pixel x_j belongs to cluster i. If most of the pixels in the x_j neighborhood belong to the same cluster i, then the value of pixel x_j in the spatial function will large. At the same time, if the membership degree is large that x_j is belongs to the cluster i, it is means that x_j is the correct distribution. On the contrary, if the membership degree is small, it is indicating that the distribution

of the pixel x_j is wrong. Combined with spatial function, we have the new membership function

$$v_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^{c} u_{kj}^p h_{kj}^q}. \tag{4.13}$$

The new clustering center can be determined according to the new membership degree of each pixel

$$C_i = \frac{\sum_{j=1}^{n} (v_{ij})^M x_j}{\sum_{j=1}^{n} (v_{ij})^M}, \tag{4.14}$$

where p and q are the parameters that control the relative importance of the two functions of u_{ij} and h_{ij}, when $p = 1$, $q = 0$, $v_{ij} = u_{ij}$. From the new membership function, we can see that the clustering result does not change for the pixels in the same class. But for the noise point, the clustering weight is reduced and the influence on the cluster center is suppressed. Therefore, the pixel that is wrong cluster by the effects of noise can easily be corrected. The implementation procedures of the SFCM algorithm are show in below [46]:

1. Given the relevant parameters c, M, p, q, ε, use FCM algorithm to segment the image;

2. Based on equation (4.7) and equation (4.11) to calculate the membership degree and clustering center of the pixel;

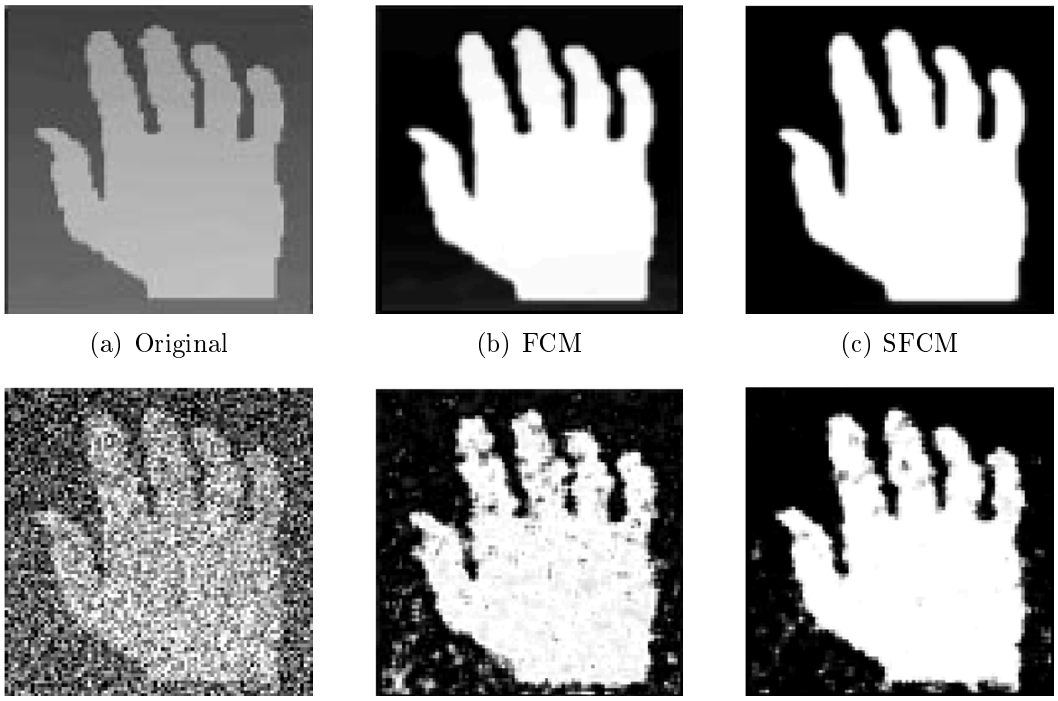
3. Perform iteration operations, if it is less than the threshold set in step 1, or it is less than a certain threshold relative to the value of the last objective function, then the algorithm stops. Otherwise go to step 2. The membership degree and the clustering center are obtained as the initial parameters of the secondary clustering;

4. Use equation (4.12), equation (4.13) and equation (4.14) to obtain the new membership degree v_{ij} and clustering center C_i;

5. Perform iteration operations, if it is less than the threshold ε which is set in step 1, or it

is less than a certain threshold relative to the value of the last objective function, then the algorithm stops. Otherwise go to step 4.

Figure 4.1(a) is a real palm image (93×93). Figure 4.1(b) and (c) show the segmentation results of the two clustering methods without adding noise, respectively. We can see that without adding noise, the two segmentation methods can achieve good segmentation results. Figure 4.1(d) showes the image with 5% Gaussian noise. Figure 4.1(e) and (f) show the segmentation results of FCM [37] and SFCM [37] after adding noise. It is easy to see that due to the palm image is simple and the target (palm) is clear, the two methods can complete segmentation. But in the details, noise has more impact for FCM method. As show in Figure 4.1(e), there is a lot of unnecessary segment in the palm. And at the edge of the palm of the hand, the segmentation result of the FCM method is not as smooth as the segmentation result of the SFCM method.



