Image Segmentation and Its Applications Based on the Mumford-Shah Model

Xiaojun Du

A Thesis

in

The Department

of

Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montréal, Québec, Canada

April 2011

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By:　　　　　Xiaojun Du

Entitled:　　　　Image Segmentation and Its Applications Based on the Mumford-Shah Model

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Computer Science)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

_____ Chair
Dr. H. Akbari

_____ External Examiner
Dr. E. Dubois

_____ External to Program
Dr. G. Vatistas

_____ Examiner
Dr. E. Doedel

_____ Examiner
Dr. A. Krzyzak

_____ Thesis Supervisor
Dr. T.D. Bui

Approved by _____
　　　　　　Dr. H. Harutyunyan, Graduate Program Director

April 7, 2011　　_____
　　　　　　　　　Dr. Robin A.L. Drew, Dean
　　　　　　　　　Faculty of Engineering and Computer Science

Abstract

Image Segmentation and Its Applications Based on the Mumford-Shah Model

Xiaojun Du, Ph.D.

Concordia University, 2011

Image segmentation is an important topic in computer vision and image processing. As a region-based (global) approach, the Mumford and Shah (MS) model is a powerful and robust segmentation technique as compared to edge-based (local) methods. In this thesis we apply the MS model to two interesting problems: image inpainting and text line detection. We further extend it by proposing a new image segmentation model to overcome some of the difficulties of the original model. As a demonstration of the new model, we apply it to the segmentation of retinal images. The results are better than the state-of-the-art approaches.

In image inpainting, the MS model is used to detect and estimate the object boundaries inside the inpainting areas. These boundaries are preserved in the inpainting results. We present a hierarchical segmentation method to detect boundaries of both the main structure and the details. The inpainting result can preserve detailed edges.

In text line detection, we use a combination of Gaussian blurring, the MS model, and morphing method. Different from other general text image detection approaches, our

method segments text documents without any knowledge of the written texts, so it can detect handwriting text lines of different languages. It can also handle different gaps and overlaps among the text lines.

Although the MS model has been used successfully in many applications, its implementation has always been based on some forms of approximation. These approximations are either inefficient computationally or applicable only to some special cases. Our new model consists of only one variable, the segmentation curve, therefore the computation is very efficient. Furthermore, no approximation is required, hence the method can segment objects with complicated intensity distribution. The new model can detect both step and roof edges, and can use different filters to detect objects of different levels of intensity. To show the advantages of the new model, we use a combination of the new model and Gabor filter to detect blood vessels in retinal images. This new model can detect objects with complicated image intensity distribution, and can handle non-uniform illumination cases effectively.

Acknowledgments

I would like to express my sincere gratitude to my advisor Professor Tien D. Bui for his guidance, help and support. His guidance gave me much inspiration for the thesis work.

I thank my fellow student Dongwook Cho, Wumo Pan, and other members of the Professor Bui's group for cooperation and useful discussions.

I thank my wife for her support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Image segmentation is a difficult problem in computer vision. A wide variety of methods have been presented, including Markov Random Field (MRF) approach [1], multi-resolution approach [2], and partial differential equations (PDEs) approach [3]. The Mumford and Shah (MS) [3] is a powerful and robust segmentation technique [4] [5]. The research work of this thesis is based on the MS model.

The MS model is based on an energy functional given by

$$E[u, C|u_0] = \int_{\Omega} (u - u_0)^2 dx + \mu \int_{\Omega \setminus C} |\nabla u(x)|^2 dx + \nu |C| \tag{1}$$

In Equation (1), u_0 is the original image; u is the smooth approximation of u_0; C is the segmentation curve; $|C|$ represents the length of the curve; Ω is the image domain; $\Omega \setminus C$ represents the image domain excluding the segmentation curve. Mumford and Shah proposed that the segmentation of an image can be obtained through the minimization of this energy functional. After the minimization, the variance of image intensity is minimized inside segmented regions. Because we assume that image intensity is smooth inside object

regions and changes abruptly across object boundaries, the MS energy functional is minimum when the segmented curves are on the object boundaries. As a region-based segmentation approach, the MS model makes use of the information not only near the boundary of objects but also inside the object regions. In this model, we try to find an optimal piecewise smooth approximation u of the original image u_0, and a set of boundaries C, such that u varies smoothly inside the homogeneous regions and discontinuously or rapidly across C. Compared with edge-based approaches, the MS model is more robust and can deal with noisy images.

In this thesis, we apply the MS model to two problems, image inpainting and text line detection. In image inpainting, the MS model is used to detect and estimate object boundaries inside the inpainting areas. These boundary edges are preserved in the inpainting results. However, it is difficult to segment complicated objects in an image. Therefore, we present a hierarchical segmentation method to deal with this situation. With this method, an image is segmented into small regions gradually until each region is smooth. We can detect boundaries of both the main structure and details. Based on the segmentation result, the inpainting result can preserve more detailed information.

In text line detection, we use Gaussian blurring to convert binary text images into gray images. Then we use the MS model to segment the text lines. Finally morphing approach is used to locate the text lines. Different from other text image processing approaches, our approach segments text image as a general image without any knowledge about the text. Our method can detect handwriting text lines of different languages. It also can handle different gaps and overlaps among the text lines.

Although the MS model was successfully used in many applications, the numerical method for solving the MS model is difficult to implement. Chan and Vese presented two important models [6, 7], the piecewise smooth approximation and the piecewise constant approximation. In the first method, we assume image intensity inside object regions can be approximated by smooth functions. The method can handle general cases. However, it requires the solution of three coupled PDEs which makes the method very inefficient. The second model approximates the intensity of each region as a constant, and it is fast because only one PDE needs to be solved. However, this method can only be used when the intensity inside each segmented region is uniform. In the case where the intensity varies inside the object, or the case of roof edges[1] the constant approximation model will fail. To deal with this case, Blake and Zisserman [8] introduced a second order derivative term into the MS model. Ambrosio, Faina, and March minimized similar energy functional by a family of elliptic functions [9]. In their approach, two coupled PDEs need to be solved. The authors of [10] used a linear approximation to minimize the MS energy functional. The linear approximation requires solving only one PDE, and it can detect non-uniform objects and roof edges. However, this approach can only detect regions whose intensity varies linearly.

In the last part of this thesis, we present a new model where the energy functional contains only two terms: the length of the segmentation curves, and the high frequency components inside segmented regions. Because only one variable, the segmentation curve, is involved, we need to solve only one PDE to minimize the energy functional. In the

---

[1]Roof edges are defined as discontinuities in the derivative rather than in the function itself

implementation of the new model, no approximations are applied. So the new model can handle any image intensity distribution inside object regions. We test the new model on some range images. In these images, there are many roof edges. The new model can detect both step and roof edges effectively. Another advantage of the new model is that we can use different filters in the model to detect different kinds of objects. We combine the new model and Gabor filter to detect blood vessels in retinal images. The Gabor filter is good to detect directional objects such as blood vessels. Because the new model can detect objects with complicated image intensity distribution, our approach can handle non-uniform illumination cases. The segmentation result is similar or better than state-of-the-art.

The organization of the thesis is as following: First, some image segmentation approaches and the level set technique are reviewed in chapter two. Implementation methods and some variations of the MS model are introduced in chapters three and four. Chapter five describes the application of the MS model on image inpainting, and chapter six describes text line detection using the MS model. The new image segmentation model is presented in chapter seven. This chapter also introduces an application of the new model on retinal image segmentation. Finally conclusions are summarized in chapter eight.

Chapter 2

Image Segmentation

This chapter reviews some image segmentation approaches and the level set approach.

2.1 Classical image segmentation approaches

Image segmentation is an important topic in computer vision and image processing. It is the bridge between low-level vision and high-level vision. Through image segmentation, objects are segmented from images, and high-level tasks, such as object detection, recognition, and tracking, can be performed. The goal of image segmentation is to divide an image into different object regions. It is assumed the image pixels in one region have similar property, such as gray level or texture. If we evaluate image segmentation result only from the image itself, then "Regions of an image segmentation should be uniform and homogeneous with respect to some characteristics such as gray tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of a segmentation should have

significantly different values with respect to the characteristic on which they are uniform. Boundaries of each segment should be simple, not ragged, and must be spatially accurate" [11].

Image segmentation approaches can be divided into two categories: region based and edge based approaches. Region based approach groups pixels of different regions according to similarity. Thresholding and region growing belong to this category. Thresholding techniques [12] are the early and popular approaches for image segmentation. Thresholding approaches are simple to implement and require very little computation. The main difficulty of this technique is the selection of threshold. Generally the threshold depends on image content. Even with adaptive threshold, it is still difficult to segment some regions with similar properties. Region growing [13] is another region based approach. Region growing algorithm starts by choosing a set of 'seed' pixels. Compare neighboring pixels with the seed according to some criteria. If they have similar property, these neighboring pixels are segmented into the same region of the seed. In this way, the region grows until all pixels are segmented into different regions. Region growing is insensitive to noise, but it needs extensive computation. The segmentation result also depends on the selection of criteria.

Edge detection is an edge based image segmentation technique. Classical edge detection approaches are based on the first and second derivatives. Some classical first derivative edge detectors include Roberts gradient operators, Prewitt operators, Sobel operators, and Canny algorithm. One classical second derivative edge detector is Laplacian operator.

These operators need very little computation, but they are sensitive to noise. Some sophisticated edge detection algorithms were developed, and good result can be obtained by wavelet based approaches. One disadvantage of edge detection is that closed segmentation curves are not guaranteed. Only edges are detected, but image is not really segmented into regions. So some post processing is needed.

In addition to the above deterministic methods for image segmentation, stochastic modeling is another main image analysis approach. This approach depends on image properties described in terms of probability distributions and prior knowledge about the underlying image defined stochastically (i.e. prior probability model). We can use Bayesian framework to formulate a posteriori probability distribution based on observations and prior probability models.

2.2 Bayesian framework for image segmentation

The goal of image segmentation is to detect boundaries of objects in an image. These boundaries are assumed to be edges in the image. In Bayesian framework, image segmentation can be described as conditional probability estimation, that is, estimate edges with condition of the observed image [14]. It can be written as $p(\Gamma|u)$ where u is the observed image; Γ is the set of edges. We can represent Γ by a function:

$$\Gamma(x) = \begin{cases} = 1 & \text{if x is an edge pixel} \\ = 0 & \text{others} \end{cases} \quad x \in \Omega \qquad (2)$$

where Ω is the image domain.

According to Bayesian formula and ignoring the prior probability of observed image, the conditional probability of edge can be estimated by following equation:

$$p(\Gamma|u) \propto p(u|\Gamma) \cdot p(\Gamma) \qquad (3)$$

This indicates that the estimation depends on two factors, the generative data model of observed image and the prior knowledge model of edge. We use Gibbs formula to estimate probability, that is,

$$p = \frac{1}{Z}e^{-E} \qquad (4)$$

where E is Gibbs formal energy and Z is the normalization parameter. Applying logarithm to the Bayesian formula, we get energy equation:

$$E(\Gamma|u) = E(u|\Gamma) + E(\Gamma) \qquad (5)$$

To get the optimal estimation of edge, we need to minimize the energy equation.

The generative data model is to model image u when the set of edges Γ is given. If we only consider edge-based model, the data energy is the sum of all edge pixels energy,

$$
\begin{aligned}
E(u|\Gamma) &= \mu \sum_{x\in\Gamma} g(x, u(x), \nabla u(x), ...) \text{ (discrete domain) or} &(6)\\
&= \mu \int_{\Gamma} g(x, u(x), \nabla u(x), ...) \text{ (continuous domain)} &(7)
\end{aligned}
$$

where g is energy of one edge pixel. Because the gradient of image intensity $|\nabla u(x)|$ is large on edges, a pixel x has larger chance to be on an edge if $|\nabla u(x)|$ is larger. For each edge pixel, probability function should be an increasing function of $|\nabla u(x)|$, and equivalently, energy function g has to be a decreasing function of $|\nabla u(x)|$. Some functions

presented in literature are: the Cauchy decay:

$$g(p) = \frac{1}{1 + ap^2}, \quad a > 0 \tag{8}$$

the Gaussian decay:

$$g(p) = e^{-ap^2}, \quad a > 0 \tag{9}$$

or simply

$$g(p) = -p^2 \tag{10}$$

To be resistant to noise, $\nabla u(x)$ can be replaced by $\nabla u_\sigma(x)$,

$$u_\sigma = g_\sigma * u \tag{11}$$

where g_σ is the Gaussian filter.

For the prior model of edges, there are two kinds of model proposed, first order Euclidean

edges and second order Euclidean edges. Length energy is a first order model, which is:

$$E_1[\Gamma] = \alpha length(\Gamma) = \alpha \int_\Gamma |dx| = \alpha \int_0^T |x'(t)| dt \tag{12}$$

The integration is contour integration, which is along edges or contours.

Elastic energy is a second order model, which was proposed in [15] as:

$$E_2[\Gamma] = \alpha \int_0^T |x'(t)| dt + \beta \int_0^T |x''(t)| dt \tag{13}$$

The minimization of the first order term in both energies eliminates extraneous loops

and ripples of edges; the minimization of second order term in elastic energy makes edges

9

smoother and less flexible. The second order term also prevents a closed edge from collapsing into a single point.

With these energy models, we can assemble some edge-based image segmentation models. If we use length energy for edge prior model, we get geodesic active contours model:

$$E_1 = \alpha \int_0^T |x'(t)|dt + \int_0^T g(|\nabla u(t)|)dt \tag{14}$$

If we use elastic energy, we get the snake model:

$$E_1 = \alpha \int_0^T |x'(t)|dt + \beta \int_0^T |x''(t)|dt + \int_0^T g(|\nabla u(t)|)dt \tag{15}$$

Both models calculate energy along edges and use edge detector function $g(|\nabla u(t)|)$ to attract segmentation boundaries to edges. They use different model to regularize edges. Both models are non-intrinsic since they depend on parameters. We can adjust parameters α and β to balance the impact of different terms in the model.

The authors of [16] presented an alternative model as:

$$E_1 = 2\sqrt{\lambda} \int_0^T g(|\nabla u(t)|)|x'(t)|dt \tag{16}$$

The model is intrinsic since it does not depend on parameters. It is demonstrated that this model is equivalent to the active contours model.

In addition to edge-based models, some region-based models were also suggested. Compared with edge-based models, region-based models are more stable because more pixels information is considered. Different from edge-based energy, data energy $E[u|\Gamma]$

10

contains pixels energy inside segmented regions:

$$E(u|\Gamma) \quad = \quad \mu \sum_{x \notin \Gamma} \phi(x, u(x), \nabla u(x), ...) \text{ (discrete domain) or} \qquad (17)$$

$$= \quad \mu \int_{\Omega \backslash \Gamma} \phi(x, u(x), \nabla u(x), ...) \text{ (continuous domain)} \qquad (18)$$

Different from the g function used in edge data energy, the potential function ϕ should be an increasing function, for example:

$$\phi(p) = p^2 \qquad (19)$$

If we combine region-based data energy and edge length energy, we get region-based active contour model:

$$E_1 = \alpha \int_0^T |x'(t)|dt + \int_{\Omega \backslash \Gamma} |\nabla u(x)|^2 dx \qquad (20)$$

where t is the variable for contour integration; x is the variable for area integration. In this model, data energy is integrated inside segmented regions. This model is very similar to the celebrated segmentation model of Mumford and Shah.

The Mumford-Shah (MS) model is used to segment and smooth images at the same time. Through the minimization of the MS model, we can obtain segmented and piecewise smoothed image. In addition to the segmentation energy, the MS model also considers the smoothed image energy. The MS model energy can be derived from the following equation:

$$E[u, \Gamma|u_0] = E[\Gamma] + E[u|\Gamma] + E[u_0|u, \Gamma] \qquad (21)$$

where u_0 is the observed image, the original image; u and Γ are the smoothed image and segmentation boundaries respectively. $E[u_0|u, \Gamma]$ is the image data energy given the

11

smoothed image and segmentation. In the MS model, it is written as:

$$E[u_0|u, \Gamma] = \int_\Omega (u - u_0)^2 dx \qquad (22)$$

Using the same energy as the region-based active contour model for the first two terms in equation (21), the MS energy functional is written as:

$$E[u, \Gamma|u_0] = \int_\Omega (u - u_0)^2 dx + \mu \int_{\Omega \backslash \Gamma} |\nabla u(x)|^2 dx + \nu|\Gamma| \qquad (23)$$

where $|\Gamma|$ is the length of the segmentation boundaries. If the MS energy functional is minimized, the image will be segmented into regions so that: (1) u is a good approximation of u_0, (2) u is smooth in each region, and (3) the boundary of each region is as short as possible. The parameters of μ and ν are used to balance the effects of different terms.

The MS model is a global region-based image segmentation model. Because information of both segmentation boundaries and image data inside segmentation regions are considered, the MS model is more stable than local information based, such as Sobel or Laplace filter, or edge-based approaches, such as active contour or snake. The MS model has been successfully used in many applications. However, the numerical method of solving the MS model is difficult to implement. Recently, the level set method has been used by many authors to facilitate the numerical solutions of the MS model. Some altenative solutions of MS model were also presented, such as the elliptic approximation to the weak formulation of the MS function [17], the active contours without edges [7, 6], and the hierarchical curve evolution based approaches in [18, 19].

One difficulty of the MS model is that the energy functional is not convex. When we use the gradient descent approach to minimize the MS energy functional, the solution is often

12

trapped by the local minima. So the segmentation result depends on the initial condition. To overcome this difficulty, [20] presents an algorithm to calculate the global minimizer. If an image is segmented into two regions, the segmented regions can be represented by a function v, where $v \in \{0, 1\}$. The length of the segmentation boundaries is replaced by the integration of the function:

$$|\Gamma| = \int_\Omega |\nabla v| \qquad (24)$$

Then v is relaxed to $v \in [0, 1]$. After this relaxation, the energy functional becomes convex, and the global minimizer can be obtained. [20] proves that the thresholded minimizer of the relaxed energy functional is the global minimizer of the original energy functional.

While the energy models are on continuous space, implementations of the models are on discrete space for digital images. In discrete space, derivative is approximated by the difference between neighboring pixels; area integration is approximated by the sum of pixels in the area; contour integration is approximated by the sum of pixels along the contour. In this thesis, all algorithms are implemented in C language on computer to process discrete digital images.

2.3   Level set method

The level set method is first introduced by Osher and Sethian [21, 22, 23, 24, 25, 26]. It is a numerical technique to analyze and compute the motion of interfaces. Many applications of the level set method are developed in image processing.

The basic idea of the level set method is to make use of a 3-dimensional surface to

13

represent the motion of a 2-dimensional curve as shown in figure 1.
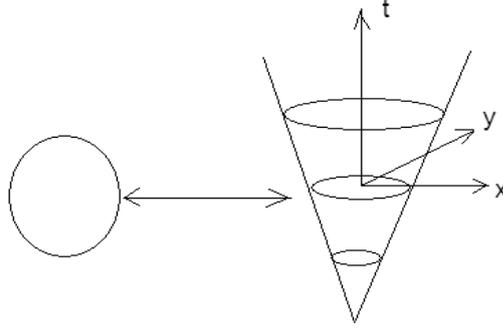
Figure 1: The 3-dimension demonstrational of level set approach [27].

In figure 1, the 3-dimension cone-shaped surface represents the motion of a 2-dimension circle. The intersection between the cone-shaped surface and the x-y plane represents the curve at different time. In this figure, the circle becomes bigger along the time axis. The 3-dimension surface can be represented by equation $\phi(x, y, z) = 0$. This surface segments the plane into two regions, inside region and outside region. We can assume $\phi(x, y, z) > 0$ inside the region and $\phi(x, y, z) < 0$ outside the region. On the x-y plane, the relationship becomes:

$$\phi(x, y, t) \begin{cases} > 0 \text{ inside the curve} \\ = 0 \text{ on the curve} \\ < 0 \text{ outside the curve} \end{cases} \tag{25}$$

As shown in figure 2, we can use the function $\phi(x, y, t)$, level set function, to represent two regions, inside region and outside region.

We also can use two level set functions to represent four regions as shown in figure 3.

14
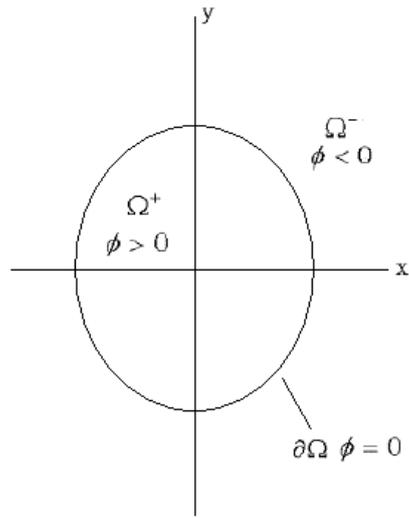
Figure 2: Representation of two regions by a level set function [28].
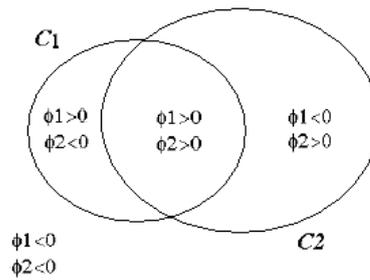
Figure 3: Representation of four regions by two level set functions [6].

The two level set functions satisfy:

$$\phi_1 > 0, \phi_2 > 0 \text{ inside region 1} \tag{26}$$

$$\phi_1 < 0, \phi_2 > 0 \text{ inside region 2}$$

$$\phi_1 > 0, \phi_2 < 0 \text{ inside region 3}$$

$$\phi_1 < 0, \phi_2 < 0 \text{ inside region 4}$$

In this way, we can use n level set functions to represent 2^n regions.

