1-1-2013

Study of Image Local Scale Structure Using Nonlinear Diffusion

Yan Wang
University of Pennsylvania, wangyan1@sas.upenn.edu

Study of Image Local Scale Structure Using Nonlinear Diffusion

Abstract
Multi-scale representation and local scale extraction of images are important in computer vision research, as in general , structures within images are unknown. Traditionally, the multi-scale analysis is based on the linear diusion (i.e. heat diusion) with known limitation in edge distortions. In addition, the term scale which is used

widely in multi-scale and local scale analysis does not have a consistent denition and it can pose potential diculties in real image analysis, especially for the proper interpretation of scale as a geometric measure. In this study, in order to overcome

limitations of linear diusion, we focus on the multi-scale analysis based on total variation minimization model. This model has been used in image denoising with the power that it can preserve edge structures. Based on the total variation model, we construct the multi-scale space and propose a denition for image local scale. The

new denition of local scale incorporates both pixel-wise and orientation information.

This denition can be interpreted with a clear geometrical meaning and applied in general image analysis. The potential applications of total variation model in retinal fundus image analysis is explored. The existence of blood vessel and drusen structures within a single fundus image makes the image analysis a challenging problem.

A multi-scale model based on total variation is used, showing the capabilities in both drusen and blood vessel detections. The performance of vessel detection is compared with publicly available methods, showing the improvements both quantitatively and

qualitatively. This study provides a better insight into local scale study and shows the potentials of total variation model in medical image analysis.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Applied Mathematics

First Advisor
James C. Gee

Subject Categories
Applied Mathematics | Radiology

STUDY OF IMAGE LOCAL SCALE STRUCTURE USING NONLINEAR DIFFUSION

Yan Wang

A DISSERTATION

in

Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

_____

James C. Gee

Associate Professor of Radiologic Science in Radiology

Graduate Group Chairperson

_____

Charles L. Epstein, Professor of Mathematics

Dissertation Committee

James C. Gee Associate Professor of Radiologic Science in Radiology

Charles L. Epstein Professor of Mathematic

Philip T. Gressman Associate Professor of Mathematics

Acknowledgments

This dissertation would not have been possible without the help, advice, and encouragement of a large group of people. To only some of whom, it is possible to give particular mention here.

I would like to thank my advisor Dr. Gee. The most valuable lesson I have learnt from him is the importance of independence in scientific research. Instead of giving every detailed instructions on what we should follow, he provided us with insightful big pictures, encouraged us and gave us the freedom to explore scientifically interesting problems. The work would not be possible without his inspirations.

I would like to thank my committee chair and program chair Dr. Epstein. Thanks for him to enroll us as the first year Applied mathematics and computational science students. If I was not given this opportunity, it would be impossible for me to get into the field of medical image analysis and participate in various projects. Thanks to him for all of his help and support in the past five years. I would also like to thank professors of all the courses that I took in graduate school.

Next, I want to thank all my collaborators in graduate school, including course

projects in medical image analysis, machine learning with Yang Liu, Wei Han, who are my AMCS classmates. Thanks to Gary, Phil, Jeff for their help during my first summer in PICSL lab. It was for the first time that I learnt to use Linux and get exposures to MRI brain imaging research. Thanks to Yuanjie for all the collaborations, guidances and discussions in all different projects. Thanks to Dr. Le as his work inspired my research in multi-scale image analysis. Thanks to all my lab members and it is an honor to be a part of PICSL lab. I would also like to thank Dr. Kontos, Dr. Keller and other lab members in CBIG lab that I have collaborated with before. Thanks to Dr. Kontos for introducing me into the field of breast imaging research and her patience in research guidance, helping me to improve my scientific writing skills. Thanks to other CBIG members for all the discussions and help.

To all of my friends, especially Lily, Yelan, Bonna and Will, thank you for all your understanding and encouragement when I am in difficult moments and I am glad to have you all by my side. Last but not least, I want to thank my parents for giving me their endless love and support and for having solid faith in me. To them, I dedicate this thesis. Every story has an end but in life every ending is a new beginning.

ABSTRACT

STUDY OF IMAGE LOCAL SCALE STRUCTURE USING NONLINEAR DIFFUSION

Yan Wang

James C. Gee

Multi-scale representation and local scale extraction of images are important in computer vision research, as in general, structures within images are unknown. Traditionally, the multi-scale analysis is based on the linear diffusion (i.e. heat diffusion) with known limitation in edge distortions. In addition, the term scale which is used widely in multi-scale and local scale analysis does not have a consistent definition and it can pose potential difficulties in real image analysis, especially for the proper interpretation of scale as a geometric measure. In this study, in order to overcome limitations of linear diffusion, we focus on the multi-scale analysis based on total variation minimization model. This model has been used in image denoising with the power that it can preserve edge structures. Based on the total variation model, we construct the multi-scale space and propose a definition for image local scale. The new definition of local scale incorporates both pixel-wise and orientation information. This definition can be interpreted with a clear geometrical meaning and applied in general image analysis. The potential applications of total variation model in retinal fundus image analysis is explored. The existence of blood vessel and drusen struc-

tures within a single fundus image makes the image analysis a challenging problem. A multi-scale model based on total variation is used, showing the capabilities in both drusen and blood vessel detections. The performance of vessel detection is compared with publicly available methods, showing the improvements both quantitatively and qualitatively. This study provides a better insight into local scale study and shows the potentials of total variation model in medical image analysis.

Contents

Chapter 1

Introduction

1.1 Motivation

This study is initially motivated by one application in medical image analysis. The medical image used in the study is retinal fundus image, a type of image commonly used in eye-disease diagnosis and screening program. With the increasing use of digital images and the needs of computer aided diagnosis, the existence of limited automatic methods makes retinal fundus image analysis an important image analysis problem serving clinical needs. In addition, what makes this problem unique and appealing are the structures of interest within the images. There are two types of geometric structures within a single image, where one structure can be categorized as blobs and the other are vessels. In the retinal fundus image, optic disk, fovea region and lesions(e.g. drusen, exudate) can be considered as blob structures(i.e. approximated by a disk region, and brighter or darker than the surroundings) and

vessel structures are blood vessels(i.e arteries and veins). One typical example of retinal fundus image for Age Related Macular Degeneration(ARMD) disease and its geometric simplification is shown in Figure 1.1. For computer aided diagnosis, the accurate detection of the location and size of both structures is crucial. For example, in ARMD grading system, numbers of drusen of various sizes are used to grade the severity of ARMD [1, 2] since the existence of drusen is a significant risk factor for ARMD. The locations and widths of the vessels are also used widely in diagnosis for diseases related to blood vessels[3, 4].

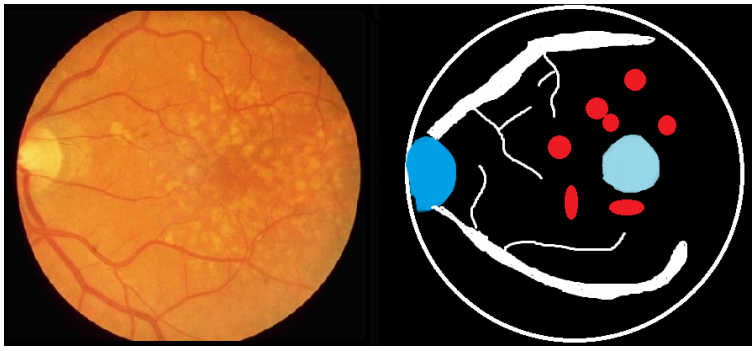


Figure 1.1: Typical example of retinal fundus image with ARMD disease. Left: color fundus image; Right: geometric simplification of eye structures

Since both types of structures exist in one image and each type of structure presents with a variety of sizes as shown in Figure1.1, it makes the automatic detection of structures a challenging but interesting problem suitable for multi-scale analysis. The multi-scale analysis started from Witkin [5] based on gaussian smoothing(or equivalently heat diffusion model), since then there have been tremendous researches focused on the multi-scale analysis based on the linear diffusion model. Among these

researches, the work of Lindeberg [6] was focused on blob detection and automatic scale selection for structures, and the work of Lowe [7], based on [6], proposed the Scale Invariant Feature Transform (SIFT). SIFT features have been used widely in image feature analysis, especially image matching, feature detection etc. Most of the multi-scale analysis on retinal fundus image is based on linear diffusion model. However, the linear diffusion model is known to cause disadvantages blurring of edge structures as the diffusion is isotropic (i.e. diffuse the image uniformly at each pixel along each direction).

Blob and vessel structures have distinct geometric properties. Blob structures can be approximated by disks, and edge information is not considered the dominating structural features. This is the main reason that the linear diffusion model can be efficient in detecting blobs. However, for vessels, the dominating structures are edges and the blurring of edges means loss of the essential structure information. One of the branches of image diffusion, based on nonlinear diffusion, makes multi-scale study more appropriate. The nonlinear diffusion equation chooses the diffusivity adaptively, i.e. to diffuse the image with different weights to preserve edges. The uniqueness of the heat equation as a linear model makes any model that can preserve edges a nonlinear problem. In general, the nonlinear diffusion equation uses the time variable t as a scale parameter(i.e. as t increases, the image is smoothed (with edges preserved/enhanced). However, compared to linear diffusion model, it is harder to interpret the meaning of t as a *scale*, especially the geometric meaning. Instead, in

linear diffusion model, the time parameter t can be related to neighborhood size σ simply as $2t = \sigma^2$ directly.

In summary, one medical image analysis problem motivated two interesting mathematical problems: (1) How can we use nonlinear diffusion equations for multi-scale analysis? What are the advantages of nonlinear diffusion models ? (2) What is the meaning of *scale* for image structures? How can we define and interpret *scale* in a proper way? As the linear diffusion model and its applications have been studied extensively by various researchers, in this study we focus on the nonlinear diffusion model limited within the total variation minimization framework, since this model is known to be able to preserve edge structures. Based on this model, we will focus on the proper construction of a multi-scale and local scale study. Finally, as one example, we will show how the model can be used in retina fundus image analysis.

1.2 Contribution

Total variation minimization model has been known to have advantages in preserving edge structures. The weight between the regularization and fidelity terms in the variational optimization function indicate that the model is related to multi-scale properties of images. However, due to the nonlinearity of the problem, the study of multi-scale properties in this model are limited and not well investigated. In this study, we address the problem of extracting image local scale information based on the total variation model. The multi-scale properties were studied with both analytical

and numerical analysis.

The multi-scale study also motivated us to revisit the real meaning of the term *scale*. Using total variation model for 1D signal filtering, a definition of local scale which uses both the spatial and directional information is proposed. Compared to definitions of scale which appeared in earlier literatures, this definition can be related to the geometric measure in an intuitive way and has potential in general image analysis applications.

The main contribution of this study can be summarized as the set up of bridges between the mathematical theory, computer vision and medical image analysis for nonlinear diffusion theory. One application on retina fundus image analysis is used to demonstrate the potential of total variation models for further applications.

1.3 Organization

In *Chapter 2*, we provide a general introduction to diffusion equations used in image analysis. In particular, based on the decomposition of diffusion operators along tangential and normal directions of the level sets, we are able to provide a clear way to interpret the meaning of image diffusion.

In *Chapter 3*, we give an overview of multi-scale study based on linear diffusion equation. Especially, applications of multi-scale space on image analysis will be shown. As one example related to the ultimate clinical application, advantages and limitations of the multi-scale space will be discussed.

In *Chapter 4*, this chapter is the theoretical background of models we study. The notion of total variation, total variation minimization and related nonlinear diffusion equation is introduced. In particular, for TV flow equation, we will discuss the wellposed-ness of the equation, regularity and asymptotic property of the solution.

Chapter 5 and *Chapter 6* are the main contributions to the theoretical part of multi-scale and local scale study. *Chapter 5* is focused on the extraction of local scale based on the TV flow equation; this work is a natural extension of the work in linear diffusion research. Previous work related to local scale extraction will be discussed and compared with the proposed definition of local scale. Applications of this model in structure detection are shown with examples. Based on the limitations of previous methods and the general lack of definition of the term *local scale*, *Chapter 6* gives a reformulation of *local scale*. The TV flow equation is used as the image filter to extract the local scale. In particular, we show the applications of the new definition on some of the general image analysis tasks.

Chapter 7 is on applications of multi-scale and local scale analysis on medical image analysis. This chapter is the main motivation for the whole research. The segmentation of blood vessels and drusen structures are compared with some state-of-art algorithms, advantages and limitations of the proposed methods are both discussed. When readers finished reading this paper, it is expected that readers can get a better idea of how nonlinear diffusion models can be used in multi-scale analysis, and how *scale*, as a term, can be extracted and interpreted in a clear way. In application

research, bridging gaps between different research areas is never a trivial task, and gaps in our study can be summarized in Figure 1.2. Our study is focused on total variation model, and it can be considered one example of how to use a mathematical model in a proper way to bridge these gaps.

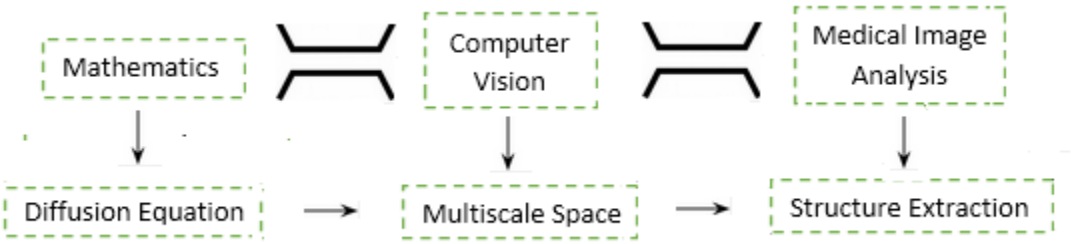


Figure 1.2: Gaps between different research areas

Chapter 2

Image Diffusion

2.1 Physical Background

In physics, the diffusion process will equilibrate concentration differences without creating or destroying mass. By Fick's law[8], the equilibration is driven by:

$$j = -D \cdot \nabla u(x) \tag{2.1.1}$$

where $x \in R^n, \nabla = (\partial/\partial x_1, \cdots, \partial/\partial x_n)$, j is a flux compensating for the concentration gradient ∇u. If j is parallel to ∇u, the diffusion process is called isotropic diffusion, and D is a positive scalar called diffusivity. If instead, j is not parallel to ∇u, it is called anisotropic diffusion. Based on the conservation of mass, the following equation can be constructed:

$$\partial_t u = -div \quad j \tag{2.1.2}$$

If we plug (2.1.1) in (2.1.2), the diffusion equation becomes

$$\partial_t u = div(D \cdot \nabla u) \tag{2.1.3}$$

This equation appears in many physical transport process. For heat transfer, Equation (2.1.2) is called heat equation written as

$$\partial_t u(t, x) = D \triangle u(t, x) \tag{2.1.4}$$

In image processing, image intensity at each location(or pixel) can be considered the concentration of mass. Given the initial condition $u(0, x) = f(x)$, the solution of Equation (2.1.2) is proven to be the convolution of initial condition $f(x)$ with the gaussian kernel with kernel size $\sqrt{2t}$. As the kernel size $\sigma = \sqrt{2t}$ increases (i.e. the time variable t increases), the image will be smoothed and evolve towards a constant image(shown in Figure 2.1). This can be considered as a physical explanation of image diffusion process, however compared to heat(or mass) transportation which may not have important underlying geometric structures, in image processing, we are more interested to study how image structures evolve with the force of diffusion. In Section 2.2, we will have a closer look at the diffusion operators, and interpret image diffusion process in a direct way.

2.2 Diffusion operator decomposition

In image analysis, a 2D image can be analyzed from three different perspectives

1. Level set: 1D curves with the same intensity level $f_L = \{(x, y)|f(x, y) = L\}$

2. 2D Image Matrix: $\{f(x, y)|(x, y) \in \Omega \subseteq R^2\}$

3. 2D image embedded in 3D space: $\{(x, y, f(x, y))|(x, y) \in \Omega \subseteq R^2\}$

Edges(or level sets) are considered the most important and interesting features in images, as they are stable under geometric transformations(e.g. translation, dilation, rotation). To understand effects of physical diffusion process on images, it is important to understand how curves, edges are changed locally under the diffusion force. In fact, the image diffusion operator can be decomposed with respect to local image structures(e.g. the tangential and normal vectors of level sets).

Figure 2.1: Example of heat diffusion on image. From left to right: original image(size 256 by 256), image smoothed with kernel size $\sigma = 2, 5, 15$ (pixel)

2.2.1 Linear Diffusion: Heat Equation

Heat equation is considered the typical and the simplest linear parabolic equation. In image processing, the heat equation or gaussian smoothing is commonly used in preprocessing. The PDE problem for the 2D image is shown as below:

$$\begin{cases} \frac{\partial u}{\partial t} = \triangle u \qquad t > 0, x \in \Omega \subseteq R^2 \\ u(0,x) = f(x) \qquad t = 0, x \in \Omega \end{cases} \tag{2.2.1}$$

The solution to Problem (2.2.1) is

$$u(t,x) = G(t,\cdot) * f(x), t > 0 \tag{2.2.2}$$

where $G(t,x)$ is the 2D-Gaussian kernel,

$$G(t,x) = \frac{1}{4\pi t} e^{-x^T x/4t}, t > 0 \tag{2.2.3}$$

In order to understand the heat diffusion process using a local frame(i.e.the tangential and normal vectors of level sets), we have the following proposition.

Proposition 2.2.1. *Given image* $f(x,y), (x,y) \in \Omega \subseteq R^2. \forall (x,y), s.t. |\nabla f| \neq 0$, *denote* $N(x,y) = \frac{\nabla f}{|\nabla f|}$ *and* $T(x,y)$ $s.t. |T| = 1, N \cdot T = 0$, *then the Laplacian operator* $\triangle$ *can be decomposed along the two directions* $N(x,y), T(x,y)$, *i.e.*

$$\triangle u(x) : u_{xx} + u_{yy} = u_{NN} + u_{TT} \tag{2.2.4}$$

Proof.

$$N(x,y) = \frac{(f_x, f_y)}{|\nabla f|} = (N_1, N_2),\ T(x,y) = \frac{(f_y, -f_x)}{|\nabla f|} = (T_1, T_2)$$

Denote n, t as the coordinates of point (x, y) along N, T direction respectively. According to the definition of directional derivative,

$$g(n) = f((x,y) + nN) = f(x + n\frac{f_x}{|\nabla f|}, y + n\frac{f_y}{|\nabla f|}) = f(x + nN_1, y + nN_2)$$

$$\begin{aligned}
g'(n) =& \lim_{t\to 0} \frac{g(n+t) - g(n)}{t} \\
=& \lim_{t\to 0} \frac{f(x + (n+t)N_1, y + (n+t)N_2) - f(x,y)}{t} \\
=& \lim_{t\to 0} \frac{f(x + (n+t)N_1, y + (n+t)N_2) - f(x + (n+t)N_1, y + nN_2)}{t} + \\
& \frac{f(x + (n+t)N_1, y + nN_2) - f(x + nN_1, y + nN_2)}{t} \\
=& f_y(x + nN_1, y + nN_2)N_2(x,y) + f_x(x + nN_1, y + nN_2)N_1(x,y)
\end{aligned}$$

$$\therefore g(0) = f(x,y), g'(0) = |\nabla f(x,y)|$$

For the second order directional derivative,

$$\begin{aligned} g^{''}(n) =& \lim_{t\to 0} \frac{g'(n+t) - g'(n)}{t} \\ =& (f_{yy}(x+nN_1, y+nN_2)N_2 + f_{xy}(x+nN_1, y+nN_2)N_1)N_2 \\ &+ (f_{xy}(x+nN_1, y+nN_2)N_2 + f_{xx}(x+nN_1, y+nN_2)N_1)N_1 \\ =& f_{yy}(x+nN_1, y+nN_2)N_2^2 \\ &+ 2f_{xy}(x+nN_1, y+nN_2)N_1N_2 + f_{xx}(x+nN_1, y+nN_2)N_1^2 \end{aligned}$$

$$\therefore g^{''}(0) = \frac{f_x^2 f_{xx} + 2f_x f_y f_{xy} + f_y^2 f_{yy}}{|\nabla f|^2}(x,y) = \frac{\partial^2 g}{\partial n^2}(0) = f_{NN}$$

Based on the new coordinates (n,t) with respect to the local system (N,T), the directional derivative along normal direction N can be written as:

$$f_N(x,y) = |\nabla f(x,y)|, f_{NN}(x,y) = \frac{f_x^2 f_{xx} + 2f_x f_y f_{xy} + f_y^2 f_{yy}}{|\nabla f|^2}(x,y) \tag{2.2.5}$$

Similarly, along the tangential direction of level set curves,

$$T = \frac{(f_y, -f_x)}{|\nabla f|} = (T_1, T_2), g(n) = f(x+nT_1, y+nT_2)$$

$$g'(n) = f_y(x+nT_1, y+nT_2)T_2 + f_x(x+nT_1, y+nT_2)T_1$$

$$g'(0) = f_y\frac{-f_x}{|\nabla f|} + f_x\frac{f_y}{|\nabla f|} = 0$$

$$g^{''}(n) = f_{yy}(x+nT_1, y+nT_2)T_2^2 + 2f_{xy}(x+nT_1, y+nT_2)T_1T_2 + f_{xx}(x+nT_1, y+nT_2)T_1^2$$

$$g''(0) = \frac{f_{yy}f_x^2 - 2f_{xy}f_xf_y + f_{xx}f_y^2}{|\nabla f|^2}$$

$$\therefore f_T(x,y) = 0, f_{TT}(x,y) = \frac{f_{yy}f_x^2 - 2f_{xy}f_xf_y + f_{xx}f_y^2}{|\nabla f|^2} \tag{2.2.6}$$

Using Equation (2.2.5) and (2.2.6), the Laplacian operator can be decomposed as

$$\triangle f(x,y) = f_{NN}(x,y) + f_{TT}(x,y) \tag{2.2.7}$$

□

In addition to Proposition 2.2.1, the other decomposition of the Laplacian operator based on the geometric property of level set curves will be introduced. Before that, a property related to level set curvature will be discussed.