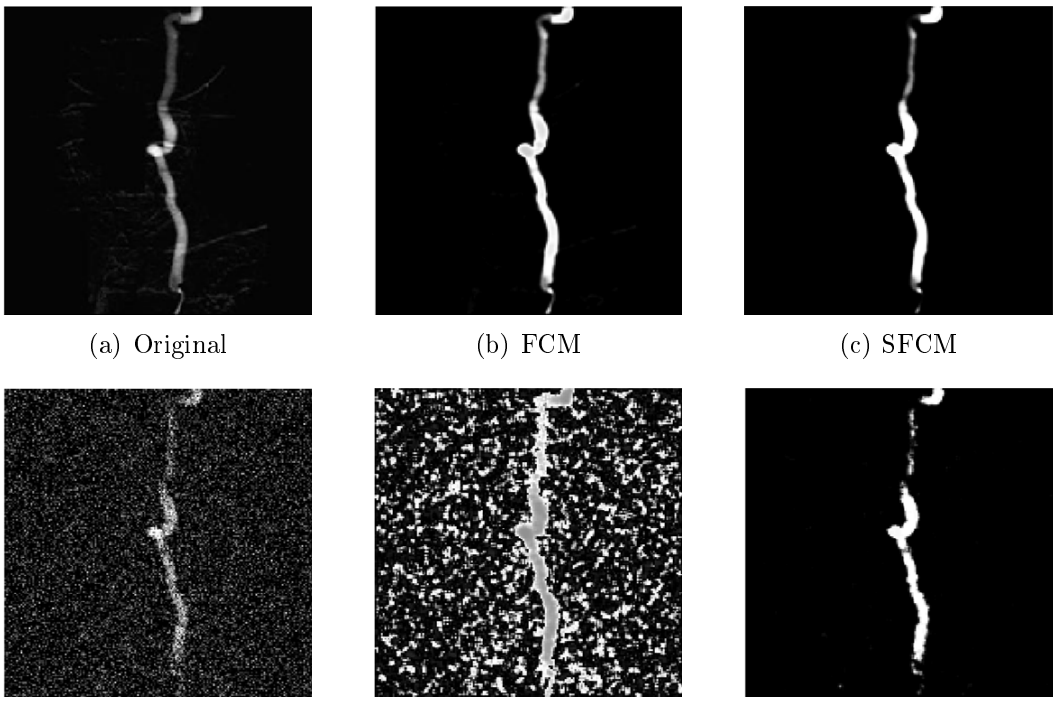
(a) Original (b) FCM (c) SFCM

(d) 5% Gaussian noise (e) FCM (f) SFCM

Figure 4.1: The comparison of segmentation results by using FCM and SFCM clustering algorithm for a real palm image with adding 5% Gaussian noise.

Figure 4.2 is a ultrasound image of blood vessel (288×287). We can see that the edge of the

target (blood vessel) in the ultrasound image is weak, and the background has a significant gray scale due to the presence of the skin tissue. Figure 4.2(b) and (c) show the segmentation results of the FCM and SFCM methods for the original image. Although the edge of the image is very weak, but the two methods have satisfactory results. Figure 4.2(d) is the original image with 5% Gaussian noise. Figure 4.2(e) and (f) are segmentation result of the image with adding noise by using two method. From the result we can see that comparing to the simple image (palms) in Figure 4.1, the FCM method is almost impossible to complete the segmentation for weaker edge vascular images. However, the SFCM method still has a good result.



(a) Original (b) FCM (c) SFCM

(d) 5% Gaussian noise (e) FCM (f) SFCM

Figure 4.2: The comparison of segmentation results by using FCM and SFCM clustering algorithm for an ultrasound image with adding 5% Gaussian noise.

The target and the background is relatively obvious in the above experiment, so the number of cluster $c = 2$, parameter $p = 1$, $q = 1$. From the above experimental results, we can see that the traditional FCM algorithm does segmentation only based on the current pixel. It is obviously that it has no strong anti-noise ability. The SFCM algorithm has considered the grey scale difference

between the pixel, and the neighborhood pixels, and the spatial location, therefore it greatly reduces the weight of the membership degree of noise. It showes that SFCM has good noise immunity. Meanwhile, the size of the parameter q has a smooth effect on the image, especially the more detail image. If the value of q is too large, it increases the weight of the spatial function, and the tiny details of the image becomes blurred. However, this condition is hard to see in this experimental.

4.3 CER image segmentation model combined with SFCM algorithm

The basic principle of the variational level set image segmentation method is to use the variational theory to minimize an energy functional which is a fusion image information about the level set function, to get the partial differential equation or the equations of the evolution of the level set function, so as to achieve the purpose of the evolution of the zero level set which is implied in it, and to make the zero level set converges to the edge of the target object in the image. The basic idea of clustering is to divide static data into homogeneous cluster. The goal is to make the similarity of the data in the cluster as large as possible and the similarity of the data out of the cluster as small as possible. Based on the discussion above, we propose an image segmentation model combined with SFCM and variational level set CER in this dissertation, referred to as SCER. This model combines the advantages of SFCM clustering and the variational level sets for image segmentation. The numerical results show that our method has the following advantages: (1) It can get a smooth segment boundary and a closed segment area; (2) Because the SFCM algorithm combined with spatial information, so the proposed model in this dissertation has strong robustness to noise images; (3) Since the image edge information clustered by SFCM is used as the initial contour condition of the CER model of the variational level set, it further improved the calculation speed, and it also solved the difficult problem of choosing the initial contour of the variational level set.

The basic idea of SCER model is: First, the image is clustering segmented by SFCM algorithm. Since the algorithm considers the spatial information of the image, so it is more robustness to the medical image containing the noise. Then we have good segmentation result. And because the traditional FCM algorithm is an unsupervised fast calculation method, so SFCM calculation speed is very fast, i.e. the calculation time is improved. Second, according to the segmentation

result of SFCM, the image edge information in the result is extracted. Third, the extracted edge information is applied as an initial contour to the variational level set which not only solve the problem of selecting the initial contours, but also reduce the number of iterations and improved the computational speed. And finally get a smooth and closed segmentation results. The specific process is show in below:

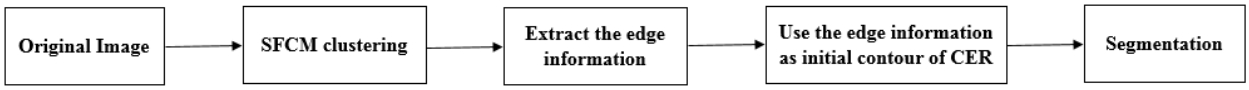


Figure 4.3: The implementation process of CER model which is combined with SFCM model.

4.4 Experimental Result

Through the above experiments, we can see that the improved SCER method not only segment the uneven gray scale image better and faster, but also has stronger robustness to the initial contour position and to strong noise than the original CER model.