15

As mentioned in the beginning, the motion of a curve can be represented by a 3-dimensional surface. Since the surface is represented by equation: $\phi(x, y, t) = 0$, at any time t the level set value of each point $\vec{x}(t)$ on the surface is always zero. This implies $d\phi(\vec{x}(t), t)/dt = 0$. By the chain rule we have:

$$\frac{\partial \phi(\vec{x}, t)}{\partial t} + \nabla\phi \cdot \frac{d\vec{x}}{dt} = 0 \qquad (27)$$

It is an initial value partial differential equation and can be written as:

$$\frac{\partial \phi(\vec{x}, t)}{\partial t} + F|\nabla\phi| = 0 \qquad (28)$$

We also need two conditions for the equation:

1. The boundary condition: we usually impose the normal derivative vanishes on the curve or the surface.

$$\frac{\partial \phi}{\partial \vec{n}} = 0 \text{ on the curve} \qquad (29)$$

   where $\vec{n}$ is the normal vector.

2. The initial condition: The initial function $\phi_0(x, y)$ usually is assumed to be an initial curve.

$$\phi(x, y, 0) = \phi_0(x, y) \qquad (30)$$

By solving equation (28), we can obtain the evolution of a 2-dimensional curve.

An example of the evolution of a curve is to assume $F = -\kappa$, where κ is the curvature of the curve. In this case, the motion equation becomes:

$$\frac{\partial \phi(\vec{x}, t)}{\partial t} = \nabla\left(\frac{\nabla\phi}{|\nabla\phi|}\right)|\nabla\phi| \qquad (31)$$
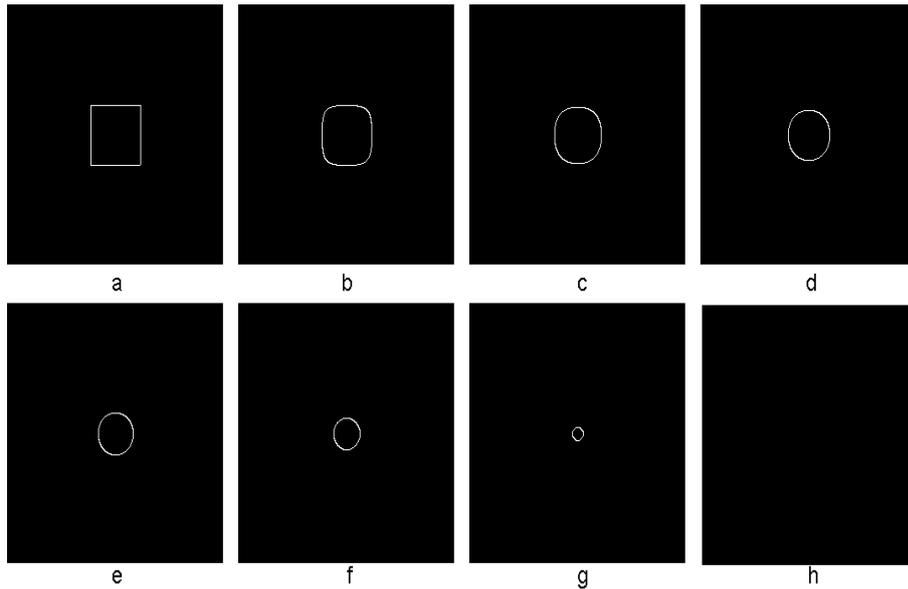
Figure 4: The evolution of a square curve.[29]

If we assume a square initial curve, the evolution of the curve is shown in figure 4.

In figure 4, the square curve first evolves to a circle and then vanishes.

There are many advantages to use the level set method to solve the evolution of curves, including:

1. The topological changes in the evolving boundary such as merging and breaking are handled naturally as shown in figure 5.

2. The numerical solution of the motion equation is easy to implement.
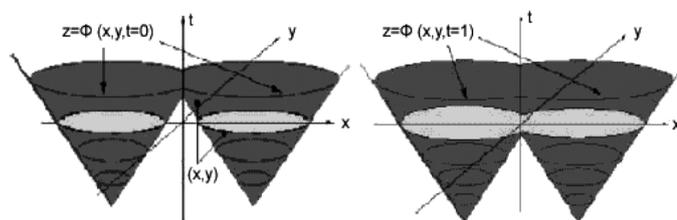
Figure 5: The merge of two curves with level-set approach [27].

The formulation of the MS model leads to the level set method as a numerical approach for solving the MS energy functional, in which the contour Γ can be presented by the level set function ϕ. This concept will be made clear in the following chapters.

Chapter 3

Implementation of the MS Model

The solution of the MS energy functional is not a trivial task. There are some alternative solutions to this problem, such as piecewise smooth approximation [6], piecewise constant approximation [7], piecewise linear approximation [30], and the elliptic approximation to the weak formulation of the MS functional [17]. These approaches are introduced in the following sections.

3.1   Piecewise smooth approximation

If we consider that a closed curve segments an image into two regions (i.e. inside and outside regions), the MS energy functional can be written as:

$$
\begin{aligned}
E(u_1, u_2, C) \;=\; & \int_{\text{inside } C} |u_1 - u_0|^2 \, dxdy + \mu \int_{\text{inside } C} |\nabla u_1|^2 \, dxdy \\
& + \int_{\text{outside } C} |u_2 - u_0|^2 \, dxdy + \mu \int_{\text{outside } C} |\nabla u_2|^2 \, dxdy + \nu \cdot |C| \quad (32)
\end{aligned}
$$

where u_1 and u_2 are smooth approximations of image inside and outside segmentation curve. The numerical solution of the MS energy functional can be implemented by the level set method. In the case that the image consists of two regions, the segmentation curve can be represented by one level set function ϕ:

$$\phi(x, y, t) = \begin{cases} > 0 & \text{if } (x, y) \text{ is inside } C \\ = 0 & \text{if } (x, y) \text{ is on } C \\ < 0 & \text{if } (x, y) \text{ is outside } C \end{cases} \tag{33}$$

Using the level set functions and the Heaviside function $H(x)$ defined by

$$H(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x < 0 \end{cases} \tag{34}$$

the piecewise smooth approximation of the MS model can be written as:

$$\begin{aligned} E(u_1, u_2, \phi) &= \int (u_1 - u_0)^2 H(\phi) dx dy + \int (u_2 - u_0)^2 (1 - H(\phi)) dx dy \\ &+ \mu \int |\nabla u_1|^2 H(\phi) dx dy + \mu \int |\nabla u_2|^2 (1 - H(\phi)) dx dy \\ &+ \nu \int |\nabla H(\phi)| dx dy \end{aligned} \tag{35}$$

The numerical form of $H(x)$ is:

$$H_\epsilon(x) = \frac{1}{2}(1 + \frac{2}{\pi} \arctan(\frac{x}{\epsilon})) \tag{36}$$

By the formula:

$$\nabla H(\phi) = \delta(\phi) \nabla \phi \tag{37}$$

20

where $\delta(\phi)$ is the Dirac function, we have:

$$E(u_1, u_2, \phi) = \int (u_1 - u_0)^2 H(\phi) dx dy + \int (u_2 - u_0)^2 (1 - H(\phi)) dx dy$$

$$+ \mu \int |\nabla u_1|^2 H(\phi) dx dy + \mu \int |\nabla u_2|^2 (1 - H(\phi)) dx dy$$

$$+ \nu \int \delta(\phi) |\nabla \phi| dx dy \tag{38}$$

The numerical form of $\delta(x)$ is:

$$\delta_\epsilon(x) = \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} \tag{39}$$

We define:

$$F(u_1, u_2, \phi) = (u_1 - u_0)^2 H(\phi) + (u_2 - u_0)^2 (1 - H(\phi))$$

$$+ \mu |\nabla u_1|^2 H(\phi) + \mu |\nabla u_2|^2 (1 - H(\phi)) + \nu [\delta(\phi) |\nabla \phi|] \tag{40}$$

Following the derivation of the Euler-Lagrange equation, we have

$$\frac{\partial F}{\partial \phi} = (u_1 - u_0)^2 \delta(\phi) - (u_2 - u_0)^2 \delta(\phi)$$

$$+ \mu |\nabla u_1|^2 \delta(\phi) - \mu |\nabla u_2|^2 \delta(\phi) + \nu [\delta'(\phi) |\nabla \phi|] \tag{41}$$

and

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial \phi_x} \right) = \nu [\delta'(\phi) \frac{\phi_x^2}{\sqrt{\phi_x^2 + \phi_y^2}} + \delta(\phi) \frac{\partial}{\partial x} \frac{\phi_x}{\sqrt{\phi_x^2 + \phi_y^2}}] \tag{42}$$

and

$$\frac{\partial}{\partial y} \left(\frac{\partial F}{\partial \phi_y} \right) = \nu [\delta'(\phi) \frac{\phi_y^2}{\sqrt{\phi_x^2 + \phi_y^2}} + \delta(\phi) \frac{\partial}{\partial y} \frac{\phi_y}{\sqrt{\phi_x^2 + \phi_y^2}}] \tag{43}$$

Thus, we have the following Euler-Lagrange equation:

$$\delta(\phi)[(u_1 - u_0)^2 - (u_2 - u_0)^2 + \mu |\nabla u_1|^2 - \mu |\nabla u_2|^2 - \nu \nabla \cdot (\frac{\nabla \phi}{|\nabla \phi|})] = 0 \tag{44}$$

21

with the boundary condition

$$\frac{\delta(\phi)}{|\nabla \phi|} \nabla \phi \cdot \hat{n} = \frac{\delta(\phi)}{|\nabla \phi|} \frac{\partial \phi}{\partial n} = 0 \tag{45}$$

Here $\hat{n}$ is the normalized normal of the boundary curve of the image.

Using the gradient projection method, we can change equation (44) to the following time dependent equation

$$\frac{\partial \phi}{\partial t} = \delta(\phi)[-(u_1 - u_0)^2 + (u_2 - u_0)^2 - \mu |\nabla u_1|^2 + \mu |\nabla u_2|^2 + \nu \nabla \cdot (\frac{\nabla \phi}{|\nabla \phi|})] \tag{46}$$

Similarly, we also can obtain the equations for the variables of u_1, u_2:

$$(u_1 - u_0) = \mu \nabla^2 u_1 \text{ inside } C \tag{47}$$
$$\frac{\partial u_1}{\partial \vec{n}} = 0 \text{ on } C$$

$$(u_2 - u_0) = \mu \nabla^2 u_2 \text{ outside } C \tag{48}$$
$$\frac{\partial u_2}{\partial \vec{n}} = 0 \text{ on } C$$

The smooth image functions u_1 and u_2 can be obtained by solving the damped Poisson equations (47) and (48), and the segmentation curve can evolve according to equation (46).

This is the piecewise smooth approximation presented by Chan and Vese [6]. Many advantages can be achieved by this approach, such as simultaneous segmentation and smoothing of noisy images, and detection of triple junctions by using multiple level set functions. However, because three PDEs equations (47), (48), and (46) are needed to be solved simultaneously, the computational cost of this approach is very large. To overcome this difficulty, Chan and Vese proposed another approach using the piecewise constant approximation [7].

3.2 Piecewise constant approximation

If image intensities inside different regions are uniform, image intensities inside different regions can be approximated by constants. In this case, the MS energy functional can be simplified to equation (49):

$$E(u, C) = \sum_k \int_{\Omega_k} (c_k - u_0)^2 dx dy + \nu |C| \tag{49}$$

where Ω_k represents the area inside each region; result image u is approximated by c_k in each region. The gradient term in the MS energy functional disappears in equation (49) because the gradient inside each region is zero.

Using the level set method [27] and the MS energy functional of the two phase segmentation, the image is segmented into two regions:

$$\begin{aligned} E(c_1, c_2, \phi) & = \int (c_1 - u_0)^2 H(\phi) dx dy \\ & + \int (c_2 - u_0)^2 (1 - H(\phi)) dx dy + \nu \int \delta(\phi) |\nabla \phi| dx dy \end{aligned} \tag{50}$$

where $H(x)$ is the Heaviside function.

Follow similiar procedure described in the previous section, we have the following time dependent equation:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) [-(c_1 - u_0)^2 + (c_2 - u_0)^2 + \nu \nabla \cdot (\frac{\nabla \phi}{|\nabla \phi|})] \tag{51}$$

Equation (51) is the evolution function for the segmentation curve. Constants c_1 and c_2 can be solved by:

$$\frac{\partial E}{\partial c_1} = 0 \quad \frac{\partial E}{\partial c_2} = 0 \tag{52}$$

23

The solutions of the above equations are:

$$c_1(\phi) = \frac{\int u_0 H(\phi) dx dy}{\int H(\phi) dx dy} \tag{53}$$

$$c_2(\phi) = \frac{\int u_0 (1 - H(\phi)) dx dy}{\int (1 - H(\phi)) dx dy} \tag{54}$$

Segmentation curve C, constants c_1 and c_2 are solved from equations (51), (53), and (54). Image u_0 is segmented into two regions $\{u = c_1\}$ and $\{u = c_2\}$.

Compared with the piecewise smooth approximation approach, the piecewise constant approximation approach is much faster because only one PDE needs to be solved. However, this approach is effective only when the intensity inside each segmented region (object) is uniform and can be approximated by a constant. When the intensities smoothly vary inside the object regions, the piecewise constant approximation will introduce bigger errors, and this approximation will not work well. Furthermore, in the case of roof edges, the intensity variance in the regions on both sides of roof edge is large but the second order derivative is small, and the intensities inside these two regions cannot be approximated by constants. The bigger the variance inside the object regions, the bigger the error introduced by constant approximation. Therefore, the piecewise constant approximation will not detect roof edges correctly.

If image consists of four regions, the segmentation curve can be represented by two level set functions ϕ_1 and ϕ_2, and the four regions are represented by signs of two level set functions:

Region 1: $\phi_1 > 0$ and $\phi_2 > 0$;

Region 2: $\phi_1 > 0$ and $\phi_2 < 0$;

Region 3: $\phi_1 < 0$ and $\phi_2 > 0$;

Region 4: $\phi_1 < 0$ and $\phi_2 < 0$.

With the piecewise constant approximation, the MS energy functional becomes:

$$E(c_{11}, c_{01}, c_{10}, c_{00}, \phi_1, \phi_2) =$$

$$\int [(u_0 - c_{11})^2 H(\phi_1) H(\phi_2) + (u_0 - c_{10})^2 H(\phi_1)(1 - H(\phi_2))$$

$$+(u_0 - c_{01})^2 (1 - H(\phi_1)) H(\phi_2) + (u_0 - c_{00})^2 (1 - H(\phi_1))(1 - H(\phi_2))] dxdy$$

$$+\nu \int |\nabla H(\phi_1)| dxdy + \nu \int |\nabla H(\phi_2)| dxdy \tag{55}$$

By solving the Euler-Lagrangian equation, we have the curve evolution equations for ϕ_1 and ϕ_2:

$$
\begin{aligned}
\frac{\partial \phi_1}{\partial t} &= \delta(\phi_1) \left[[-(c_{11} - u_0)^2 - (c_{01} - u_0)^2] H(\phi_2) - \right. \\
&\qquad \left. [(c_{10} - u_0)^2 - (c_{00} - u_0)^2](1 - H(\phi_2)) + \nu \nabla \cdot \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) \right] \\
\frac{\partial \phi_2}{\partial t} &= \delta(\phi_2) \left[[-(c_{11} - u_0)^2 - (c_{10} - u_0)^2] H(\phi_1) - \right. \\
&\qquad \left. [(c_{01} - u_0)^2 - (c_{00} - u_0)^2](1 - H(\phi_1)) + \nu \nabla \cdot \left(\frac{\nabla \phi_2}{|\nabla \phi_2|} \right) \right] \tag{56}
\end{aligned}
$$

Constants c_{11}, c_{10}, c_{01}, and c_{00} can be solved by:

$$\frac{\partial E}{\partial c_{11}} = 0 \quad \frac{\partial E}{\partial c_{10}} = 0 \quad \frac{\partial E}{\partial c_{01}} = 0 \quad \frac{\partial E}{\partial c_{00}} = 0 \tag{57}$$

Thus

$$c_{11}(\phi) = \frac{\int u_0 H(\phi_1)H(\phi_2)dxdy}{\int H(\phi_1)H(\phi_2)dxdy} \tag{58}$$

$$c_{10}(\phi) = \frac{\int u_0 H(\phi_1)(1 - H(\phi_2))dxdy}{\int H(\phi_1)(1 - H(\phi_2))dxdy} \tag{59}$$

$$c_{01}(\phi) = \frac{\int u_0(1 - H(\phi_1))H(\phi_2)dxdy}{\int (1 - H(\phi_1))H(\phi_2)dxdy} \tag{60}$$

$$c_{00}(\phi) = \frac{\int u_0(1 - H(\phi_1))(1 - H(\phi_2))dxdy}{\int (1 - H(\phi_1))(1 - H(\phi_2))dxdy} \tag{61}$$

If the image consists of more regions, we can use n level set functions to represent 2^n regions. So it is supposed that any number of regions can be segmented if enough level set functions are used. However, this is not true. First of all, the number of regions in an image is unknown before segmentation. We may use as many level set functions as possible to segment image to avoid missing any regions. In this case, the evolutions of many curves are coupled, and the computation cost will be very large. The most important problem is that the MS energy functional is not convex. If many curves evolve simultaneously, the problem of initial condition will be more difficult to handle. To overcome these problems, hierarchical segmentation schemes are presented.

3.3   Piecewise linear approximation

The piecewise constant approximation can be used only when the intensities of objects in an image are uniform and can be approximated by constants. When the variance of intensity inside an object is large, the piecewise constant approximation will introduce larger errors. To adapt to the distribution of intensity inside object regions, we can use a linear function $u(x, y) = a + bx + cy$, instead of a constant, to approximate the intensity of an object [30]. Here a, b, and c are constants.

For the case of two phases, we use two linear functions to approximate image intensity inside and outside reagions respectively:

$$u_1(x, y) = a_1 + b_1 x + c_1 y \tag{62}$$

$$u_2(x, y) = a_2 + b_2 x + c_2 y$$

Using level set functions and Heaviside function $H(x)$, the MS energy functional is written as:

$$
\begin{aligned}
E(a_i, b_i, c_i, \phi) &= \int (a_1 + b_1 x + c_1 y - u_0)^2 H(\phi) dx dy \\
&+ \int (a_2 + b_2 x + c_2 y - u_0)^2 (1 - H(\phi)) dx dy \\
&+ \mu (b_1^2 + c_1^2) \int H(\phi) dx dy + + \mu (b_2^2 + c_2^2) \int (1 - H(\phi)) dx dy + \\
&\nu \int |\nabla H(\phi)| dx dy \tag{63}
\end{aligned}
$$

We define

$$
\begin{aligned}
F(\phi) &= (a_1 + b_1 x + c_1 y - u_0)^2 H(\phi) + (a_2 + b_2 x + c_2 y - u_0)^2 (1 - H(\phi)) \\
&+ \mu (b_1^2 + c_1^2) H(\phi) + \mu (b_2^2 + c_2^2)(1 - H(\phi)) + \nu \delta(\phi) |\nabla \phi| \tag{64}
\end{aligned}
$$

Following the derivation of the Euler-Lagrange equation, we have

$$
\begin{aligned}
\frac{\partial F}{\partial \phi} =\ & (a_1 + b_1 x + c_1 y - u_0)^2 \delta(\phi) - (a_2 + b_2 x + c_2 y - u_0^2)\delta(\phi) + \\
& \mu(b_1^2 + c_1^2)\delta(\phi) - \mu(b_2^2 + c_2^2)\delta(\phi) + \nu\delta'(\phi)|\nabla\phi|
\end{aligned}
\tag{65}
$$

and

$$
\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \phi_x}\right) = \nu[\delta'(\phi)\frac{\phi_x^2}{\sqrt{\phi_x^2 + \phi_y^2}} + \delta(\phi)\frac{\partial}{\partial x}\frac{\phi_x}{\sqrt{\phi_x^2 + \phi_y^2}}]
\tag{66}
$$

and

$$
\frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \phi_y}\right) = \nu[\delta'(\phi)\frac{\phi_y^2}{\sqrt{\phi_x^2 + \phi_y^2}} + \delta(\phi)\frac{\partial}{\partial y}\frac{\phi_y}{\sqrt{\phi_x^2 + \phi_y^2}}]
\tag{67}
$$

Thus, we have following Euler-Lagrange equation:

$$
\begin{aligned}
\delta(\phi)[-\nu\nabla \cdot \frac{\nabla\phi}{|\nabla\phi|} + (a_1 + b_1 x + c_1 y - u_0)^2 + \mu(b_1^2 + c_1^2) \\
-(a_2 + b_2 x + c_2 y - u_0)^2 - \mu(b_2^2 + c_2^2)] = 0
\end{aligned}
\tag{68}
$$

with the following boundary condition

$$
\frac{\delta(\phi)}{|\nabla\phi|}\nabla\phi \cdot \hat{n} = \frac{\delta(\phi)}{|\nabla\phi|}\frac{\partial\phi}{\partial n} = 0
\tag{69}
$$

Here $\hat{n}$ is the normalized normal of the boundary curve of the image.

Using the gradient projection method, we change equation (68) to the following time dependent equation for $\phi(x, y, t)$

$$
\begin{aligned}
\frac{\partial \phi}{\partial t} =\ & \delta(\phi)[\nu\nabla \cdot \frac{\nabla\phi}{|\nabla\phi|} - (a_1 + b_1 x + c_1 y - u_0)^2 - \mu(b_1^2 + c_1^2) \\
& +(a_2 + b_2 x + c_2 y - u_0)^2 + \mu(b_2^2 + c_2^2)]
\end{aligned}
\tag{70}
$$

This is the evolution equation for segmentation curve.