Proposition 2.2.2. *If we denote $\kappa(x,y)$ as the signed curvature of the level set curve at point (x,y) in image $f(x,y)$, with $|\nabla f(x,y)| \neq 0$, then*

$$\kappa(x,y) = \nabla \cdot (\frac{\nabla f}{|\nabla f|})$$

Proof. Given

$$N = \frac{\nabla f}{|\nabla f|} = (N_1, N_2), T = \frac{(f_y, -f_x)}{|\nabla f|} = (T_1, T_2), \kappa = \frac{dN}{ds} \cdot T$$

Here, s is the arc length and use the change of variable with respect to arc length,

$(x_s, y_s) = T$,

$$\kappa = \frac{dN}{ds} \cdot T = (\frac{\partial N}{\partial x}\frac{dx}{ds} + \frac{\partial N}{\partial y}\frac{dy}{ds}) \cdot T$$

$$\begin{aligned}
\kappa =& (\partial_x N_1 x_s + \partial_y N_1 y_s, \partial_x N_2 x_s + \partial_y N_2 y_s) \cdot (T_1, T_2) \\
=& (\partial_x N_1 T_1 + \partial_y N_1 T_2, \partial_x N_2 T_1 + \partial_y N_2 T_2) \cdot (T_1, T_2) \\
=& (T_1, T_2) \begin{pmatrix} \partial_x N_1 & \partial_x N_2 \\ \partial_y N_1 & \partial_y N_2 \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = T \nabla N T^T \\
=& \frac{f_{xx} f_y^2 + f_{yy} f_x^2 - 2 f_x f_y f_{xy}}{|\nabla f|^3} = \nabla \cdot (\frac{\nabla f}{|\nabla f|})
\end{aligned}$$

□

Signed curvature v.s. unsigned curvature

Remark 2.2.3. Curvature of a curve or surface has been used widely in image analysis, and it is crucial to use the proper definition of curvature in numeric implementations. In fact, curvature of a curve can be defined as a *unsigned*(i.e. nonnegative), or *signed*(i.e. positive or negative) scalar. The notion of unsigned curvature is consistent with the intuition, as curvature describes how fast the tangential direction of a curve changes. However, as curve can only be studied meaningfully with a certain parameterization, it is expected the sign of the curvature should also depend on the parametrization, and can possibly have changing signs.

In real applications, the choice of curvature definition might be a confusing problem.

On one hand, curvature is merely a descriptor about changes of tangential direction, a positive scalar is preferred. However, there are also many examples using the signed curvature definition, for example the mean curvature diffusion introduced later in this section. A simple example below will be used to illustrate the reason that in image analysis, the signed curvature definition is preferred. Given a curve Γ shown in Figure 2.2, parametrized with Equation (2.2.8),

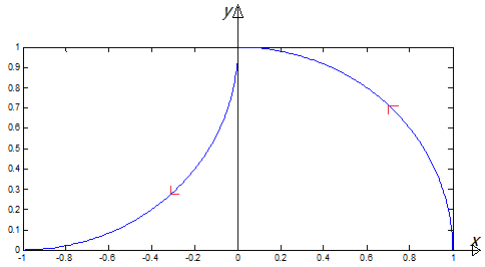


Figure 2.2: Parametrized Curve Γ

$$r(t) = \begin{cases} (\cos t, \sin t) & 0 \le t \le \frac{\pi}{2}, \\ (-1 + \sin t, 1 + \cos t) & \frac{\pi}{2} \le t \le \pi \end{cases} \tag{2.2.8}$$

1, For unsigned curvature,

$$\kappa(t) = \frac{|r'(t) \times r''(t)|}{|r'(t)|^3} = 1, T(t) = r'(t), T'(t) = k(t)N(t)$$

$$T(t) = \begin{cases} (-\sin t, \cos t) & 0 \le t \le \frac{\pi}{2}, \\ (\cos t, -\sin t) & \frac{\pi}{2} \le t \le \pi \end{cases}, N(t) = \begin{cases} (-\cos t, -\sin t) & 0 \le t \le \frac{\pi}{2}, \\ (-\sin t, -\cos t) & \frac{\pi}{2} \le t \le \pi \end{cases}$$

2, For signed curvature,

$$\kappa(t) = \frac{r'(t) \times r''(t)}{|r'(t)|^3} = \begin{cases} 1 & 0 \le t \le \frac{\pi}{2}, \\ -1 & \frac{\pi}{2} \le t \le \pi \end{cases}$$

$$T(t) = \begin{cases} (-\sin t, \cos t) & 0 \le t \le \frac{\pi}{2}, \\ (\cos t, -\sin t) & \frac{\pi}{2} \le t \le \pi \end{cases}, N(t) = \begin{cases} (-\cos t, -\sin t) & 0 \le t \le \frac{\pi}{2}, \\ (\sin t, \cos t) & \frac{\pi}{2} \le t \le \pi \end{cases}$$

In Figure 2.2, starting from point $(1, 0)$, and travel to $(-1, 0)$ along Γ, it can be

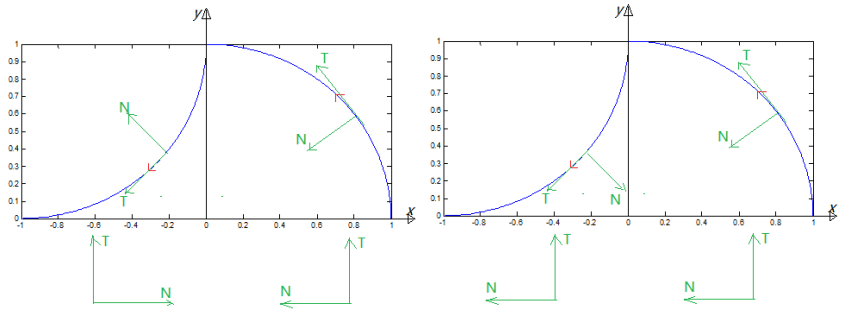


Figure 2.3: (N, T): Left: Unsigned curvature; Right: Signed curvature

observed that traveling from $(1, 0)$ to $(0, 1)$, the angle between tangential direction and x axis is increasing with rate 1, and starting from $(0, 1)$ to $(-1, 0)$, the angle is decreasing with rate 1, consistent with the calculation for signed curvature. With a simple rotation, the two segment of curves i.e. $(1, 0)$ to $(0, 1)$ and $(0, 1)$ to $(-1, 0)$ can be considered as the same based on the shape. However, under a certain parametrization, it would be better if the curvature can characterize the difference between the two segments. It is preferred to use the definition of signed curvature. In addition, according to the explicit calculation, (N, T) is plotted shown in Figure 2.3. The orien-

tation between (N, T) is consistent for signed curvature and changes(or inconsistent) for unsigned curvature. It can also support the usage of signed curvature framework, using (N, T) as local coordinate system is more appropriate. In the following study, *curvature* will mean *signed curvature*.

Proposition 2.2.4. *Given image* $f(x), x \in \Omega \subseteq R^2$. $\forall x, s.t. |\nabla f(x)| \neq 0$,*we denote* $N(x) = \frac{\nabla f(x)}{|\nabla f(x)|}$ *and the Laplacian operator* $\triangle$ *can be decomposed as*

$$\triangle f(x, y) : f_{xx} + f_{yy} = f_{NN} + \kappa(x, y)|\nabla f(x, y)| \tag{2.2.9}$$

Proof. By direct calculation,

$$\begin{aligned}
(\nabla \cdot \frac{\nabla f}{|\nabla f|})|\nabla f| =& (\partial_x(\frac{\partial_x f}{|\nabla f|}) + \partial_y(\frac{\partial_y f}{|\nabla f|}))|\nabla f| \\
=& (\frac{\triangle f}{|\nabla f|} - \frac{f_x^2 f_{xx} + 2 f_x f_y f_{xy} + f_y^2 f_{yy}}{|\nabla f|^3})|\nabla f| \\
=& \triangle f - \frac{f_x^2 f_{xx} + 2 f_x f_y f_{xy} + f_y^2 f_{yy}}{|\nabla f|^2} \\
=& \triangle f - f_{NN}
\end{aligned}$$

Using Proposition 2.2.2

$$\triangle f = f_{NN} + (\nabla \cdot \frac{\nabla f}{|\nabla f|})|\nabla f| = f_{NN} + \kappa(x, y)|\nabla f(x, y)|$$

□

Remark 2.2.5. One of the reasons that motivated us to discuss about signed and unsigned curvature is [9]. This paper was published on the same journal of [10] as a correction of the proof and conclusion of [10]. In [9], the author used the definition of unsigned curvature, and claimed that the decomposition of Laplacian operator should have the following form

$$\triangle f = f_{NN} \pm \kappa |\nabla f|, |\nabla f| \neq 0 \tag{2.2.10}$$

In fact, without further analysis, the existence of $\pm$ signs in Equation 2.2.10 can pose problems in real applications, as it does not indicate explicitly which sign should be used in the decomposition formula in [9].

2.2.2 Nonlinear Diffusion Equation

Within the scope of research interests, in this part, discussions are focused on nonlinear diffusion equations that have already been used widely in image analysis. Among many equations [11, 12, 13, 14, 15], the Perona-Malik equation[11] is considered the first nonlinear diffusion equation proposed to overcome the limitations of the linear diffusion equation.

Perona-Malik Model

The general Perona-Malik model [11] is given below

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \nabla \cdot (g(|\nabla u(t,x)|^2)\nabla u(x,t)), & t > 0, x \in \Omega \subseteq R^2 \\ u(0,x) = f(x) & x \in \Omega \end{cases} \tag{2.2.11}$$

where the diffusivity $g : [0, +\infty) \to (0, +\infty)$ is a smooth and decreasing function, adaptive to the gradient of the image $u(t, x)$. It is a direct improvement to preserve the image structures, and reduce the blurring effects at edges. In [11], one example of the diffusivity $g(\cdot)$ functions is given as

$$g(|\nabla u|^2) = \frac{1}{1 + |\nabla u|^2/\lambda^2}, \lambda > 0 \tag{2.2.12}$$

Diffusion operator decomposition

Similar to Laplacian operator, the nonlinear diffusion operators can also be decomposed with respect to the two local directions N, T. Recall that

$$u_{NN} = \frac{u_{xx}u_y^2 - 2u_x u_y u_{xy} + u_{yy}u_x^2}{|\nabla u|^2}, u_{TT} = \frac{u_{xx}u_x^2 + 2u_x u_y u_{xy} + u_{yy}u_y^2}{|\nabla u|^2}$$

Proposition 2.2.6. *The Perona-Malik diffusion operator can be decomposed as :*

$$\nabla \cdot (g(|\nabla u|^2)\nabla u) = g(|\nabla u|^2)u_{NN} + h(|\nabla u|^2)u_{TT}, \textit{here } h(s) = g(s) + 2sg'(s) \tag{2.2.13}$$

Proof.

$$\begin{aligned}\nabla \cdot (g(|\nabla u|^2)\nabla u) =& g(|\nabla u|^2) \triangle u + g'(|\nabla u|^2)(\frac{\partial |\nabla u|^2}{\partial x} u_x + \frac{\partial |\nabla u|^2}{\partial y} u_y) \\ =& g(|\nabla u|^2) \triangle u + 2g'(|\nabla u|^2)(u_{xx}u_x^2 + 2u_x u_y u_{xy} + u_{yy}u_y^2) \\ =& g(|\nabla u|^2)(u_{TT} + u_{NN}) + 2g'(|\nabla u|^2)|\nabla u|^2 u_{NN} \\ =& g(|\nabla u|^2)u_{TT} + (g(|\nabla u|^2) + 2g'(|\nabla u|^2)|\nabla u|^2)u_{NN} \\ =& g(|\nabla u|^2)u_{TT} + h(|\nabla u|^2)u_{NN}\end{aligned}$$

□

Remark 2.2.7. Functions g, h are considered as weight coefficients of image diffusion. In order to smooth along tangential direction T instead of normal direction N(i.e. avoid blurring edges), it is desirable that $g(|\nabla u|^2) > h(|\nabla u|^2)$. Consider the $g(s)$ as shown in the example, $h(s) - g(s) = 2sg'(s)$,

$$sg'(s) = -\frac{s}{\lambda^2(1 + \frac{s}{\lambda^2})^2} < 0, s < 0. \therefore g(|\nabla u|^2) > h(|\nabla u|^2)$$

It is also natural to require that the diffusivity along both directions to be positive i.e. $g(|\nabla u|^2) > 0, h(|\nabla u|^2) > 0$, if $g(s)$ is given as in Equation 2.2.12.

$$h(s) = g(s) + sg'(s) = \frac{1}{1 + \frac{s}{\lambda^2}} - \frac{s}{(1 + \frac{s}{\lambda^2})^2\lambda^2} = \frac{1}{(1 + \frac{s}{\lambda^2})^2\lambda^2}(\lambda^2 - s) > 0$$

If $|\nabla u(t, x)| \leq \lambda$, $h(|\nabla u|^2) \geq 0$ (*). In fact, under this condition(*), Perona-Malik

model is considered as forward parabolic type. Instead, if $|\nabla u(t,x)| > \lambda$, it is considered backward parabolic type which is closely related to an well known ill-posed differential equation $u_t = - \triangle u$. The detailed analysis of Perona-Malik model can be found in [11, 16, 17, 18]

Mean curvature motion

In [19], authors proposed the Alvarez-Lions-Morel pure anisotropic diffusion, also commonly known as mean curvature motion(MCM). The PDE is motivated by Perona-Malik equation with slight differences:

$$\frac{\partial u}{\partial t} = |\nabla u| \nabla \cdot (\frac{\nabla u}{|\nabla u|}), u(0,x) = f(x), x \in \Omega \subseteq R^2 \tag{2.2.14}$$

According to Proposition 2.2.2, it is not difficult to conclude

$$|\nabla u| \nabla \cdot (\frac{\nabla u}{|\nabla u|}) = \kappa(x,y)|\nabla u| = u_{TT} \tag{2.2.15}$$

The mean curvature diffusion can than decomposed as

$$\frac{\partial u}{\partial t} = u_{TT} + 0 \cdot u_{NN} \tag{2.2.16}$$

Authors [19] claimed that "it(Equation 2.2.15) represents a degenerated diffusion term, which diffuses u in the direction orthogonal to its gradient ∇u, and doesn't diffuse at all in the direction of ∇u. The aim of the degenerated diffusion term is

to make u smooth on both sides of an edge with a minimal smoothing of edge itself. This means that for such a theory, edges are nothing but the boundaries of the level lines of the image".

Total Variation Flow Diffusion

In Chapter 4, we will present a detailed analysis of total variation flow diffusion, here we include the diffusion equation as comparisons with other models.

$$\frac{\partial u}{\partial t} = \nabla \cdot (\frac{\nabla u}{|\nabla u|}), u(0,x) = f(x), x \in \Omega \subseteq R^2 \tag{2.2.17}$$

Based on the previous analysis, this equation can be written in the following two forms, for $x, s.t. |\nabla u(t,x)| \neq 0$

$$\frac{\partial u}{\partial t} = \kappa(t,x) \tag{2.2.18}$$

$$\frac{\partial u}{\partial t} = \frac{1}{|\nabla u|} u_{TT} \tag{2.2.19}$$

The decomposition implies that the total variation flow diffusion of the image is completely decided by the curvature property of the level set curves.

2.3 Numerical implementations

2.3.1 Linear Diffusion Equation

Notations: $\delta_x^{-1}, \delta_y^{-1}$: backward difference; $\delta_x^{+}, \delta_y^{+}$: forward difference;

dim = 1, $u_t = u_{xx}, u_n^m(t,x) = u(nh, mk)$

1, Explicit Scheme: $u_i^{n+1} = u_i^n + \frac{k}{h^2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n)$

2, Implicit Scheme: $u_i^{n+1} = u_i^n + \frac{k}{h^2}(u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1})$

Using the stability analysis,

$$g = \frac{1}{1 + 4\frac{k}{h^2} sin^2 \frac{\theta}{2}} \Rightarrow |g| \leq 1$$

If the 1D signal has length N, then at each iteration, we need to solve the following linear system

$$\begin{pmatrix} 1+\lambda & -\lambda & & 0 \\ -\lambda & 1+2\lambda & \ddots & \\ & \ddots & \ddots & -\lambda \\ & & -\lambda & 1+\lambda \end{pmatrix} \begin{pmatrix} u_1^{n+1} \\ u_2^{n+1} \\ \vdots \\ u_N^{n+1} \end{pmatrix} = \begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_N^n \end{pmatrix}, k/h^2 = \lambda \tag{2.3.1}$$

The implicit scheme for this parabolic equation is unconditionally stable, and in general, it is true for parabolic equation. Explicit scheme is easy to evolve at each iteration, however restricted by the small time step size, and it can be time consuming. Implicit scheme has no restriction on time step size, however at each step, it needs to solve a linear system even a nonlinear system for nonlinear differential equation. It is needed to balance the cons/pros when choosing a numerical scheme. *3, Semi-implicit Scheme* - Besides the two schemes above, there is a combination of these

two: $u_i^{n+1} = u_i^n + \frac{1}{2}\frac{k}{h^2}(u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1} + u_{i+1}^n - 2u_i^n + u_{i-1}^n)$ This scheme is called Crank-Nicholson scheme, the growth factor

$$g = \frac{1 - 2\frac{k}{h^2}sin^2\frac{\theta}{2}}{1 + 2\frac{k}{h^2}sin^2\frac{\theta}{2}} \Rightarrow |g| \leq 1$$

dim = 2, $u_t = \triangle u, u_{i,j}^n = u(nk, ih, jh)$

1, Explicit Scheme: $u_{i,j}^{n+1} = u_{i,j}^n + k[\delta_x^{\pm}(\delta_x^{\pm}u_{i,j}^n) + \delta_y^{\pm}(\delta_y^{\pm}u_{i,j}^n)]$

There are 2^4 different choices of combination of backward or forward difference schemes, and the scheme we choose to use is:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{k} = \frac{\delta_x^+ u_{i,j}^n - \delta_x^+ u_{i-1,j}^n}{h} + \frac{\delta_y^+ u_{i,j}^n - \delta_y^+ u_{i,j-1}^n}{h} \tag{2.3.2}$$

The stability condition for this scheme is $\frac{k}{h^2} \leq 1/4$. This stability condition depends on the dimension of spacial variable, and can be generalized to $\frac{k}{h^2} \leq 1/(2\dim(x))$

2, Rotation invariant scheme: Different from 1D operator, the above schemes don't take 2D property into consideration. As we know that $\triangle$ operator is rotation invariant, to use this property, instead we use the following scheme

$$\begin{aligned}\triangle u_{i,j}^n = &\lambda\frac{u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}}{h^2} \\ &+ (1-\lambda)\frac{u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1.j-1} - 4u_{i,j}}{2h^2}\end{aligned}$$

λ is the coefficient needs to be decided by enforcing the "rotation invariant" property, and in [2], λ is chosen to be $\frac{1}{3}$. Compared to previous schemes, this approximation uses the 3×3 neighborhood of each pixel. The advantages are (1) it becomes rotation invariant; (2) it will be less sensitive to noise.

2.3.2 Nonlinear Diffusion

For nonlinear diffusion equations, there are numerous numerical methods studied. For example, finite difference, finite element, neural network approximate etc. [16]. However, among different methods, finite difference scheme is preferred by many researchers. One of the reasons is that the discrete representation of the image provides a natural structure for a finite difference scheme.

The first numerical scheme for total variation model was explicit([20],[21]). Later, many studies found faster and more accurate schemes, for example [22, 23, 24]. For the model in our study, there are two important aspects in numerical implementation to be aware of when using the finite difference scheme.

1, Singularity: The curvature term κ is replaced with a regularized version to avoid the singularity at $x, s.t. |\nabla f| = 0$, $\epsilon > 0$ can be chosen arbitrarily small:

$$\kappa(t, x) = \nabla \cdot (\frac{\nabla f}{|\nabla f|}) \Rightarrow \nabla \cdot (\frac{\nabla f}{\sqrt{|\nabla f|^2 + \epsilon^2}})$$

2, Stability: The stability condition in finite difference scheme is considered important. If the step size for time t is not chosen under the stability condition, the numer-

ical solution can be highly oscillated. In 2D image analysis, the stability condition can be summarized as

$$\frac{\triangle t}{\triangle x \triangle y} \leq \frac{\epsilon}{4} \tag{2.3.3}$$

To make the numerical solution close to the true solution of the original equation, ϵ should be chosen close to 0, and the step size $\triangle t$ should be small enough accordingly, making the diffusion process very slow. In general, there should be a balance between speed and accuracy of the numerical simulation. Taking Equation 2.2.17 as an example for both 1D and 2D cases:

1D Explicit Scheme

$$\partial_t u(x,t) = \partial_x(\frac{\partial_x u}{|\partial_x u|}), u(x,0) = f(x) \tag{2.3.4}$$

Using regularization, we will solve the following problem

$$\partial_t u(x,t) = \partial_x(\frac{\partial_x u}{\sqrt{|\partial_x u|^2 + \epsilon^2}}), u(x,0) = f(x) \tag{2.3.5}$$

Denote $u_m^n = u(mh, nk)$

$$\delta_t^{-(+)} u_m^n = \frac{u_m^{n(n+1)} - u_m^{n-1(n)}}{k}, \delta_x^{-(+)} u_m^n = \frac{u_{m(m+1)}^n - u_{m-1(m)}^n}{h}$$

$$\frac{u_m^{n+1} - u_m^n}{k} = \frac{1}{h}(\frac{u_{m+1}^n - u_m^n}{\sqrt{(u_{m+1}^n - u_m^n)^2 + \epsilon^2 h^2}} - \frac{u_m^n - u_{m-1}^n}{\sqrt{(u_m^n - u_{m-1}^n))^2 + \epsilon^2 h^2}}) \tag{2.3.6}$$

2D Explicit Scheme

$$\frac{u^{n+1} - u^n}{k} = \nabla \cdot (\frac{\bigtriangledown u^n}{\sqrt{|\bigtriangledown u^n|^2 + \epsilon^2}})$$

The finite difference scheme is

$$u^{n+1} = u^n + k[\delta_x^-(\frac{\delta_x^+ u}{\sqrt{\delta_x^+ u^2 + m(\delta_y^- u, \delta_y^+ u)^2 + \epsilon^2}} + \delta_y^-(\frac{\delta_y^+ u}{\sqrt{\delta_y^+ u^2 + m(\delta_x^- u, \delta_x^+ u)^2 + \epsilon^2}})]$$

$m(a,b) = minmod(a,b) = \min(|a|,|b|)(\text{sign } a + \text{sign } b)/2$, the reason *minmod* used here is to balance the bias of using merely forward or backward difference.