4.4.1 Comparison of calculate time

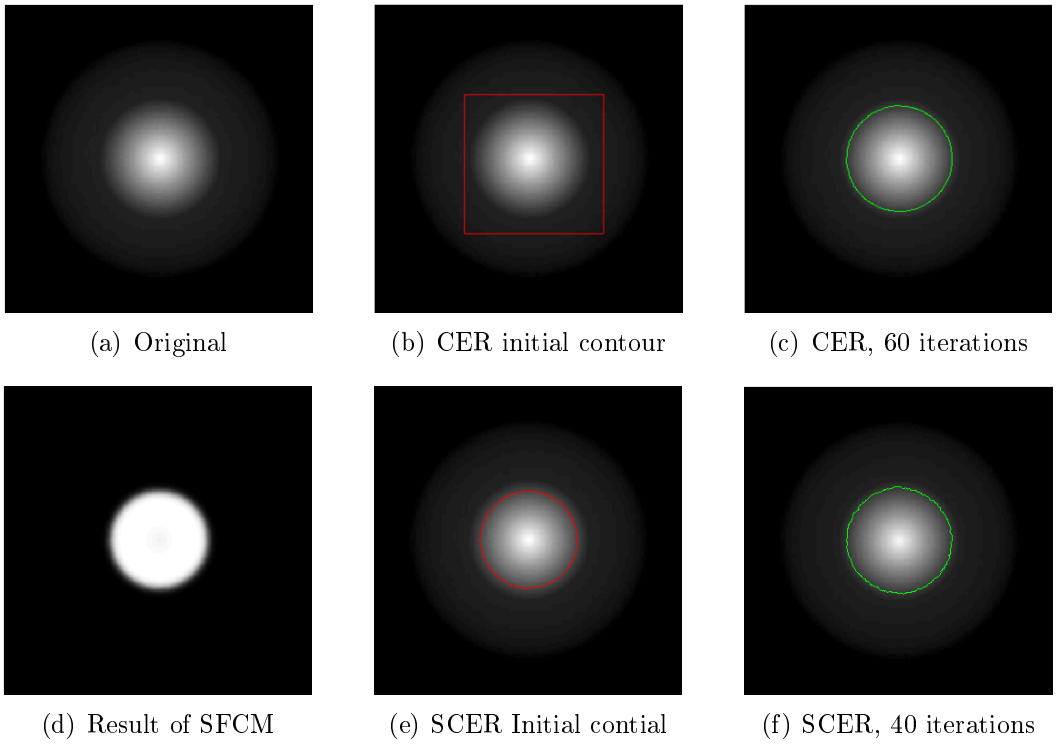
Table 4.1 showes the number of iterations and the time required to segment an image by using the CER model and the SCER model. It is more intuitive to show that the SCER model has certainly improved the calculation time.

Table 4.1: Comparison of the number of iterations and calculation time required to segment an image by using two models.

	CER		SCER	
	iteration times	time	iteration times	time
Folding cone image (Figure 4.4)	60	4.44s	40	2.75s
Vascular access image (Figure 4.5)	40	1.32s	20	0.63s

Figure 4.4(a) is a folding cone image (290×290), where the edge between the cone and the ground plane belongs to the ridge edge. For these deep images, since the gray scale value of each point is related to the distance between the point and the observation point, so it is difficult for a

common segmentation method to detect and segment the target in those images [21]. But our two models can get the correct segmentation results. The CER model needs 60 iterations for 4.44s. The improved SCER model just need 40 iterations for 2.75s. It also showes that the improved SCER model has certainly improved the calculation time.

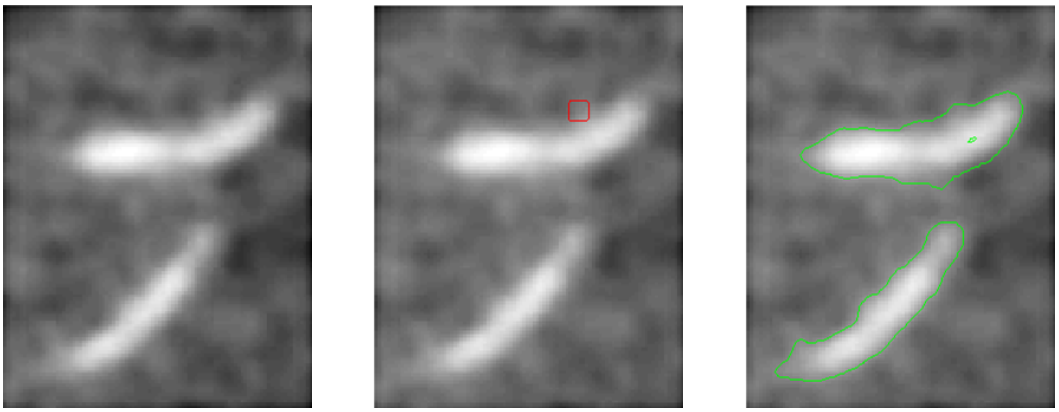


(a) Original (b) CER initial contour (c) CER, 60 iterations

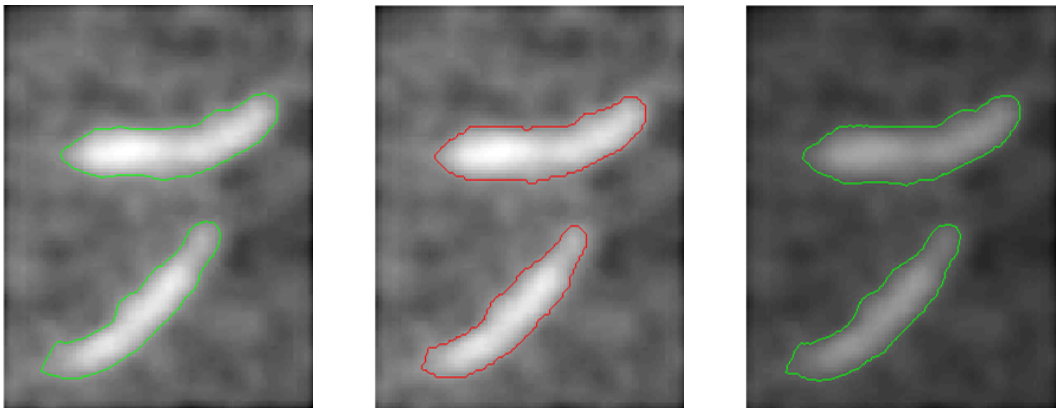
(d) Result of SFCM (e) SCER Initial contial (f) SCER, 40 iterations

Figure 4.4: The comparison of the number of iterations (time) by using two models to segment a folding cone image.

Figure 4.5 showes an original vascular access image and the segmentation result by using two models. Figure 4.5(c) and (d) show the segmentation results of CER model with iterating of 20 times and 40 times (1.32s). It is obvious to see from the figure below that SCER model only need 20 iterations for 0.63s. Only from the number of iterations, we can see that the SCER model has improved segmentation time.