28

We calculate a_1, b_1, and c_1 via the following equations:

$$\frac{\partial E}{\partial a_1} = 0 \quad \frac{\partial E}{\partial b_1} = 0 \quad \frac{\partial E}{\partial c_1} = 0 \tag{71}$$

or

$$a_1 \int H(\phi)dxdy + b_1 \int H(\phi)xdxdy$$

$$+c_1 \int yH(\phi)dxdy = \int u_0 H(\phi)dxdy$$

$$a_1 \int xH(\phi)dxdy + b_1 \int [x^2 + \mu]H(\phi)dxdy$$

$$+c_1 \int xyH(\phi)dxdy = \int xu_0 H(\phi)dxdy \tag{72}$$

$$a_1 \int yH(\phi)dxdy + b_1 \int xyH(\phi)dxdy + c_1 \int [y^2 + \mu]H(\phi)dxdy$$

$$= \int yu_0 H(\phi)dxdy$$

Similar to above equations, we obtain the equations for a_2, b_2, and c_2:

$$\frac{\partial E}{\partial a_2} = 0 \quad \frac{\partial E}{\partial b_2} = 0 \quad \frac{\partial E}{\partial c_2} = 0 \tag{73}$$

or

$$a_2 \int (1 - H(\phi))dxdy + b_2 \int (1 - H(\phi))xdxdy$$

$$+c_2 \int y(1 - H(\phi))dxdy = \int u_0(1 - H(\phi))dxdy$$

$$a_2 \int x(1 - H(\phi))dxdy + b_2 \int [x^2 + \mu](1 - H(\phi))dxdy$$

$$+c_2 \int xy(1 - H(\phi))dxdy = \int xu_0(1 - H(\phi))dxdy \tag{74}$$

$$a_2 \int y(1 - H(\phi))dxdy + b_2 \int xy(1 - H(\phi))dxdy +$$

$$c_2 \int [y^2 + \mu](1 - H(\phi))dxdy = \int yu_0(1 - H(\phi))dxdy$$

Through solving equations (70), (72), and (74), we obtain the evolution of level set curves and final segmentation. In this process, only one PDE equation (70) needs to be solved. In every step of the iteration to solve equation (70), we use equations (72), and (74) to update constants a_1, b_1, c_1, a_2, b_2, and c_2.

To segment images with more complicated structures, we need use two or more level set functions. For the four phase case, we use two level set functions to segment an image into four regions. The four regions can be approximated by four planar functions:

$$u_{11}(x, y) = a_{11} + b_{11}x + c_{11}y \tag{75}$$

$$u_{10}(x, y) = a_{10} + b_{10}x + c_{10}y$$

$$u_{01}(x, y) = a_{01} + b_{01}x + c_{01}y$$

$$u_{00}(x, y) = a_{00} + b_{00}x + c_{00}y$$

The MS energy functional becomes:

$$E_4^{linear}(a_{ij}, b_{ij}, c_{ij}, \phi_1, \phi_2) =$$

$$\int (u_0 - a_{11} - b_{11}x - c_{11}y)^2 H(\phi_1)H(\phi_2)dxdy$$

$$+ \int (u_0 - a_{10} - b_{10}x - c_{10}y)^2 H(\phi_1)(1 - H(\phi_2))dxdy$$

$$+ \int (u_0 - a_{01} - b_{01}x - -c_{01}y)^2 (1 - H(\phi_1))H(\phi_2)dxdy$$

$$+ \int (u_0 - a_{00} - b_{00}x - c_{00}y)^2 (1 - H(\phi_1))(1 - H(\phi_2))dxdy$$

$$+ \mu(b_{11}^2 + c_{11}^2) \int H(\phi_1)H(\phi_2)dxdy$$

$$+ \mu(b_{10}^2 + c_{10}^2) \int H(\phi_1)(1 - (H(\phi_2))dxdy$$

$$+ \mu(b_{01}^2 + c_{01}^2) \int (1 - H(\phi_1))H(\phi_2)dxdy$$

$$+ \mu(b_{00}^2 + c_{00}^2) \int (1 - H(\phi_1))(1 - (H(\phi_2))dxdy$$

$$+ \nu \int |\nabla H(\phi_1)|dxdy + \nu \int |\nabla H(\phi_2)|dxdy \tag{76}$$

To minimize this MS energy functional, we need to solve the PDEs for two curves, and the solutions of these PDEs are coupled. Therefore the computation cost will be large. To avoid this problem, we use the segmentation approach proposed in [19] to minimize the MS energy functional. In this approach, we first apply one initial curve, then evolve the curve using equation (77):

$$\frac{\partial \phi_1}{\partial t} = \delta(\phi_1)[-(a_1 + b_1x + c_1y - u_0)^2 - \mu(b_1^2 + c_1^2)$$
$$+ (a_2 + b_2x + c_2y - u_0)^2 + \mu(b_2^2 + c_2^2) + \nu \nabla \cdot \frac{\nabla \phi_1}{|\nabla \phi_1|}] \tag{77}$$

Through curve evolution, the curve will segment image into two phases: $\phi_1 > 0$ for inside region and $\phi_1 < 0$ for outside region of the curves. Then we apply the second level

set curve ϕ_2 on the image. With the second curve, each of the two phases segmented by the first curve is segmented by the second level set function. For the inside region 1 ($\phi_1 > 0$), we can solve the curve evolution of the second curve ϕ_2 by:

$$
\begin{aligned}
\frac{\partial \phi_2}{\partial t} =\ & \delta(\phi_2)[-(a_{11} + b_{11}x + c_{11}y - u_0)^2 - \mu(b_{11}^2 + c_{11}^2) \\
& + (a_{10} + b_{10}x + c_{10}y - u_0)^2 + \mu(b_{10}^2 + c_{10}^2) + \nu\nabla \cdot \frac{\nabla\phi_2}{|\nabla\phi_2|}]
\end{aligned}
\tag{78}
$$

And for the outside region 2 ($\phi_1 < 0$)), the curve evolution equation of the second curve ϕ_2 is:

$$
\begin{aligned}
\frac{\partial \phi_2}{\partial t} =\ & \delta(\phi_2)[-(a_{01} + b_{01}x + c_{01}y - u_0)^2 - \mu(b_{01}^2 + c_{01}^2) \\
& + (a_{00} + b_{00}x + c_{00}y - u_0)^2 + \mu(b_{00}^2 + c_{00}^2) + \nu\nabla \cdot \frac{\nabla\phi_2}{|\nabla\phi_2|}]
\end{aligned}
\tag{79}
$$

We obtain the segmentation through the solution of three PDEs (77), (78), and (79). Because the solutions of the three PDEs are decoupled, the computation will be much faster than the direct solution of equation (76).

3.4   Elliptic approximation

In addition to the level set approach, we can use elliptic approximation to solve the MS model [17]. The segmentation curve can be represented by a signature function:

$$
z : \Omega \to [0, 1]
\tag{80}
$$

which is nearly 1 almost everywhere except on a narrow tubular neighborhood of the segmentation curve, where it is close to 0. The MS model is approximated by equation (81):

$$
\begin{aligned}
E &= \int_\Omega |u - u_0|^2 \, dxdy + \mu \int_\Omega z^2 \, |\nabla u|^2 \, dxdy \\
&\quad + \nu \int_\Omega \left(\varepsilon \, |\nabla z|^2 + \frac{(1-z)^2}{4\varepsilon} \right) dxdy
\end{aligned}
\tag{81}
$$

When $\varepsilon \to 0$, the signature function z represents the segmentation curve.

Taking the variation on u and z separately, we can yield the Euler-Lagrange equations as following equations:

$$
(u - u_0) - \mu \nabla \cdot (z^2 \nabla u) = 0
\tag{82}
$$

$$
\mu |\nabla u|^2 z + \nu \left(-2\varepsilon \nabla z + \frac{z-1}{2\varepsilon} \right) = 0
\tag{83}
$$

with the boundary conditions:

$$
\frac{\partial u}{\partial n} = 0, \frac{\partial z}{\partial n} = 0, \text{on the curve}
\tag{84}
$$

Equations (82) and (83) can be solved by efficient elliptic solver and iterative scheme.

Chapter 4

Variations on the MS Model

In this chapter, we introduce different variations of the MS model. These are the ROF MS model, the MS model with second order derivative term, and hierarchical segmentation approaches

4.1 ROF Mumford-Shah model

In the MS model, the gradient of the image intensity is small in object regions, but is large across the boundaries. Therefore the MS model can be used to detect discontinuities in image surface. While the MS model works well in many applications, there are two cases that are difficult for the MS approach. One is the detection of low contrast edges. This is because L^2-norm of the gradient term in the MS energy functional is too large for edges. The minimization of the MS energy functional will remove low contrast edges. Different from the classical MS energy model, Rudin, Osher and Fatemi (ROF) [31] changed the

L^2-norm to the L^1-norm and modeled the energy functional as follows:

$$E(u, C) = \int_{\Omega \backslash C} |u - u_0|^2 dxdy + \mu \int_{\Omega \backslash C} |\nabla u| dxdy + \nu |C| \qquad (85)$$

In this model, the L^1-norm of the gradient can preserve low contrast edges.

Zhang in [30] presented a piecewise constant approximation approach to solve this model. Because the gradient term disappears in the constant approximation, [30] added the gradient of the original image u_0 in the MS energy functional to detect low contrast edges. The MS energy functional used in [30] is:

$$E(u, C) = \int_{inside\ C} (c_1 - u_0)^2 dxdy + \int_{outside\ C} (c_2 - u_0)^2 dxdy$$
$$+ \mu \int_{inside\ C} |\nabla u_0| dxdy + \nu |C| \qquad (86)$$

In this model, [30] only calculated the gradient inside C. The advantage of this approximation compared to the original ROF model is that the calculation is faster since we do not need to solve two coupled PDEs for u inside and outside.

The level set function equation becomes:

$$\frac{\partial \phi}{\partial t} = -\delta(\phi)[(c_1 - u_0)^2 - (c_2 - u_0)^2 + \mu|\nabla u_0| - \nu \nabla \cdot (\frac{\nabla \phi}{|\nabla \phi|})] \qquad (87)$$

4.2 Modified MS model with second order derivative term

Another difficulty with the classical MS model is the detection of roof edges. For this kind of edges, the gradient of u is not small in the regions on both sides of the edge, but the second order derivatives of u is small, and the gradient of u is discontinuous across boundary. That is there is a step edge in the first order derivative functional space. To

35

minimize the classical MS energy functional, the gradient of u inside the segmented regions is forced to be small, but this will make it difficult to detect roof edges. To overcome this difficulty, we modify the MS model to the following form [32]:

$$E(u, C) = \int_{\Omega \backslash C} (u - u_0)^2 dx dy + \lambda \int_{\Omega \backslash C} |\Delta u|^2 dx dy + \int \nu |C| \qquad (88)$$

In the following, we will denote equation (88) as the modified Mumford-Shah (MMS) model. In MMS we ignore the gradient term because this term forces the gradient of u to be small inside each region, but this is not true for roof edges. If we include this term, we will miss the roof edges. On the other hand, we add the term of the second order derivative of u. Because the second order derivative of u is small inside regions and large across both step edges and roof edges, minimizing the second order derivative term in MMS model can detect both types of edges.

The authors of [9] introduced an energy with second order derivative term as following:

$$\begin{aligned} E(u, K_0, K_1) = & \int_{\Omega \backslash (K_0 \cup K_1)} (|\Delta u|^2 + \phi(x, u)) dx \\ & + \alpha \mathcal{H}^{n-1}(K_0 \cap \Omega) + \beta \mathcal{H}^{n-1}((K_1 \backslash K_0) \cap \Omega) \end{aligned} \qquad (89)$$

where α, β are positive parameters; $\mathcal{H}^{(n-1)}$ is the Hausdorff $(n-1)$-dimensional measure. K_0 represents the set of jump points for u (step edges), and $K_1 \backslash K_0$ is the set of crease points (roof edges).

Actually, the MMS model is a special case of equation (89). When $\phi(x, u) = \mu |u - u_0|^2$ and $\alpha = \beta$, equation (89) is equivalent to the MMS model.

Finding the solution of MMS model for an arbitrary image is more difficult than for the MS model. The authors in [9] used elliptic functional approximation to solve equation

(89). In this thesis, we use level set function and piecewise linear approximation to solve equation (88). Compared with the approach in [9], our approach is simpler and faster.

For roof edges, the gradient of u is not small inside the segmented regions, and the image intensity varies inside these regions. So the piecewise constant approximation is not suitable for this case. It could segment one region into several regions. To conserve the information of the gradients, we use the piecewise linear approximation shown by equation (62). The MMS model can be written as:

$$
\begin{aligned}
E(a_i, b_i, c_i, \phi) \;=\; & \int (a_1 + b_1 x + c_1 y - u_0)^2 H(\phi) dx dy \\
& + \int (a_2 + b_2 x + c_2 y - u_0)^2 (1 - H(\phi)) dx dy \\
& + \nu \int |\nabla H(\phi)| dx dy
\end{aligned}
\tag{90}
$$

In equation (90), the second derivative term of u disappears because of the linear approximation.

Using the gradient projection method, we can obtain time dependent equation for the level set function $\phi(t)$, which represents image segmentation.

$$
\frac{\partial \phi}{\partial t} \;=\; \delta(\phi)[\nu \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - (a_1 + b_1 x + c_1 y - u_0)^2 + (a_2 + b_2 x + c_2 y - u_0)^2] \tag{91}
$$

This is the evolution equation for the segmentation curve. Equation (91) is very similar to equation (70) but simpler.

Similar to section 3.3, we can calculate a_i, b_i, and c_i via equations (71) and (73) to

obtain:

$$a_1 \int H(\phi)dxdy + b_1 \int H(\phi)xdxdy$$

$$+c_1 \int yH(\phi)dxdy = \int u_0 H(\phi)dxdy$$

$$a_1 \int xH(\phi)dxdy + b_1 \int x^2 H(\phi)dxdy$$

$$+c_1 \int xyH(\phi)dxdy = \int xu_0 H(\phi)dxdy \tag{92}$$

$$a_1 \int yH(\phi)dxdy + b_1 \int xyH(\phi)dxdy + c_1 \int y^2 H(\phi)dxdy$$

$$= \int yu_0 H(\phi)dxdy$$

and

$$a_2 \int (1 - H(\phi))dxdy + b_2 \int (1 - H(\phi))xdxdy$$

$$+c_2 \int y(1 - H(\phi))dxdy = \int u_0(1 - H(\phi))dxdy$$

$$a_2 \int x(1 - H(\phi))dxdy + b_2 \int x^2(1 - H(\phi))dxdy$$

$$+c_2 \int xy(1 - H(\phi))dxdy = \int xu_0(1 - H(\phi))dxdy \tag{93}$$

$$a_2 \int y(1 - H(\phi))dxdy + b_2 \int xy(1 - H(\phi))dxdy +$$

$$c_2 \int y^2(1 - H(\phi))dxdy = \int yu_0(1 - H(\phi))dxdy$$

Note the slight difference between equations (72), (74) and (93), (94)

The segmentation of an artificial roof edge is shown in figure 6.

In figure 6, the piecewise constant approximation of the MS model segments the image into three regions and cannot detect the roof edge. By contrast, the linear approximation of the MMS model can detect the roof edge. This is because the gradient inside each region is a constant, that is, the second order derivative is small inside each region, and the second
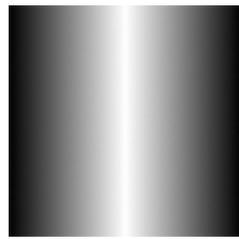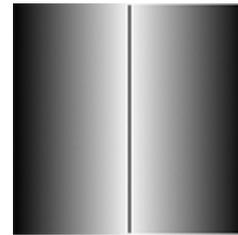
(a) Original image
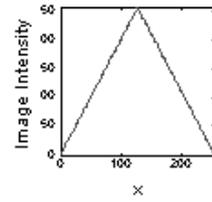
(b) Image intensity distribution along horizontal direction

(c) Segmentation result by the piecewise constant approximation of the MS model

(d) Segmentation result by linear approximation of the MMS model

Figure 6: Segmentation of an artificial roof edge

order derivative is large across the roof edge.

4.3 Hierarchical segmentation

With the level set method, the MS model works well in many applications. However, generally only one or two level set functions are used to segment an image into two or four phases. Therefore, most experiments are done on simple images. For images with complicated structures, the regions in the images cannot be represented by one or two level set functions. We need n level set functions to represent 2^n regions. There are three difficulties for this approach. First, we cannot define the number of regions before hand. We may use many level set functions to segment an image to avoid missing any region.

39

However this approach requires large computational cost. The most important problem for the MS model using the level set method is the initial condition. Because the MS energy functional is not convex, the minimization result often is trapped by a local minimum. Consequently the segmentation result depends on the initial conditions. This problem is more serious for the cases with many level set functions. As a result, generally the MS model can only detect the main structure of an image rather than detailed segmentation.

To segment images with complicated structures, [18] and [19] presented hierarchical segmentation approaches. The basic idea of the hierarchical segmentation approach proposed in [18] is as follows: First, put an arbitrary initial closed curve in the image. Through the evolution, the curve will segment the image into two or more regions. Because in the evolution, the closed curve could split into two or more closed curves, the segmentation result could consist of more than two regions. In the second stage, put another arbitrary initial closed curve in each region obtained from the previous segmentation. Then through curves evolution, the curves inside each region will segment the region into more regions. We can put more curves into regions obtained from the last segmentation stage and segment the image into more regions. In this way, we can segment the image into any number of regions. The difference between [18] and [19] is the following: after one segmentation stage, [19] puts a second curve over all regions, instead of one curve in each region, and segments all the regions simultaneously.

In the approach in [19], we cannot decide if one region needs more additional segmentations, and there is no stopping criterion. On the other hand, the stopping criterion mentioned in [18] suggests that this hierarchical segmentation process will stop when the

new segmentation will not change the MS energy functional. Because the segmentation results depend on the initial conditions, whether the MS energy of the new segmentation is smaller than the original MS energy depends on the initial conditions. Even when the MS energy of the new segmentation is not smaller than the original MS energy, it is still possible that the region needs additional segmentations. So this stopping criterion cannot guarantee the final segmentation result. Although the hierarchical segmentation approaches in [18] and [19] can segment an image into many regions, these approaches cannot guarantee the segmentation is correct. Because the MS energy functional is not convex, it is possible that the segmentation curves do not reach the boundaries of objects, and one object may be segmented into several regions, i.e. the image is over segmented.

For images with complicated or detailed structures, although the MS model can produce segmentation results for the main structures, local segmentation approaches can produce better results for detailed structures. To combine the advantages of these two kinds of approaches, we present a new hierarchical segmentation scheme which makes use of both global and local information [33]. In the process of the hierarchical segmentation, we use a local window (a small area of image intensity) over a segmentation region to detect whether the region needs additional segmentations or not. If the region needs more segmentation, we put an initial segmentation curve in the region and do more segmentation. Otherwise, we stop the segmentation of the region. In this way, we can segment the image hierarchically until each object region is smooth. The advantages of this approach are that the segmentation result does not depend on the initial conditions and the method is relatively fast.

The algorithm for hierarchical segmentation is given below:

Algorithm: Hierarchical Segmentation

Input: image to be segmented.

Output: segmented image.

Procedure:

1. Define an initial close curve with area threshold T_a.

2. Move the initial curve to different positions in one region (in the beginning, the region is the whole image), and calculate the average intensities inside the curve and the whole region. If at any position, the absolute difference between the two average intensities is smaller than the contrast threshold T_c the region needs no more segmentation. Otherwise, position the initial curve at the location where the absolute difference is the largest.

3. In the region, use piecewise constant approximation of the MS model to solve the curve evolution and obtain the segmentation result.

4. Number different regions in the image.

5. In each region, repeat steps 2 and 3 to segment the region into smaller regions. In this step, all the calculations are inside one region.

6. Repeat 5 for each region obtained from different segmentation stages until all regions need no more segmentation.

7. Calculate the average intensity of each region. If the absolute difference between the average intensities of two neighboring regions is smaller than the contrast threshold T_c, the two regions are merged.

8. Calculate the area of each region. If the area is too small, for example less than 10 pixels, the region is regarded as noise, and is merged into its neighboring region.

To detect all the details of an image, we can define the minimum object area as one pixel. If there is noise, however, noise can also be detected as an object. Therefore, a minimum object area is defined to avoid the effect of noise in step 1.

In step 2, it is obvious that if the absolute differences between the probe areas and the background are smaller than the contrast threshold, there is no object in the region, and the region needs no more additional segmentation. If at some positions, the difference is bigger than the contrast threshold, we position the initial level set curve at the location where the difference is the largest. At this location, the MS energy of the initial curve is the smallest compared to all other locations.

Repeat steps 2 and 3 on all regions until no more segmentation is needed in each region. At this point, each region is smooth enough. However, this is not the final result. Because the curve evolution of segmentation does not guarantee that the curve reaches the boundary of an object. It is possible that an object is segmented into more than one region. In this case, we merge those neighboring regions whose average intensities are similar (step 7). After merging, the image intensity changes smoothly inside objects and the boundaries are kept because neighboring regions with large differences are not merged.

In step 8, we remove those regions with very small area that are assumed to be noise instead of object.

In this segmentation scheme, there are two important parameters: the area threshold and the contrast threshold of objects. We use these two parameters to define objects in image, that is, the object to be segmented in image should be bigger than an area threshold, and the contrast between object and background should be bigger than a contrast threshold. These two parameters can be obtained from observations of an image. The minimum object area T_a is smaller than the minimum object in image. This value can be roughly estimated. As long as it is smaller than the minimum object and bigger than noise speckle, the exact value won't affect segmentation result.

To define T_c, we first find the object with the lowest contrast to background in image through observation. Then we calculate the average intensities of the object and its neighboring background respectively. T_c is estimated as the absolute difference between these two values.