The way we analyze the stability condition is based on discrete maximum-minimum principle [25], it is closely related to the causality property when we define a scale space, and it is a more restrictive stability than the von Neumann stability. The von Neumann stability is necessary and sufficient for stability for linear PDE models. The PDE must be constant-coefficient with periodic boundary conditions and have only two independent variables, and the scheme must use no more than two time levels. In a nonlinear problem, the maximum-minimum principle is more proper. According to [25], the stability condition is $k/h^2 \leq \epsilon/(2N), x \in R^N$, for image processing, $N = 2, h = 1, k \leq \epsilon/4$. As $\epsilon \to 0$, the time step size k is very small making the iteration slow.

Implicit Scheme

$$\frac{u^{n+1} - u^n}{k} = \nabla \cdot (\frac{\bigtriangledown u^{n+1}}{\sqrt{\epsilon^2 + |\bigtriangledown u^{n+1}|^2}})$$

At $(n+1)$ th step of the diffusion process, given u^n

$$u^{n+1} - k \bigtriangledown \cdot (\frac{\bigtriangledown u^{n+1}}{\sqrt{\epsilon^2 + |\bigtriangledown u^{n+1}|^2}}) = u^n$$

Using the fixed point iteration method to solve the above equation for u^{n+1} At kth fixed point iteration , we solve for u_k^{n+1}

$$u_{k+1}^{n+1} - k \bigtriangledown \cdot (\frac{\bigtriangledown u_{k+1}^{n+1}}{\sqrt{\epsilon^2 + | \bigtriangledown u_k^{n+1}|^2}}) = u^n$$

It will result in solving a linear system for u_{k+1}^{n+1}, the fixed point iteration will stop when the difference between right and left hand side is within an error toleration.

If we denote $w = u_{k+1}^{n+1}, v = w_k^{n+1}, f = u^n$,the finite difference scheme for 1D is:

$$w(i) - k\delta_x^{-1}(\frac{\delta_x^+ w(i)}{\sqrt{\epsilon^2 + |\nabla v(i)|^2}}) = f$$

$$V = \frac{k}{h^2}(\frac{1}{\sqrt{\epsilon^2 + |\nabla v(1)|^2}}, \cdots, \frac{1}{\sqrt{\epsilon^2 + |\nabla v(N)|^2}}), i = 1, ..., N$$

$$\begin{pmatrix} 1+V_1 & -V_1 & & & 0 \\ -V_1 & 1+V_1+V_2 & -V_2 & & \\ \ddots & & \ddots & & \ddots \\ & -V_{N-2} & 1+V_{N-2}+V_{N-1} & -V_{N-1} & \\ & & -V_{N-1} & 1+V_{N-1} & \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{pmatrix} = \begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_N^n \end{pmatrix} \tag{2.3.7}$$

The finite difference scheme for 2D case is

$$w(i,j) - k\delta_x^{-1}(\frac{\delta_x^+ w(i,j)}{\sqrt{\epsilon^2 + |\nabla v(i,j)|^2}}) - k\delta_y^{-1}(\frac{\delta_y^+ w(i,j)}{\sqrt{\epsilon^2 + |\nabla v(i,j)|^2}}) = f$$

By vectorizing the matrix, we can get a similar coefficient matrix as in (2.3.7), however due to the vectorizing scheme, the matrix will not be tridiagonal.

For 1D case, we can solve for w,where the coefficient matrix is tridiagonal. This computation is well defined because we know that strictly diagonally dominant matrices($A = (a_{ij}), \forall i, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$) are invertible. In each iteration, we compare the difference between two sides of the equality to check the stopping condition. However, for 2D image, to treat 2D image as a vector, if image has size $N \times N$, the matrix size would be $N^2 \times N^2$, which can be a heavy memory load for Matlab,even if $N = 128$. In our study, we use explicit finite difference scheme for 2D problem.

2.4 Interpretation of image diffusion

With the introduction of diffusion equation, decomposition of diffusion operators and basic numerical schemes, we can simulate the diffusion process for 1D signals and 2D images. Especially, we can simulate effects of diffusion operator along the local tangential and normal directions of level sets. One important question is how we can interpret the image diffusion process. The proper interpretation of the process will help us to understand the diffusion better and be helpful for analysis in multi-scale space. Before going directly to the construction of diffusion based multi-scale space, we take a step back to interpret the diffusion process. Within the scope of our study, we focus on diffusion operator in 1D and 2D space.

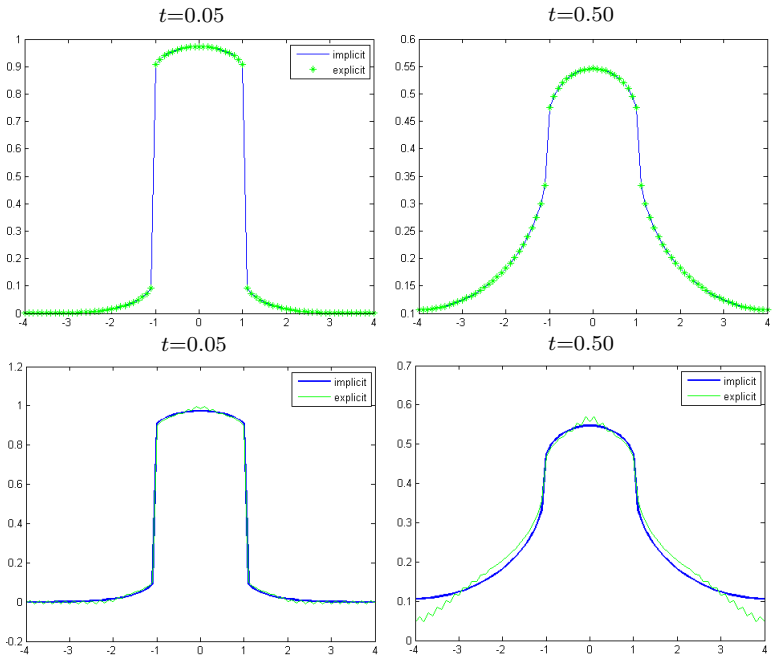


Figure 2.4: In this experiment, we use explicit and implicit difference schemes for 1D TV flow diffusion, with the initial condition $x = -4 : 0.1 : 4, \Delta x = 0.1, f(x) = 1, |x| \leq 1; 0,$ else $, \epsilon = 0.1,$ in the four plots, explicit and implicit results are green and blue curves respectively. In the first row, $\Delta t = 0.25\epsilon\Delta x^2$ satisfying the stability condition, we can see that the evolving results are very close between these two schemes; while in the second row, $\Delta t = \epsilon\Delta x^2$ disobeys the stability condition, we can find for the explicit scheme, there are high oscillations, and implicit schemes are stable compared to both the explicit scheme and the first row.

1D Signal Diffusion

Diffusion for 1D signal can be formulated in a general form:

$$\frac{\partial u}{\partial t} = a(t,x)\partial_{xx}u, u(0,x) = f(x), x \in I \subseteq R \tag{2.4.1}$$

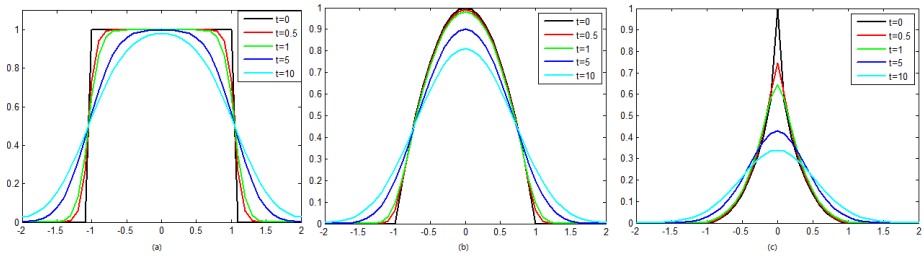


Figure 2.5: Heat diffusion for three types of 1D signal(*i.e.* $a(t,x)=1$)

2D image diffusion

For 2D images, we are interested to study effects of diffusion along local directions ξ, η (may not necessarily be the local tangential or normal directions),

$$\frac{\partial u}{\partial t}(t,x) = a(x,t,u)\partial_{\xi\xi}u(x,t) + b(x,t)\partial_{\eta\eta}u(x,t) \tag{2.4.2}$$

We will start from the following diffusion process.

$$\frac{\partial u}{\partial t} = \partial_{\xi\xi}u(x,t), u(0,x) = f(x), x \in \Omega \subseteq R^2 \tag{2.4.3}$$

The ξ used in the numerical experiments are N and T. Recall that

$$u_{NN} = \frac{u_{xx}u_y^2 - 2u_x u_y u_{xy} + u_{yy}u_x^2}{|\nabla u|^2}, u_{TT} = \frac{u_{xx}u_x^2 + 2u_x u_y u_{xy} + u_{yy}u_y^2}{|\nabla u|^2}$$

In the example, we will compare the effects of the two diffusion equations.

$$(1)\frac{\partial u}{\partial t} = u_{NN}; (2)\frac{\partial u}{\partial t} = u_{TT}$$

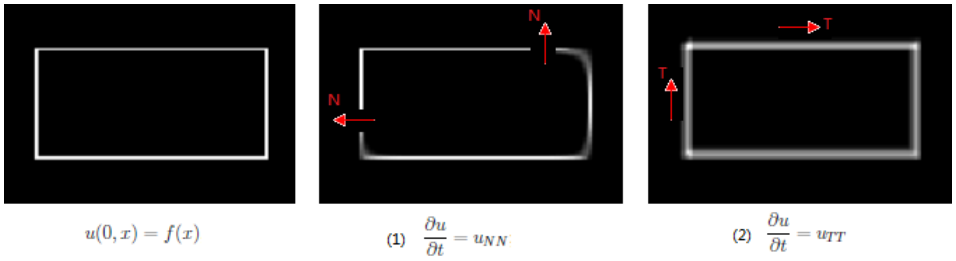


Figure 2.6: Image diffusion along direction N, T

Remark 2.4.1. Equation (2) in above is the mean curvature diffusion discussed in the previous section. Compare (1) with (2) in Figure 2.6, it can be observed that the diffusion force enforced by u_{NN} tends to diffuse the image intensity along the normal direction N of the level set. This force tends to distort the edge information, as observed in Figure 2.6(1), edge structures are distorted to be round corners instead. The diffusion along tangential direction tends to diffuse the image intensity along this direction only, shown in Figure 2.6(2). In Figure 2.6, there are still slight blurring of edges along N direction. It can be explained by the fact that a regularization term $1/\sqrt{\epsilon^2 + |\nabla u|^2}$ was used in the numerical implementation to avoid the singularity for $1/|\nabla u|$ occurs as $|\nabla u| \to 0$.

Remark 2.4.2. From the two examples, it becomes clear that image diffusion that is normally expressed using (x, y) coordinates can be better interpreted based on a decomposition of diffusion operators, using a local coordinate system. In general, this local coordinate system uses the natural tangential and normal vectors of the level sets. In regard to the purpose of edge preserving/enhancing, it shows that image

diffusion along tangential direction is preferred to the normal direction. This is the reason why mean curvature diffusion, total variation flow attracts more interests in image analysis.

Chapter 3

Scale: Linear Diffusion

3.1 What is scale?

In computer vision research, multi-scale analysis and texture analysis are two branches of study that have overlaps and both attract researchers' interests. In addition, there are applications in medical image analysis using multi-scale analysis and texture analysis. For example, in [26], authors use multi-scale analysis for blood vessel detection and drusen detection in retina fundus image analysis. In [27], texture features of digital mammogram images are extracted for breast tumor study.

The two terms *Scale* and *Texture* haven been used widely in theory and applications, however, neither *scale* nor *texture* has a uniform definition across the research community[28, 29]. It is important and natural that if a term needs to be studied quantitatively, the term should be based on a well-founded definition.

In general, *scale* is considered as a term for geometric measure of an image, and

texture is the description of geometric patterns within an image. For example, in [30], authors define "texture" as repetitive patterns within the image(though, may not be true for most of natural images). It is important to give a definition for these terms which are used repeatedly in the analysis, otherwise complicated features are generated from natural images, which cannot be interpreted properly. In our study, the focus is the term *scale*. Based on the established mathematical theory, we aim to establish the definition for scale that has better geometric interpretations.

How to define the term *Scale*?

Proper definition of a term can be related and decided for the purpose of a certain application. Considering the definition of scale, *First*, if *scale* is a term used to describe the coarseness of an image, it can be any parameter monotonically changing. For example, t, the time parameter used in the diffusion equation, is commonly treated as a scale parameter. Given image $f(x)$, the solution to the equation at time t is denoted as $\{u(x,t)\}_{t>0}$. However, in general, it is not easy to relate t to any geometric measure, especially for a nonlinear diffusion equation. It can only be said that the parameter t provides a general sense that, as t increases gradually(for example), the image is smoothed or the details within the image are gradually removed. *Second*, if *scale* is the term used in the definition of local scale, it would be natural to expect that this *scale* measure the height, width, length, etc. i.e. scale defined pixel-wisely should have a direct geometric meaning. The usage of the first interpretation(coarseness) can be seen in many examples of multi-scale or multi-resolution analysis[31], and the

second interpretation can be seen in the detection of *interesting* or *dominate* scale. For example, Lindeberg [6], Lowe [7], it is defined as the local extreme of a certain edge response functions. In order to study multi-scale analysis based on nonlinear diffusion, we begin with spaces based on linear diffusion.

3.2 Review: Linear Scale space

3.2.1 History of multi-scale analysis

The idea of using multiple resolution or multiple scales in image analysis is motivated by the underlying natural multi-level structures of images. The study of multi-scale analysis based on diffusion equations can be traced back to the work of Witkin, 1983 [31]. However, the idea of multi-scale analysis in image analysis is not a completely new idea. From the early days, the simplest, yet still challenging image analysis task is edge detection. This task partly motivated the usage of multi-scale analysis. Examples of these work, include 1) Rosenfeld and Thurston(1971) [32]. 2) Klinger (1971) [33]. The author used the idea of *quad tree*. The image is recursively divided into smaller regions with a ratio of 2. Given image with size $2^N \times 2^N, N \in Z$, and defines a measure V to describe the gray level variation of the subregions. Denote $f^0 = f$, if at stage i, the image variation $V(f^i)$ is larger than certain threshold v, the image f^i will be divided into smaller region into f^{i+1}. The process will ends when the overall variation converges below the predefined threshold. This method

can be treated as "merge and split" image segmentation task. The detailed example is explained in [33]. 3) In the work of Burt(1981) [34] and Crowley(1991)[35], they both proposed the *pyramid* representation of images. Instead of dividing the image into subregions, the pyramid representation of an image of size $2^N \times 2^N, N \in Z$ is based on certain method of sub-sampling and smoothing. Naturally, from level f^i to f^{i+1}, the image is represented at a coarser level. By representing the image in this way, the size of the image decreases exponentially. 4) Wavelet theory is also widely used in multi-resolution analysis, including the work of Mallat et.al.[36, 37, 38].

3.2.2 Linear multi-scale space

In 1983, Witkin proposed the multi-scale study based on gaussian smoothing in [31]. The basic idea is to embed the given image $f(x)$ into a family of solutions $\{u(t,x)\}_{t>0}$ to the heat diffusion equation, parameter t is considered as the descriptor for scale. Given $f(x), x \in \Omega \subseteq R^N$ as the original image, the solution to the heat equation is

$$u(t,x) = G(t,x) \star f(x), G(t,x) = \frac{1}{(4\pi t)^{N/2}} e^{-\sum x_i^2/4t}, t > 0 \tag{3.2.1}$$

When $N = 1$, the conventional representation of a gaussian kernel is

$$G(\sigma, x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \sigma^2 = 2t \tag{3.2.2}$$

The kernel size σ is directly related to the neighborhood size of each pixel x, and $\sqrt{2t}$ can be interpreted as the measure of neighborhood size. In order to construct multi-scale analysis in a systematic way, there is some research on axioms for multi-scale spaces. In the work of Alvarez[19], the multi-scale analysis based on the heat diffusion is proven to satisfy one axiom system.

Gaussian kernel, Derivatives and Scale

Figure 3.1 visualizes the property of gaussian kernel and its derivatives, and the analytical results are summarized in Table 3.1.

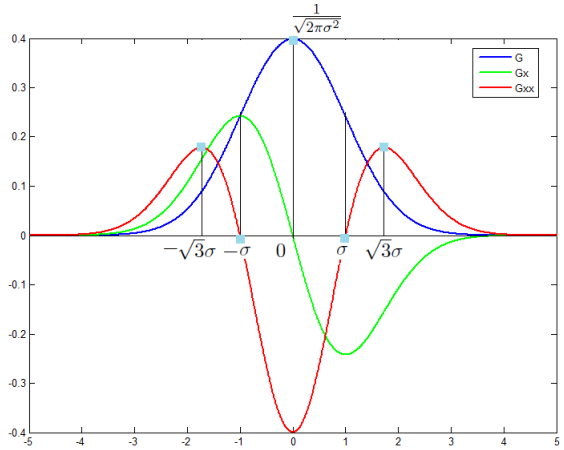


Figure 3.1: Gaussian, Derivatives and Local Extreme Points

Table 3.1: Gaussian Kernel and Derivatives

$f(x)$	Forumla	$\{x\|f'(x)=0\}$
$G(x,\sigma)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{\sigma^2}}$	0
$G'(x,\sigma)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\frac{-x}{\sigma^2}e^{-\frac{x^2}{\sigma^2}}$	$\pm\sigma$
$G''(x,\sigma)$	$\frac{-1}{\sqrt{2\pi\sigma^2}}(\frac{1}{\sigma^2}-\frac{x^2}{\sigma^4})e^{-\frac{x^2}{\sigma^2}}$	$\pm\sqrt{3}\sigma$

3.2.3 Scale Selection

Construction of multi-scale space enables the analysis of a given image at different scales, however what is more important is how to choose the proper scale for analysis. Questions may include if local scale should be chosen adaptively for each pixel and what criteria should be set up in order to select a meaningful scale. Compared to literatures on multi-scale representation, there is less work on the discussion of scale selection. One approach proposed by Schiele (1997) [39] is to use all information from a range of different scales. This approach hugely increases the requirement of data storage. Here, instead, we will mention the seminal work of Lindeberg(1993, 1994) [6, 40], Lindeberg suggested the following principle for the automatic scale selection:

> In the absence of other evidence, assume that a scale level, at which some (possibly non-linear) combination of normalized derivatives assumes a local maximum over scales, can be treated as reflecting a characteristic length of a corresponding structure in the data.

3.2.4 Feature Detection: SIFT

One of the most important applications of the idea of scale selection is the scale invariant feature transform(SIFT). It is a family of features that are invariant under the change of scale and rotation transformations. It was first proposed by David

Lowe in 1999 [41], and then it is extended in a more systematic way in 2004 [7]. Lowe's work provided a method to extract and describe distinctive and stable image features that can be used in object recognition, object matching etc. Steps in feature extraction can be divided into four main steps.

1, Scale-space extrema detection

In [41], the difference-of-Gaussian function(DOG) convolved with the image is used to detect the stable key-point locations in the scale space. For 2D image,

$$L(x, y, \sigma) = G(x, y, \sigma) \star f(x, y), G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3.2.3}$$

The idea of DOG is shown in the left plot of Figure 3.2. Local extrema of $D(x, y, \sigma)$

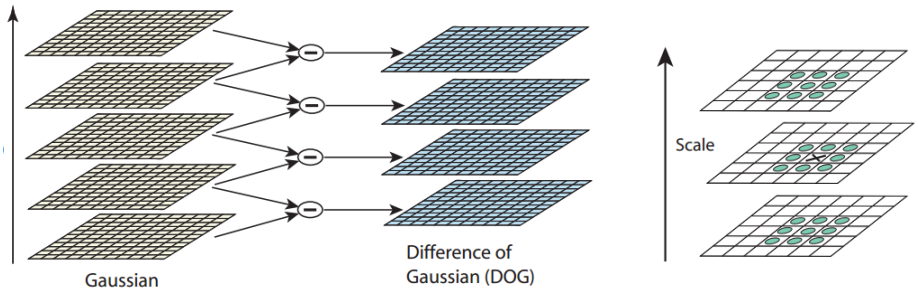


Figure 3.2: Left: Calculation of DOG; Right: Key-point Localization

are detected by comparing the eight neighborhood points in the current image and nine neighbors in the scale above and below. (x, y, σ) will be selected if $D(x, y, \sigma)$ is largest (or smallest) in 3D neighborhood shown in right plot of Figure 3.2

2, Key-point localization

Given the local extrema point (x, y, σ) in Step 1, the next step is to perform a detailed fit to the nearby data for location, scale, and ratio of principal curvatures. This step can reject points with low contrast, sensitive to noise etc.

$$G(v) = D((x, y, \sigma) + v) = D(x, y, \sigma) + \frac{\partial G}{\partial v} v + \frac{1}{2} v^T \frac{\partial^2 G}{\partial v^2} v$$

The local extrema of $G(v)$ is

$$\hat{v} = -\frac{\partial^2 G}{\partial v}^{-1} \frac{\partial G}{\partial v}$$

$\hat{v}$ is the offset with respect to (x, y, σ), if $\hat{v}$ is larger than 0.5 in any dimension, the sample point is changed and this step is performed again for the new sample point. In addition, the response $G(\hat{v})$ is used to reject the unstable extrema point. In [41], all extrema points with $|G(\hat{v})| < 0.03$ were discarded.