(a) Original (b) CER initial contour (c) CER, 20 iterations



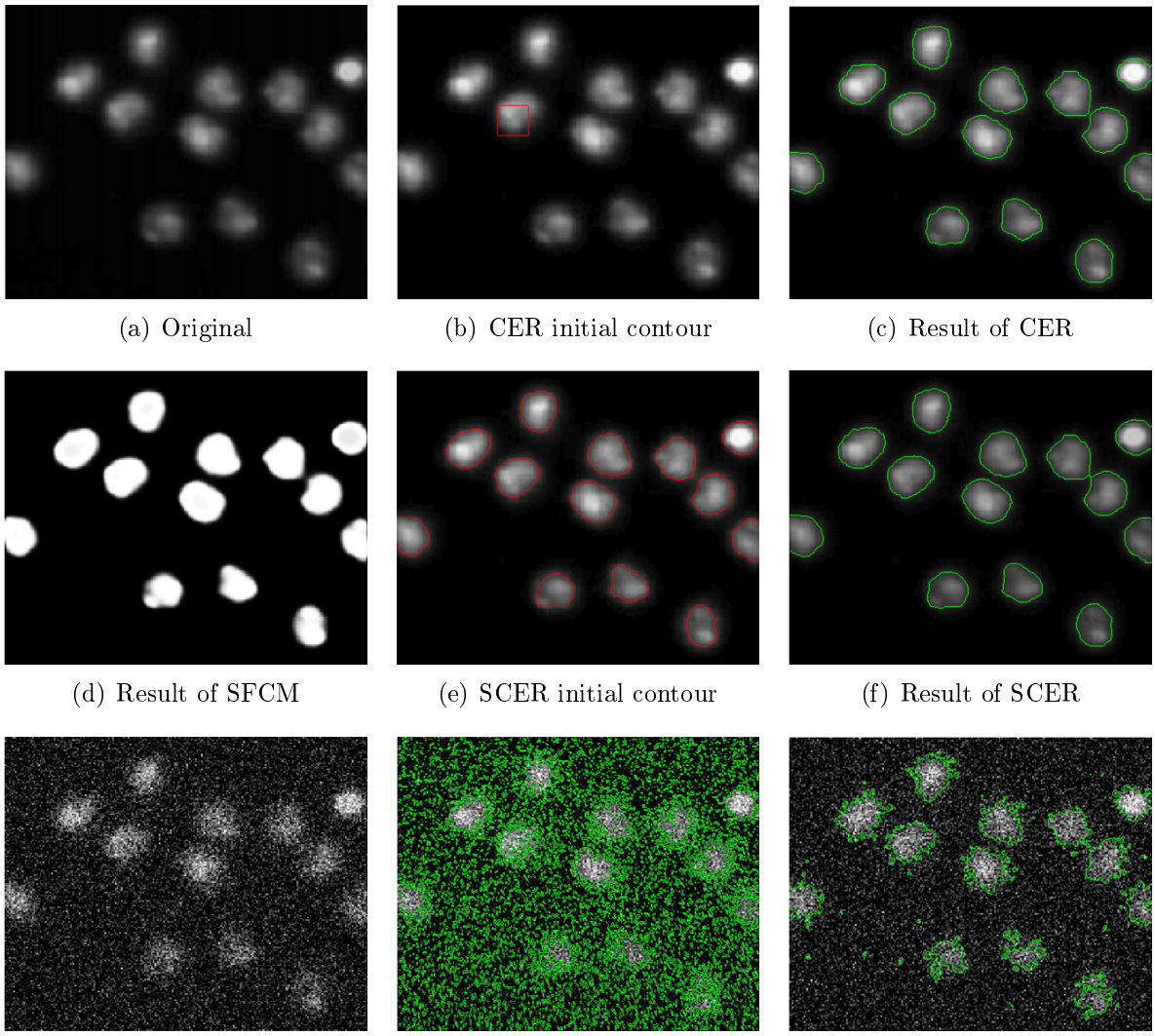
(d) CER, 40 iterations (e) SCER Initial contial (f) SCER, 20 iterations

Figure 4.5: The comparison of the number of iterations (time) by using two models to segment a vascular access image.

4.4.2 Comparison of noise sensitivity

Figure 4.6(a) is a multi-target cell image with partial target crossing the boundary (250×203). For this type of image, we cannot select a curve to include all the objects in the image. Those segmentation models that are sensitive to the initial contours, using only one initial contour, are difficult to segment such images. In contrast, our model is insensitive to the initial contour, and accurately extract all the targets in the image after several iterations, as show in Figure 4.6(c) and (f). Of course, our improved SCER model, using the SFCM segmented image boundary as the initial contour, can have multiple initial contours at the same time and contain all the targets. Figure 4.6(g) is an image with 5% Gaussian noise. Although the CER proposed in Chapter 3 has strong

robustness to noise, but it is still powerless for strong noise. From the above experimental results (h) and (i), we can see that the CER model is almost impossible to get the right segmentation result. While the 5% Gaussian noise has influence on the SCER model, but the model produced relatively correctly segmented result. We can see that the improved SCER model has more robustness to noise.