Instead of defining these two thresholds directly, the parameters ν and μ are defined in the MS model. Since μ disappears in the piecewise constant approximation, ν becomes the only parameter. Gao and Bui [19] suggest that ν is proportional to the variance σ^2 of an image, and suggest a proportional ratio range. In this thesis, we prefer to calculate ν from the contrast threshold in the object definition: $\nu < T_c^2/2$ [29]. While most papers define ν directly, we believe that it is more reasonable to estimate ν from the minimum object contrast because the contrast can be observed from the image itself. However, it is difficult to define T_c for images with low contrast and large noise. In this case, T_c must be defined
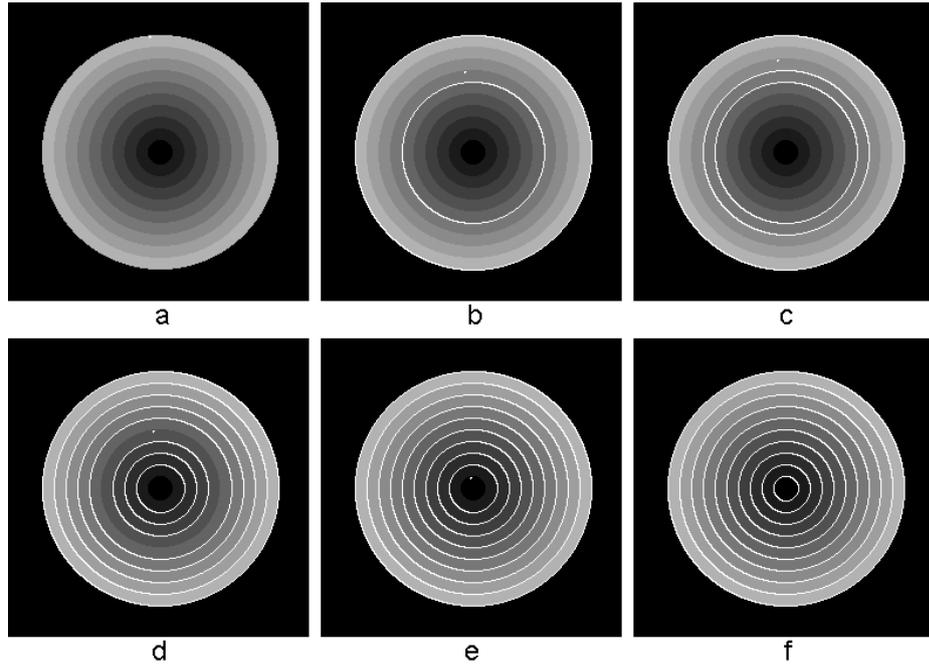
Figure 7: Hierarchical segmentation of circles

accurately; otherwise, some objects with low contrast could be missed, or some noise could be detected as objects. We can only roughly estimate T_c through observation. The accurate value sometimes has to be defined through some segmentation experimentations.

Figure 7 shows the hierarchical segmentation process. Figure 7 (a) is the initial curve and the original image; Figures 7 (b - e) show intermediate stages of hierarchical method. Figure 7 (f) is the final result. The initial curves of different stages can be located automatically, and after evolution, the image is hierarchically segmented into regions. Finally, the segmentation stops until reach region is smooth.

Chapter 5

Image Inpainting Using the MS model

This chapter introduces application of the MS model in image inpainting.

5.1 Image inpainting

Image inpainting is originally an artistic procedure to recover a damaged painting or picture. It has been introduced in [34] and received attention from many researchers in computer vision and image processing.

From a technical point of view, digital image inpainting can be described as a procedure to fill a defined *inpainting domain* (i.e. a set of damaged pixels in a given image). In this case, we can recall some related topics in image processing and computer vision, namely image interpolation, object removal or disocclusion, texture synthesis, film restoration, and damaged block recovery of compressed digital image or video.

A straightforward approach for image inpainting is interpolation. However, interpolation appraches can not handle complicated content of inpainting image areas, especially for big inpainting area. Chan and Shen [35] first employed the Bayesian framework for image inpainting. They presented the total variation (TV) inpainting model. The model is as following:

$$E(u, C) = \int_\Omega |\nabla u|^2 \, dxdy + \frac{\lambda}{2} \int_{\Omega \setminus C} |u - u_0|^2 \, dxdy \tag{94}$$

The advantages of TV model include lower complexity and easy digital implementation. Because the model only considers the variance of image intensity, it often over smoothes the image. It works well for local inpainting problems, including digital zooming and superresolution. However the over smoothing effect is obvious for big inpainting area. At coarse scale, images can be approximated by smooth functions. But in small scale, images contain a lot of details, including edges, corners, T-junctions, and textures. These singular features can't be approximated by smooth functions, but they are crucial for human perception. Recently, many image inpainting approaches consider these factors for human perception.

Masnou and Morel [36] addressed a disocclusion approach using level lines to achieve strong discontinuities. Bertalmio et al. [34] formulated a third-order partial differential equation (PDE) which considers the continuity of the isophotes at the boundary and a measure of the change in the information to be propagated. However, the method leads to a number of iterative operations and fails in regions with more complicated image structures. They proposed an improved method for simultaneous filling-in of texture and structure by combining an image decomposition algorithm and texture synthesis [37].

In [38], Criminisi et al. proposed a comparatively fast inpainting algorithm for filling a large region occluded by an object. Their method fills the region with a selected exemplar (a small patch in the image) based on an edge-driven reconstruction criterion. However, the method was designed only for estimated edges with straight lines in the inpainting area. In order to overcome this problem, graph-based structure and texture propagation by either confidence map or interactive guidance could be applied for better results as proposed in [39]. Methods using transform-based sparse representation such as discrete cosine transform or wavelet transform were also proposed for the approximation of unknown regions [40, 41, 42, 43].

In this thesis, we use the Mumford-Shah model and the level set method to detect edges in the inpainting domain. In our approach, image structure of the damaged region is estimated by our hierarchical method. For better region segmentation and edge estimation in the inpainting domain, we adopt a hierarchical approach which uses multi-level set functions. Although multi-level set functions can be used to segment an image into many regions, classical approaches cause extensive computations and the solutions also depend on initial conditions. This is a critical issue for practical applications. Our proposed approach utilizes a faster and stable hierarchical level set method. Because we detect both the main structure and detailed edges, our approach preserves edges in the inpainting area. In addition, our approach adapts a textured region filling procedure based on estimated segmentation map. Since texture plays a critical role to express detailed information in an image, a variational approach to fill in the inpainting region may not be enough as pointed out in [37].

5.2 The MS model for inpainting

Tsai, Yezzi, and Willsky [18], and Chan and Shen [14] are the first to present the idea of applying the MS model to inpainting. For inpainting, they modified the MS model as follows.

$$E(u, C) = \int_\Omega \lambda(x, y) \left| u - u_0 \right|^2 dxdy + \mu \int_{\Omega \backslash C} \left| \nabla u \right|^2 dxdy + \nu \cdot |C| \qquad (95)$$

where $\lambda(x, y) = 0$ if (x, y) is inside the inpainting area and 1 otherwise. The above equation indicates that only the variance of image and the length of segmentation curve are considered inside inpainting area.

The solution of the MS energy functional is not a trivial task. We can use piecewise smooth approximation of the Mumford-Shah model alternatively. If we consider that a closed curve segments an image into two regions (i.e. inside and outside regions), the MS energy functional can be written as:

$$
\begin{aligned}
E(u_1, u_2, C) &= \int_{\text{inside } C} \lambda(x, y) \left| u_1 - u_0 \right|^2 dxdy + \mu \int_{\text{inside } C} \left| \nabla u_1 \right|^2 dxdy \\
&+ \int_{\text{outside } C} \lambda(x, y) \left| u_2 - u_0 \right|^2 dxdy + \mu \int_{\text{outside } C} \left| \nabla u_2 \right|^2 dxdy \\
&+ \nu \cdot |C| \qquad (96)
\end{aligned}
$$

where u_1 and u_2 are smooth approximations of the image inside and outside the curve. Following the same procedure in chapter 3, we can get piecewise smooth approximation PDEs for image inpainting.

$$\lambda(x, y)(u_1 - u_0) = \mu \nabla^2 u_1 \text{ inside } C, \text{ and } \frac{\partial u_1}{\partial \vec{n}} = 0 \text{ on } C \qquad (97)$$

$$\lambda(x, y)(u_2 - u_0) = \mu \nabla^2 u_2 \text{ outside } C, \text{ and } \frac{\partial u_2}{\partial \vec{n}} = 0 \text{ on } C \qquad (98)$$

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\left[\nu\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \lambda(x,y)(u_1 - u_0)^2 - \mu|\nabla u_1|^2\right.$$
$$\left. + \lambda(x,y)(u_2 - u_0)^2 + \mu|\nabla u_2|^2\right] \tag{99}$$

The smooth image functions u_1 and u_2 can be obtained by solving the damped Poisson equations (97) and (98), and the segmentation curve can evolve according to equation (99). This is the piecewise smooth approximation presented by Chan and Vese. To reduce the computational cost of this approach, we also can use the piecewise constant approximation.

If the image intensities inside different regions are uniform, the image intensities inside different regions can be approximated by constants. In this case, the MS energy functional for image inpainting can be written as Equation (100):

$$E(c_k, C) = \sum_k \int_{\Omega_k} \lambda(x,y)(c_k - u_0)^2 dxdy + \nu|C| \tag{100}$$

where Ω_k represents the area inside each region. The gradient term in the MS energy functional disappears in equation (100) because the gradient inside each region is zero.

Using the level set method and the MS energy functional of the two phase segmentation, the image is segmented into two phases:

$$E(c_1, c_2, \phi) = \int \lambda(x,y)(c_1 - u_0)^2 H(\phi)dxdy \tag{101}$$
$$+ \int \lambda(x,y)(c_2 - u_0)^2(1 - H(\phi))dxdy + \nu\int \delta(\phi)|\nabla\phi|dxdy$$

where $H(x)$ is the Heaviside function.

To minimize the energy functional with respect to c_1, c_2, and ϕ, we obtain the following equations:

$$c_1(\phi) = \frac{\int \lambda(x,y)u_0 H(\phi)dxdy}{\int H(\phi)dxdy} \tag{102}$$

$$c_2(\phi) = \frac{\int \lambda(x,y)u_0(1 - H(\phi))dxdy}{\int (1 - H(\phi))dxdy} \tag{103}$$

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\left[\nu\nabla \cdot \left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \lambda(x,y)(u_0 - c_1)^2 + \lambda(x,y)(u_0 - c_2)^2\right] \tag{104}$$

After solving these equations, we obtain the information on c_1, c_2, and C. The image u_0 will then be segmented into two phases $\{u = c_1\}$ and $\{u = c_2\}$.

5.3 Image inpainting with hierarchical segmentation and texture filling

With the level set method, the MS model works well in many applications. However, generally only one or two level set functions are used to segment an image into two or four phases. Therefore, most experiments have been performed on relatively simple images. For images with complicated structures, this approach can only detect main structures. As mentioned in the beginning of this chapter, detailed structures are critical for human perception. To detect and preserve details, we adopt hierarchical method to segment images. This approach can detect image structures gradually to detailed structures [44].

Before image inpainting, the inpainting area is manually defined. For color images, the color image is converted to gray level. First we use diffusion approach to fill inpainting regions marked by inpainting mask. This result is only used for segmentation. Then the image is segmented by the hierarchical segmentation method described in section 4.3. After segmentation, the boundaries of different objects are detected and estimated inside the inpainting area. Finally we use a diffusion-based technique or texture patch to fill the

inpainting area. When the inpainting area is filled, the edges inside the inpainting area are preserved because the diffusion is not across the boundaries.

In segmentation, a piecewise constant approximation is used. Compared with piecewise smooth approximation, it is not necessary to solve three coupled PDEs and it is relatively fast. We can detect boundaries of most regions. However, the piecewise constant approximation does not work well for regions with high frequency, image intensity variance, and strong noise. For this kind of regions, we define minimum segmentation region area as the whole region instead of the high frequency detailed structure or noise speckles. For noisy images, we define the minimum segmentation regions to be larger than 3σ, that is bigger than most noise speckles. The segmentation result only detects region boundaries. The high frequency details or noise will be inpainted by the texture filling approach as described in the following.

One of the main problems for image inpainting algorithms based on variational or PDE models [34], [45], [46], [35], [47], [48] is that texture, which expresses detailed patterns in an image, is not considered.Our proposed approach can directly fill in inpainting areas using diffusion-based technique. In this case, textured region to be filled may lose detailed information because of too much smoothness. In order to avoid this problem, we propose a region filling approach using segmentation map result and similar texture patches.

Segmentation map estimated in the hierarchical segmentation suggests the most probable image structure of the inpainted area. On top of the structure layer, it is desirable to overpaint using the most plausible texture patches. Recently some inpainting approaches

utilize texture synthesis algorithm to obtain close similarity between the unknown inpainting region and surrounding image regions [37], [38]. We also use a similar approach to apply plausible example texture patches. This procedure can be summarized as follows:

Algorithm: Inpainting by Similar Texture Patches

Input: image to be inpainted, segmentation map and inpainting region mask.

Output: inpainted image.

Procedure:

1. Segmentation map is obtained using the hierarchical approach. The map includes estimated edges and partitioned regions in the inpainting domain.

2. Patch priorities in the inpainting domain are computed for deciding the filling orders of inpainting region [38]. In our proposed method, edges or boundaries of segmentation map in the inpainting area have higher priorities and are considered first. Also inpainting boundary has higher priority. For example, a center point of a patch P in figure 8 can have the highest priority since the point is on the boundary ($\partial\Gamma$) and the segmented edge (e).

3. Find the most similar patch on the edge and copy it into the patch region. An example patch P' can be selected by measuring the similarity with P in terms of pixel values and segmented regions in figure 8. Similarity measurement between P and P' could be performed by various different ways. In our case, we simply used Euclidean distance measure between two patches. To speed up the searching step, multiple searching ranges could be defined. With a low similarity in a smaller range, the next

searching range is extended.

4. Finally, fill in non-edge inpainting area. Similar example patch is first detected in the same segmentation region.
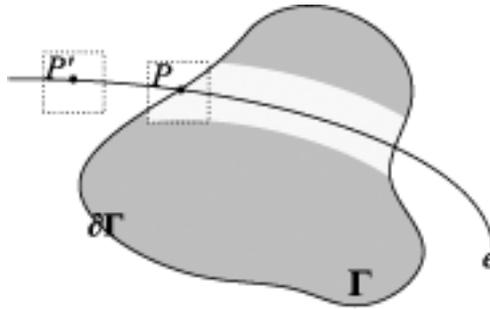
Figure 8: Texture patch on the edge

5.4   Experimental results

To evaluate and compare our proposed approach with other approaches, we implement the inpainting algorithm with total variation (TV) approach, the MS model approach with one level set function, and the MS model approach with hierarchical multi level set functions respectively.

Figure 9 is the inpainting results on the image of peppers. For the MS model approach, after segmentation of the image, we can detect boundaries of objects in inpainting area, and the edges of objects are kept very well. In the detailed images (c) and (f), it is obvious that the inpainting result of the MS model approach is better than that of the TV approach.

Figure 10 is the inpainting results on the image of a street scene. Because the segmentation with hierarchical multi level set functions detects more edges than the segmentation
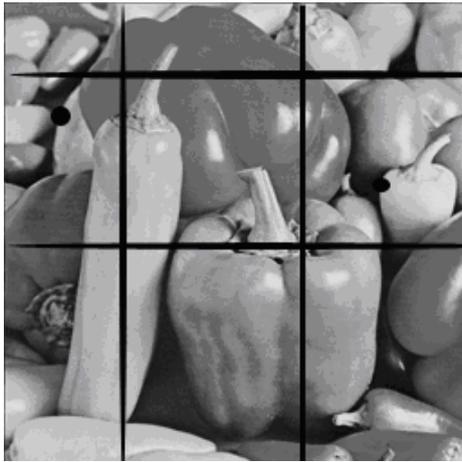
(a) image with inpainting lines

(b) inpainting result with TV
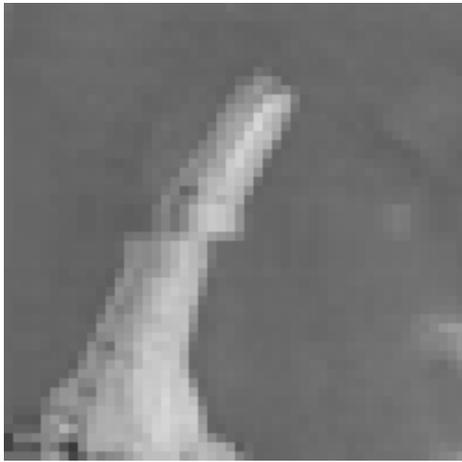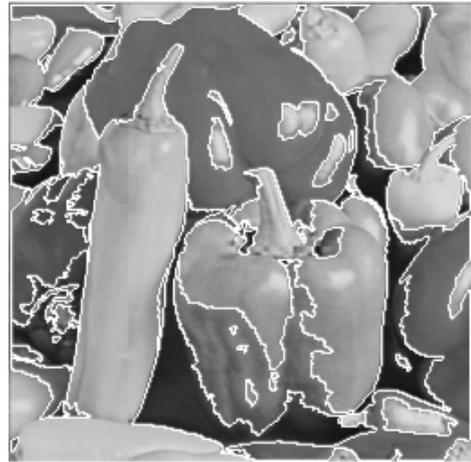
(c) detail of (b)

(d) segmentation result with MS

(e) inpainting result with MS

(f) detail of (e)

Figure 9: The inpainting results on the image of peppers.

(a) image with inpainting lines   (b) inpainting result with TV   (c) detail of (b)

(d) segmentation result with MS   (e) inpainting result with MS   (f) detail of (e)

(g) segmentation result with MS model and hierarchical multi-level set functions   (h) inpainting result with MS model and hierarchical multi-level set functions   (i) detail of (h)

Figure 10: The inpainting results on the image of street.

with only one level set function, more edges inside the inpainting area are preserved. While the edge of the car is blurred in detailed images (c) and (f), the edge is preserved in image (i).

To show the advantages of the proposed algorithm, we tested the algorithm on some city images. In figure 11, (c) is the TV inpainting result. We can see that the upper boundary of the central black building is not recovered very well. (d) is the segmentation result. (e) is inpainting result based on the segmentation result. Because the boundary is detected, the inpainted boundary of the building is preserved. (f) is the result using both segmentation and texture filling approaches. In figure 11(f), both large and small structures, and even detailed textures are recovered very well.

In figure 12, (c) is the segmentation result. In this figure, we detect most edges. But there are also some edges not detected. They are high frequency part inside segmented regions. (d) is diffusion inpainting result based on the segmentation map. We can see that the missing edges are not recovered well, such as the straight horizontal lines on the central building. In this case, we consider those missing small edges as texture. We use texture filling approach to recover them. (f) is inpainting result using both segmentation and texture filling approaches. We can see that both large boundaries and detailed edges are recovered. Specifically, the straight horizontal lines on the central building are exactly the same as the original image. The segmentation approach can detect and preserve most edges. It also segments an image into different regions for texture filling. Texture filling is only applied inside the inpainting regions but not across segmentation boundaries. In this way, different regions will not be filled by wrong textures. This approach can recover details missed by
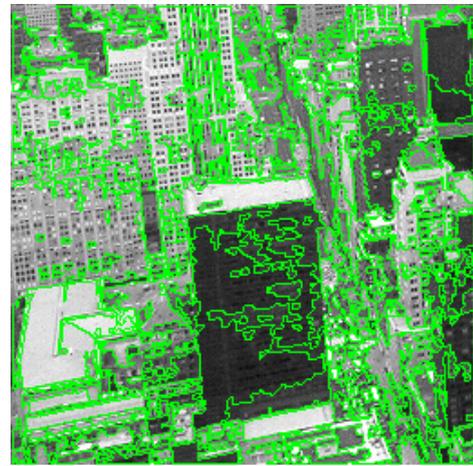
(a) image with inpainting lines

(b) inpainting mask

(c) inpainting result with TV

(d) segmentation result

(e) inpainting result with segmentation

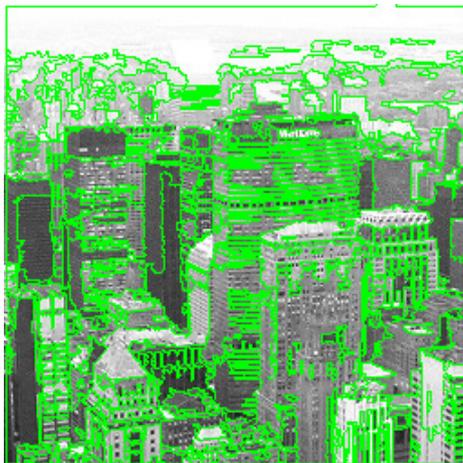(f) inpainting result with segmentation and texture filling

Figure 11: The inpainting results on the image of city 1.
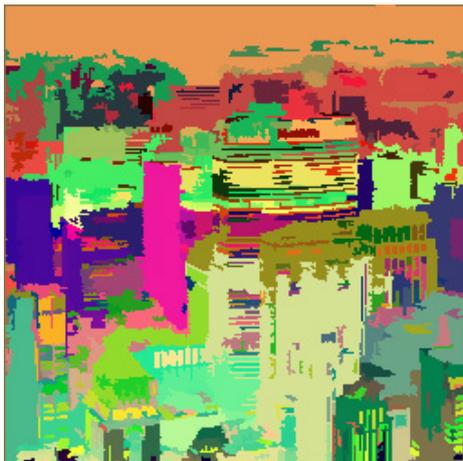
(a) image with inpainting lines

(b) inpainting mask

(c) segmentation result

(d) inpainting result with segmentation

(e) Segmentation map

(f) inpainting result with segmentation and texture filling

Figure 12: The inpainting results on the image of city 2.

the segmentation step. As a result, almost all high frequency details can be recovered.