3, Orientation assignment

At each key point (x^*, y^*, σ^*) after Step 2, for each point (x, y) within the neighborhood of the key-point,

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1))/(L(x+1, y) - L(x-1, y))$$

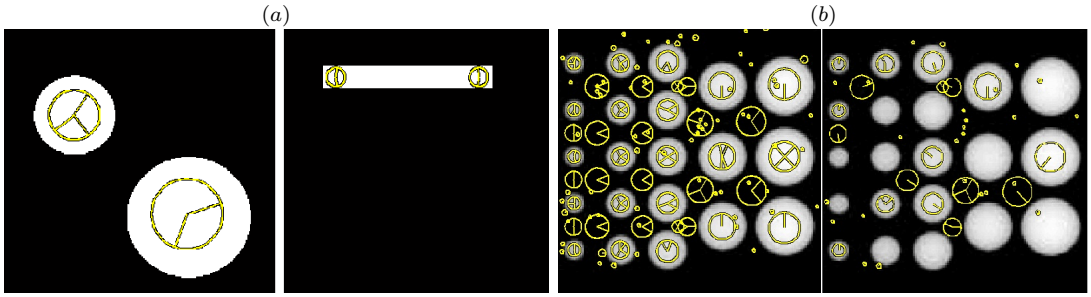
A histogram of $\theta(x, y)$ within this neighborhood(window size chosen to be 1.5σ) is weighted by the gradient magnitude $m(x, y)$. Peaks of this weighted histogram is regarded as the dominant direction. The peak within 80% of the highest peak of the histogram is used as the local orientation of the key-points.

4, Key-point descriptor

Key-point descriptor is composed of location (x, y), scale σ and local orientation θ_i.

Example

SIFT features have been used widely in image matching, feature analysis, since the focus of our study is the image structure detection, it is better to start with analysis on images containing simple shapes. In the following examples, we chose images with typical blob(close to disk shape) and vessel(i.e. elongated shape) structures from both synthetic and real clinical images. Center of the yellow circle in the image indicates the location of key-point (x, y) and the radius of the circle indicates the *scale* of the key-point is denoted as σ.



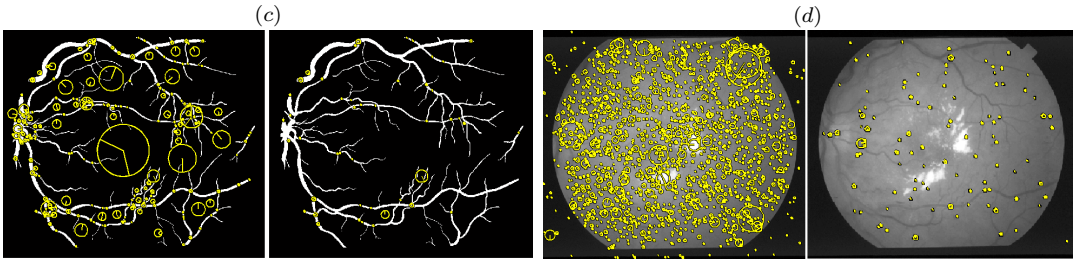


Figure 3.3: Examples of SIFT Detection: For (b), (c), (d), Left figure: All key-points; Right figure: Randomly selected key-points

Remark 3.2.1. Based on Figure (3.3), it can be observed that (1) SIFT features works best for blob detection, shown in (a). This can be partly explained by the fact that heat diffusion is an isotropic diffusion; (2) For elongated structures, only points at the two ends are detected, right figure of (a); In (c), junction points along the vessels are detected. The fact that large part of edge structures is not detected can be partly explained by the fact that heat diffusion is known to blur edge structures; (3) The detection of scale can represent the relative relationship between scales of structures shown in (a), (b), however in general, it is not absolutely correct; (4) In (b), besides the bright blobs, the dark regions between blobs are also detected as key-points. This can be explained by the fact that the blob detection was defined as regions brighter or darker than the surrounding regions[42]. Additionally, there are many small spots detected that may be due to the existence of oscillations within the image. These oscillations can be removed by different choices of threshold(e.g. some response function), however, in automatic image analysis, the choice of parameters poses an additional parameter-tuning problem; (5) In (d), a real medical image(from

public dataset STARE for retina image analysis), the existence of many key-points can pose a problem for additional key-point selection. The remarks we made here do not mean that SIFT features are not helpful, but that the theoretical basis of SIFT features can limit SIFT in some image analysis tasks, especially for elongated structures(i.e. vessel structures). It is a motivation for the intensive research on nonlinear diffusion equations in the image analysis.

Chapter 4

Total Variation model

4.1 Total Variation Flow Minimization

4.1.1 Total Variation and Bounded Variation

Images are usually considered in $L^p(\Omega)$ space, here $\Omega \subseteq R^n$, however $L^p(\Omega)$ spaces fail to capture the local oscillatory patterns within images. With the introduction of Sobolev space $W^{1,1}(\Omega)$, the irregularities of images are also taken into consideration. One limitation is that $W^{1,1}(\Omega)$ space does not allow edges within an image. In computer vision, edges are considered crucial clues for object pattern recognition. In between the $L^p(\Omega)$ and $W^{1,1}(\Omega)$ spaces, there is one function space that can balance the regularity characterization and accept edge structures at the same time. This space is called the Bounded Variation(BV) space.

Definition 4.1.1. Given $u \in L^1(\Omega), \Omega \subseteq R^n$,

$$\|u\|_{TV(\Omega)} : \int_\Omega |Du| = \sup\{\int_\Omega u\nabla \cdot (\vec{g})dx : \vec{g} \in C_0^1(\Omega; R^n), |\vec{g}(x)|_{l_2} \leq 1, \forall x \in \Omega\}$$

$$\|u\|_{BV(\Omega)} := \|u\|_{L_1(\Omega)} + \|u\|_{TV(\Omega)}$$

$$BV(\Omega) = \{u \in L^1(\Omega) : \|u\|_{BV(\Omega)} < \infty\}$$

Here, for convenience, $\|u\|_{TV(\Omega)}$ is also denoted as $\|Du\|(\Omega)$. The total variation norm $\|\cdot\|_{TV}$ is a semi-norm; $(BV(\Omega), \|\cdot\|_{BV(\Omega)})$ is a Banach space. The introduction of BV spaces in imaging processing is in the work of Rudin, Osher and Fatemi [20, 21].The relationship between BV space and the two function spaces mentioned before can be summarized as the following:

$$W^{1,1}(\Omega) \subseteq BV(\Omega) \subseteq L^1(\Omega)$$

4.1.2 ROF Model

The total variation minimization model appeared in the seminal work of Rudin-Osher-Fatemi [20]. The model proposed for the image denoising problem as the following: given an observed image $u_0(x, y) = u(x, y) + n(x, y), (x, y) \in \Omega,$ where $n(x, y)$ is the additive noise, and $u(x, y)$ is the original image that needs to be reconstructed. Different from previous researches [43, 44, 45] in which the least squares criteria (i.e. L_2 norm) was imposed, the authors in [20] proposed as an optimization problem mini-

mizing the total variation norm of the image. It has been suggested in [46] that the total variation norm(i.e. TV norm) is a better norm to describe an image than the L_2 norm, especially for images with edges as discontinuities. The optimization problem is proposed with constraints shown as follows

$$\begin{aligned} \min_{u \in BV(\Omega)} \quad & \int_\Omega \sqrt{u_x^2 + u_y^2} dxdy \\ \text{subject to} \quad & \int_\Omega u(x,y)dxdy = \int_\Omega u_0(x,y)dxdy, \\ & \int_\Omega \frac{1}{2}(u-u_0)^2 dxdy = \sigma^2, \sigma > 0. \end{aligned}$$

This model is proposed with the prior information that the noise $n(x, y)$ has zero mean and standard deviation σ. In [20], this optimization problem is solved using the gradient descent method.

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}}\right) \tag{4.1.1}$$

$$-\lambda(u - u_0),\ t > 0, (x, y) \in \Omega \tag{4.1.2}$$

$$u(x, y, 0) \text{ is given} \tag{4.1.3}$$

$$\frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \tag{4.1.4}$$

The model can also be reformulated as an equivalent unconstrained optimization problem, commonly knows as ROF Model:

$$\min_{u \in BV(\Omega)} \lambda \int_\Omega |Du| + \int_\Omega (f-u)^2 dx, \tag{4.1.5}$$

Here, $\lambda > 0$ is the weight between regularity and the fidelity term. The numerical experiments in [20] demonstrate the ROF model (4.1.5) can perform best in denoising very noisy images, especially yielding sharp edges in images. However, there is one important limitation of (4.1.5) that is the loss of contrast in solutions, even when the image $f(x)$ is noise-free. Detailed discussions on the limitations of the ROF model are shown in [47].

4.1.3 TV/L_1 Model

There have been many studies and variations of the ROF model (4.1.5), and TV/L_1 is the model with special interests. In this model, the total variation norm is used in the regularization term, and the fidelity term is replaced with L^1 norm, the problem is proposed as (see [47]):

$$\min_{u \in BV(\Omega)} \lambda \int_{\Omega} |Du| + |f - u|_{L^1(\Omega)} \tag{4.1.6}$$

here $f(x)$ is the image within region $\Omega \subseteq R^2$, Ω is a rectangle, $|u|_{TV(\Omega)}$ is the regularization term, $|f - u|_{L^1(\Omega)}$ is the fidelity term, λ is the weight between regularity and fidelity term. As λ increases, the minimizer of Problem (4.1.6) is a more regularized version of the original image, and smaller feature including noise that has high oscillation is gradually removed.

Considering a 2D image $f(x) = k\mathbf{1}_{\mathbf{B_r(0)}}(\mathbf{x}), \mathbf{k} > \mathbf{0}$. It is shown in [47] that the

minimizer to Problem (4.1.6) is:

$$u(x,\lambda) = \begin{cases} k\mathbf{1}_{\mathbf{B_r(0)}}(\mathbf{x}) & \lambda < 2r \\ \{a\mathbf{1}_{\mathbf{B_r(0)}}(\mathbf{x})\} & \lambda = 2r, 0 < a < k, \\ 0 & \lambda > 2r \end{cases} \tag{4.1.7}$$

In general, Problem (4.1.6) cannot be solved analytically and this simple example demonstrates several important properties of TV/L_1 model. *First*, the critical point $\lambda^* = 2r$ at which the solution $u(x,\lambda)$ changes from 0 to f depends on the radius of the disk (i.e. the scale of geometric structure) only, and does not dependent on the intensity level of the disk. It indicates that TV/L_1 model can maintain and reveal the inherent geometric properties of the image. In image analysis, it is desired that image structures be identified without any affects of the transformation of image intensity values. *Second*, considering the solution to TV/L_1, except at the point $\lambda = 2r$, the solution is either 0 or f, and the general image contrast remains constant. On the contrary, one of the known limitations of the ROF model is that the image contrast will be decreased even for images without noise.

The effect of λ for ROF model and TV/L_1 models are very different. For the ROF model, objects of distinct scale with edges in the image gradually lose contrast and merge with their neighbors as the parameter λ increases. In TV/L_1 model, objects can maintain their contrast with respect to their neighbors and their boundaries are gradually smoothed out. This process continues until a critical value of λ is reached,

a value related to the scale of certain structure, and at this point the object suddenly merges with a neighboring one. Below one example of the TV/L_1 solution as a function of λ is shown. In Figure 4.1, *fig1* and *fig2* are images with the same structures and different contrasts. It is shown that the critical value of λ (at which a sudden change of intensity occurs) is not affected by contrast and positively correlated with *scale* of structures. Properties of TV/L_1 motivate us to use this model for fundus image analysis shown in Chapter 7.

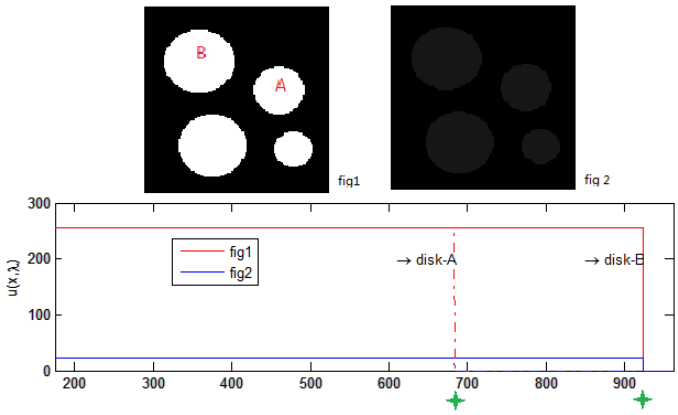


Figure 4.1: Effects of λ in TV/L_1 model with different image contrast. The x-axis is index i of λ_i, y-axis is the image intensity when $\lambda = \lambda_i$ at the centers A, B of the two disks.

4.2 Total Variation Flow Equation

Total variation flow appears when solving the unconstrained total variation minimization problem:

$$\min_{u \in BV(\Omega)} \int_{\Omega} |Du| \tag{4.2.1}$$

The total variation flow equation with the Neumann boundary condition is formulated as the following:

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (\frac{\nabla u}{|\nabla u|}) & (t,x) \in (0,\infty) \times \Omega \\ \frac{\partial u}{\partial n} = 0 & (t,x) \in (0,\infty) \times \partial\Omega \\ u(0,x) = f(x) & x \in \Omega \end{cases} \tag{4.2.2}$$

where Ω is a bounded set in R^N with Lipschitz boundary $\partial\Omega$ and $f \in L^1(\Omega)$. This partial differential equation is used in the steepest descent method to minimize the total variation (4.2.1) of the image. The Neumann boundary condition means that the image is reflected across the boundary, which is natural in image processing.

To use TV flow equation (4.2.2) for multi-scale analysis, it is important to consider (1) the well-posedness of Equation 4.2.2, i.e. existence, stability and uniqueness of the solution; (2) the asymptotic properties of the solution, i.e. the behavior of $u(t,x)$ as $t \to \infty$ considering the families of solutions of $\{u(t,x)\}_{t>0}$.

4.2.1 Well-posedness of the PDE

In the study, we are interested in constructing a multi-scale space $\{u(t,x)\}_{t>0}$ based on solutions to Problem (4.2.2). The details of the well established theory on the well posedness of the PDE problem are discussed in [48]. Some of the terms that are used in order to introduce the results are listed as below.

Definition 4.2.1. (Complete accretive operator) Let $M(\Omega)$ be the space of measur-

able functions in Ω. Given $u, v \in M(\Omega)$, we denote that

$$u \ll v \Leftrightarrow \int_\Omega j(u)dx \leq \int_\Omega j(v)dx, \forall j \in J_0$$

where $J_0 = \{j : R \to [0, \infty], j(0) = 0, \text{convex and lower semicontinous}\}$. Let A be an operator in $M(\Omega)$ $(A \subseteq M(\Omega) \times M(\Omega)), A$ is *completely accretive* if $u - \hat{u} \ll u - \hat{u} + \lambda(v - \hat{v}), \forall \lambda > 0, (u, v), (\hat{u}, \hat{v}) \in A$. A completely accretive operator in $L^1(\Omega)$ is said to be *m-completely accretive* if $R(I + \lambda A) = L^1(\Omega), \forall \lambda > 0$

Definition 4.2.2. (Abstract Cauchy problem) Given initial condition $f(x) : [0, T] \to X$, the abstract Cauchy problem is

$$\frac{du}{dt} + \mathcal{A}u \ni 0, u(0, x) = f(x) \tag{4.2.3}$$

If $u(t, x) \in C([0, T]; X) \cap W^{1,1}_{loc}((0, T); X)$and satisfies (4.2.2), $u(t, x)$ is called a *strong solution* of (4.2.2) ; If $u(t, x) \in C([0, T]; X)$ and for each $\epsilon > 0$, there is a discretization of $[0, T]$ denoted as $0 \leq t_0 < \cdots < t_N \leq T$, a solution of the discretization v satisfying $v : [0, T] \to X, v(t_i) = v_i, v(t) = v_i$ for $t \in (t_{i-1}, t_i), i = 1, ..., N, (v_i - v_{i-1})/(t_i - t_{i-1}) + Av_i \ni f_i, i = 1, \cdots, N$. If such v satisfies $\|u - v\| \leq \epsilon, 0 \leq t \leq T$, $u(t, x)$ is called a *mild solution* of the abstract problem (4.2.2).

Semigroup solution for abstract cauchy problem

To prove the existence of solution to Problem (4.2.2), we consider an abstract Cauchy problem, where $-\nabla \cdot (\frac{\nabla u}{|\nabla u|})$ is associated with a completely accretive operator $\mathcal{A}$.

$$X(\Omega)_1 = \{z \in L^\infty(\Omega, R^2) : div(z) \in L^1(\Omega)\}$$

Denote $T_k(x) = (k - (k-x)^+)sgn(x), k \geq 0, x \in R$, $L^1_\omega(0,T; BV(\Omega) = \{\omega : \forall \phi \in BV(\Omega)^*, t \in [0,T] \to < \omega(t), \phi >$ is measurable and $\int_0^T \|\omega(t)\| < \infty\}$

Definition 4.2.3. $(u,v) \in \mathcal{A}$ if and only if $u, v \in L^1(\Omega), T_k(u) \in BV(\Omega), \forall k > 0$ and there exists $z \in X(\Omega)_1, \|z\|_\infty \leq 1, v = -div(z)$ *in* $\mathcal{D}'(\Omega)$, such that

$$\int_\Omega (w - T_k(u))v dx \leq \int_\Omega z \cdot \nabla w dx - \int_\Omega \|DT_k(u)\|,$$

for all $w \in W^{1,1}(\Omega) \cap L^\infty(\Omega), k > 0$

Theorem 4.2.4. *The operator* $\mathcal{A}$ *is a m-completely accretive in* $L^1(\Omega)$ *with dense domain. Given* $f \in L^1(\Omega)$*, the semigroup solution* $u(t) = e^{-t\mathcal{A}}f$ *is a strong solution of the Problem* (4.2.2)*.*

Remark 4.2.5. The existence of the solution is based on the Crandall-Liggett's theorem [48], which states that given $\mathcal{A}$ a m-completely accretive in $L^1(\Omega)$, $\mathcal{A}$ generates a contraction semigroup in $L^1\Omega$ given by an exponential formula

$$e^{-t\mathcal{A}}f = \lim_{n\to\infty} (I + \frac{t}{n}\mathcal{A})^{-n} f, \forall f \in L^1(\Omega)$$

$u(t) = e^{-t\mathcal{A}}$ is the mild solution of Problem (4.2.2). In addition, denote by $S(t) = e^{-t\mathcal{A}}$, due to the homogeneity of the operator $\mathcal{A}$, it can be proved that given $t > 0$, for a sequence $t_n \downarrow 0$, the sequence of functions

$$\{\frac{S(t+t_n)f - S(t)f}{t_n}\}_{n=1}^{\infty}$$

is weakly convergent in $L^1(\Omega)$, implying that $S(t)$ is the strong solution to Problem (4.2.2). In addition, it can be proved that as a consequence that

$$|u'(t)| \leq \frac{2}{t}|f|$$

Existence of weak solution to Problem 4.2.2

Definition 4.2.6. A measurable function $u : (0.T) \times \Omega \to R$ is a weak solution of (4.2.2) in $(0,T) \times \Omega$, if $u \in C([0,T], L^1(\Omega)) \cap W^{1,1}_{loc}(0,T; L^1(\Omega)), T_k(u) \in L^1_w(0,T; BV(\Omega))$for all $k > 0$, and there exists $z \in L^\infty((0,T) \times \Omega)$with $\|z\|_\infty \leq 1, u_t = div(z)$ in $D'((0,T) \times \Omega$ such that

$$\int_\Omega (T_k(u(t)) - w)u_t(t)dx \leq \int_\Omega z(t) \cdot \nabla w dx - \int_\Omega |DT_k(u(t))| \tag{4.2.4}$$

for every $w \in W^{1,1}(\Omega) \cap L^\infty(\Omega)$ and a.e. on $[0,T]$.

Theorem 4.2.7. *Let $f \in L^1(\Omega)$. Then there exists a unique weak solution of* (4.2.2) *in $(0,T) \times \Omega$ for every $T > 0$ such that $u(0,x) = f(x)$. Moreover, if $u(t), \hat{u}(t)$are weak*

solutions corresponding to initial data $f_0, \hat{f}_0$, *respectively, then for all* $t \geq 0$

$$\|(u-\hat{u})^+\|_1 \leq \|(f-\hat{f})^+\|_1, \|u-\hat{u}\|_1 \leq \|f-\hat{f}\|_1 \tag{4.2.5}$$

Equivalence between weak solution and strong solution

Based on the existence of strong solution to the abstract Cauchy problem, we will prove that the weak solution to (4.2.2) is a strong solution to the abstract problem.

Proof. If u is the weak solution of (4.2.2), in $(0,T)\times\Omega$, , then according to the definition, we have t $(1), u \in C([0,T], L^1(\Omega)) \cap W^{1,1}_{loc}((0,T), L^1(\Omega))$

$(2), T_k(u) \in L^1([0,T], BV(\Omega)), \forall k > 0$

$(3) \exists z \in L^\infty((0,T)\times\Omega), \|z\|_\infty \leq 1, u_t = div(z), \mathcal{D}'((0,T)\times\Omega)$

$$\int_\Omega (T_k(u)-w)u_t dx \leq \int_\Omega z\cdot\nabla w dx - \int_\Omega \|DT_k(u)\|$$

$$\forall w \in W^{1,1}(\Omega)\cap L^\infty(\Omega), a.e.[0,T]$$

We can have that a.e. $t \in [0,T], u_t(t) = div(z(t)), in\ \mathcal{D}'(\Omega)$. a.e. $t \in [0,T], z(t) \in X(\Omega)_1, (u(t), -u_t(t)) \in \mathcal{A}, i.e. u'(t) + \mathcal{A}u(t) \ni 0, a.e.[0,T]$. We then have it is actually a strong solution in the semigroup sense. The uniqueness can be proved by the uniqueness of the strong solution and the stability can be proved by the complete accretivity of operator $\mathcal{A}$. □

In addition, we have the similar conclusions when the initial condition $f(x)$ is in

$L^2(\Omega)$.

Definition 4.2.8. A function $u(t,x) \in C([0,T];L^2(\Omega))$ is a strong solution of (4.2.2) if $u \in W^{1,2}_{loc}(0,T;L^2(\Omega)) \cap L^1_\omega((0,T);BV(\Omega))$, $u(0) = f(x)$, and there exists $z \in L^\infty((0,T)\times\Omega;R^n)$ such that $\|z\|_\infty \leq 1$, $\forall t \in [0,T], \partial z/\partial n = 0, x \in \partial\Omega$ satisfying $u_t = div(z)$ in $D'((0,T)\times\Omega$ and

$$\int_\Omega (u(t)-w)u_t(t)dx = \int_\Omega (z(t),Dw)dx - \int_\Omega \|Du\| \tag{4.2.6}$$

for every $w \in L^2(\Omega)\cap BV(\Omega)$ and a.e. on $[0,T]$.