(a) Original (b) CER initial contour (c) Result of CER

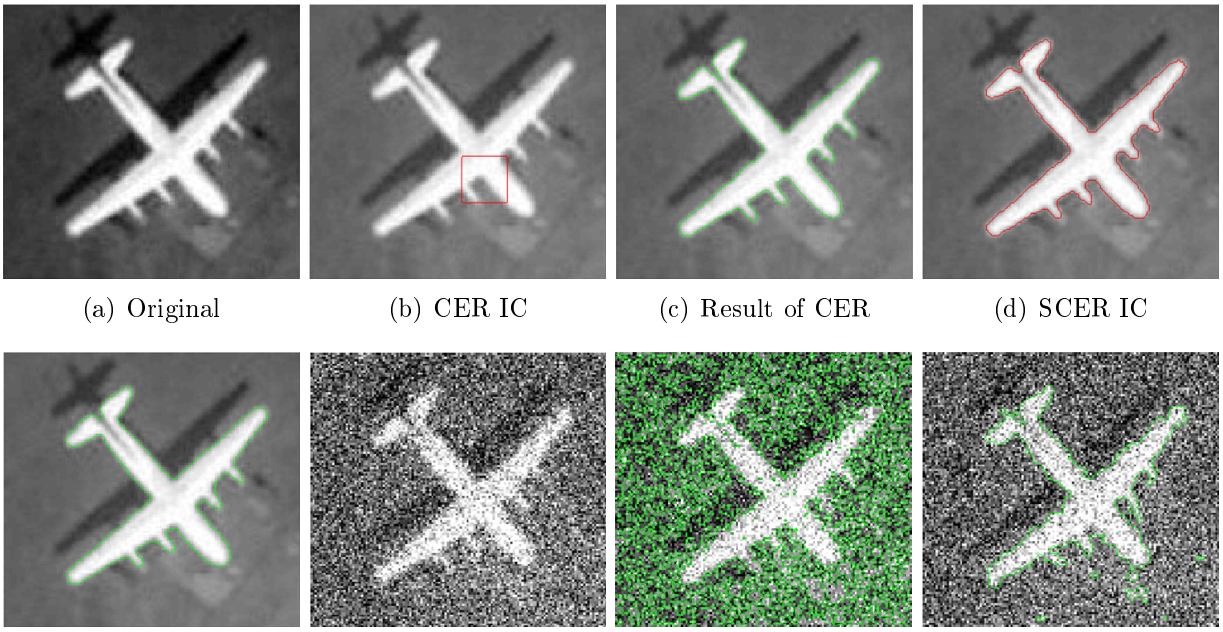
(d) Result of SFCM (e) SCER initial contour (f) Result of SCER

(g) 5% Gaussian noise (h) Result of CER (i) Result of SCER

Figure 4.6: The segmentation results of the improved SCER model and the CER model for a cell image with noise respectively.

Figure 4.7(a) is an actual aircraft image (135125) with pseudo boundary (shadow). Figure

4.7(c) and (e) show the segmentation results of CER model and SCER model for the original image. Both methods get accurate segmentation results. Figure 4.7(f) is the image with 5% Gaussian noise added in the original image. Figure 4.7(g) is the segmentation results of CER for an image with added noise. As we can see from the figure, the model cannot complete the segmentation. However, SCER can still get a relatively correct segmentation result as show in Figure 4.7(h). The above experiment once again shows that the SCER model has a stronger anti-noise property.



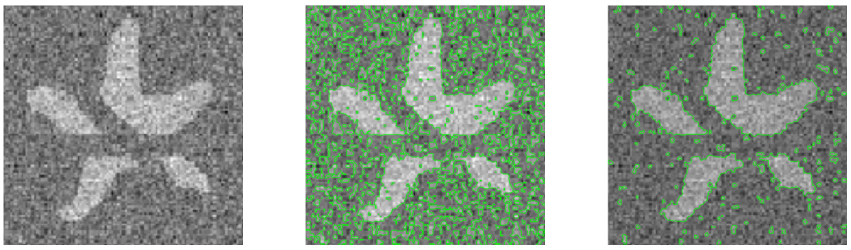
(a) Original (b) CER IC (c) Result of CER (d) SCER IC

(e) Result of SCER (f) 5% Gaussian noise (g) Result of CER (h) Result of SCER

Figure 4.7: The segmentation results of the improved SCER model and the CER model for an actual aircraft image with noise.

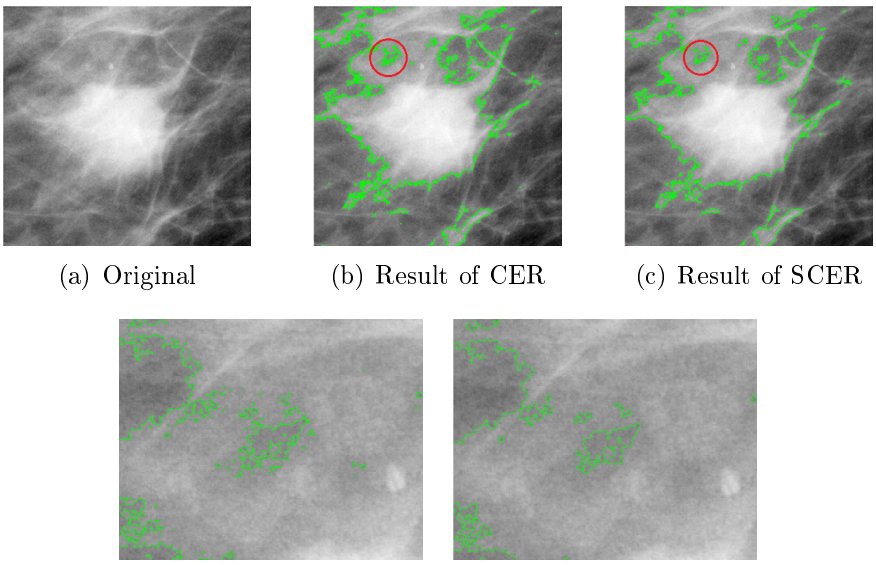
4.4.3 Low contrast image and medical image segmentation

Figure 4.8 shows the segmentation results of two models for an CT image (200×200) with noise. This image has more serious noise, but from the experimental results we can see that our improved SCER model is still able to segment the target very well. This also shows that the SCER model has a strong robustness to noise. Here we compare this kind of noise images, because real CT and MRI images always have certain noise at different degree of pollution.



(a) Original (b) Result of CER (c) Result of SCER

Figure 4.8: The segmentation results of two models for an CT image with noise.