We also applied our approach to a real inpainting example. Figure 13 shows the artificial inpainting result on a small part of the 'Last Judgement', Michaelangelo's fresco located in Sistine Chapel. The algorithm helps estimating the right edge structure discontinued by wall cracks.

Finally, Figures 14 and 15 show examples of object removal and image completion. After inpainting regions are decided manually, segmentation map is constructed for edge estimation and the exemplar-based filling method has been applied to these images. The estimated boundary between different textured regions is defined in the segmentation map. For example, the ridge of a mountain occluded by the rightmost pillar in figure 15(d) is recovered based on edge information of figure 15(c).

Our inpainting method contains two steps: segmentation and texture filling. Segmentation step requires an iterative approach to solve the PDE of segmentation curve. Unlike the piecewise smooth approximation approach, piecewise constant approximation only requires to calculate the average intensity of each segmentation region for each iteration instead of solving the PDEs in equations (97) and (98) to calculate the smoothed image intensity inside segmentation regions. The computation complexity of segmentation depends on segmentation curve evolution. For our proposed hierarchical piecewise constant approximation, each curve evolution iteration takes constant computations to calculate equations (102), (103), and (104). So the complexity is $O(n)$. For the piecewise smooth approximation, we need iteratively solve equations (97) and (98) to compute one curve evolution iteration in equation (99). Each iteration to solve equations (97) and (98) takes constant
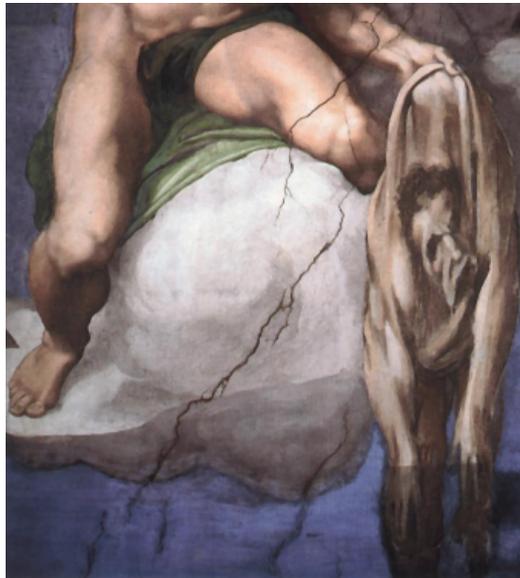
(a) Original image
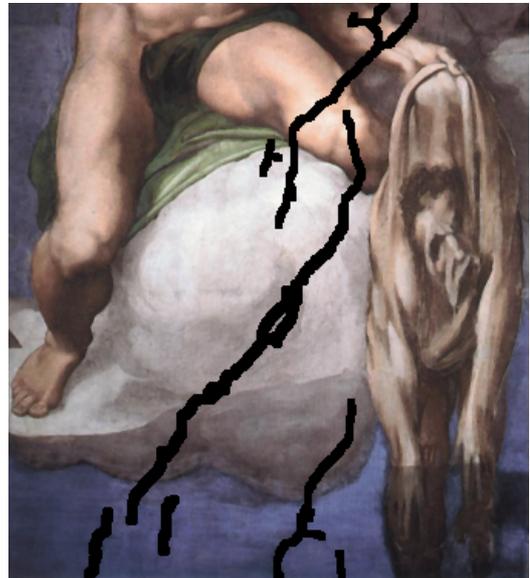
(b) Masked image
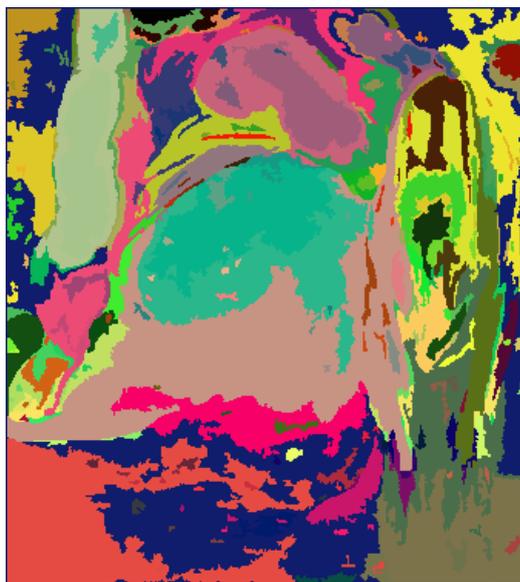
(c) Segmentation map

(d) Inpainted image

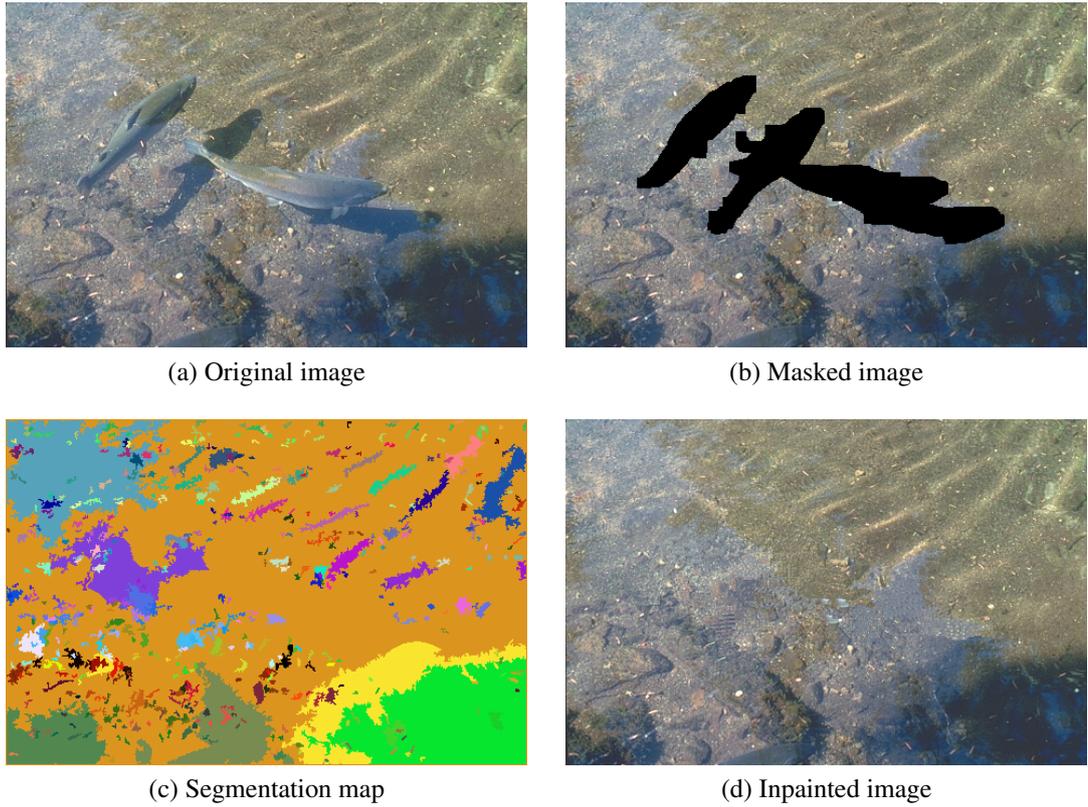Figure 13: Inpainting results using Michaelangelo's Last Judgement image (partial)

(a) Original image   (b) Masked image

(c) Segmentation map   (d) Inpainted image

Figure 14: Inpainting results using 'fish' image

computations. So the complexity is $O(n^2)$. This is caused by the coupled 3 PDEs. The complexity of texture filling is $O(n)$. The piecewise constant approximation approach is much faster than piecewise smooth approximation. The actual computation time depends on image size and computation platform. For our experiments using piecewise constant approximation, it takes about 20s to finish one level segmentation of an image of the size 256 by 256. It is faster than the piecewise smooth approximation, which takes more than 1 minute. To complete the hierarchical segmentation of an image, several levels of segmentation need to be applied. The total computation time depends on image structure. The main calculation in texture filling is patch matching. Because texture patch is small, the matching calculation is fast. The actual computation time depends on the area of inpainting

(a) Original image

(b) Masked image

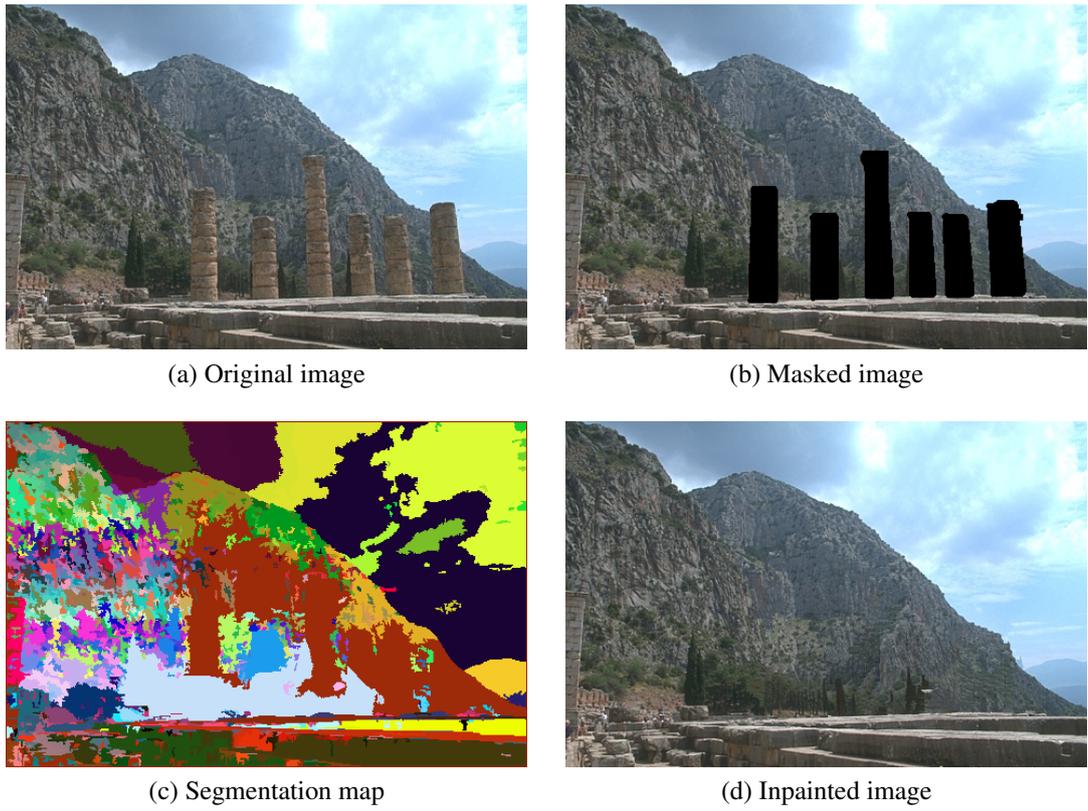(c) Segmentation map

(d) Inpainted image

Figure 15: Inpainting results using 'pillar' image

region. With our non-optimized Matlab and C++ codes, the total processing time, including

segmentation and texture filling, is about 2 to 3 minutes.

Chapter 6

Text Line Detection in Handwritten

Documents Using the MS Model

This chapter introduces application of the MS model in text line detection.

6.1 Text line detection

Text line segmentation in digital documents is an important step in document processing. Many methods have been developed in the document image processing community, which usually work fine for printed documents. Complete review of these methods can be found in [49], [50], and [51]. These text line segmentation approaches usually make two assumptions: 1) the gap between two neighboring lines is significant; and 2) the lines are reasonably straight. However, these assumptions are not always valid for handwritten documents. Some of these methods have been extended to handle handwritten documents. In

[52], a cut text minimization (CTM) technique was proposed to segment text lines from handwritten English documents, which varies the cutting angle and the start location to minimize the text pixels cut while tracking between two text lines. A graph model, which is similar to the document spectrum [53], was proposed in [54] and a production system was applied to this model for text line extraction. Projection profile analysis was applied in [55] to identify the boundaries of the text lines. Instead of splitting those touched text lines with a straight cut, a droplet line segmentation method was proposed to simulate the behavior of a water-drop that fell vertically along the boundary of a text line and identify the actual touching points that need to be cut.

All these techniques of text line detection are successful in limited situations and cannot handle free style handwritten documents. Processing of such handwritten documents is still an open problem. The authors of [56] and [57] proposed to use the level-set method to segment the text lines. In [56], the binary text image is first converted to a gray scale image, and the text line segmentation is achieved by using a general image segmentation approach: the level set method. More comprehensive development of this idea was given in [57] together with a thorough evaluation of this method on documents written in nine scripts. Comparisons between classical text line segmentation methods are also given. The method in [56] and [57] obtains very good results. However, the approach in [56] and [57] also faces some difficulties: Because the boundary based level set method is used, the segmentation result depends on the number of boundary evolution steps. In addition, the failure cases shown in [56] and [57] indicate that the approach is sensitive to the overlaps between neighboring text lines. In this thesis, we present a new algorithm based on the

Mumford-Shah (MS) model. Because the text image only consists of two uniform regions: text region and background region, the piecewise constant approximation of the MS model is very appropriate for the segmentation of text lines. Different from the approach in [56] and [57], the MS model is a region based approach. Segmentation is achieved by minimizing the MS energy functional. Therefore, the segmentation result does not depend on the number of evolution steps. We also use morphing to remove overlaps between neighboring text lines and connect the broken lines. Compared with the text line localization approach in [56] and [57], our approach takes advantages of the level set representation to morph the segmentation regions. Our method is straightforward and easy to implement. The approach also can overcome the problem of overlaps between neighboring text lines.

6.2 Text line segmentation algorithm

The level set approach is first used to solve the text line segmentation problem by the authors in [56] and [57]. They proposed a three-steps algorithm for text line segmentation: (1) blurring the text image to enhance text lines; (2) segmentation of text lines by the level set method; (3) text line localization. We adopt the first step of their algorithm, but we use different approaches in the other two steps. Our algorithm consists of the following steps: (1) blurring the text image to enhance text lines; (2) segmentation of text lines by the Mumford-Shah model; (3) text line detection by morphing approach [58].

1. Blurring text lines

   This step is the same as the first step in [56] and [57]. To enhance the text lines, the

66

text image is blurred by a Gaussian filter. The filter window is a rectangle. Its width is larger than its height. Assuming the text lines are almost horizontal, a wide window blurs words and fills the horizontal gaps between the words. As the result, isolated words are connected into horizontal lines. On the other hand, a too narrow window may not be able to close the gaps between the words on the same line. In the blurred image, we cannot distinguish detailed structure of characters. The only information is the blurred lines across the image. Different from the original binary text image, the blurred image is a gray scale image. The intensity of each pixel represents its probability to be on the text lines: the darker the pixel, the more likely the pixel is on the text lines. We segment the text lines according to the image intensity.

2. Text line segmentation

We use piecewise constant approximation of the Mumford-Shah model to segment text lines. Because the text image can be represented by two phases: text line region and background region, we use one level set function to represent the regions, i.e.

$$\phi(x, y, t) = \begin{cases} > 0 & \text{if } (x, y) \text{ is inside } C \\ = 0 & \text{if } (x, y) \text{ is on } C \\ < 0 & \text{if } (x, y) \text{ is outside } C \end{cases} \tag{105}$$

We use two constants to represent image intensities of two regions. In segmentation process, an initial curve is set to segment image into two regions: inside and outside regions. Then segmentation curve evolves to the boundaries of text lines according to the minimization of the MS energy functional. In each evolution step, the constant c_1 or c_2 of each region is calculated by the average of the region intensity of the region,

67

as in Equations (53) and (54). The segmentation curve evolves to the boundary of text lines according to the Euler-Lagrange equation (51). Finally, the image is segmented into two regions. Different from the level set approach in [56] and [57], our approach does not depend on the number of evolution steps, and the segmentation can automatically stop after the MS energy is minimized.

3. Text line detection

After the image segmentation step, image is segmented into two regions: text line region and background region. Because of the overlaps among different text lines, neighboring text lines may be connected. Another problem is the broken text lines due to large horizontal gaps between different words. We use morphing approach to overcome these problems.

First, we shrink (erode) the text line region along the horizontal direction. If the overlap between two neighboring (vertical) text lines is not large, the connection of the two text lines is thin along the horizontal direction. In this case, the shrinkage of the text line region will remove the overlaps.

Second, we prolong (dilate) the text line region horizontally. The prolongation length is longer than the shrinkage length. If the gaps between neighboring (horizontal) broken text lines is small, the horizontally broken text line will be connected and merged into a long text line across the image.

Third, we shrink the text line regions back to the original horizontal length.

Finally, we check the length and height of every segmentation region. If the length or

68

height is too small, this segmentation region is considered as noise and is removed.

The level set method is a good approach to obtain the segmentation regions, it is also easy to implement the morphing based on the level set approach. For the morphing process, the evolution equation of the level set function is given by:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)F \tag{106}$$

where F is the morphing speed.

6.3   Experimental results

We test the algorithm on different text images. These images were provided by Dr. Doermann and his team at the University of Maryland. In figure 16, the original image (a) first is blurred as (b); then it is segmented as (c) by the Mumford-Shah model. In the segmentation, we need not fix the number of iteration steps, and we can obtain good segmentation results automatically. After this step, there are many overlaps between neighboring text lines, and some text lines are broken as seen in figure 16(c). As shown in (d), (e), and (f), we shrink text lines to remove overlaps among neighboring text lines, and prolong text lines to connect broken cases. The final step removes the noise, and the final result is shown in (g). For comparison, we use the level set method as in the algorithm of [56] to segment text image of figure 16(a). Figure 17(a) is the initial segmentation curve; b, c, and d are the segmentation results with different iteration steps. As shown in the figure, the segmentation results with different iteration steps are quite different. In the algorithm of [56], we need to fix the number of iteration steps to get the best result. In addition, even in the best result, there are
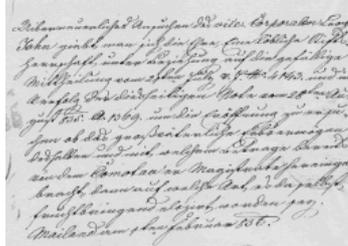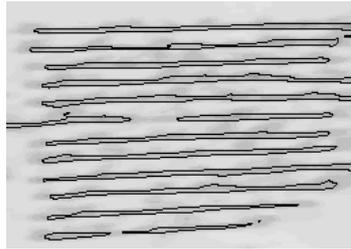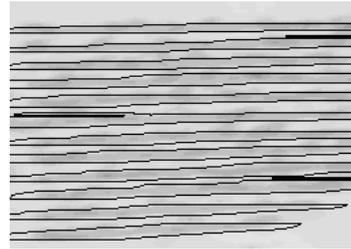
(a) Original image
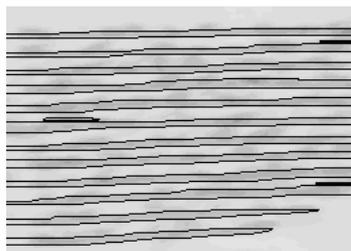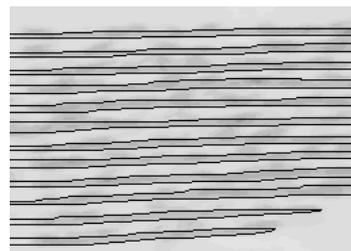

(b) blurred image
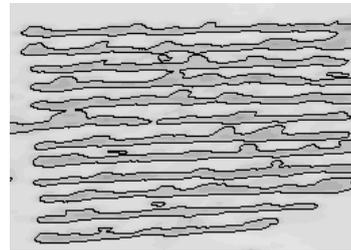
(c) segmented image

(d) shrunk text lines

(e) prolonged text lines

(f) shrunk back to original text line length

(g) final segmentation result after removing noise

Figure 16: The process of text line segmentation algorithm

a lot of overlaps among different text lines, and these overlaps can not be removed by the text line localization in [56]. As a result, those text lines with overlaps can not be separated. The approach in [57] improves on the approach in [56]. The major modification is that in each boundary evolution step, the algorithm detects new overlaps created by the boundary evolution and removes these overlaps. In their algorithm, we need to scan all segmented regions and analyze the relationship among them to detect overlaps in each evolution step. So this algorithm is expensive computationally. In addition, if the overlaps already exist in the initial condition, the overlaps will not be removed by this approach. Figure 18 shows the experimental results on different language text images. Especially, figure 18(g) is a mix of Chinese and mathematical equations. The results indicate that our algorithm is script in-dependent. As comparison, we also give in this figure the segmentation results of a method that is based on connected component analysis.

Figure 19 is the segmentation results of figure 18(a) using the algorithm in [56]. Figure 19a is the initial segmentation curve; b, c, and d are the segmentation results with different iteration steps. In Figures 19(c) and d, more iteration steps create more overlaps among text lines, and many text lines can not be separated. This indicates that the choice of the iteration steps is critical to the algorithm in [56].