Theorem 4.2.9. *Let $f \in L^2(\Omega)$. There exists a unique strong solution $u(t,x)$ of (4.2.2) in $[0,T]\times\Omega$ for every $T>0$. Moreover, if u and v are the strong solutions of (4.2.2) corresponding to the initial conditions $u_0, v_0 \in L^2(\Omega)$, then*

$$\|u(t)-v(t)\|_2 \leq \|u_0-v_0\|_2, \forall t > 0$$

4.2.2 Asymptotic behavior of solutions

As an image diffusion process used for multi-scale analysis, it is also important to study the asymptotic behaviors of solutions as time approaches infinity.

Theorem 4.2.10. *Given $f \in L^2(\Omega), \Omega \subseteq R^2, u(t,x)$ is the unique strong solution of*

(4.2.2), *there exists a finite time* T_0 *such that*

$$u(t,x) = (f)_\Omega = \frac{1}{L^N(\Omega)} \int_\Omega f(x)dx, \forall t \geq T_0.$$

Proof. The proof is based on energy method. Given $u(t,x)$ as the strong solution to (4.2.2), there exists $z \in L^\infty, \|z\|_\infty \leq 1, s.t.$

$$\int_\Omega (u(t) - w)u_t dx = \int_\Omega (z(t), Dw) - \int_\Omega \|Du(t)\|, \forall w \in BV(\Omega) \cap L^\infty(\Omega)$$

If we replace w with constant function $(f)_\Omega$,

$$\int_\Omega (u - (f)_\Omega)u_t dx = -\int_\Omega \|Du\|$$

Using Sobolev-Poincare inequality which states that

$$\|u - (u)_\Omega\| \leq C\|Du\|(\Omega), \forall u \in BV(\Omega), 1 \leq p \leq \frac{N}{N-1}$$

Let $N = 2$, it gives that $\|u - (f)_\Omega\|_2 \leq C(\Omega) \int_\Omega \|Du\|$. It is then derived that

$$\frac{1}{2}\frac{d}{dt}\int_\Omega (u - (f)_\Omega)^2 dx + \frac{1}{C}\|u - (f)_\Omega\|_2 \leq 0$$

Denote $M(t) = \|u - (f)_\Omega\|_2^2 \geq 0$, we have that $\frac{1}{2}M'(t) + \frac{1}{C}M(t)^{1/2} \leq 0$. It implies that $\frac{d}{dt}(M(t)^{1/2}) \leq -\frac{1}{c}$ i.e. $M(t)^{1/2} - M(0)^{1/2} \leq -\frac{1}{C}t$.

$\therefore M(t) = 0, \forall t \geq C \cdot (M(0)^{1/2}). \therefore \exists T_0(f, \Omega), u(t, x) = (f)_\Omega, a.e.\Omega, \forall t \geq T_0$ $\square$

It can be observed from the proof that T_0 depends on the region Ω and the image $f(x), x \in \Omega$. Theorem (4.2.10) show that the solution to Problem (4.2.2) will converge to a constant function equals to the mean within the region Ω after T_0. For multi-scale analysis, Theorem (4.2.10) implies that the proper choice of T_{max} to construct the multi-scale space $\{u(x,t)\}_{0<t<T_{max}}$ is critical. If T_{max} is not large enough, then the multi-scale profiles construed may not contain enough information. However if $T_{max} >> T_0$, the multi-scale profiles of the image may be dominated by $\{u(x,t\}_{t>T_0}$, which also depends on how the multi-scale analysis is implemented.

4.3 Applications

The main applications of total variation minimization model is image restoration and denoising, for example in the seminal work of Rudin [20, 49]. It is shown that for image denoising using total variation as regularization term performs better. In the work of Nikolova [50], where fidelity term is the L_1 norm, the TV/L_1 decomposition model is used to remove impulsive noise (i.e. salt and pepper noise) and identify the outliers of 1D signals.

In addition, these models have also been used in texture analysis. In the work of Chen [51, 52], the model is used to remove nonuniform illumination in face images; In [53], texture features from natural images are extracted and different total variation-based texture extraction are also compared; In [15], a definition of region-based local scale

is proposed based on total variation flow diffusion, and the local scale feature was used in texture discrimination for image segmentation. In [54], the inhomogeneous background from cDNA microarray images are removed based on the TV/L_1 model. Considering the interpretation of λ as a weight between the fidelity and the regularity term, where λ can be related to the local scale of image structures, the multi-scale decomposition of images was proposed in previous work. For example, in [55], method for 2-dimensional/3-dimensional images for multi-scale magnetic resonance image registration is proposed. In [56], TV/L_1 model is used in multi-scale feature selection. These interesting applications motivate us to investigate more potential applications of total variation based models in image analysis problems.

4.4 Discussion

In general, due to the nonlinearity of total variation minimization model, it is not easy to establish a meaningful analytic theory to describe effects of total variation regularization. Besides, there is relatively limited literature on why and how TV regularization works. The existing work is generally abstract and theoretical, not easy to be applied directly. Considering the analysis of images based on TV minimization, it is important to know that the limited understanding of TV regularization can lead to unclear numerical results.

Chapter 5

Scale space: TV Flow (I)

5.1 Related Work

5.1.1 Brox-Weickert Scale

The study of total variation flow in multi-scale analysis is not completely new. In the work of Brox and Weickert [15], TV flow equation is used for the region based local scale measure in texture discrimination. Given an image $f(x, y), (x, y) \in \Omega \subseteq R^2$, f evolves according to the TV flow equation

$$\frac{\partial u}{\partial t} = \nabla \cdot (\frac{\nabla u}{|\nabla u|}), u(t = 0) = f(x, y)$$

In [15], authors claimed that the evolution of u will bit-by-bit lead to larger regions in the image, as TV flow diffusion tends to yield piece-wise constant, segmentation-like

results. They proposed a way to measure the size of the regions inside which the pixels have the same gray-level value in the following way:

$$\frac{1}{m} = \frac{1}{2}\frac{\int_0^T |\partial_t u| dt}{\int_0^T (1 - \delta_{\partial_t u,0}) dt} \tag{5.1.1}$$

where m is the number of pixels within a certain region(i.e. the local scale measure), $\delta_{a,b} = 1, \text{if } a = b$, and 0 otherwise. The choice of T should be careful, otherwise interesting image structures maybe removed after certain time of diffusion(Theorem (4.2.10)). In [15], the authors claimed that T_{max} the finite extinction time satisfying $T_{max} \leq \frac{1}{4}\text{size}(f) \cdot c_{max}$, c_{max} is the maximum contrast within the image f. With regarding to the application of the local scale measure, the authors claimed that "the local scale measure can be combined with texture features, and the texture features and the image intensity can distinguish even very similar textures", i.e. the local scale measure can have a potential applications in texture discrimination. To understand the meaning of region-based local scale measure, one example is shown in Figure 5.1,

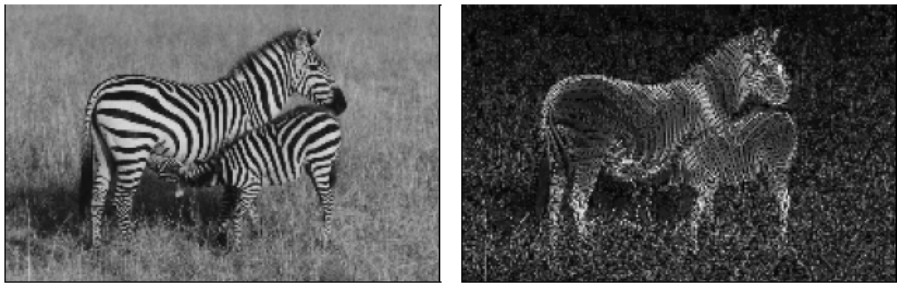


Figure 5.1: Left: Original Image; Right: Region based local scale measure. Dark regions correspond to large scales, and bright regions to fine scales

Remark 5.1.1. Though shown with good performance on texture discriminations and image segmentation,there are a few remarks we want to point out here.

1, *Scale*: The scale defined in [15] for 2D images is the size of the piecewise constant region, i.e. it only counts the number of pixels within a certain region, without any geometric information of the region. In image segmentation, the geometric property of a region is important. When defining scale of a region, the definition for example of the width along a given direction is more important than just the area of the region. We are more interested to extend the *scale* definition.

2, *Formula*: As a region-based local scale measure, m equals to

$$m = \frac{2\int_0^T (1 - \delta_{\partial_t u,0})dt}{\int_0^T |\partial_t u|dt} \tag{5.1.2}$$

The scale $m(P)$ is defined for each pixel P, it can be confusing that m describes the region-based property of an image. *First,* m is independent of the diffusion time t, and dependent on the pixel location $P = (x, y)$. Considering the description of region size, what region size this definition captures is not very clear. *Second*, Suppose m captures the region property of the initial image, and there is a region Ω_1 with constant intensity, $P_1, P_2 \in \Omega_1$, it is expected that $m(P_1) = m(P_2)$ according to the definition of this local scale measure. However, if P_1 is at the center of Ω_1 and P_2 is along the edge of Ω_2, according to the property of TV flow diffusion, it's expected that $\partial_t u(t, P_1), \partial_t u(t, P_2)$ is different as t changes. It contradicts the idea that m captures a region based property; *Third*, Equation (5.2.6) can be ill-posed, e.g. for

(x, y) $s.t.$ $\partial_t u(x, y) = 0$. It poses a problem for scale definition at such points.

3, *Parameter*: In real applications, the robustness of a model can be highly dependent on parameters used in the model. The definition of local scale measure of a region depends on parameter including $T(\text{size}(f)\ , c_{max})$ given the assumption that the extinction time T_{max} is decided by the image f completely. In addition, the authors claimed that choosing $T_{max} = 20$ is robust without further validation. Considering that T depends on the number of pixels within the region, if structure size or image size varies, it is expected that the choice of T should be changed. In fact, T_{max} is closely related to the scale of structures, which should be hard to choose uniformly across different images.

5.1.2 Strong-Chan Scale

In the work of Chan[57], authors studied the scale-dependent properties of total variation regularization. Authors claimed that the change in image intensity due to TV regularization is inversely proportional to the scale of the image feature. To be more specific, they gave a definition of scale based on the variational problem

$$min_u \frac{1}{2}\|u - f\|_{L_2}^2 + \alpha\|u\|_{TV} \tag{5.1.3}$$

where f is the original image, α is the a weight for fidelity and regularization term.

Definition 5.1.2. (Scale of piecewise constant image feature [57])

Let Ω represent a piecewise constant image feature, the scale of the region Ω is defined

as $|\Omega|/|\partial\Omega|$, where $|\Omega|$ is the area of Ω in R^2(or: volume in R^3), $|\partial\Omega|$ is the boundary length in R^2(or: surface area in R^3).

Remark 5.1.3. In [57], there are some remarks on this definition: *First*, if Ω is a circle with radius r in R^2, scale $= \frac{\pi r^2}{2\pi r} \propto r$; *Second*, the change in intensity δ can be rewritten as $\delta = \frac{\alpha}{\text{scale}}$, indicating the property of TV regularization: TV regularization causes smaller-scaled features(such as noise) to be partially or entirely removed(i.e. large δ while larger-scaled features are relatively unaffected(i.e. small δ); *Third*, if the change of image intensity can be measured locally, the local scale can be found $\text{scale}(\vec{x}) = \frac{\alpha}{\delta(\vec{x})}$. It is shown in [58], $\text{scale}(\vec{x})$ can be used to construct an automatic scale recognition scheme.

Remark 5.1.4. Based on the focus of our study, we will make some additional remarks on Strong-Chan scale. *First*, as shown above, the authors proposed two definitions of scale: δ(piecewise constant region), and $\delta(x)$(pixel-wise local measure). For the region-based scale δ, it is shown in the example that the scale δ is proportional to the radius of the disk r. However, in general, regions or structures are not radially symmetric, this definition of scale may not be able to capture the property of a region completely. For example, if the region Ω is a rectangle with size length $a, b, 0 < a \leq b$,

$$scale(\Omega) = \frac{ab}{2(a+b)} = \frac{a}{2(\frac{a}{b}+1)} \sim \begin{cases} \frac{a}{2} & \text{if } \frac{a}{b} \sim 0 \\ \frac{a}{4} & \text{if } \frac{a}{b} \sim 1 \end{cases}$$

The scale of the rectangle tends to capture the smallest scale over the two directions.

However, if $0 < \frac{a}{b} < 1$, it can be difficult to have a direct interpretation as to the region-based definition of scale. *Second*, for the pixel-wise based measure of scale(x), considering $P_1 \in \Omega_1, P_2 \in \Omega_2, \Omega_1, \Omega_2$ are two piecewise constant regions with the same shape(for example two circles with the same radius), the image intensities at P_1, P_2 satisfy that $f(P_1) \neq f(P_2)$. As a result, $\delta(P_1) \neq \delta(P_2)$, thus scale$(P_1) \neq$ scale(P_2). It is shown that the local definition of $scale(x)$ is dependent on image intensity $f(x)$ combined with the local geometric property, contradicting the original intention for the definition that depends solely on the geometric property of local structures. *Third*, last but not the least important, either the region Ω or the change of intensity $\delta(x)$ depends on the time point t in the evolution $u(t, x)$ or the weighting coefficient α for optimization problem $u(\alpha, x)$. How to decide the parameter can be important and difficult for automatic image analysis. If t or α is chosen too large, the important image structures can be removed; if t or α is too small, even the small noisy structures may remain in the images. Finding the optimal parameter poses a potential question for this local scale study.

5.2 Local Scale: TV Flow Diffusion

Given the previous analysis, TV Flow diffusion model still has several interesting and useful properties. As a nonlinear diffusion model, compared to heat diffusion, it can preserve important image features, including edges, corners, etc. The diffusion process based on Equation (5.2.6) tends to yield a piecewise constant, segmentation-like

images. Instead of extracting the region-based scale measure, we are more interested to extend the idea of local scale selection used in linear diffusion model(Section 3.2.3), and study how this information can be applied in the real image analysis.

Given an image $f(x) \in L^2(\Omega)$, Problem (4.2.2) has unique solution $u(t,x) \in BV(\Omega)$, and particularly, for almost all t,

$$\left|\frac{\partial u}{\partial t}(t,x)\right| \leq \frac{2}{t}|f(x)|, a.e.\ x$$

Similarly to DoGs used in heat diffusion based scale selection(see Equation (3.2.3)), we define the response function $Sf(t,x) : (0,\infty) \times \Omega \to R$ in the following way:

$$Sf(t,x) = t\frac{\partial u}{\partial t}(t,x) = t\nabla \cdot (\frac{\nabla u}{|\nabla u|}) \tag{5.2.1}$$

$$\|Sf(t)\|_{L^\infty} \leq \|f\|_{L^\infty}, \text{ for all } t > 0$$

Following [59], we give the definition of local scales based on TV Flow Diffusion.

Definition 5.2.1 (Local Scales)**.** For each $f \in L^2(\Omega) \cap L^\infty(\Omega)$, let $Sf(t,x)$ be defined as in (5.2.1). The local scales of f at $x \in \Omega$ is defined as

$$T_f(x) = \{t \geq 0 : |Sf(t,x)| \text{ is a local maximum}\}$$

For each $\beta > 0$, we say $t \in T_f(x)$ is β-visible if $|Sf(t,x)| > \beta$.

Example 5.2.2. Let $f(x) = k\chi_{B_r(0)}$ for some $r > 0$ and $k > 0$. From [48], we have the

solution to Equation (4.2.2) is

$$u(t,x) = \begin{cases} (k - \frac{2}{r}t)\chi_{B_r}(x) & \text{if } t < t^*, \\ 0 & \text{if } t \geq t^*. \end{cases}$$

$$Sf(t,x) = t\frac{\partial u}{\partial t}(t,x) = \begin{cases} t\left[-\frac{2}{r}\chi_{B_r}\right] & \text{if } t < t^*, \\ 0 & \text{if } t \geq t^*. \end{cases}$$

Here, $t^* = k\frac{r}{2}$. This solution shows that for all $x \in B_r(0)$, $|Sf(t,x)|$ has one local maxima at $t = t^*$, and $|Sf(t,x)| = 0$ for all $x \in R^2 \setminus B_r(0)$. Thus, the local scale of f at $x \in B_r(0)$ is proportional to both the radius r and intensity value k. $|Sf(t^*,x)| = k, x \in B_r(0)$, implying that the visibility of local scales of f at $x \in B_r(0)$ is k.

Denote $d_\delta f(x) = f(\delta x)$, $\delta > 0$, be the dilating operator. When defining local scale, an important requirement is that the local scale satisfies the invariance of the number of local scales under dilation i.e. the dilation transformation should not introduce new local scales and it is called dilation consistency property:

Definition 5.2.3 (Dilation Consistency Property). Denote by $T_f(x)$ the set of local scales of f at x. We say the local scales $T_f(x)$ have the Dilation Consistency Property if

$$T_{d_\delta f}(x) = \{\delta^s t : t \in T_f(\delta x)\}, \text{ for some } s \in R. \tag{5.2.2}$$

In order to prove the dilation consistency property, we need to firstly prove a property of the solution $u(t,x)$ to Equation 4.2.2.

Proposition 5.2.4. *Let $u(t) \in BV(\Omega)$ be the solution of (4.2.2) with the initial condition $u(0) = f$. Then $d_\delta u(\delta t)$ is the solution of (4.2.2) with $d_\delta f$ being the initial condition. Here we denote $d_\delta u(\delta t, x) = u(\delta t, \delta x)$.*

Proof. Denote the domain $\Omega_\delta = \{x : x \in \Omega_\delta \Leftrightarrow \delta x \in \Omega \subseteq R^2\}$. Given the initial condition $d_\delta f$, we will prove the solution to the following problem is $v(t,x) = u(\delta t, \delta x)$, where $u(t,x)$ is the solution to Problem (4.2.2)

$$\begin{cases} \frac{\partial v}{\partial t} = \nabla \cdot (\frac{\nabla v}{|\nabla v|}) & (t,x) \in (0,\infty) \times \Omega_\delta \\ \frac{\partial v}{\partial n} = 0 & (t,x) \in (0,\infty) \times \partial\Omega_\delta \\ v(0,x) = f(\delta x) & x \in \Omega_\delta \end{cases} \tag{5.2.3}$$

It is not difficult to prove that $v(t,x)$ satisfies the initial condition. In addition, $\delta > 0, \forall\, T > 0, T_\delta = T/\delta$, assume $u(t,x)$ is the solution to Equation (4.2.2), then with the dilation transformation, $v(t,x) \in C([0,T_\delta], L^1(\Omega_\delta)) \cap W^{1,1}_{loc}((0,T_\delta); L^1(\Omega_\delta)), T_k(u) \in L^1(0,T_\delta; BV(\Omega_\delta)), \forall k > 0$. Next, we need to prove the existence of $z \in L^\infty((0,T_\delta) \times \Omega_\delta)$ satisfying (4.2.4). For $u(t,x)$, there exists such a function z satisfying (4.2.4), define $z_1(t,x) = z(\delta t, \delta x), z_1(t,x) \in L^\infty((0,T_\delta) \times \Omega_\delta)$, $\forall\, \omega_1(x) \in W^{1,1}(\Omega_\delta) \cap L^\infty(\Omega_\delta), w(x) := w_1(x/\delta), \forall x \in \Omega$, denote $\delta t := s, \delta x := y$

$$\begin{aligned} \int_{\Omega_\delta} (T_k(v(t,x)) - \omega_1(x)) v_t dx &= \int_{\Omega_\delta} (T_k(u(\delta t, \delta x)) - w(\delta x)\partial_t u(\delta t, \delta x)) dx \\ &= \int_\Omega (T_k(u(s,y)) - w(y))\partial_s u(s,y) \frac{1}{\delta} dy \end{aligned}$$

$$\int_{\Omega_\delta} z_1(t,x)\nabla\omega_1(x) = \int_{\Omega_\delta} z(\delta t,\delta x)\nabla\omega(\delta x)dx = \int_{\Omega} z(\delta t,y)\nabla\omega(y)\frac{1}{\delta}dy$$

$$\begin{aligned}\|DT_k(v(t))\| =& \sup\{\int_{\Omega_\delta} T_k(v(t,x))\nabla\phi dx : \phi \in C_0^\infty(\Omega_\delta, R^n), |\phi(x)| \le 1, x \in \Omega_\delta\}\\ =& \sup\{\int_{\Omega} T_k(u(s,y))\nabla\phi\frac{1}{\delta}dy : \phi \in C_0^\infty(\Omega, R^n), |\phi(x)| \le 1, x \in \Omega_\delta\}\\ =& \frac{1}{\delta}\|DT_k(u)\|\end{aligned}$$

$\therefore \forall\omega_1(x) \in W^{1,1}(\Omega_\delta) \cap L^\infty(\Omega_\delta)$

$\int_{\Omega_\delta}(T_k(v(t,x)) - \omega_1(x))v_t dx \le \int_{\Omega_\delta} z_1(t,x)\nabla\omega_1(x) + \|DT_k(v(t))\|$, In summary, the solution $u(\delta t,\delta x)$ is the weak solution of (4.2.2), and according to Theorem (4.2.4) i.e. the uniqueness of solution, it is proven that $v(t,x) = u(\delta t,\delta x)$ is the solution of (4.2.2) with initial condition $f(\delta x)$. □

Proposition 5.2.5. *Let $T_f(x)$ be the local scales of f at x. Then, for all $\delta > 0$,*

$$T_{d_\delta f}(x) = \{\delta^{-1}t : t \in T_f(\delta x)\}. \tag{5.2.4}$$

Proof. Let $u(t)$ and $v(t)$ be solutions to (4.2.2) with initial conditions f and $d_\delta f$ respectively. Using Proposition 5.2.4, we have $v(t,x) = u(\delta t,\delta x)$, and

$$S(d_\delta f)(t,x) = t\frac{\partial}{\partial t}(v(t,x)) = t\frac{\partial}{\partial t}(u(\delta t,\delta x)) = s\frac{\partial u}{\partial s}(s,\delta x) = Sf(s,\delta x),$$

where $s = \delta t$. Let $s^* \ge 0$ be a local maxima for $|Sf(s,\delta x)|$. Then $t^* = \delta^{-1}s^*$ is a

local maxima for $|S(d_\delta f)(t,x)|$. Thus (5.2.2) holds. □

Remark 5.2.6. In [15], Brox-Wickert proposed to use the same nonlinear diffusion model in defining the region-based local scale using

$$m_f(x) = \frac{2T}{\int_0^T |\partial_t u| \, dt}, \tag{5.2.5}$$

where $u(t,x)$is the solution to (4.2.2), and T is a global parameter. Since $u(\delta t, \delta x)$ is the solution of 4.2.2 with $d_\delta f$ being the initial condition, we have, if $x = 0$,

$$m_{d_\delta f}(x) = \frac{2T}{\int_0^T \left|\frac{\partial}{\partial t}(u(\delta t, 0))\right| \, dt} = \frac{1}{\delta}\frac{2\delta T}{\int_0^{\delta T} \left|\frac{\partial u}{\partial t}(t,0)\right| \, dt}. \tag{5.2.6}$$

The relation between $m_{d_\delta f}(0)$ and $m_f(0)$ is given by

$$\frac{1}{m_{d_\delta f}(0)} = \frac{1}{m_f(0)} + \frac{1}{2T}\int_T^{\delta T} |\partial_t u(t,0)| \, dt,$$

In general, this definition of scale does not satisfy dilation invariant property, since $m_{d_{\delta f}}$ and m_f does not differ only by a scale factor, and the difference also depends on the local properties of $u(t,x)$.