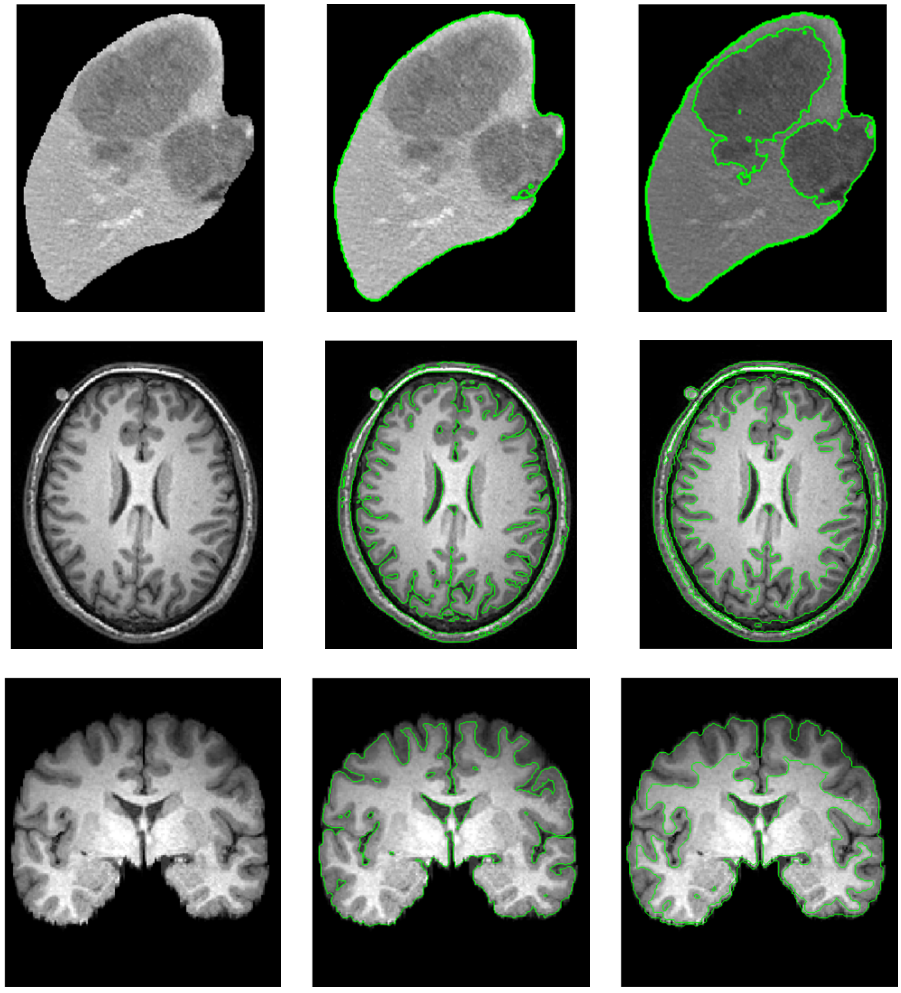
(a) Original (b) Result of CER (c) Result of SCER

(d) CER detail (e) SCER detail

Figure 4.9: The segmentation results of two models for a low contrast texture image respectively and the comparison of details.

Figure 4.9 is a low contrast texture image (705×681). The property of this type of image is the low contrast with weak edges. And it usually has noticeable grayscale variation in the target edge. Figure 4.9(b) and (c) show that we can get more satisfactory segmentation results. Overall, the segmentation results of the two models are similar, and not much difference. Figure 4.9(d) and (e) are the enlarged detail images with the red circle marked part of the segmentation result. From the details of the image, we can see, although there is over segmentation of two models, but CER has more. And the segmentation result of SCER model is better than CER model in detail. Therefore, we can also see that the improved SCER model is more smooth and can complete the segmentation

in more detail.



(g) Original (h) Result of CER (i) Result of SCER

Figure 4.10: The segmentation results of two models for a real heart image and the MRI of human brain image.

In Figure 4.10 from the first line to the third line are three real MRI images and the segmentation results of the two models. Figure 4.10 includes a real heart image (200×205), and two human brain images ($169 \times 207, 141 \times 157$). The first column showes the original images, the second column showes segmentation results of CER, the third column showes segmentation results of SCER. The experimental results once again show that which the same number of iterations and the same parameters, both models can get relatively good segmentation results. But in the details, the

improved SCER model segmentation results are better in terms of smoothness and completeness. For example, the CER model only showes the edge of the contour of the heart. While the SCER model not only showes the edge of the contours of the heart, but also extracts the two atriums.

Chapter 5
CONCLUSION AND FUTURE WORK

With the development of imaging technology and the expansion of image applications, image segmentation becomes more and more important in real life. Therefore, the research of image segmentation is of important theoretical value and practical significance. The method of partial differential equations has advantages such as the complete mathematical theoretical basis and the good extensibility. Therefore, in recent years, its application in image processing and computer vision has become the focus of research worldwide. In this dissertation, partial differential equation method is used, combined with information entropy and clustering analysis. Mathematical models and algorithms are used to study image segmentation technology. And satisfactory results have been obtained. The specific work and conclusion are as follows.

5.1 Summary of the Main Work

1. This dissertation studies various methods and models of image segmentation theory of partial differential equations based on the existing analysis, analyzes the theoretical basis and the mathematical principle of these methods, and compares their advantages and disadvantages. We build two new models and apply finite difference method and conduct digital simulation by MATLAB programming. The numerical results show that the new model has obvious advantages.

2. As we all know, with the development of imaging technology, the complexity of images is increasing, especially in the medical field, where images show extremely complex features. Weak edge, strong noise and uneven brightness often appear in the images. When one or more of these phenomena occur, the traditional active contour model cannot achieve the correct segmentation of target boundaries, especially for medical magnetic resonance images and ultrasonic images. To solve this problem, this dissertation introduces the concept of information entropy and improves the existing active contour RSF model. On this basis,

a new model CER combining CV model and RSF model is proposed in this dissertation. This model takes into account the influence of image region and boundary contour at the same time. By minimizing the energy function and taking into account the internal and external energy of the target boundary, we obtain the segmentation of the target boundary in the uneven intensity image. Numerical experiments show that the algorithm can extract the target with weak edge and uneven brightness. At the same time, the robustness of the original RSF model to noise and the sensitivity to the initial contour is also improved.

3. Image segmentation method based on clustering analysis has been one of the popular segmentation methods in recent years. The main feature of the traditional FCM clustering algorithm is an unsupervised segmentation method, which is fast in computation but sensitive to outliers. Therefore, this dissertation considers an improved SFCM algorithm, which combines the spatial information of the image, and reduces the sensitivity to noise of the traditional FCM algorithm. Using this advantage, this dissertation proposes to preprocess images by SFCM, and takes the edge information of the image after clustering as the initial contour of the CER model. This method not only improves CER model's robustness to noise, but also solves the problem in the initial contour selection. In addition, the improved rationality of the initial contour selection also leads to the improvement in computational efficiency to a certain extent. A large number of experiments show that this method can not only accurately segment strong noise images, but also improve the problem in initial contour selection and reduce computation time, thus enhance the CER model's robustness to noise.

5.2 Conclusion

This dissertation studies image segmentation based on the partial differential equation. It improves the problems of sensitivity to initialization, lack of strong noise robustness and slow evolution. Through solving specific problems, this dissertation studies the causes of the above problems and proposes CER model combining CV model, and improves RSF model based on information entropy. The experimental results show that CER model can improve the above problems to some

extent, and obtain good segmentation results.