Figure 20 is the text line detection results. The different colors represent different text lines. Figures 20(a) and 20(d) are the text line detection results using our algorithm. Our algorithm can detect all text lines. Figures 20(b), 20(c), 20(e), and 20(f) are the text line detection results using the algorithm in [56]. Figures 20(b) and 20(e) are the results of Figures 17(b) and 19(b). Because of the overlaps among text lines, some text lines can

71

(a) initial segmentation curve

(b) segmentation result of 8 iteration steps

(c) segmentation result of 10 iteration steps

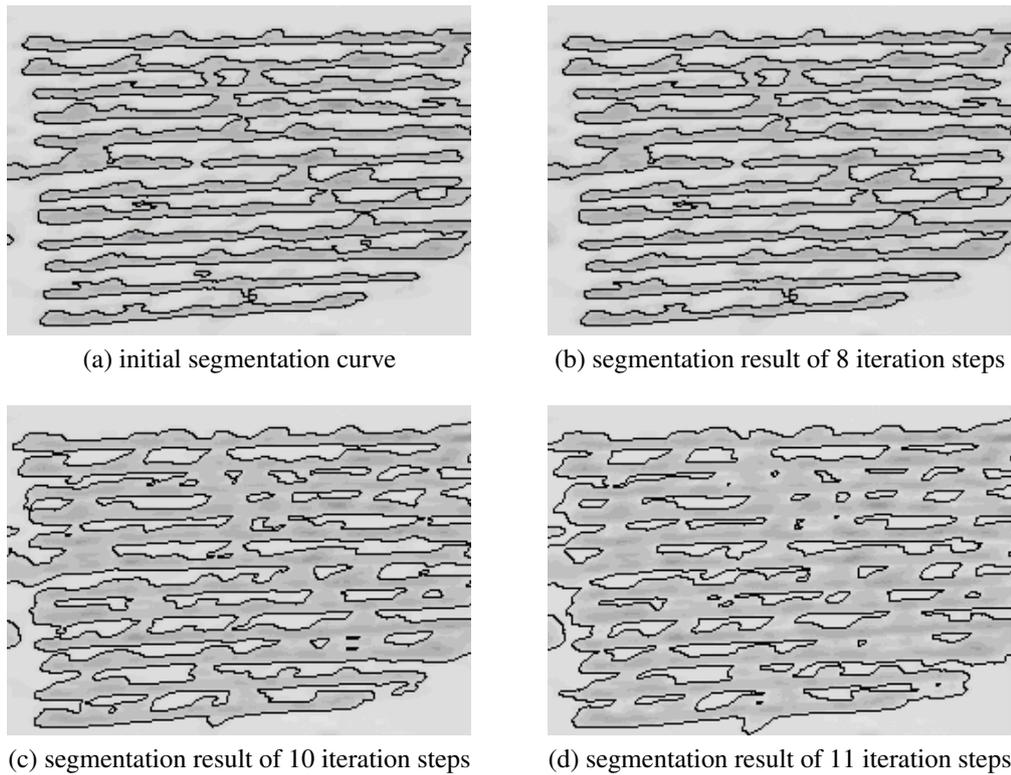(d) segmentation result of 11 iteration steps

Figure 17: The segmentation results of figure 16(a) using the level set method as in the algorithm of [4]

not be separated. Figures 20(c) and 20(f) are the results of Figures 17(c), 17(d), 19(c) and 19(d). Because more iteration steps cause more overlaps among text lines, all text lines can not be separated.

We test our algorithm on a document database of 3 different scripts. Specifically, we have constructed a database of 100 Chinese documents, 96 Hindi documents and 100 Korean documents. The ground truths of these documents are obtained manually, where each text line in the document images is delineated by a polygon. The evaluation method is the same as that in [56] and [57]. The test results are shown in Figures 21 and Table 1.
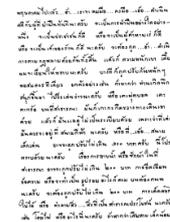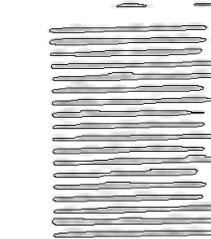
(a) text images
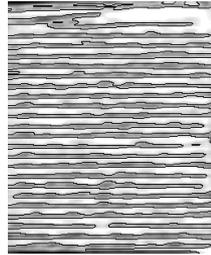
(b) segmentation results of the proposed method
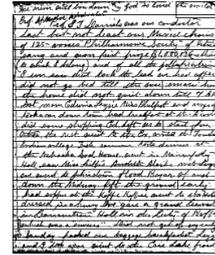
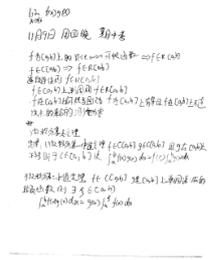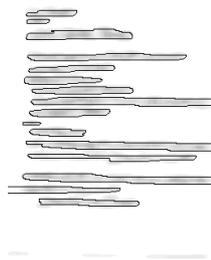(c) segmentation results of a connected component based method

(d) text images

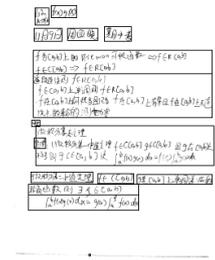(e) segmentation results of the proposed method

(f) segmentation results of a connected component based method
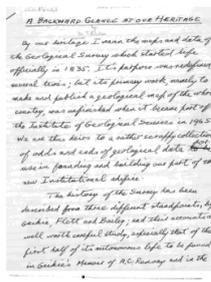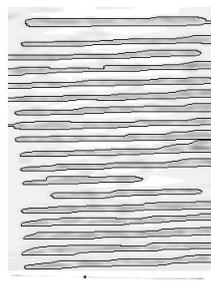

(g) text images

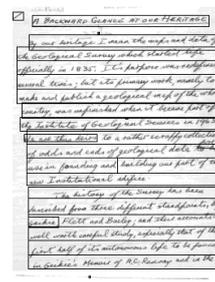(h) segmentation results of the proposed method

(i) segmentation results of a connected component based method

(j) text images

(k) segmentation results of the proposed method

(l) segmentation results of a connected component based method

Figure 18: Experiments on different languages

73

Table 1: Correct detection ratio

	Chinese	Hindi	Korean
Average correct detection ratio	98%	98%	86%
Standard deviation	0.03	0.03	0.07

The correct detection ratio is 98% for Chinese and Hindi. This indicates that our algorithm can produce good detection results. Some typical test samples are shown in figure 22.

For the Korean documents, the correct detection ratio is 86%, which is lower than that of the other two scripts. Through analysis, we find that this is because of the quality of Korean documents. Compared with Chinese and Hindi test samples, some Korean documents are much more messy. On those Korean documents with similar quality as Chinese and Hindi documents, we can obtain similar correct detection ratio. Some difficult cases of the Korean documents are shown in figure 23.

Figure 23 indicates some reasons for failure. The drawings between neighboring text lines can connect two lines. The broken text lines with large horizontal gaps can not be merged. Also our algorithm fails on those oblique text lines. This is because our algorithm assumes that text lines are along horizontal orientation.

6.4   Discussions

In the following we will discuss some of the issues related to our method.

1. The size of blurring window

   The width of the blurring window depends on the horizontal gaps between words

in one text line. The height of the blurring window depends on the vertical gaps between text lines. Too small width will produce broken text lines. On the other hand too large height will increase overlaps between two neighboring text lines. However, this factor is not critical because the overlaps and broken text lines can be corrected by morphing.

2. The initial condition for segmentation

Because the MS energy functional is not convex, minimization result depends on the initial conditions. However, because the content of text images is simple (only text and background), this problem is not an issue. We use different initial conditions and obtain the same results.

3. The shrinkage length of text regions

The shrinkage length depends on the thinness of the overlaps between neighboring text lines. In the case of broken text lines, it also depends on the minimum length of the broken text segments. If the shrinkage length is too small, the overlaps will not be removed. On the other hand, too large shrinkage length will remove some short broken text segments (as shown in figure 24). In the case that the overlaps are large and the broken segments are short, it is tricky to select the shrinkage length.

4. The prolongation length of text regions

The prolongation length depends on the horizontal gaps between neighboring words in one text line. Too small prolongation length will produce many broken text lines. Consider all these factors, we select the same parameters for all experiments, i.e.

blurring window width = 20 pixels, height = 5 pixel, shrinkage length = 50 pixels, and prolongation length = 200.

Figure 24 is a failure example of our algorithm. Figure 24(a) is the text image. Figure 24(b) is the segmented image before morphing. In this figure, there are some broken text lines and overlaps between neighboring text lines. Figure 24(c) is the final segmentation result with shrinkage length of 50. In this case, some texts are missed in the upper right corner. If we use smaller shrinkage length, say 20 pixels, these texts can be kept, but some overlaps between neighboring text lines cannot be removed as shown in figure 24(d).
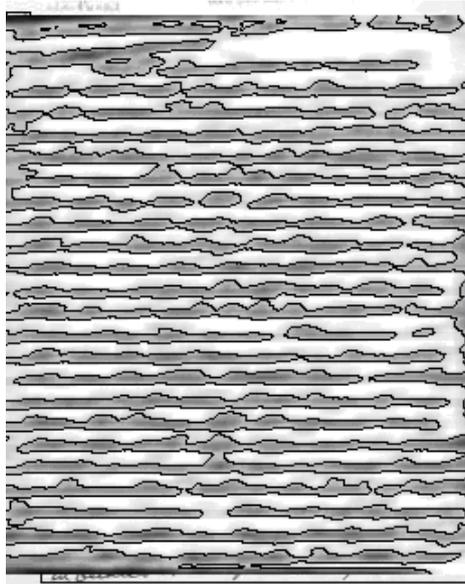
(a) initial segmentation curve
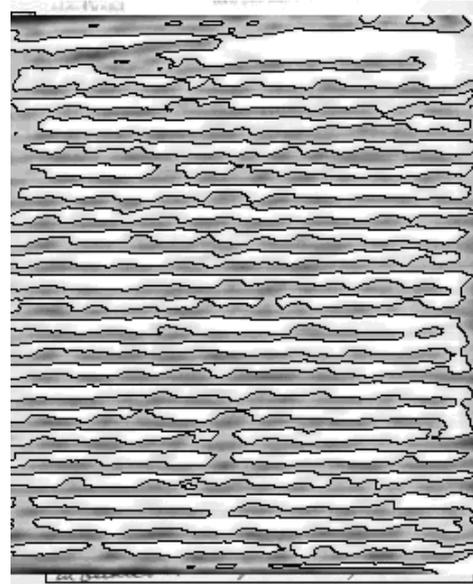
(b) segmentation result of 10 iteration steps
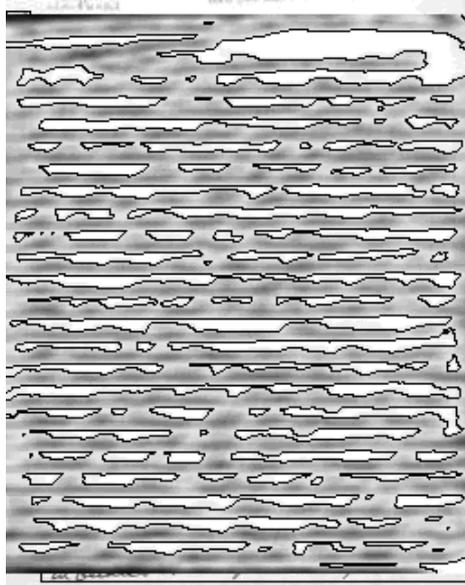
(c) segmentation result of 20 iteration steps
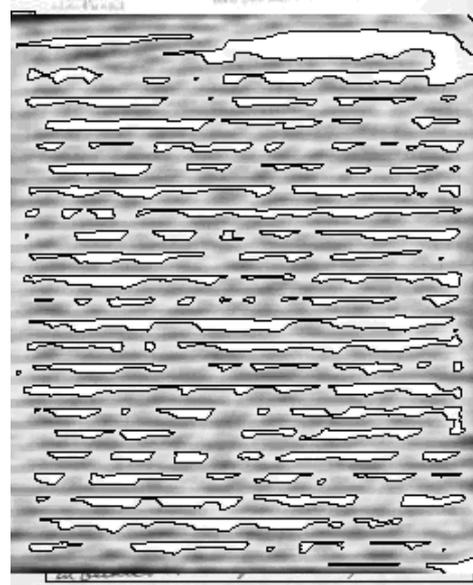
(d) segmentation result of 30 iteration steps

Figure 19: The segmentation results of figure 18(a) using the level set method as in the algorithm of [4]

77
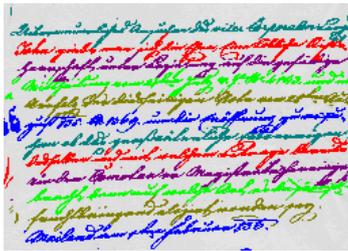
(a) result of figure 16   (b) result of figure 17b   (c) result of Figures 3c and 3d
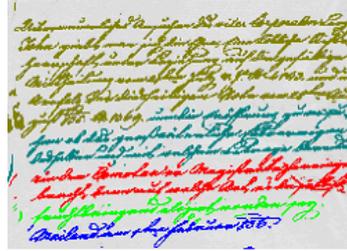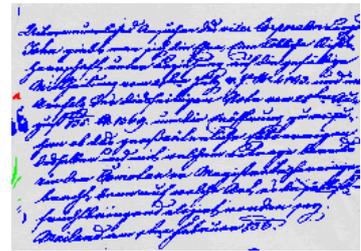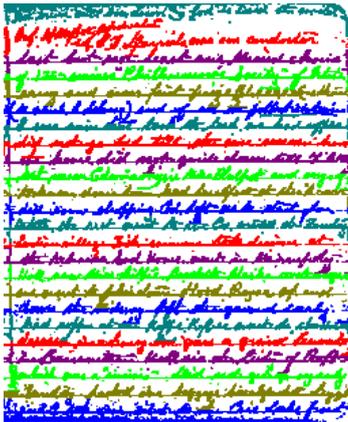
(d) result of figure 18b   (e) result of figure 19b   (f) result of Figures 19c and 19d
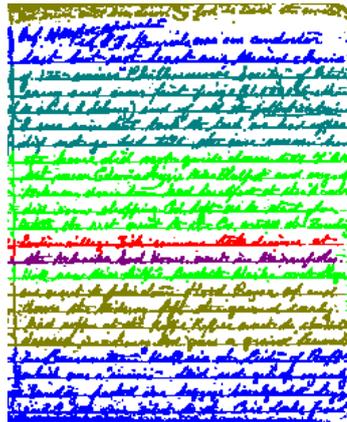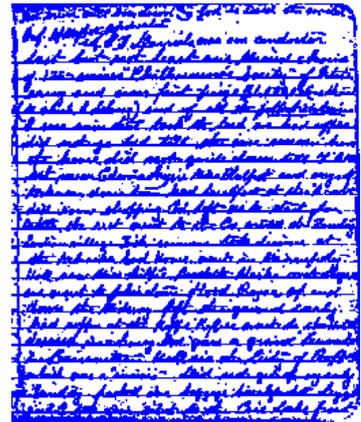
Figure 20: Text line detection results
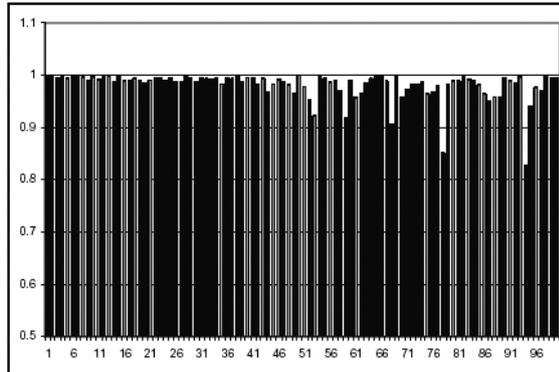
(a) Test results on Chinese
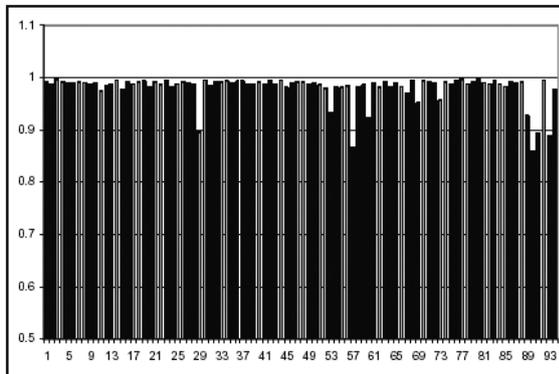
(b) Test results on Hindi
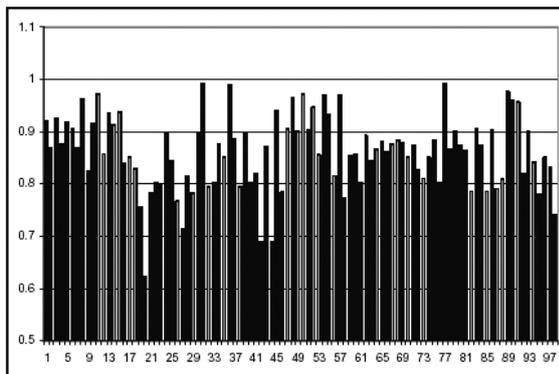
(c) Test results on Korean

Figure 21: Test results on document database

Figure 22: Chinese and Hindi documents

Figure 23: Difficult cases of Korean documents
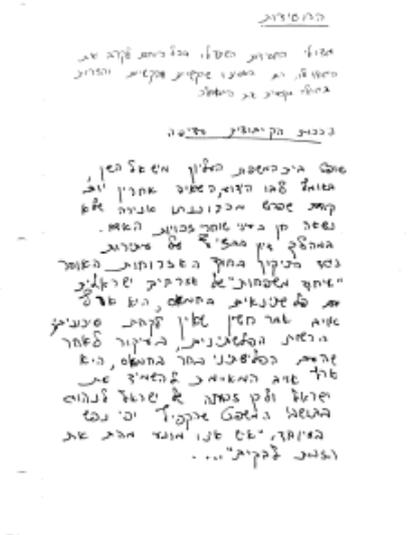
(a) Text image
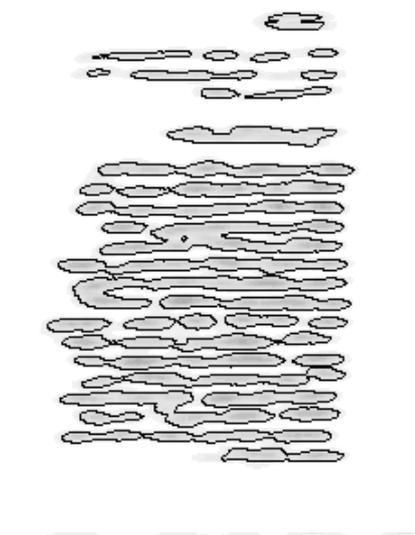
(b) segmented image before morphing
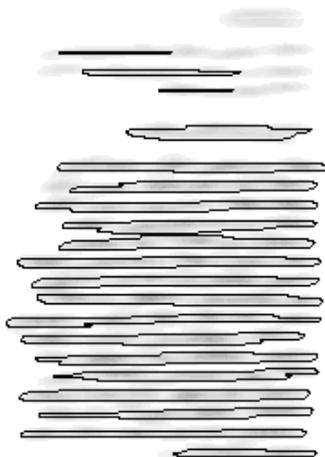
(c) final segmentation result with shrinkage length of 50
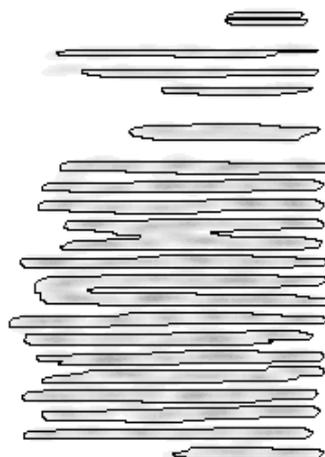
(d) final segmentation result with shrinkage length of 20

Figure 24: The segmentation results with different parameters

Chapter 7

A New Image Segmentation Model

In this chapter, we present a simpler and more efficient segmentation model based on the MS model.

7.1 Difficulties of the MS model

The region-based (global) approach of the MS model is a powerful and robust segmentation technique as compared to pixel-based (local) methods such as Canny or Laplace. However, the minimization of the MS energy functional is not a trivial task. There are some alternative solutions to this problem, including piecewise smooth approximation [6], piecewise constant approximation [7], and piecewise linear approximation [30]. In the piecewise smooth approximation, smooth functions are used to approximate image intensity inside object regions, and the MS model can be solved by three coupled PDEs: two PDEs for

smoothed image intensity and one PDE for segmentation curve. This is a general approximation and can be used for most cases, however, this solution is very time consuming. To overcome this difficulty, Chan and Vese also presented piecewise constant approximation. If image intensity inside object regions can be approximated by constants, we do not need solve two PDEs for the image. Instead, the image intensity inside object regions can be represented by the average intensity of object, hence needing only one PDE to solve the segmentation curve. So the approach is fast. For more complicated cases, piecewise linear approximation was proposed to approximate image intensity by linear functions. This approach is fast and can handle some image intensity changes in object regions. These approaches can only be applied when the image intensity inside object regions is uniform or lineal distributed. They will fail for complicated image intensity distribution.

Another effort is to apply the MS model on different kinds of edges. Because the gradient of image intensity in the MS model can only detect step changes of image intensity, the MS model only works well for step edge. For roof edge, the gradient of image intensity is not small in the regions on both sides of the roof edge, but the second order derivative of image intensity is small, and the gradient of image intensity is discontinuous across the boundary. There is a step edge in the first order derivative functional space. In this case, the MS model often segments one region into several sub-regions. To detect roof edges, second derivative [9] is used to replace gradient in the MS model. Although roof edge can be detected by the approach, the second derivative is sensitive to noise. To overcome these difficulties, we present a new model for image segmentation.

7.2 The new image segmentation model

The energy functional for the new model is given by [59]:

$$E[C] = \int_{\Omega \setminus C} |F(u_0)|^2 dx dy + \nu |C| \qquad (107)$$

where u_0 is the original image; $|C|$ is the length of the segmentation curves; F is a high frequency filter; $F(u_0)$ is the high frequency components of the image; $\Omega \setminus C$ is the image domain excluding the segmentation curves.

The new energy functional consists of two terms: the high frequency components inside regions and the length of segmentation curves. If the energy functional is minimized, the image intensity inside regions only contains low frequency components, and the high frequency components are preserved on boundaries. The constraint on the curve lengths is used to suppress the effect of noise. The result of the new model is equivalent to that of the MS model: image intensity inside regions is smoothed, and image intensity changes abruptly across the segmentation curves.