5.3 Experiments

Local Scale

In the experiments, we assume that images are defined on $\Omega = [0,1]^2$, and the (logarithmic) local scales are computed over a discrete set $T \subseteq [0,1]$, where $T = \{t = a^{\tau_i} : a = 1.05, \tau_i = -100, \cdots, 100\}$.

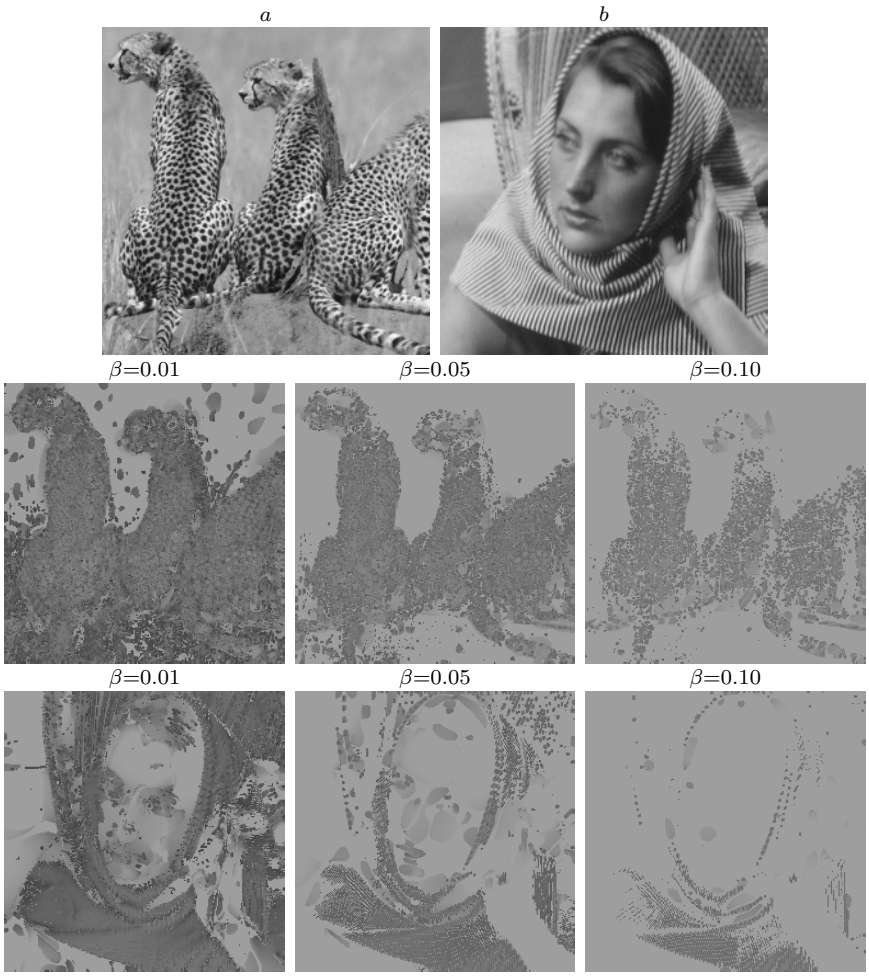


Figure 5.2: The first (logarithmic) β-visible local scale τ at x, where $|Sf(\tau, x)| > \beta$, for various β's. Here, Sf are defined in (5.2.1) using the TV flow.

Remark 5.3.1. Figure 5.2 shows the basic meaning of local scale in Definition 5.2.1.

The uniform regions (e.g. background of grass in (a), background, forehead, hair, hand in (b)) have relatively larger scale, appearing brighter in the images. Texture regions(e.g. smaller dots in (a), strips in (b)) have smaller scales appearing darker in the images. As the visibility threshold β is increased, structures with lower scale visibility were removed gradually. Figure 5.2 indicates that Definition 5.2.1 describes the relative size of local structures.

Key Point Detection

Motivated by the selection of scale and detection of key points based on linear diffusion model, we are interested to study what key points can be detected using the nonlinear diffusion model: total variation flow diffusion. Key points are expected to indicate the location of interesting structures features e.g. edges, corners, center of blobs etc. For blob detection based on heat diffusion, the key points are detected around the center of blobs. To detect key points, we give the following definition,

Definition 5.3.2. (Key point)

$$(t^*, x^*) = \arg \max_{0<t<T, x\in\Omega} |Sf(t,x)| \tag{5.3.1}$$

The point (t^*, x^*) is called *scale-space maxima* with respect to $Sf(t,x)$.

The image containing blob structures with various sizes are used as comparisons with the results of SIFT Features.

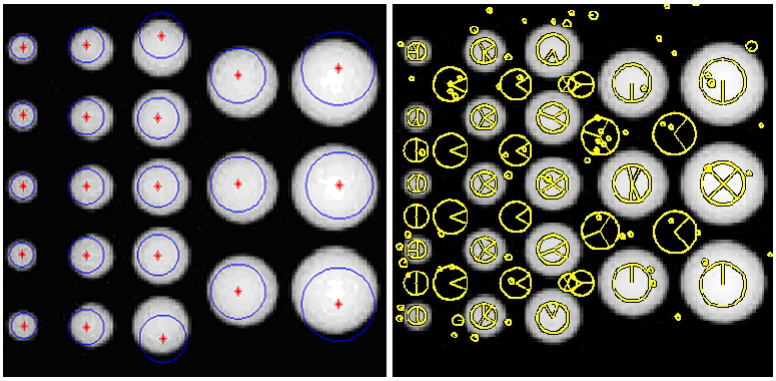


Figure 5.3: Key point detection for blob structures: Left, Key points based on TV flow diffusion; Right, Key points based on Heat diffusion. Image Source([60])

Remark 5.3.3. The above example is a typical problem in blob detection with various sizes. Due to the simplicity of this problem, it is a good example that can be used to understand different models for both advantages and disadvantages. In the left plot of Figure 5.3, we detected the key points first, and thresholded on $|Sf(x,t)|$ to remove effects of edge responses(as done in SIFT [7]). The red point stands for the key point location x^*, and the radius of the blue circle is proportional to the scale t^*. For visual convenience, we picked a spatial length r^*(unit: pixel) for the smallest blob, $s.t. r^* = \lambda t^*$, and use λ as the coefficients for other t^*. Compared to results of SIFT, (1) The result based on TV Flow model can capture the key points(i.e. centers of blob) better in the sense that all and only central of blobs are detected. SIFT is based on the understanding blob as regions brighter or darker than the surroundings, as a result, the darker regions in the image surrounded by bright disks are also detected using SIFT; (2) TV flow model however is worse in detecting the blob center accurately. Observing Left figure 5.3, for example the right column

with three largest blobs. The center of top blob is lifted up toward the edge, while the bottom blob is pulled down toward the bottom line. One of the possible explanations is the nature of numerical implementations. The evolution of images will be strongly affected by the neighborhood regions.

5.4 Discussion

Interpretation of Local Scale

Results in Experiment (5.3) show that the local scale definition based on TV flow diffusion can capture the relative scale property of structures within a given image. Compared with the region-based local scale proposed by Brox-Weickert [15], the new definition is better in preserving the scale under scaling; compared with SIFT features [7], the proposed method for key-point detection can detect more stable and correct key-points(for blob detection). However, one obvious limitation is the interpretation of t^* defined in Equation (5.3.1). In general, for TV flow diffusion equation, it is hard to get the analytical solution and to relate t^* to a spatial measure. For real applications, including blob detection, vessel detections, the local scale information is considered important. We need to reconsider the real meaning of *scale* and how we can improve current definitions.

Parameter Dependent

The automatic choice of parameters is always a problem in many applications, and commonly models are proposed with unavoidable parameters. For some models, parameters may not have huge effects on the performance of analysis, and for other models the appropriate choice of parameters can be critical. Either in SIFT key-point detection, or key-point detection based on TV Flow diffusion, there are a series parameters. For TV Flow diffusion, in analytical representation, the parameters to detect the key-point include 1) Diffusion time T_{max}; 2) Visibility of scale β; in numerical implementation, factors include 1) the regularization coefficient ϵ; 2) the time step size $\triangle t$ used in finite difference scheme. Within these parameters, the optimal choice of T_{max} and visibility of scale β can directly affect the results of key-point detection, as T_{max} decides the end coarseness of the image, and β decides how many points can be selected. In general, it is not easy to decide parameters analytically. However, given prior knowledge of image structures, it is possible to estimate these parameters based on machine learning methods.

Computational Cost

One other limitation of multi-scale analysis is the computational cost. In order to construct the multi-scale space based on nonlinear diffusion model, $\{u(t_i, x)\}_{i=1}^{N}$ need to be stored for local maximum detection, the number of N, samples of t can be large. In real image analysis, the size of the image can also be large, for example,

the retina fundus image has size is 768×584(pixel$\times$pixel) or 650×700(pixel$\times$pixel), digital mammogram image usually has size 1914×2294(pixel$\times$ pixel) or even larger. The storage of $\{Sf(t,x)\}$ can be a huge memory cost. In addition, solving Equation 4.2.2 can be time consuming depending on the numerical method used. The idea of using TV Flow diffusion in image analysis is proven to be valuable, however in order to gain broader applications, there are many spaces for further improvements.

Chapter 6

Scale space: TV Flow (II)

6.1 Revisit: What is local scale?

People use multi-scale analysis to understand the underlying multi-scale levels of natural image structures. It implies that finding a proper scale t^* for a single pixel or region to describe the geometric property is the focus of multi-scale analysis. In current studies as well as shown in Section(5.3), image $f(x)$ is embedded into a families of evolved images $\{u(x,t)\}_{t>0}$. In general, it is not a easy task to find a direct connection between the time parameter t and the geometric measure analytically and numerically.

Based on a well-posed image diffusion process, as t increases, structures within images become simpler or coarser. However, in image analysis applications, especially medical image analysis, extracting the detailed and accurate image structure information is important for object recognition and disease diagnosis. Though multi-scale analysis

has been developed for a long time, how image structure complexity is changed at different time point t remains less clear. The problem becomes even more complicated when nonlinear diffusion models are introduced.

Limitations in current multi-scale analysis and the needs in medical image analysis motivate us to find a proper way to define image local scale. In order to do that, there are two points worth mentioning: *First*, multi-scale spaces are constructed to extract multi-level structures. Improper multi-scale model(i.e. edge information changed at different scales) will distort the information used for further analysis, and can potentially affect the conclusions on the image analysis; *Second*, In many applications in image analysis, it is important to extract geometric properties of structures for analysis. It is thus desired that the definition of local scale has a clear geometric interpretation.

6.1.1 Scale: pixel-wise

A new definition of local scale is proposed to overcome the limitations of current local scale definitions. It is desirable that the local scale can have better interpretation of geometric structures, and remain broadly applicable in image analysis. The two important components emphasized in the definition are pixel-wise and orientation. *Pixel-wise* means the local scale is defined for each pixel, and *Orientation* information can be used to describe the scale region-based or pixel-wise based.

Why is *pixel-wise* important?

Consider two pixels A and B in Figure 6.1(1), with the radius of the disk in the image equals to R. It can be claimed that the region-based scale for this disk is R(i.e. radius), however, it may be inappropriate to define the local scale at A, B to be R. In fact, the distance from A, B to the edge of the disk is different : $d_A \neq d_B < R$. If the image is filtered with a Gaussian kernel, with size R, as $R > d_B > d_A$, the filter can potentially distort the boundary information close to A, B which is a well known disadvantage of Gaussian filtering. Accurate extraction of local scale pixel-wisely can be helpful for adaptive image smoothing. In addition, the study of image pixel in a local patch pattern become popular in recent studies. However, in general, the patch size is chosen in a relatively arbitrary way, which may introduce potential bias in patch comparisons. It can be explained by the fact that structures have an inherent scale property. The extraction of local scale information can be helpful for a proper choice of patch size as well.

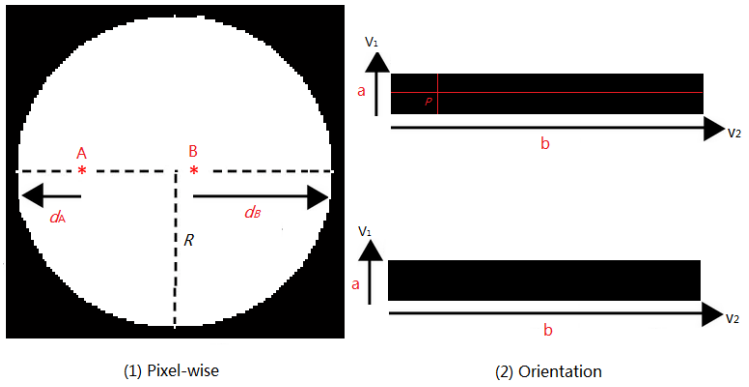


Figure 6.1: Necessity of defining scale based on pixel-wise and orientation information

6.1.2 Scale: Orientation

Considering the two rectangles in Figure 6.1(2). *First*, for pixel-wise based local scale, the behavior of the pixel along different orientations can be different. In Figure 6.1(2), point P has different cross-sectional profiles along two directions $\vec{v_1}, \vec{v_2}$ respectively. It can be hard to describe the local scale of P without considering the orientation information; *Second*, for region-based scale definition, different from disk shape, which is radially symmetric, it is more meaningful to describe scale of the region(especially elongated regions) along certain direction. For example, we can claim that the scale of the rectangle along $\vec{v_1}$ is a and is b along $\vec{v_2}$. The reason that it is important lies in the fact that most of the structures studied in natural images cannot be simply parameterized i.e. the number of degrees of freedom needed to decide a structure can be large. Using only one parameter as a scale descriptor for a region may not contain enough information.

6.2 Local Scale: pixel-wise and orientation

6.2.1 General Framework

Given an image $\{f(x,y)|(x,y) \in \Omega \subseteq R^2, 1 \leq x \leq M, 1 \leq y \leq N\}$, Figure 6.2 is shown as an example to explain the general framework for local scale extraction, here the coordinates follow the conventions in MATLAB for the convenience of numerical implementations. At each point $P = (x,y) \in \Omega$, with a given orientation θ, in order to

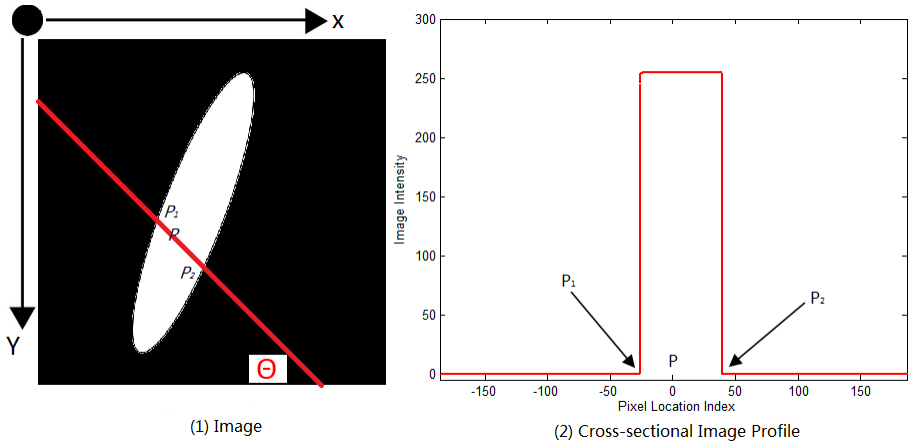


Figure 6.2: Simplified example of Local Scale Detection

detect the local scale, the cross-section profile of P along θ is the signature for P, and the detection of two jumping points P_1, P_2 in the neighborhood of P(shown in Figure 6.2 (2)) is used to calculate the local scale at P. Compared to real image analysis, Figure 6.2 is one simplified version of the analysis. In the following, a systematic method for 1D signal analysis will be discussed.

From 2D Image to 1D Cross-section

Before explaining in detail the process of extracting image local scale, there are some points that needed to be discussed. Based on the framework shown in Figure 6.2, the extraction of local scale at point P given an orientation θ in 2D image is formulated as a 1D signal analysis problem. To detect the local scale(which will be defined later), the focus of the 1D signal analysis is the detection of edges(or jump) points(e.g. P_1, P_2 in Figure 6.2) of the 1D signal $\{h(l), l_{min} < 0 \leq l \leq l_{max}\}$. Piecewise constant 1D signals are the simplest cases for edge detection, and TV flow diffusion is a well

known candidate for removing small oscillations and signal noises, generating piecewise constant signals. The definition of local scale explores a perfect connection with the 1D total variation flow diffusion at this point.

Advantages: *First*, total variation model has been shown to preserve edges, however currently no clear analytical proof can be used to validate this claim, especially for 2D image. In comparison, the analysis of 1D signals is easier and more manageable. A clearer understanding of 1D TV Flow model can help with better numerical analysis on 1D signals; *Second*, in numerical experiments, the 1D TV flow model has a simpler implementation. Finite difference(FD) is used commonly due to the simplicity of implementation. In general 2D image analysis, using the speed of explicit FD scheme is limited by a stability condition and the implicit FD schemes are limited by the memory requirement. For 1D signals with smaller sizes compared to 2D image, the numerical limitation will be removed; *Third*, there can be better interpretations for image local scale. For applications on image analysis, direct interpretation of the diffusion process is considered important. By adding orientation information, the description of local scale provides comprehensive information of image structures.

Potential disadvantages: *First*, one potential disadvantage is the computational cost, especially in real image analysis. The computational cost depends on image size(M, N), 1D TV flow diffusion calculation(T_{max}), the number of orientations ($\{\theta_i\}_1^K$) etc.. In some applications, the computational cost can be reduced;*Second*, in 2D image analysis, the formation of a 3D object into a 2D image captures the

structures of natural image. Reducing from $2D$ space for (x, y) into $(1D + \{\theta\})$ can possibly lose some information. More discussions are left at the end of this chapter.

Local Scale Definition

Definition 6.2.1. (Cross-section profile) Given an image $f(x, y), (x, y) \in \Omega \subseteq R^2$, for any point $P = (x, y) \in \Omega$, with orientation $\theta \in [0, \pi]$, the cross-section profile of (x, y) along θ is defined as $\{g(P, \theta, l)\}$

$$g(P, \theta, l) = f(P_{l,\theta}) = f((x, y) + l(\cos\theta, \sin\theta)), l_{min} \leq l \leq l_{max}$$

where $l_{min}(l_{max})$ is defined as the maximum $l, l < 0$(minimum $l, l > 0$) such that $P_{l,\theta}$ reaches the boundary of image domain $\partial\Omega$.

In Figure 6.3, examples of different cross-section profiles of natural images are presented. These examples provide general patterns of cross section profiles, and can give potential hints as how to pre-process cross-section 1D signals. Given cross-section profile $g(P, \theta, l)$, using Equation (6.2.1) with initial condition i.e. $f(l) = g(P, \theta, l)$,

$$\begin{cases} \frac{\partial u}{\partial t} = \partial_x(\frac{\partial_x u}{|\partial_x u|}), t > 0, x \in [a, b] \subseteq R^1 \\ u(0, x) = f(x), x \in [a, b] \\ \partial_x^- u(t, a) = 0, \partial_x^+ u(t, b) = 0, t > 0 \end{cases} \tag{6.2.1}$$

The solution of Equation (6.2.1) i.e. the preprocessed signal is denoted as $u(P, l, t^*)$, the selection of t^* will be discussed in the later section. The absolute curvature of

curve $(l, u(P, l, t^*))$ will be denoted as $h(l)$ as in Definition 6.2.2.