The main characteristic of clustering algorithm is the segmentation according to similarity between pixels. It can not only improve the efficiency of computation, but also solve the image segmentation problem with complex content (strong noise and low contrast). Based on the advantages of the clustering algorithm, SCER model, an improved CER model, is proposed. The basic idea is to preprocess the image by SFCM algorithm, and then extract the edge information of the image after clustering as the initial contour of the CER model. This model not only reduces computational time, but also increases the robustness of the CER model to noise. At the same time, the problem in the initial contour selection is solved.

5.3 Main innovations of this dissertation

1. The concept of image information entropy is introduced into the image segmentation of partial differential equation method, and an improved scheme combining the famous CV model and RSF model is proposed. At the same time, we consider the internal and external energy of the target area to make CV model and the RSF model complement each other in the defects of algorithm. This improved CER model enhances the original model's robustness to contour initialization and to strong noise, and achieves satisfactory segmentation results.

2. The SFCM clustering algorithm combining spatial information is introduced into the CER model proposed in this dissertation. The clustering algorithm has the advantages of fast calculation and strong noise immunity, which improves the computational efficiency of CER model and its robustness to noise. This method also obtained satisfactory segmentation results and solved the problem in initial contour selection to a certain extent.

5.4 Future Work

The research of this dissertation is only a small part of the image segmentation of partial differential equations. The traditional methods are discussed and improved, and some achievements have been made. However, there is still much work to be done in the research of the image segmentation techniques for the active contour models of partial differential equations. In the future, I will study:

1. When an object is included by other objects, how to combine the prior geometry information of the target image with the active contour model?

2. How to preprocess medical images with strong noise so as to reduce the influence of noise on partial differential equations?

3. How to segment 3D images with stereo images?

4. For the numerical implementation of the model, only a simple finite difference scheme is considered in this dissertation, which will definitely affect the validity and the experiment of the model. Some more complex differential schemes or other numerical implementation methods will be developed for better performance.

REFERENCES

[1] R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994.

[2] M. Aizerman, E. Braverman, and L. Rozonoer. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25:821–837, 1964.

[3] S. Akaho. A kernel method for canonical correlation analysis. *In Proceedings of the International Meeting of the Psychometric Society, Spring-Verlag*, 2001.

[4] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel et.al. Axioms and fundamental equations of image processing. *Archive for Rational Mechanics and Analysis*, 16(9):200–257, 1993.

[5] G. Aubert and P. Kornprobst. Mathematical problems in image processing:Partial Differential Equations and the Calculus of Variation. *Springer, Applied Mathematical Sciences*, 147, 2002.

[6] E. Backer and A. Jain. A clustering performance measure based on fuzzy set decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(1):66–75, 1981.

[7] O. Bernard, B. Touil, A. gelas, R. Prost, and D. Friboulet. A RBF-based multiphase level set method for segmentation in echocardiography using the statistics of the radiofrequency signal. *IEEE Transactions on Image Processing*, 3:157–160, 2007.

[8] Y. Boykov and M. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Proceedings of the Eighth International Conference on Computer Vision (ICCV)*, 1:105–112, 2001.

[9] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. Osher. Fast global minimization of the active contour/snake mode. *Journal of mathematical Imaging and Vision*, 28:151–167, 2007.

[10] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[11] V. Caselles, F. Catte, T.coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathmatik*, 66:1–31, 1993.

[12] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contour. *International Conference on Computer Vision (ICCV)*, pages 694–699, 1995.

[13] T. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *UCLA CAM Report 04-54*.

[14] T. Chan and S.Esedoglu. Aspects of total varation regularized L^1 function approximation. *UCLA CAM Report 04-07*.

[15] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[16] Guangnan chen. Theorietic and test research on medical image processing based on partial differential equations. *A Thesis Submitted to Huazhong University for the Doctor's Degree of Science and Technology*, May, 2009.

[17] K. S. Chuang, H. L. Hzeng, S. Chen, J. Wu, and T. J. Chen. Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics*, 30:9–15, 2006.

[18] L. Cohen and I. Cohen. Finite-element methods for active contour models and ballons for 2-D and 3-D image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131–1147, 1993.

[19] L. D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218, 1991.

[20] R. Crandall. Image segmentation using the Chan-Vese algorithm. *ECE Progect*, Fall, 2009.

[21] X. Du and T. D. Bui. A new model for image segmentation. *IEEE Signal Processing Letters*, 15:182–185, 2008.

[22] C. L. Epstein and M. Gage. The curve shortening flow in wave motion: theory, modeling, and momputation. *Springer-Verlag, New York*, 1987.

[23] Bo Feng. Image segmentation based on multi-agent. *A Thesis Submitted to Xidian University for the Master's Degree of Sciences*, Jan, 2009.

[24] B. R. Frieden. Restoring with maximum likelihood and maximum entropy. *Journal of the Optical Society of America*, 62(4):511–518, 1972.

[25] D. Gabor. Information theory in electron microscopy. *Laboratory Investigation*, 14:801–807, 1965.

[26] T. Goldstein, X. Bresson, and S. Osher. Geometric applications of the split bregman method: Segmentation and surface reconstruction. *Journal of Scientific Computing*, 45:272–293, 2010.

[27] J. Gomes and Faugeras O. Reconciling distance functions and level sets. *Visiual Communication and Image Representation*, 11:209–223, 2000.

[28] R. C. Gonzalez and R. E. Woods. Digital image processing. *Prentice Hall, 3 edition*, Aug, 2007.

[29] Yucui Guo. Methods of mathematical physics. *Tsinghua University Press, 2 edition*, 2006.

[30] A. K. Jain. Partial differential equation and finite-difference methods in image processing, part 1: Image representation. *Optimization Theory and Applications*, 23:65–91, 1977.

[31] B. K. Jeon, Y. B. Jung, and K. S. Hong. Image segmentation by unsupervised sparse clustering. *Pattern Recognition Letters*, 27(14):1650–1664, 2006.

[32] M. Kass, A. Witriw, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1:321–369, 1988.

[33] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Conformal curvature flows: from phase transitions to active vision. *Archive for Rational Machanics and Analysis*, 134:275–301, 1996.

[34] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. *IEEE International Conference in Computer Vision (ICCV)*, page 1995, 810-815.