In the new model, only one variable, the segmentation curve, needs to be solved. Different from the MS model, the image intensity inside regions is smoothed by minimizing high frequency components in the new model. Because high frequency filter can be easily implemented by a convolution, PDEs are not needed for the smoothed image intensity in the MS model. As the result, we need to solve only one PDE for segmentation curve, and it is faster than the piecewise smooth approximation.

Secondly, because no constant approximation, linear approximation, or any other kinds of approximations are used, the new model can deal with various types of image intensity

distribution. We can use it to segment images with complicated image intensity distribution.

The third advantage of the new model is that we can use different high frequency filters to calculate the high frequency components inside the object regions. Different high frequency filters can detect different features. In this thesis, we use the complement of the Gaussian filter and the Gabor filter as high frequency filter. The complement of the Gaussian filter is:

$$F(u_0) = u_0 - G(u_0) \tag{108}$$

where G is the Gaussian filter. The Gabor filter is described in section 7.4.

If we consider that a closed curve segments an image into two regions (i.e. inside region and outside region) the new energy functional in equation (107) can be written as:

$$E[C] = \int_{insideC} |F(u_0)|^2 dxdy + \int_{outsideC} |F(u_0)|^2 dxdy + \nu|C| \tag{109}$$

where $F(u_0)$ is the high frequency filter result. In equation (109), the high frequency filter is only applied in segmented regions, and the filtering doesn't cross segmentation boundaries. Thus the filtering result depends on the segmentation curve, the only unknown variable.

Using the level set function and the Heaviside function, the equation (109) can be written as:

$$E[\phi] = \int |F(u_0)|^2 H(\phi) dxdy + \int |F(u_0)|^2 (1 - H(\phi)) dxdy + \nu \int \delta(\phi)|\nabla\phi| dxdy \tag{110}$$

where ϕ is the level set function to represent the segmentation curve. Minimizing the energy functional with respect to ϕ, we obtain the following Euler-Lagrange equation:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)[\nu \nabla \cdot (\frac{\nabla\phi}{|\nabla\phi|}) - |F_{\text{inside } \phi}(u_0)|^2 + |F_{\text{outside } \phi}(u_0)|^2] \tag{111}$$

86

The segmentation curve will evolve according to equation (111). In equation (111), the last two terms attract segmentation curve to the object boundaries; the first term is curvature of the segmentation curve, which keeps the segmentation curve smooth and removes noise regions.

7.3 Experimental results with the complement of Gaussian filter

We tested the new model on some range images using the complement of the Gaussian filter as high frequency filter. Due to the variations of the image intensity throughout objects corresponding to the distance from the viewer, it is difficult for the constant approximation to detect and segment this kind of objects.

(a) 3D figure   (b) Range image   (c)   Segmentation result
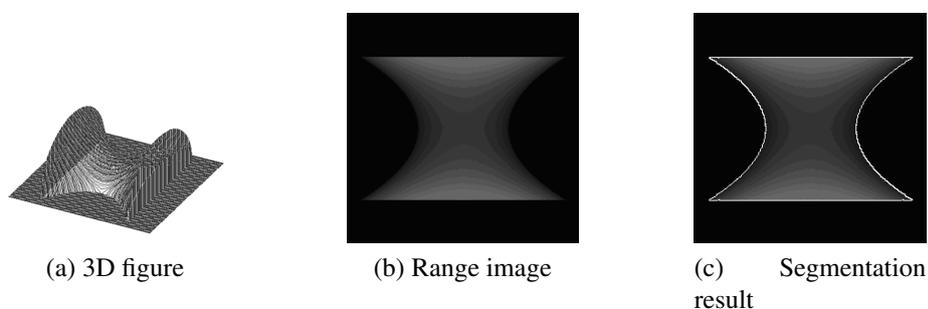
Figure 25: The segmentation of a saddle

Figure 25 is the segmentation of a saddle. The end boundaries of the saddle are step edges, and the side boundaries are roof edges. Figure 26 is the segmentation of a cone. The boundary between the cone and the bottom plane is a roof edge. Figure 27 and figure 28 are the segmentations on a broken cone and a creased cone. The creases on the cones are

(a) 3D figure	(b) Range image	(c) Segmentation result
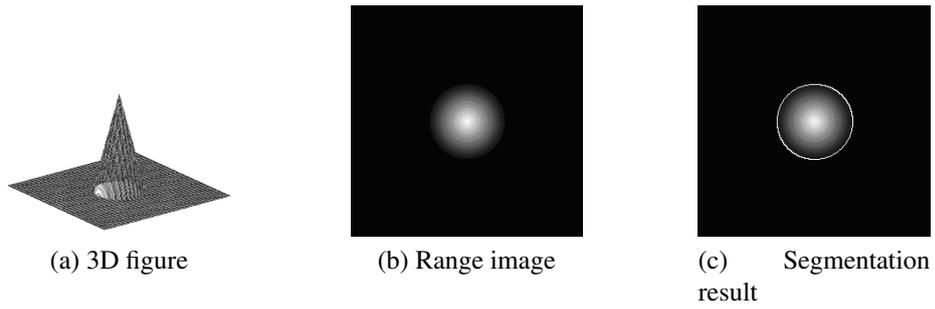
Figure 26: The segmentation of a cone

(a) 3D figure	(b) Range image	(c) Segmentation result
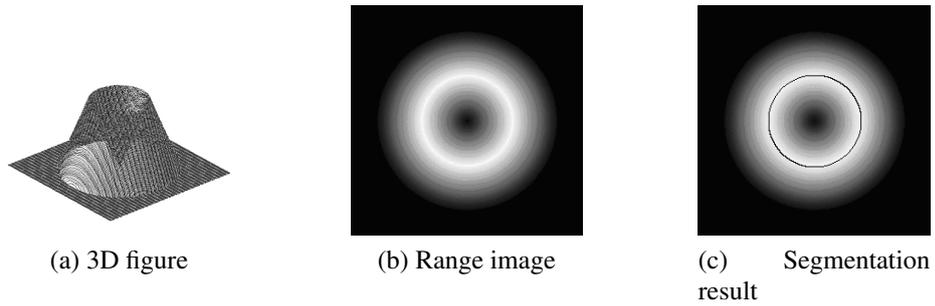
Figure 27: The segmentation of a broken cone

detected as roof edges. These experiments indicate that the new model can detect both step edges and roof edges. Figure 25 also indicates that the new model can segment objects with complicated image intensity distribution. Figure 29 is the segmentation of a heart vessel. Figure 29(b) is the segmentation result of the constant approximation approach. In this case, the vessel is segmented into several regions because the image intensity varies across the image. On the contrary, the segmentation result of the new model in figure 29(c) is very clean. This indicates that the new model can segment objects with different image intensity distributions.

Because only one PDE needs to be solved in the new model, the optimization of the new model is much faster than the MS model. For comparison, we implement the piecewise

(a) 3D figure   (b) Range image   (c)   Segmentation result
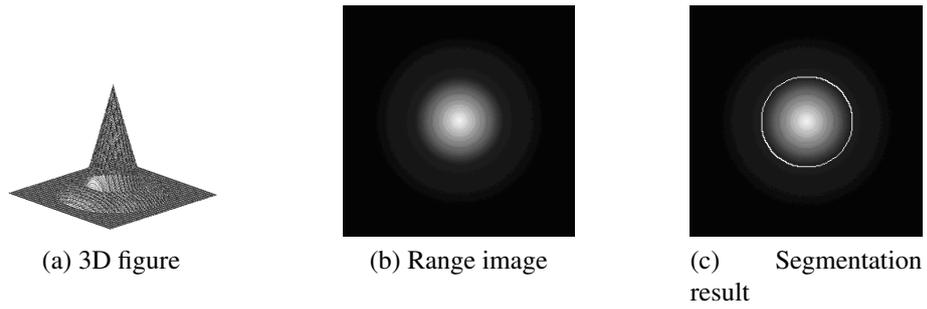
Figure 28: The segmentation of a creased cone

(a) Original image   (b) result by constant approximation approach   (c) Segmentation result by the new model
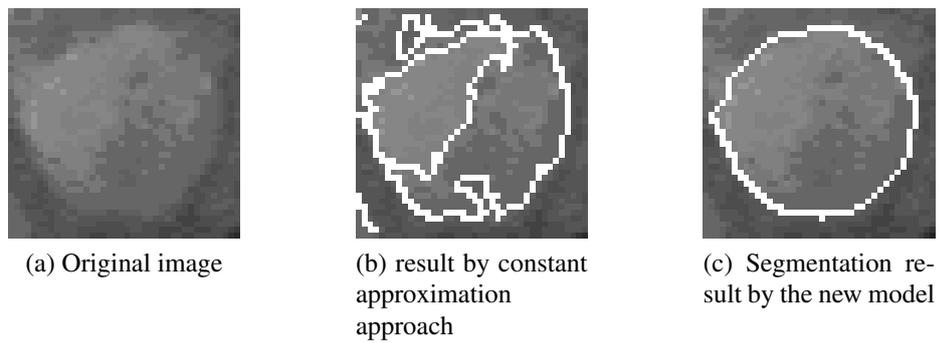
Figure 29: The segmentation of a heart vessel

smooth approximation of the MS model. Its computation time for the experiment in figure 25 is 76s, while the computation time is 36s with the new model. For piecewise smooth approximation of the MS model, the image is smoothed by solving iteratively the PDEs for each curve evolution step. It takes about 9 iterations to solve these PDEs for the experiment in figure 25. The computation time could be different for different images, different initial conditions, or different computers; but the computation time of the new model is always much less than that of the MS model.

7.4   Retinal image segmentation and Gabor filter

For general images, the piecewise constant approximation of the MS model can achieve good results because most objects can be approximated by constants. This is proved in the work of image inpainting and text line detection. The new image segmentation model can be used segment some difficult images. To show the advantage of the new model, we combine the new model and Gabor filter to detect blood vessels in retinal images. Because the new model can detect objects with complicated image intensity distribution, our approach can handle non-uniform illumination in retinal images.

Retinal image analysis is widely used in medical community to diagnose many diseases affecting retina and choroids behind it, including hypertension, diabetes, arteriosclerosis, cardiovascular disease, and stroke. The measurement of blood vessels, such as diameter, branch angle, and length, can be used to diagnose disease, assess the severity of disease and the treatment progress. Blood vessels segmentation is an important step in this process. While manual segmentation requires trained specialist and a lot of labor, the automatic segmentation is always desirable. Recently, the blood vessel segmentation in retinal image has received much attention, and several approaches were proposed, including vessel tracking approach [60] [61], deformable model approach [62] [63], matched filter approach [64], and supervised classification approach [65] [66]. In this thesis we use the new image segmentation model and Gabor filter [67] to detect blood vessels [68]. We produced similar or better segmentation results compared with the state-of-the-art approaches. Our method need not training and is relatively fast.

Gabor filter contains a group of functions to detect features with different frequencies and orientations. Gabor filter has strong orientation selectiveness capacity and can be fine tuned to specific frequency [67]. It has been successfully applied to edge detection, object detection, face recognition, and tracking. Especially, because Gabor filter is efficient to detect directional feature, it was proposed to enhance and detect blood vessels in retinal images [66]. In this thesis, we combine the new image segmentation model and Gabor filter to improve blood vessels detection.

In the spatial domain, a 2-D Gabor filter is a Gaussian kernel modulated by a sinusoidal function as following:

$$\psi_G = exp(jk_0 x)exp(-\frac{1}{2\sigma^2}|Ax|^2) \tag{112}$$

where

$$A = diag[\varepsilon^{1/2}, 1] \tag{113}$$

is diagonal matrix that defines the anisotropy of the filter. k_0 defines frequency of the filter. To detect blood vessel, we define ε as 4 which elongates filter along blood vessel; $k_0 = [0, 3]$ with significant oscillations perpendicular to the large axis of the filter. We use imaginary part of Gabor filter to detect boundary (high frequency part) of blood vessels.

To detect different directional vessels, the Gabor filter spans different orientations. The high frequency filter is as:

$$S = cos(\omega y sin\theta)exp(\frac{\varepsilon(xcos\theta)^2 + (ycos\theta)^2}{2\sigma^2}) \tag{114}$$

Where θ is the orientation angle. We use maximum modulus over all possible orientations as the filter result. To limit computation, θ spans from 0^0 to 180^0 with step of 45^0. As a

wavelet filter, Gabor filter can be applied with different scales. We can use different scales to detect objects of different sizes. In this thesis, to accelerate computation, we only use one scale, and the segmentation result is good. The shape of the filter is shown in figure 30.

Figure 30: Imaginary part of Gabor filter

As shown in figure 30, when we use the filter to detect blood vessels, the image intensity of vessels is smoothed along the vessels direction to remove noise, and the differential is calculated perpendicular to the vessels direction to detect vessel boundaries. Compared with the isotropic differential filter, the filter is efficient to detect directional boundaries.

7.5 Retinal image segmentation with the new model and Gabor filter

To segment a retinal image, an initial segmentation is set in the image. For the segmentation of retinal image, because the illumination is not uniform over image and image intensity is complicated, we use adaptive threshold to create initial segmentation. First we define a local area. Each pixel is compared with its local area average. If the pixel value is lower than the average minus a fix threshold, the pixel is classified as vessel region; otherwise, it is background. As shown in figure 31(b), the adaptive threshold segmentation is not good. The result contains noise regions, and the segmentation boundary is not on the vessel boundary. However, as long as the initial segmented vessel regions contain most part of vessels, the final segmentation result will detect vessels. In the segmentation process, the segmentation curve will evolve to reach the boundaries of vessels. Those noise regions will become smaller and smaller until they disappear. This is caused by the curvature term in the segmentation model. In each iteration step, the image intensity is filtered by the Gabor filter, which is implemented by a convolution. Actually, we only update the filtered intensity in a narrow band around the segmentation curves because the filtered result will remain the same outside this band. The image segmentation result after iteration is shown in figure 31(c). In the last step, we remove segmented regions with small area, which is considered as noise. The final segmentation result is shown in figure 31(d).

We tested our method on two publicly available databases: DRIVE [65] and STARE [64]. The STARE database consists of 20 RGB color images. The images are of size

(a) initial segmentation   (b) initial segmentation

(c) segmentation result with noise regions   (d) final segmentation result
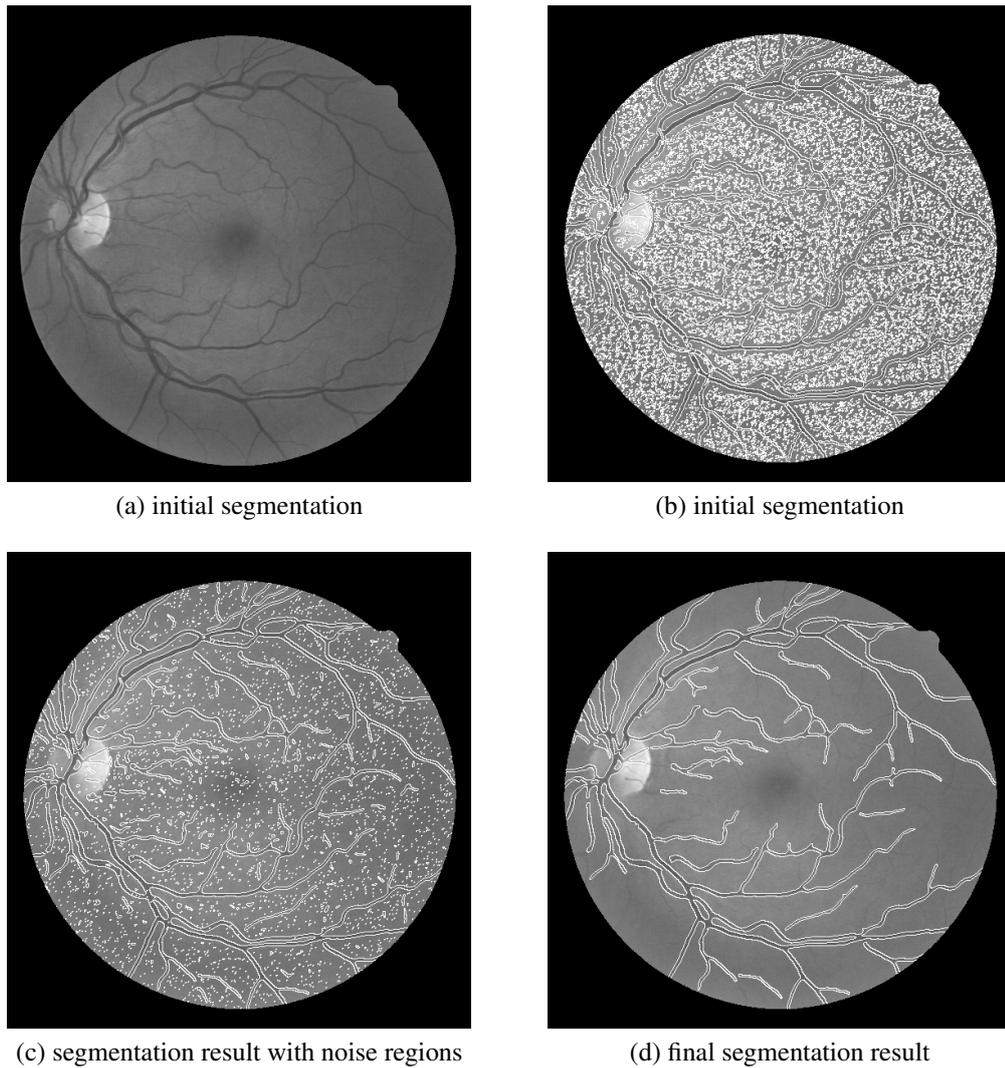
Figure 31: Retinal image segmentation process

605×700 pixels, 24 bits per pixel. There are two hand-labeling available for the 20 images made by two different human observers. DRIVE database consists of 40 images, including 20 training images and 20 test images. The images are of size 768×584 pixels, 24 bits per pixel. There are also two hand-labeling available made by two different human observers. Because our method does not need training, we only use 20 test images. For STARE database, we also test our method on 20 images. We only use green channel of color

94

images for segmentation.

To show the advantage of the new segmentation model, we first use piecewise constant approximation MS model to segment retinal image as shown in figure 32(a). Because the illumination of retinal image is not uniform, the piecewise constant approximation completely fail on retinal image. Due to the complexity of the image intensity, piecewise linear approximation also can not handle these images. Figure 32(b) is the segmentation result of our proposed method. In the result, most blood vessels are correctly detected. This indicates that our method can handle images with non-uniform illumination.
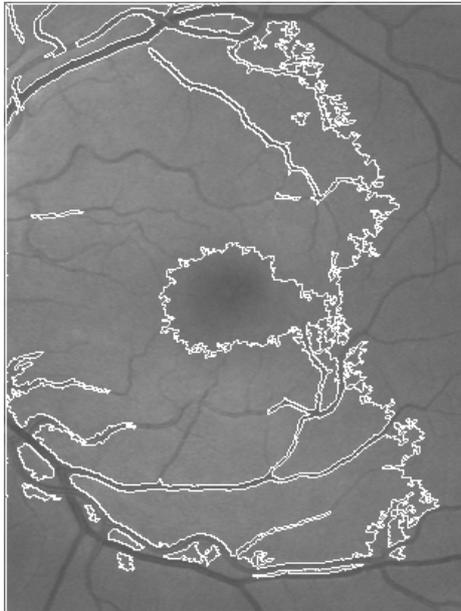
(a) Segmentation results of piecewise constant approximation
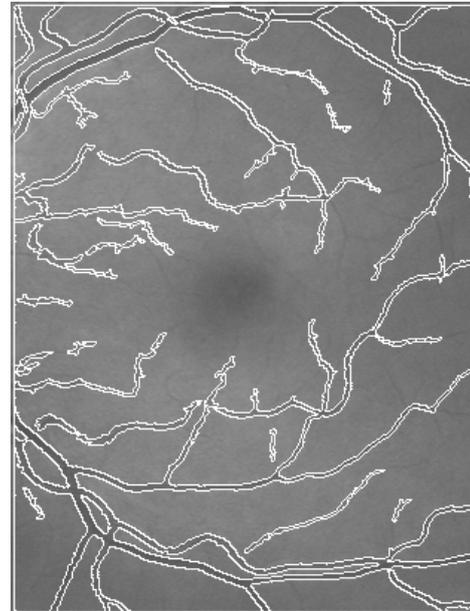
(b) Segmentation results of the new model

Figure 32: Segmentation results of piecewise constant approximation and the new model

Segmentation results of DRIVE and STARE databases are shown in figure 33 and figure 34. From the figures we can see most blood vessels are correctly segmented. The method can overcome difficulties caused by the central darkness and bright optical disc. While

95

the segmentation result reaches state-of-the-art level, some small blood vessels are missed. This will be discussed in next section.
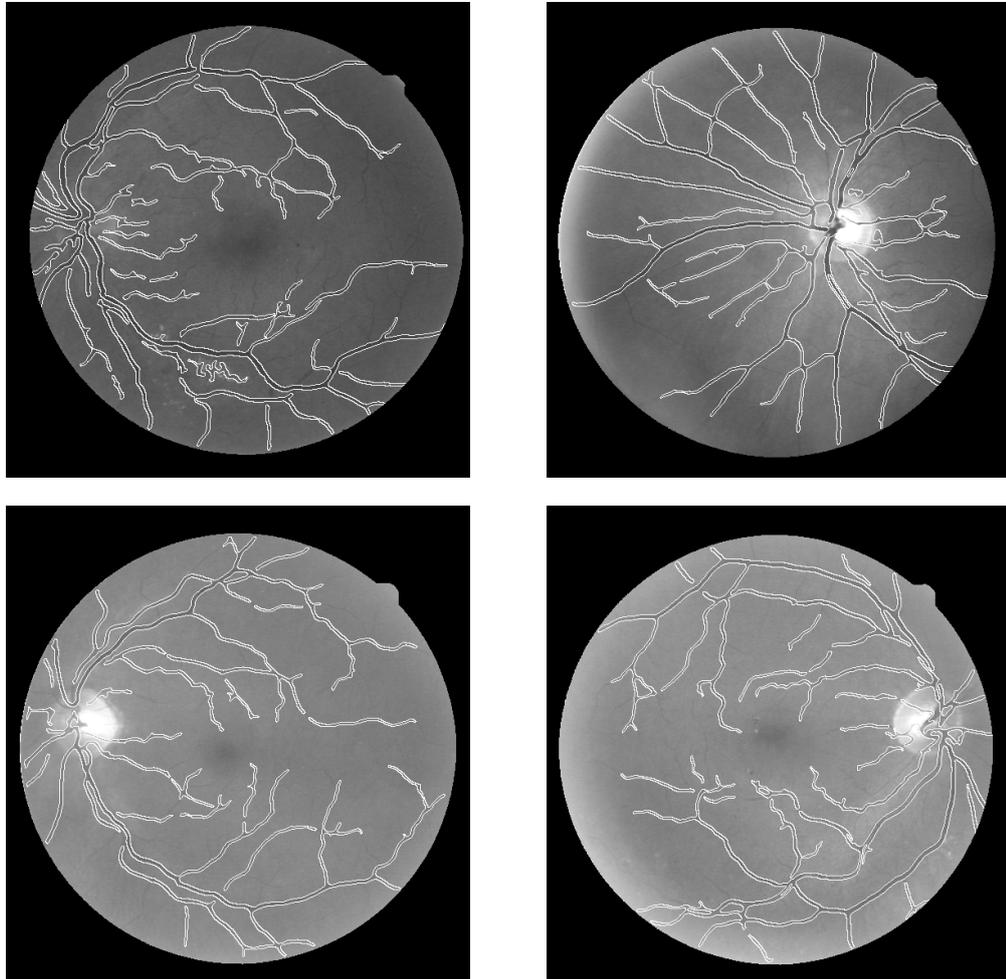
Figure 33: Segmentation result of DRIVE database

DRIVE database contains two groups of manual segmentation result, and they are similar. STARE database also contains two groups of manual segmentation result. The second result detects much more thinner blood vessels. We use manual segmentation result 2 of DRIVE database and manual segmentation result 1 of STARE database as ground truth.