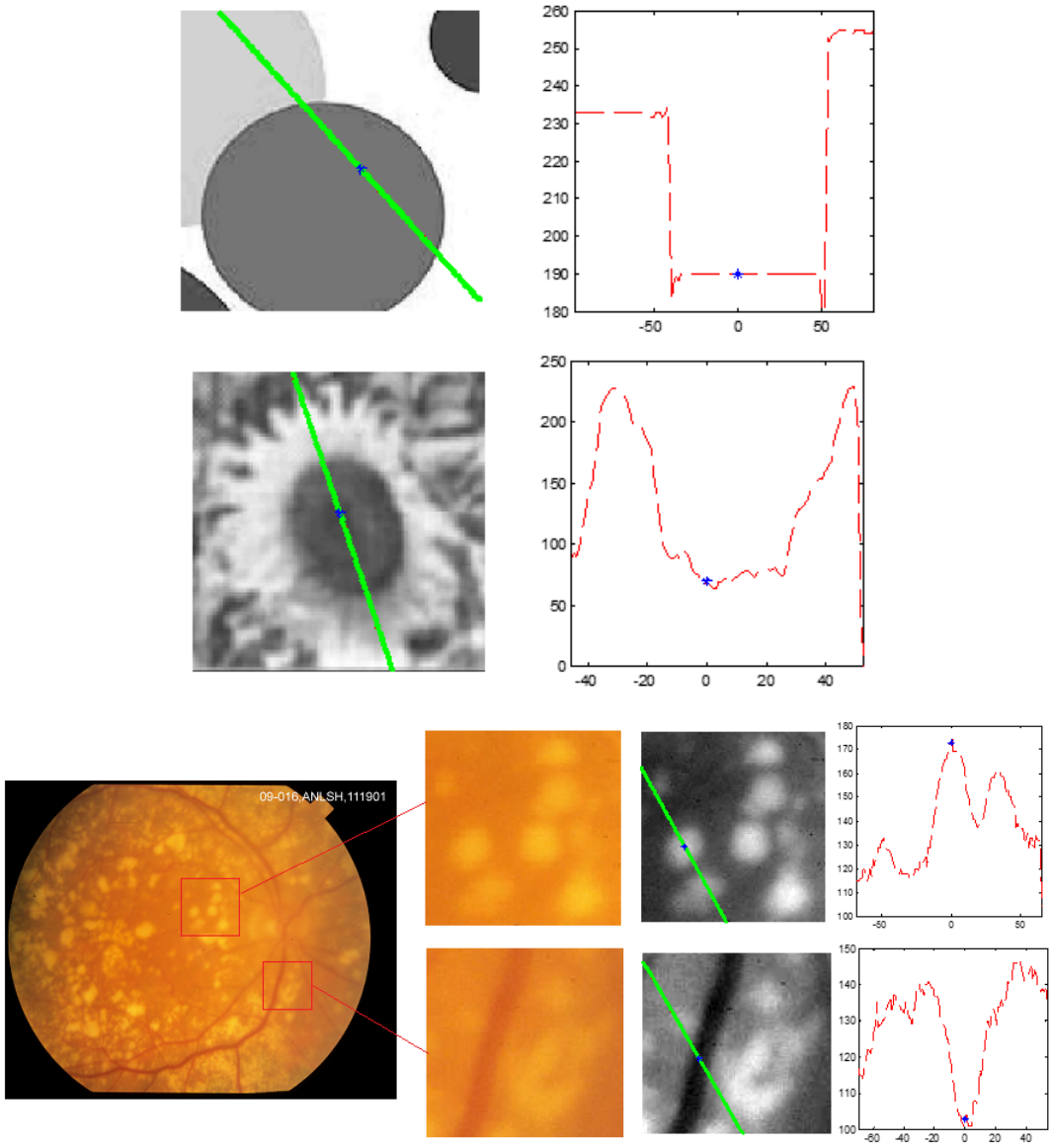


Figure 6.3: Examples of cross-section profiles in natural images

Definition 6.2.2. (Local Scale with orientation)

$$l_-(P,\theta) = \arg\max_{l_{min}\le l<0}\{l|h'(l)=0\}, l_+(P,\theta) = \arg\min_{0\le l\le l_{max}}\{l|h'(l)=0\}$$

In addition, if P lies within a certain region, $S(P,\theta) = l_-(P,\theta) + l_+(P,\theta)$ is defined as the *cross-section length* of the region along θ; if the local scale information is used to decide the adaptive diffusion size, $D(P,\theta) = \min(l_-(P,\theta), l_+(P,\theta))$ is defined as the adaptive diffusion size along orientation θ.

In addition to Definition 6.2.2, if P is in the background(i.e. not within any region of interest), $S(P,\theta)$ can be used as an indication of the distance between different subregions within the image. In more specific application problems, by adding prior knowledge of the region of interest, it is possible to extract information such as the point inside or outside the region. In summary, Definition 6.2.2 can be used as a general framework for local scale related study and be adjusted for different specific applications. Definition 6.2.2 is shown in Figure 6.4.

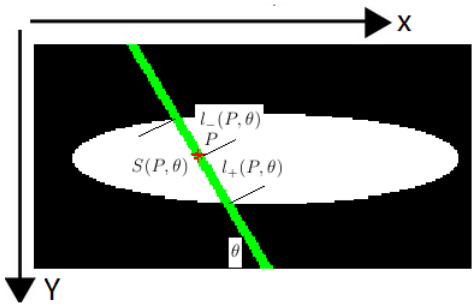


Figure 6.4: Local Scale with orientation information

6.2.2 Total Variation Flow in 1D

Definition 6.2.2 is based on the preprocessed signal $h(l)$ of the cross-section profile $\{g(P, \theta, l)\}$. The choice of preprocessing method is critical. In this study, we use the total variation flow model in 1D space. 1D total variation flow is formulated in Equation (6.2.1). Compared to 2D TV flow, 1D TV flow equation has been well studied in several papers [61, 62, 63, 64], both analytically and numerically. It is well-defined, can preserve edge structures, and be easily solved numerically.

The numerical algorithm to solve 1D TV flow has been discussed in Section 2.3. Parameters associated with the numerical algorithm include the regularization term ϵ, the time step size Δt and the diffusion time T_{max}. These parameters are closely related to the property of the multi-scale profile $\{u_\epsilon(t, x)\}$ as an approximation to $\{u(t, x)\}$. In the experiments, $\epsilon = 0.01, \Delta t = 0.5(\Delta x)^2$. The two parameters are chosen as a balance between solution accuracy and numerical efficiency. The only parameter that needs to be selected is T_{max}, and in some literature [65, 66, 67, 15, 68, 69], it is defined as the problem to choosing the stopping time T_{max}. In [65], the time for stopping the evolution was mentioned, however compared to heat diffusion model, there is no general agreement on a criterion for the stopping time. In [69], it is mentioned that for the nonlinear diffusion model, a stopping time is needed otherwise the signal will be smoothed out. One way the author proposed is to use constrained evolution instead, however the convergence and uniqueness of the solution to the constrained PDE problem is not always guaranteed. In [15], authors discuss the choice of T_{max},

and provide an upper bound for $T_{max} \leq \frac{1}{4} size(f) \cdot c_{max}$ and how the value is derived is not stated explicitly.

Effects of image filters on 1D signals

In general, nonlinear diffusion equations cannot be solved analytically. Considering the singularity that may appear in some nonlinear problems, it makes the analytical analysis even more difficult. In this section, numerical results for image filters on 1D signals are discussed. Results are organized by three factors. First(*Signal type*), two types of signals are used in the analysis: (1) piecewise constant signal; (2) smooth signal(e.g. gaussian function); The two types of signals are chosen based on observations from cross section profiles of natural images(Figure 6.3); Second(*Transformation type*), for each type of signal, two transformations are considered namely (1) $(f(x), f(kx))$; (2) $f(x), kf(x)$; Third(*Filter type*), three filters based on (1) Heat diffusion; (2) Mean curvature diffusion; (3) TV flow diffusion.

I, Piecewise constant signal

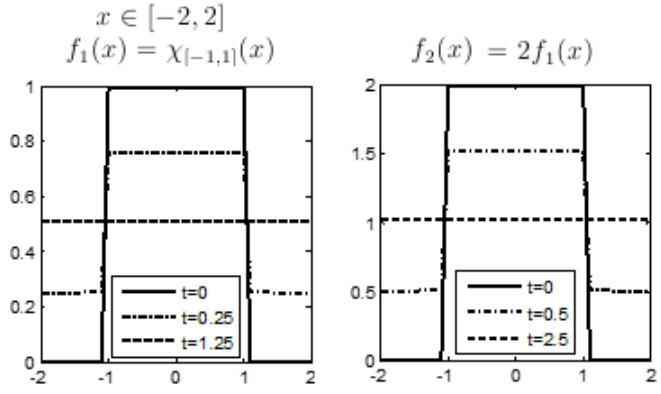


Figure 6.5: TV diffusion for signal f_1, f_2, f_3

Remark 6.2.3. Figure 6.5 demonstrates three properties of TV flow diffusion: (1) TV Flow diffusion tends to decrease the image contrast, and converges to a constant signal within a finite time; (2) with the introduction of the regularization term ϵ, the solution $u_\epsilon(x,t)$ is a regularized version of $u(x,t)$ i.e. u_ϵ may not be a strictly piecewise constant signal; (3) The locations of edges are preserved in the image diffusion process.

To describe properties of 1D signals in multi-scale space, features such as minimum, minimum, maximum, mean, energy and total variation of signals are used for characterization. The relationship between signals $f(x), f(kx), kf(x)$ evolving with time t is also compared. The multi-scale profiles are shown below.

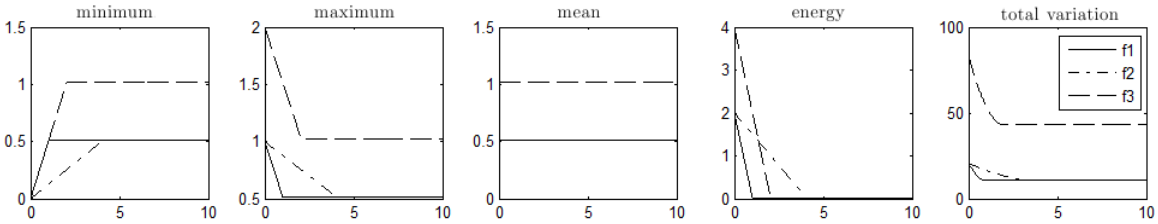


Figure 6.6: Five features for multi-scale signals. Solid line, dash-dot line and dashed line stands for multi-scale profiles of f_1, f_2, f_3 respectively. Here $t_1^* = 0.5301 \sim t_2^* = 0.5302 \sim 0.5t_3^* = 1.0428$ indicating the time that the sharp change occurs in the multi-scale space.

Figure 6.6 validates that (1) TV flow diffusion obeys the maximum-minimum principle(i.e. the maximum(minimum) value of the signal would not increase(decrease)) and the conservation law(i.e. the mean value of the signal remains the same); (2) Total variation of the signal decreases as time t increases; (3) The time at which the features converges to stable state is proportional to the scale factor k, between f_1, f_3. For example, it takes t^*, kt^* for f_1, f_3 to converge to constant signal respectively.

II, Gaussian signal

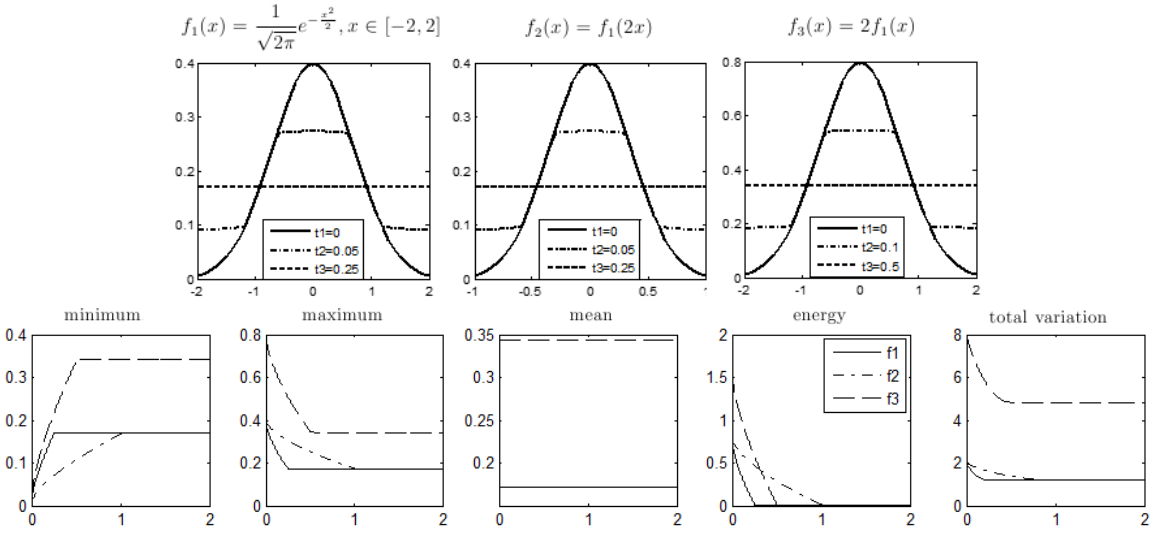


Figure 6.7: TV Flow diffusion for gaussian signal and multi-scale features

Figure 6.7 shows that TV Flow diffusion tends to evolve the smooth signal to piecewise constant signal. Considering the property of cross-section profiles of vessels, the tv flow diffusion can be used to pre-process the signals and the edges of vessel can be detected based on the piecewise constant signal.

III, Effects of diffusion filters

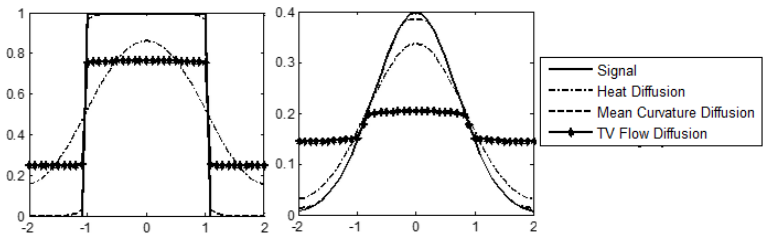


Figure 6.8: Effects of diffusion filters on signals

Remark 6.2.4. It can be observed from Figure 6.8 that heat diffusion, MCM diffusion and TV diffusion are different types of diffusion models. Heat diffusion tends to

smooth the signal regardless of edge structures, resulting in distortion of edges; TV Flow diffusion is able to preserve edge points of the signal, however it decreased image contrast. Compared to heat diffusion and TV flow diffusion, MCM diffusion acts as a diffusion model between the two models. It blurs the signal around the region where the curvature is large, but the blurring effects are less compared with heat diffusion. The image contrast is not changed as much compared to TV flow diffusion. Studying the behaviors of filters based on these signals shows that in order to detect the edge points, TV flow filter is the best choice.

Preprocess of cross-section profile

Based on the comparisons between three selected image filters, TV Flow diffusion is selected as the 1D signal filter. The purpose of pre-processing is to remove small oscillations and preserve the edge information. In the work of Canny [70] etc., gaussian filter is used and the derivatives of gaussian filtered signal are used to detect edges. As shown in Figure 6.8, compared to MCM and TV flow diffusion, improper choice of gaussian kernel can distort the edge information. The disadvantage of gaussian filtering makes the edge detection based on gaussian filtering tricky. Compared to MCM diffusion, though signal contrast is decreased in the TV flow diffusion, the edge information is well preserved.

In the experiments, it is observed that if the 1D signal $f(x)$ is considered as curve $(x, f(x))$, curvature of the curve can capture the edge points with a clearer pattern as a sharp jump of curves happens at the edge point. Schemes for the cross-section

profile pre-processing are summarized as below: given pixel P, along direction θ, the cross-section profile is denoted as $f(x) = g(P, \theta, x), x_{min} \leq x \leq x_{max}$. Using TV flow diffusion, signal $f(x)$ is embedded in a multi-scale profile $\{u(t, x)\}$. The total variation of signal $\{u(t, x)\}$ is calculated at each time t, and the time t^* is chosen such that (1) $t < T_{max} = \arg\min\{t : TV(u(t, x)) = 0\}$; (2) $TV(u(t^*, x)) = \theta \cdot TV(u(0, x)), 0 < \theta < 1$. The introduction of θ as an additional parameter is inevitable at the current research stage, since without prior knowledge of signal structures and it can be hard to set up a criterion to choose t^*. In our experiments, we chose $\theta = 0.25$ uniformly instead. Curvature $\kappa(x, t)$ of curve $(x, u(t^*, x))$ is used to extract edge information for $x < 0, x > 0$ respectively. Here, the total variation of a discrete 1D signal and curvature of curve$(x, y = u(t, x))$ is calculated as:

$$TV(u(t, x)) = \sum_{i=1}^{N-1} |u(t, x_{i+1}) - u(t, x_i)|, x = \{x_i\}_1^N$$

$$\kappa(t, x) = \frac{u''(t, x)}{(1 + u'(t, x)^2)^{3/2}} = \frac{x'(s)y''(s) - y'(s)x''(s)}{|(x'(s), y'(s))|^3}$$

$\kappa(t, x)$ is signed curvature as discussed in Section 2.2.1, s is the arc length, and $\kappa(t, x)$ is better for curve characterization. Magnitude of curvature is used for edge detection and $h(x)$ in Definition 6.2.2 is defined as $h(x) = |\kappa(t^*, x)|$. Three examples of the proposed preprocessing method are shown, the first example is a piecewise signal added with gaussian noise, SNR = 10; the second example is a gaussian signal. The third example is from a real retina fundus image and the signal is a cross-section profile of

a blood vessel structure.

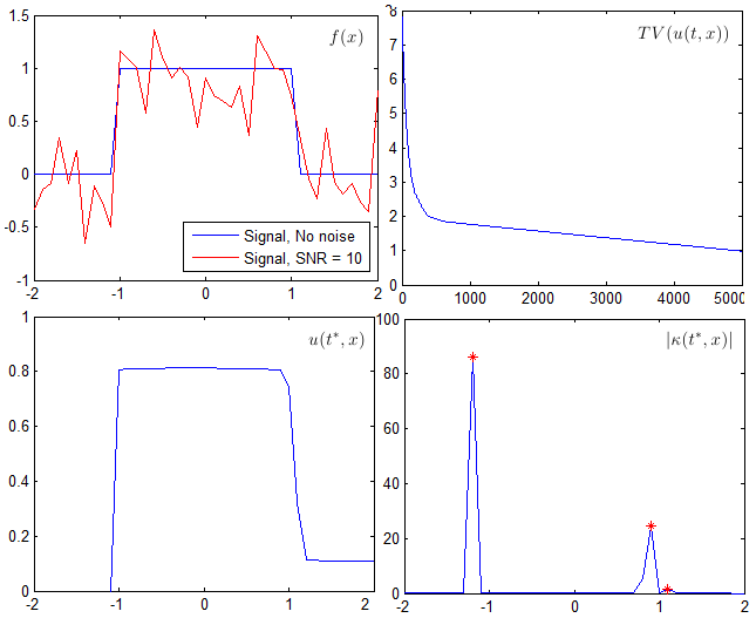


Figure 6.9: Example 1: signal with gaussian noise, SNR=10

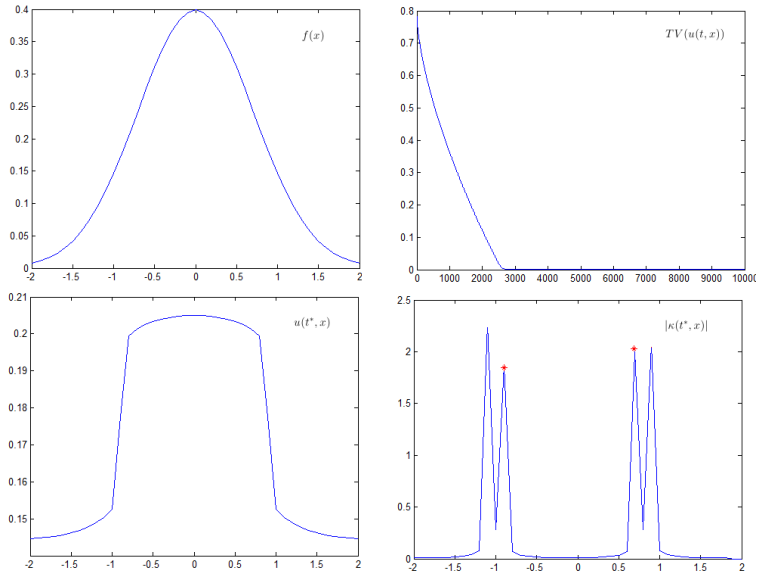


Figure 6.10: Example 2: edge detection for gaussian signal

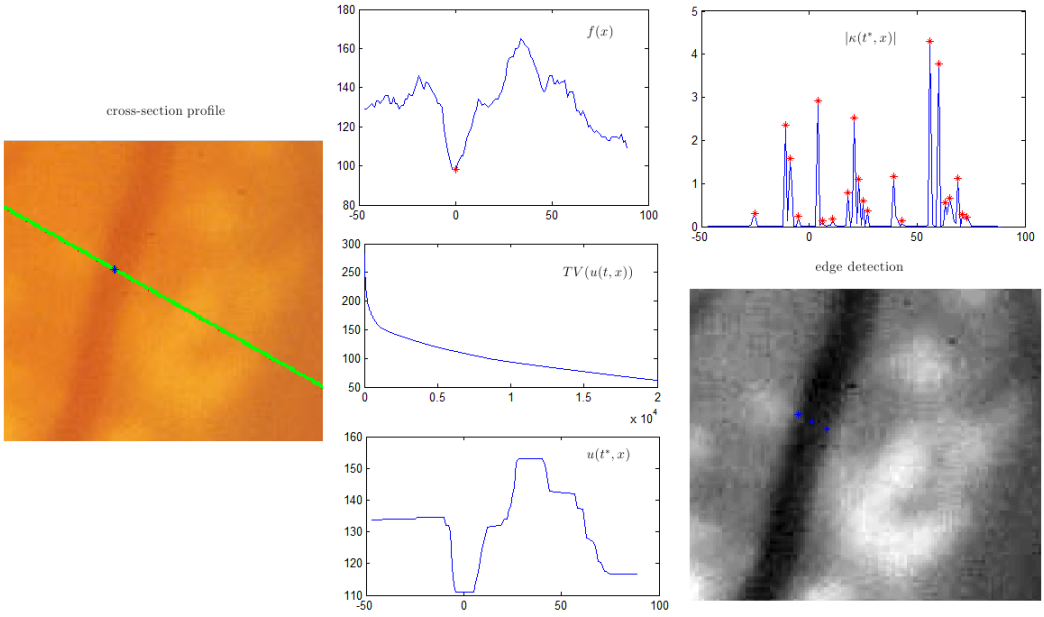


Figure 6.11: Example 3: edge detection for fundus image

6.2.3 Examples

Example 6.2.5. (Analytical v.s. Numerical) Given a disk with radius R in a square with side length a, image intensity outside the disk equals to 0, and inside equals to 1. $P = (a, a)/2 + r(\cos\beta, -\sin\beta), 0 \leq r \leq R, 0 \leq \beta \leq \pi$ shown in Figure 6.12.

$$l_{\mp}(P,\theta) = \sqrt{R^2 - r^2\sin^2(\beta+\theta)} \mp r\cos(\beta+\theta), S(P,\theta) = 2\sqrt{R^2 - r^2\sin^2(\beta+\theta)}$$

If P is the original point O i.e. $r = 0$, $S(P,\theta) = 2R$, consistent with the intuition that the cross-section length through the center of a disk is equals to the diameter.