[35] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[36] Dingke Kong. Research on geometric active contour models for image segmentation. *A Thesis Submitted to Zhejiang University for the Doctor's Degree of Applied Mathematics*, April, 2010.

[37] Bing Nan LI. FCM & SFCM algorithm computer code.

[38] C. Li, C.Xu, C. Gui, and M. D. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Tansactions on Image Processing*, 19(12):3243–3254, 2010.

[39] C. Li, C. Kao, C. Gore, and Z. Ding. Implicit active contours driven by local binary fitting energy. *IEEE Computer Vision and Pattern Recognition*, 6:17–22, 2007.

[40] C. Li, C. Kao, J. C. Gore, and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE Transactions on Image Processing*, 17(10):1940–1949, 2008.

[41] C. Li, C. Xu, C. Gui, and M. D. Fox. Level set evolution without re-initialization: a new variational formulation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:430–436, 2005.

[42] Chunming Li. CV model computer code.

[43] Chunming Li. RSF model computer code.

[44] Jun Li, Xin Yang, and Pengfei Shi. A fast level set approach to image segmentation based on Mumford-Shah model. *CHINESE COMPUTERS*, 25(1):1176–1183, 2002.

[45] Yujie Li. Research for image segmentation based on cluter analysis and level set. *A Thesis Submitted to the School of Information and Engineering of Yangzhou University for the Master's Degree of Computer Sciences*, May, 2002.

[46] Yunsong Li and Ming Li. Fuzzy c-means clustering based on gray and spatial feature for image segmentation. *Computer Engineering and Design*, 28(6):1358–1363, Mar, 2007.

[47] J. Lie, M. Lysaker, and X. C. Tai. A bariant of the level set method and applications to image segmentation. *Mathematics of Computation*, 75(255):1155–1174, 2006.

[48] J. Lie, M. Lysaker, and X. C. Tai. A binary level set model and some applications to mumford-shah image segmentation. *IEEE Transactions on Image Processing*, 15(5):1171–1181, 2006.

[49] B. Liu, H. D. Cheng, J. Huang, J. Tian, X. Tang, and J. Liu. Probability density difference-based active contour for ultrasound image segmentation. *Pattern Recognition*, 43:2028–2042, 2010.

[50] R. Malladi, J. A. Setian, and B. C. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.

[51] D. P. Mukherjee, P. Pal, and Das J. Sonar image segmentation by fuzzy c-means. *Signal Processing*, 54(3):295–301, 1996.

[52] D. Mumford and J. shah. Optimal approximations by piecewise smooth functions and variational problems. *Communications on Pure and Applied Mathematics*, 42(2):577–685, 1989.

[53] J. T. Orlando. Image segmentation by histogram thresholding using fuzzy sets. *IEEE Transactions on Image Processing*, 11(2):1457–1465, 2002.

[54] S. J. Osher and L. I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 27:919–940, 1990.

[55] N. Paragios and R. Deriehe. Coupled geodesic active regions for image segmentation: a level set apporach. *Proceedings of Euorpean Coneference in Computer Vision, Dublin*, (2):224–240, 2001.

[56] N. Paragios and R. Deriehe. Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46:223–247, 2002.

[57] B. Romeny. Geometry driven diffusion in computer vision. *Boston, MA: Kluwer*, 1994.

[58] Qiuqi Ruan and Jiying Wu. Partial differential equation (PDE) method on digital image processing. *Signal Processing*, 28(3):301–314, Mar, 2012.

[59] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen. A survey of thresholding techniques. *Computer Vision Graphing Image Processing*, 41:233–260, 1988.

[60] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–165, 2004.

[61] Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 1948.

[62] D. W. Shattuck, S. R. Sandor-Leahy, K. A. schaper, D. A. Rottenberg, and R. M. Leahy. Magnetic resonance image tissue classification using apartial volume model. *Neuroimage*, 13:256–876, 2001.

[63] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[64] A. Shiozaki. Edge extraction using entropy operator. *Computer Vision, Graphics, and Image Processing*, 36(1):1–9, 1986.

[65] H. Sun, S. Wang, and Q. Jiang. FCM-Based model selection algorithms for determining the number of cluster. *Pattern Recognition*, 34(10):2027–2037, 2004.

[66] X. C. Tai, O. Christianse, P. Lin, and I. skjaelaaen. Image segmentation using some piecewise constant level set methods with MBO type of project. *International Journal of Computer Vision*, 73(1):61–76, 2007.

[67] L. A. Vese and T. F. Chan. A multiphase level set famework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293, 2002.

[68] Dakai Wang, Yuqing Hou, and Jinye Peng. Partial differential equation method on digital image processing. *Science Press, 1 edition*, 2008.

[69] L. Wang, C. Li, Q. sun, D. Xia, and C. Kao. Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation. *Computerized Medical Imaging and Graphics*, 33:520–531, 2009.

[70] S.Y. Wang, M.Q. Zhou, and G.H. Geng. Application of fuzzy cluster analysis for medical image data mining. *Mechatronics and Automation*, 2:631–636, 2005.

[71] X. Wang, D. Huang, and H. Xu. An efficient local Chan-Vese model for image segmentation. *Pattern Recognition*, 43(3):603–618, 2010.

[72] Wenying Wen. Partial differential equations based models for image segmentation. *A Thesis Submitted to Chongqing University for the Doctor's Degree of Science*, April, 2013.

[73] A. P. Witkin. Scale-space filtering. *Proceedings of the 8th International Joint Conference Artificial Intelligence (IJCAI), Karlsruhe Germany*, 2:1019–1021, 1983.

[74] Wang Yan. Studies on image segmentation based on partial differential equations. *A Thesis Submitted to Chongqing University for Doctor's Degree of Science*, March, 2012.

[75] Runling Yang, Xinbo Gao, and Jun Jie. A fast image segmentation algorithm based on fuzzy clustering. *Xi'an University of Arch and Tech (Natural Science Edition)*, 39(2):289–294, April, 2007.

[76] Jianjun Yuan. Image segmentation technology based on partial differential equation. *A Thesis Submitted to Chongqing University for the Doctor's Degree of Engineering*, May, 2012.

[77] H.K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level-set approach to multiphase motion. *Comput Phys*, 127:179–195, 1996.

[78] S. Zhu and A. Yuille. Region competition: unlyfing snkaes, region growing ang Byaes/MDL for multi-bnad image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.