96
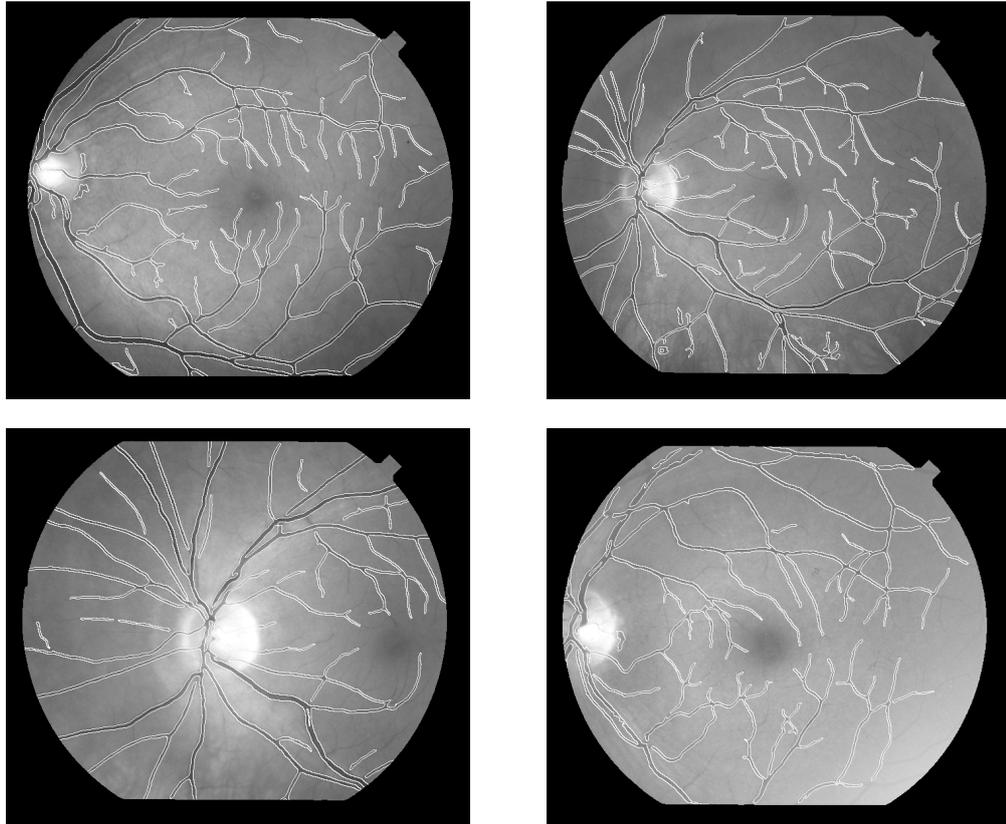
Figure 34: Segmentation result of STARE database

We calculate segmentation accuracy of our method based on the percentage of the correctly segmented pixels, the pixels correctly segmented to blood vessels and background, over all pixels. For comparison, we also collect the segmentation accuracies of some other approaches in table 2. Table 2 indicates that the result of our method is similar or slightly better than other approaches.

Table 2: Retinal image segmentation accuracy

Segmentation method	DRIVE accuracy	STARE accuracy
Our method	94.66%	94.94%
[66]	94.66%	94.80%
[69]	89.11%	90.09%
[65]	94.41%	95.16%

In the process of image segmentation, we use narrow band approach to accelerate computation. In each iteration step, Gabor filter result is calculated only in a narrow band beside the segmentation curves. We need not calculate filter result in other area. So the segmentation process is relatively fast. The actual speed depends on the length of segmentation curves. Most image segmentation can be completed in around 1 minute. [66] uses Gabor filter to generate features. Different from our method, [66] uses classification approach to segment blood vessels. So [66] needs a training process, which needs training images and it takes 9 hours. The real time Gabor filter feature generation takes 3 minute in [66]. Compared with [66], our method is relatively fast and need no training. In [69], adaptive threshold is use. It is simple, but the result is worse. The algorithm in [65] is based on ridge, which contains three steps: ridge extraction, patch feature computation, and feature classification. Although it can produce good result, it is more complicated than our approach.

As mentioned in the previous section, although most blood vessels can be detected by our method, some thinner vessels are missed. This is a general problem for most approaches. In our method, this is caused by the curvature term. In the new image segmentation model, the last term minimizes the length of segmentation curve. This term will smooth the segmentation curve and minimize the curvature to suppress noise. At the end of thinner blood vessels, the curvature is big and the image intensity of vessels is weak. So the segmentation curve will fall to inside at the end and the thinner blood vessel will be shortened. If we use smaller μ and decrease the force caused by the curvature, we can overcome this difficulty. But the segmentation curve will not be smooth in this case, and

more errors will appear on other places.

Another problem is the initial segmentation curve. Because the MS model is not convex, neither is our image segmentation model; the segmentation result depends on initial condition. In our test, if the main part of one blood vessel is not roughly segmented into vessel regions, there will be no segmentation curve to evolve around the vessel, and the blood vessel will be segmented as background. We use an adaptive threshold to generate initial segmentation. In the initial segmentation result, we try to make the blood vessel region to contain most part of blood vessels. At the same time, a lot of noise regions also created. Although the segmentation evolution can make these noise regions smaller or disappear, we still need remove those noise regions with small area. Sometimes, this removal process is tricky. In this thesis, to emphasize the advantage of our new image segmentation model, we only use a simple approach to generate initial segmentation. In real applications, we can use a more sophisticated approach. The combination of our approach with some state-of-the-art approaches will produce better result.

Chapter 8

Conclusion

As a region-based (global) approach, the Mumford and Shah (MS) model is a powerful and robust segmentation technique as compared to edge-based (local) methods. This thesis presents our research on some aspects of the MS model. The contributions of the thesis include:

1. Application of the MS model in image inpainting. It is a challenge for image inpainting to preserve edges in inpainting area. In this thesis, we make use of the MS model to detect object boundaries inside the inpainting areas; then we use diffusion or texture patch approach to fill the inpainting regions. This filling does not cross the object boundaries, so the boundaries and edges are preserved. We present a hierarchical segmentation method to detect boundaries of both the main structure and detailed objects. As the result, more detailed edges are preserved in the final inpainting results. In addition to edges, texture is also difficult to restore. We utilize texture synthesis algorithm to fill object regions where texture needs to be preserved.

2. Application of the MS model in text line detection. It is difficult to detect text lines in handwritten documents. Most approaches are sensitive to topological changes in handwritten documents. Our approach consists of three steps: first we blur text image by Gaussian filter; then we use the MS model to segmentation text lines; finally text lines are detected by morphing approach. Because our approach segments text image as general image without any knowledge about text, the method is script independent and is not sensitive to topological changes. Our morphing approach can also handle broken text lines and text line overlaps.

3. New image segmentation model based on the MS model. Because direct solution of the MS model is difficult, some alterative approximated approaches were proposed. Although some approaches, such as piecewise constant approximation, have been successfully used in many applications, these approaches are either inefficient, such as piecewise smooth approximation, or limited to special image intensity distributions. We present a new model for image segmentation based on the MS model. Because it consists of only one variable, the segmentation curve, the solution is efficient. Because no approximations are required in the new model, it can segment objects with complicated image intensity distributions. The new model can segment both step and roof edges. In addition, we can use different high frequency filters to produce better denoise result or detect special features.

4. Retinal image segmentation by the new image segmentation model. The thin blood

vessels and non-uniform illumination are two difficulties in retinal image segmentation. We combine the new model and Gabor filter to detect blood vessels in retinal images. The Gabor filter is good to detect directional objects. We use this special filter in the new model to detect blood vessels. Because the new model can detect objects with complicated image intensity distribution, our approach can also handle non-uniform illumination in retinal images. As the result, segmentation results are similar or better than state-of-the-art.

The new image segmentation model is a varitional approach framework. We can use different high frequency filters in the new model to detect different features. In this thesis, we only test two kinds of filters: complement of Gaussion filter and Gabor filter. In the future work, we will try to use more sophisticated filters, such as wavelet filter, to produce better image segmentation results. In this thesis, the advantages of the new model are demonstrated by some experiments. For a solid theoretical analysis, a mathematical proof of the new model is also necessary. We plan to do this in the future work.

Bibliography

[1] S. Z. Li, Markov Random Field Modeling in Computer Vision, New York, Springer-Verlag, 2001.

[2] M. Unser, "Texture classification and segmentation using wavelet frames," IEEE Transactions on Image Processing, Vol. 11, no. 4, pp. 1549-1560, 1995.

[3] D. Mumford, and J. Shah, "Optimal approximation by piecewise smooth functionals and associated variational problems," Comm. Pure Appl. Math, Vol. 42, pp.577-685, 1989.

[4] M. Vatsa, R. Singh, and A. Noore, "Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 38, No. 4, pp 1021-1035, 2008.

[5] S. Mahmoodi, "Shape-based active contours for fast video segmentation," IEEE Signal Processing Letters, Vol. 16, No. 10, pp. 857-860, 2009.

[6] L. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," International Journal of Computer Vision, Vol. 50, no. 3, pp.271-293, 2002.

[7] T. F. Chan and L. A. Vese, "Active contours without edges," IEEE Trans on Image Processing, Vol. 10, no. 2, pp.266-277, 2001.

[8] A. Blake and A. Zisserman, Visual Reconstruction, The MIT Press, Cambridge, Massachusetts, 1987.

[9] L. Ambrosio, L. Faina,and R. March, "Variational approximation of a second order free discontinuity problem in computer vision," SIAM J. Math. Anal., Vol. 32, No. 6, pp. 1171-1197, 2001.

[10] T. D. Bui, S. Gao, and Q.H. Zhang, "A generalized Mumford-Shah model for roof-edge detection," IEEE International Conference on Image Processing, Vol. 2, pp. 1214-1217, 2005

[11] R. M. Haralick, L.G. Shapiro, "Image segmentation techniques," Computer Vision, Graphics, and Image Processing, Vol. 29 , No. 1, pp. 100-132, 1985.

[12] R. C. Gonzalez, R. E. Woods, Digital Image Processing, 2nd ed., pp. 567-642, Upper Saddle River, N. J., Prentice Hall, 2002.

[13] S. A. Hojjatoleslami and J Kittler, "Region growing: a new approach," IEEE Transactions on Image Processing, Vol. 7, no. 7, pp. 1079-1084, Jul 1998.

[14] T. Chan and J. Shen, Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods, SIAM, 2005.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snake: active contour models," International Journal of Computer Vision, Vol. 1, pp.321-331, 1988.

[16] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," Proceedings of the 5-th International Conference on Computer Vision, pp. 694-699, MA, June 1995, IEEE Computer Society Press.

[17] G. Aubert and P. Kornprobst, "Mathematical problem in image processing: partial differential equations and the calculus of variations," Applied Mathematical Sciences, Vol. 147, Springer, 2002.

[18] A. Tsai, A. Yezzi, and Alan S. Willsky, "Curve evolution implementation of the Mumford-Shah functionalal for image segmentation, denoising, dnterpolation, and magnification," IEEE Trans. on Image Processing, Vol. 10, no. 8, pp.1169-1186, 2001.

[19] S. Gao and T. D. Bui, "Image segmentation and selective smoothing by using Mumford-Shah model," IEEE Trans on Image Processing, Vol. 14, no. 10, pp.1537-1549, 2005.

[20] M. Nikolova, S. Esedoglu, and T. F. Chan, "Algorithms for finding global minimizers of image segmentation and denoising models," SIAM J. Appl. Math. Vol. 66, No. 5, pp. 1632-1648, 2006.

[21] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed : algorithm based on the Hamilton-Jacobi formulation," Journal of Computational Physics, 79 pp12-49, 1988.

[22] S. Osher and R. P. Fedkiw, "Level set Methods: an overview and some recent results," Journal of Computational Physics, Vol. 169, pp.463-502, 2001.

[23] D. Adalsteinsson, and J. A. Sethian, "A Fast level set methods for propagating interfaces," Journal of Computational Physics, Vol. 118, no. 2, pp.269-277, 1995.

[24] H. K. Zhao, T. F. Chan, B. Merriman, S. Osher, "A variational level set approach to multiphase motion," Journal of Computational Physics, Vol. 127, pp. 179-195, 1996.

[25] S. Osher and R. P. Fedkiw, Level set Methods and Dynamic Implicit Surfaces, Springer Verlag, 2002.

[26] H. K. Zhao, S. Osher, B. Merriman, and M. Kang,"Implicit and nonparametric shape recognition from unorganized data using a variational level set method," Computer Vision and Image Understanding, Vol. 80, pp. 295-314, 2000.

[27] J. A. Sethian, Level Set Methods and Fast Marching Methods, Cambridge University Press, 1999.

[28] S. Gao, A New Segmentation and Smoothing Method Based on the Mumford-Shah Variational Model, Thesis for Master's Degree, Concordia University, 2003.

[29] X. J. Du, A New Hierarchical Method for Image Segmentation and Inpainting Using Mumford-Shah Model, Thesis for Master's Degree, Concordia University, 2006.

[30] Q. H. Zhang, Mumford-Shah Model and its Application in Image Processing, Thesis for Master's Degree, Concordia University, 2005.

[31] L. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D, Vol. 60, 259-268, 1992.

[32] X. J. Du and T. D. Bui, "Image segmentation based on the Mumford-Shah model and its variations," 5th IEEE International Symposium on Biomedical Imaging, ISBI'08, pp. 109-112, 2008

[33] X. J. Du and T. D. Bui, "A new hierarchical image segmentation method," 18th International Conference on Pattern Recognition, ICPR 2006, pp. 108-122, 2006

[34] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), pp. 417-424, 2000.

[35] T. F. Chan and J. Shen, "Mathematical models for local non-texture inpainting," SIAM J. Appl. Math., Vol. 62, no. 3, pp. 1019-1043, 2002.

[36] S. Masnou and J. M. Morel, "Level lines based disocclusion," IEEE International Conference on Image Processing (ICIP), Vol 3, pp.259-263, 1998.

[37] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting,: IEEE Transactions on Image Processing, Vol. 12, No. 8; pp.882C889, 2003.

[38] A. Criminisi, P. Péz, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," IEEE Transactions on Image Processing, Vol. 13, No. 9; pp. 1200-1212, 2004.

[39] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, "Image completion with structure propagation," ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Vol. 24, pp. 861-868, 2005.

[40] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: Part I - theory," IEEE Transactions on Image Processing, Vol. 15, pp. 539-554, 2006.

[41] D. Cho and T. D. Bui, "Image inpainting using wavelet-based inter- and intra-scale dependency," International Conference on Pattern Recognition (ICPR), pp. 1-4, 2008.

[42] J. Miral, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," IEEE Transactions on Image Processing, Vol. 17, No. 1, pp. 53-69, 2008.

[43] G. Peyre, "Texture synthesis with grouplets," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 8, pp. 733-746, 2009.

[44] X. J. Du, D. W. Cho, and T. D. Bui, "Image segmentation and inpainting using hierarchical level set method," Signal Processing, accepted.

[45] M. Bertalmio, A. Bertozzi, and G. Sapiro. "Navier-stokes, fluid-dynamics and image and video inpainting," International Conference on Computer Vision and Pattern Recognition (CVPR), Vol 1, pp. 355-362, 2001.

[46] T. F. Chan and J. Shen, "Non-texture inpainting by curvature-driven diffusions (CDD)," J. Visual Comm. Image Representation, Vol. 12, No. 4, pp. 436-449, 2001.

[47] S. Esedoglu and J. Shen, "Digital inpainting based on the Mumford-Shah-Euler image model," European J. Appl. Math., Vol. 13, pp. 353-370, 2002.

[48] D. Tschumperle and R. Deriche, "Vector-valued image regularization with PDEs: A common framework for different applications," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 4, pp. 506-517, 2005.

[49] G. Nagy, "Twenty years of document image analysis," PAMI, IEEE Trans.Pattern Anal. Mach. Intell. Vol. 22, No. 1, pp. 38-62, 2000.

[50] D. Doermann, "Page decomposition and related research," Proceedings of the Symposium Document Image Understanding Technology, pp.39-55, 1995.

[51] Y. Y. Tang, D. Y. Chang, C. Y. Suen, "Document processing for automatic knowledge acquisition," IEEE Trans. Knowl. Data Eng. Vol. 6, No. 1 pp. 3-21, 1994.

[52] C. Weliwitage, A.L. Harvey, and A.B. Jennings, "Handwritten document offline text line segmentation," Proceedings of Digital Imaging Computing: Techniques and Applications, pp. 184-187, 2005.

[53] L. O'Gorman, "The document spectrum for page layout analysis," IEEE Trans. Pattern Anal. Machine Intell., Vol. 15, no. 11, pp. 1162-1173, 1993.

[54] S. Nicolas, T. Paquet and L. Heutte, "Text line segmentation in handwritten document using a production system," Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition, pp. 245-250, 2004.

[55] M. Bulacu, R. Koert, L. Schomaker and T. Zant, "Layout analysis of handwritten historical documents for searching the archive of the Cabinet of the Dutch Queen," Proc. ICDAR, pp. 357-361, 2007.

[56] Y. Li, Y. F. Zheng, D. Doermann, and S. Jaeger, "A new algorithm for detecting text line in handwritten documents," Proceedings of the Tenth International Workshop on Frontiers in Handwriting Recognition, pp. 35-40, 2006.

[57] Y. Li, Y. F. Zheng, D. Doermann, and S. Jaeger, "Script-independent text line segmentation in freestyle handwritten documents," IEEE Trans. Pattern Anal. Machine Intell., Vol. 30, no, 8 pp. 1313-1329, 2008.

[58] X. J. Du, W. M. Pan, and T. D. Bui, "Text line segmentation in handwritten documents using Mumford-Shah model," Pattern Recognition, Vol. 42, No. 12, pp. 3136-3145, 2009.

[59] X. J. Du and T. D. Bui, "A new model for image segmentation," IEEE Signal Processing Letters, Vol. 15, pp. 182-185, 2008.

[60] Y. A. Tolias and S. M. Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," IEEE Trans. Med. Imag., Vol. 17, no. 2, pp. 263-273, 1998.

[61] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," IEEE Trans. Inf. Technol. Biomed., Vol. 3, no. 2, pp. 125-138, 1999.

[62] T. McInerney and D. Terzopoulos, "T-snakes: topology adaptive snakes," Med. Image Anal., Vol. 4, pp. 73-91, 2000.

[63] A. Vasilevskiy and K. Siddiqi, "Flux maximizing geometric flows," IEEE Trans. Pattern Anal. Mach. Intell., Vol. 24, no. 12, pp. 1565-1578, 2002.

[64] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Trans. Med. Imag., Vol. 19, no. 3, pp. 203-210, 2000.

[65] J. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," IEEE Trans. Med. Imag., Vol. 23, no. 4, pp. 501-509, 2004.

[66] J. V. B. Soares, J, J, G. Leandro, R. M. Cesar Jr., H. F. Jelinek, and M, J, Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," IEEE Trans. Med. Imag, Vol. 25, no. 9, pp. 1214-1222, 2006.

[67] J. P. Antoine, P. Carette, R. Murenzi, and B. Piette, "Image analysis with two-dimensional continuous wavelet transform," Signal Process., Vol. 31, pp. 241-272, 1993.

[68] X. J. Du and T. D. Bui, "Retinal image segmentation based on Mumford-Shah model and Gabor wavelet filter," 2010 20th International Conference on Pattern Recognition, ICPR 2010, pp.3384-3387, 2010.

[69] X. Jiang and D. Mojon, "Adaptive local thresholding by verification based multi-threshold probing with application to vessel detection in retinal images," IEEE Trans. Pattern Anal. Mach. Intell., Vol. 25, no. 1, pp. 131-137, Jan. 2003.