Example 6.2.6. (Synthetic image: with different level of gaussian noise)

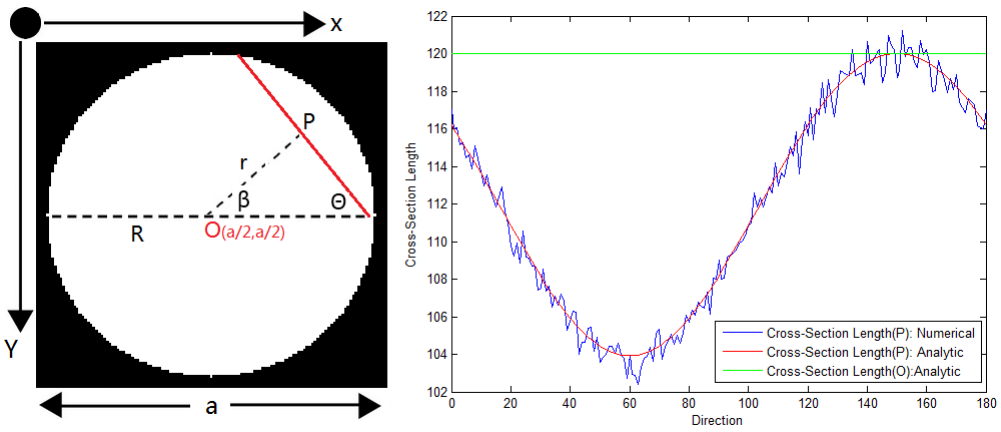


Figure 6.12: x-axis: θ, y-axis: $S(P, \theta)$. Blue and red curves: cross-section length through P numeric and analytic result, green curve: cross-section length through O.

Noise	Level	Analytic Length	Numeric Length
Gaussian	0.01	196	197
	0.1	196	198
	0.25	196	198

Table 6.1: Cross-section length detection with different gaussian noise level

6.3 Applications

6.3.1 Adaptive image diffusion

In most of the image analysis tasks, gaussian smoothing (i.e. $f \star G_\sigma$) is used as the preprocessing step to eliminate noise exists in the images.

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{6.3.1}$$

There is a dilemma for the proper choice of σ: if σ is too small, the smoothing process can hardly remove any noise; on the contrary, if σ is too large(e.g. larger than the

scale of certain image structures), the important structure information(e.g. edge, corner) can be distorted. Example shown in Figure 6.13.

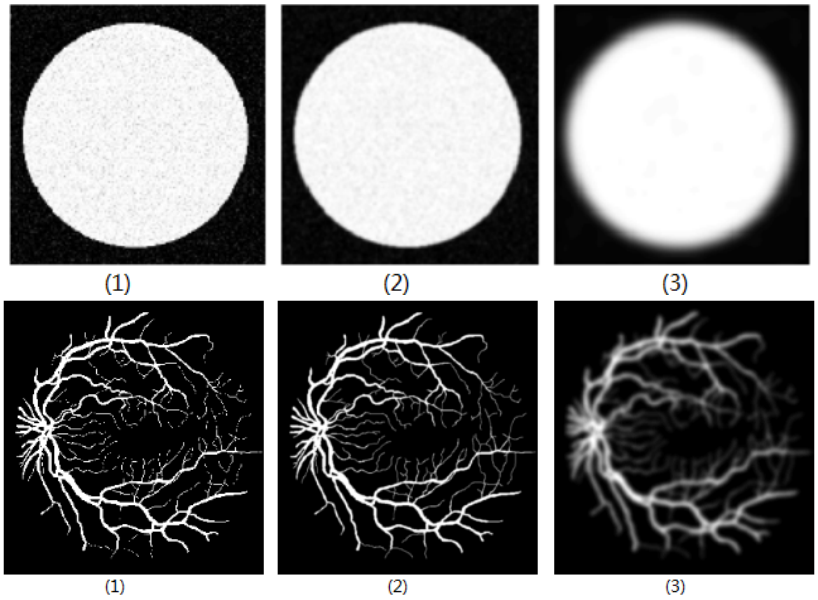


Figure 6.13: First row: (1) disk with gaussian noise; (2), (3) σ is small, large for smoothing; Second row: (1) vessel structure; (2),(3) σ is small, large for smoothing

The direct application of local scale is the adaptive heat diffusion, since the gaussian kernel size σ is related to the neighborhood size. With the introduction of $D(P, \theta)$, if the direction θ for certain diffusion is given, the adaptive diffusion size $D(P, \theta)$ can provide a prior knowledge of diffusion step size. If an isotropic diffusion (e.g. gaussian diffusion) is needed, define

$$D(P) = \min_{0 \leq \theta \leq \pi} D(P, \theta)$$

Compared to empirical choice of the kernel size, the local scale can be more accurate compared to other methods. $D(P), D(P, \theta)$ can also be useful in deciding the size of local patches around a pixel, depending on the shape of the local patch.

Example 6.3.1. (Adaptive Gaussian Smoothing)

Given an image $f(x)$, the adaptive gaussian smoothing $f \star G_{D(x)}$ is formulated as

$$f \star G_{D(x)}(x) \sim \int_{\Omega} f(y) e^{-\frac{(y-x)^2}{2D(x)^2}} dy, x \in \Omega \tag{6.3.2}$$

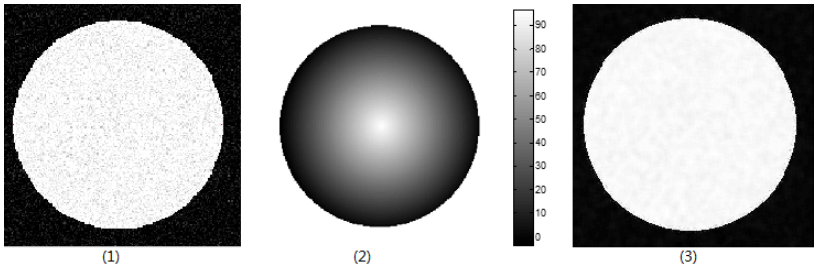


Figure 6.14: (1) Binary disk with gaussian noise level 0.01; (2) Local scale of each pixel $D(x)$ within the disk; (3) The denoised image

6.3.2 Vessel Tracking

In natural images, besides shapes close to disk, elongated shapes with length longer than width exist widely, e.g. stalks in leaves, blood vessels in lung or retina fundus images and breast tissue flow structures in digital mammograms etc. These structures are termed as *vessel* in our study. Examples are shown in Figure 6.15. These structures have dominate edge structures. In medical image analysis, accurate detection of edges can be crucial for disease diagnosis and treatment. In Section 7.2.2, different methods in vessel detection will be reviewed. Here, vessel tracking is one branch of methods that captures the interest. Vessel tracking is not a fully automatic method

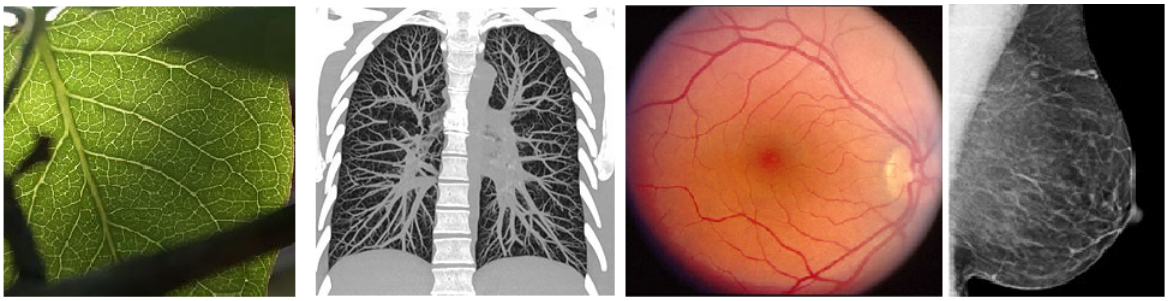


Figure 6.15: Examples of Vessel Structures in natural and medical images

as it requires a seed point to initialize the tracking procedure. However, compared to other methods, vessel tracking can provide better detection performance provided with an accurate initial seed and the algorithm is more efficient as the region of interest(ROI) will be restricted within the vessel regions. The necessary information used for vessel tracking includes the main (local) direction of a vessel, and step size traveling along the main direction(shown in Figure 6.16).

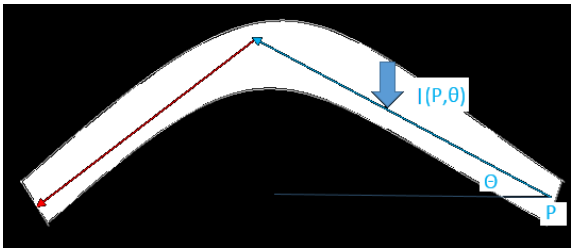


Figure 6.16: Vessel Tracking: main direction θ and step size $l(P, \theta)$

Detection of main vessel direction

A vessel can travel with varying directions within a long distance, and each segment of vessel can be approximated by a rectangle. In order to utilize the cross-section profile to decide the main vessel direction, a simplified example based on the analytic

calculation is shown. In this example, the vessel segment is approximated by a rectangle with side length $a = a_1 + a_2 < b$,where the distance between seed point P and corner point is a_1, a_2. The cross-section length through P along θ can be calculated as the following

$$l(P,\theta) = \begin{cases} a_1/\cos\theta, & 0 \le \theta < \theta_1 \\ b/\sin\theta, & \theta_1 \le \theta < \theta_2 + 90^\circ \\ -a_2/\cos\theta, & \theta_2 + 90^\circ \le \theta < 180^\circ \end{cases}$$

Here $\theta_1 = \arctan(a_1/b), \theta_2 = \arctan(a_2/b)$. The numerical and analytical result for cross-section length is shown in Figure 6.17. It can be observed that due to the existence of vessel width(in respect to vessel length), the cross-section length does not achieve its maximum along the main vessel direction, and instead the main vessel direction is a local minimum point in between the two local maximum points. Theoretically, the main vessel direction θ_{vessel} at P is defined as a θ satisfying

$$(1) l'(P,\theta) = 0, l''(P,\theta) > 0; (2) \theta_1 < \theta < \theta_2, l'(P,\theta_i) = 0, l''(P,\theta_i) \le 0 (i = 1,2) \tag{6.3.3}$$

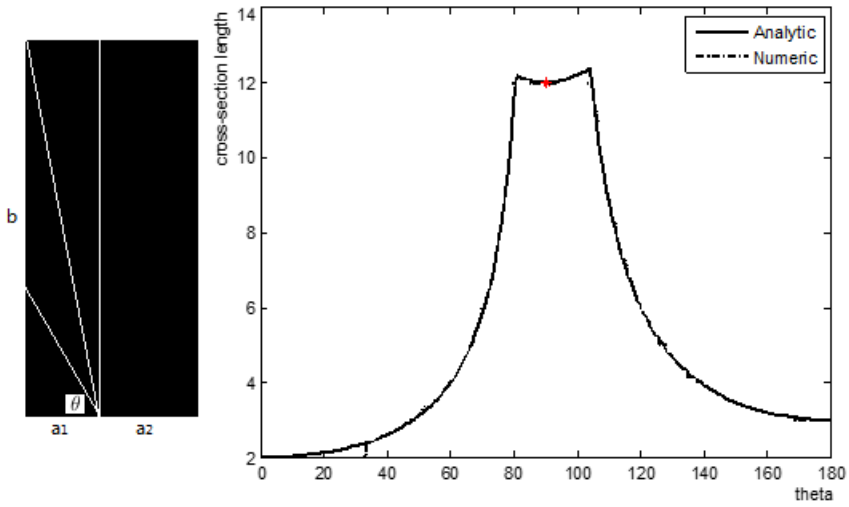


Figure 6.17: Example of cross-section profile of vessel

Scheme for vessel tracking

Definition 6.2.2 can provide the cross-section lengths along different orientations. Based on the assumption that the cross-section length achieves its maximum along the main direction of the vessel θ_P, $l(P, \theta_P)$ indicates the step size along the direction θ_P. In summary, Definition 6.2.2 provides a natural way for vessel tracking. The main procedure is explained in the following, shown in Figure 6.20.

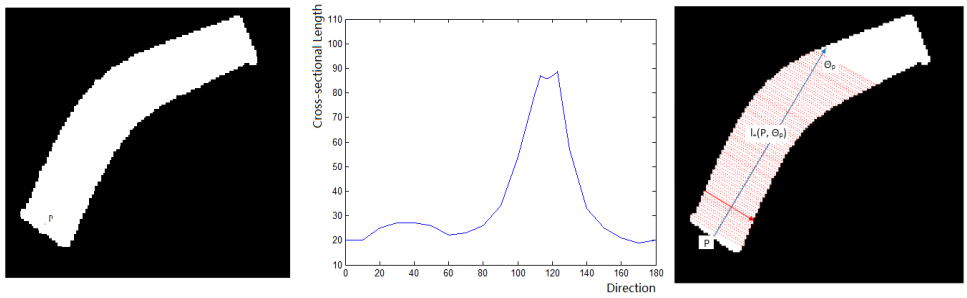


Figure 6.18: Vessel tracking: Left: image given starting point P for vessel tracking; Middle: x-axis: orientation; y-axis: cross-sectional length; Right: vessel tracking from P

Step 1, Seed Initialization: given seed P within the region for vessel tracking;

Step 2, Vessel Direction: Search main vessel direction: search for the orientation θ_P with the maximal cross-section length $l(P, \theta)$ based on Equation 6.3.2.

Step 3, Vessel Tracking: track the vessel along direction θ_P, with tracking length $l(P, \theta_P)$. To avoid the possibility of tracking outside of the vessel region, it can track back 1-2 pixels, and repeat the previous steps. The iteration process will be terminated if the tracking length along vessel direction is zero. This step provides a single line track of the vessel. Considering the width of the vessel, the next step is necessary;

Step 4, Cross-section of vessel:The width of a vessel can be decided using the perpendicular orientation $\theta_P + sgn(90° - \theta_P)90°$, and the vessel width along this direction is $S(P, \theta_P + sgn(90° - \theta_P)90°)$. The points tracked along the orientation $\theta_P + sgn(90° - \theta_P)90°$ will be labeled as pixels within the vessel.

Step 5, Update seed point: $P_{new} = P_{old} + (\cos\theta_P, \sin\theta_P)l_{\theta_P}$ and repeat above steps. Stop the iteration until step size l_{θ_P} is zero.

Example 6.3.2. (Blood vessel tracking)

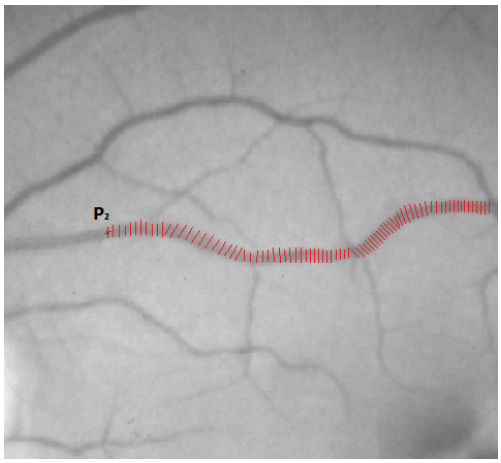


Figure 6.19: Subregion from STARE database [3]. Vessel tracking with seed P_1 .

Remark 6.3.3. (Is the correct detection of main vessel direction important?) It can be seen from the scheme that pixels within vessel region will all be detected. In fact, if the vessel region needs to be segmented instead of merely the central line, the condition that the vessel direction should be detected correctly can be relaxed(Figure 6.20). Both directions in $(a), (b)$ can be used to detect all vessel pixels, however in general, θ_P that is closer to the main vessel direction((b)) can be more efficient in the whole vessel tracking.

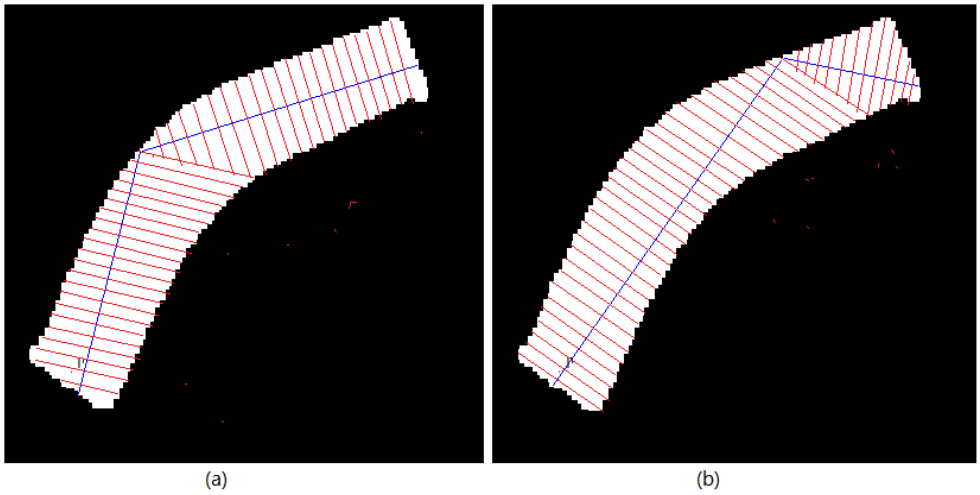


Figure 6.20: Left: track deviates from vessel direction; Right: track along vessel direction

6.4 Discussion

Scale, multi-scale, local-scale terms that appear in multi-scale analysis commonly are used with certain ambiguity. Depending on different application purposes, these terms can be defined in different ways. In order to serve the purpose of image structure detection, we give a definition of local scale by incorporating the spatial and

directional information at the same time. To the best of our knowledge, it is the first definition of local scale that incorporate both types of information. The motivation of the definition is that *local scale* can only be meaningful given a pixel with specific direction. Several applications utilizing advantages of this definitions are discussed, including adaptive diffusion, vessel tracking etc. The example of vessel tracking applied in real clinical images is also shown.

As stated in Introduction, gaps between theory, algorithms and real applications exist. One of the most important limitation for the application of local scale definition is computational cost. The computational cost depends on size of the image, time used for pre-process the cross-section signal(i.e. TV Flow diffusion time), number of directions θ used to analyze cross-section profiles etc. In some applications, certain limitations can be removed. For example, in vessel tracking, when a main vessel direction is detected, the others pixels along the vessel directions can be labeled as vessel automatically, making the tracking process faster instead of searching all points in the image. Further work needs to be done to better connect the bridge between algorithms and real application problems, especially how to implement the algorithms in a more efficient way.

Chapter 7

Retina Fundus Image Analysis

7.1 Clinical Background

7.1.1 ARMD

Age related macular degeneration(ARMD) is the leading cause of permanent visual impairment among the elderly in western countries. ARMD accounts for almost 50% of those registered as blind or partially sighted. Nowadays, effective treatment is limited to the management of sub-retinal neovascularization(SRNV), and early recognition and prevention of potential disease is not as yet applicable to diseases other than SRNV. In [1], ARMD are defined as a disorder of the macular area most often apparent after 50 years of age, and it can be characterized by three aspects: *First*, discrete whitish-yellow spots identified as drusen; *second*, increased pigment or hyper-pigmentation associated with drusen;*third*, sharply demarcated area of de-

pigmentation or hypo-pigmentation of the retinal pigment epithelium and associated drusen. Here, *druse* is a German word, and 'drusen' is a plural for 'druse'. Drusen accumulated underneath the retina and will affect human vision.

ARMD can be divided into two stages i.e 1, early ARMD and 2, late ARMD. The early stage of ARMD are dry, meaning that it does not involve any blood vessel issue. People at the early stage of ARMD may not development into vision loss, depending on different situations. For late ARMD patients, it may develop into wet ARMD since there is one symptom called Choroidal neovascularization (CNV): the creation of new blood vessels in the choroid layer of the eye; and they may also develop into dry ARMD with geographic atrophy. Examples of fundus images at different phases are shown in Figure 7.1.

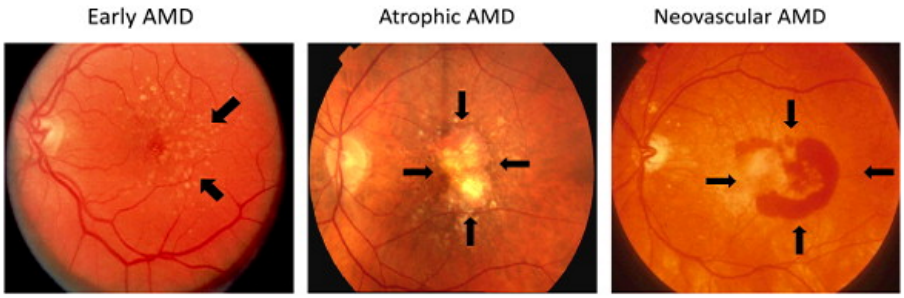


Figure 7.1: Fundus images of ARMD at different phases (Image source [71])

7.1.2 Risk factors of ARMD

Age is counted as one of the important risk factors for ARMD[72], and it is not a controllable factor. One of the other important risk factors that has proven to be important ARMD modifiable is the smoking habit[1]. In addition, genotype(e.g.

gene CFH, ARMS2) and ethnic origin are considered important risk factors. In one study of [1], the heritability of ARMD is 2/3. Nutrition, exogenous post menopausal estrogen, and light are also included as the risk factors.

7.1.3 Diagnosis of ARMD

In retina image analysis, the color fundus image and fundus autofluorescence are two modalities used commonly, examples are shown in Figure 7.2. The later is used to measure the progression of eye diseases. In our study, the analysis is based on color fundus images(left of Figure 7.2). Left in Figure 7.3 shows normal structures in a retina fundus image, including optic disk, blood vessels, fovea-macula region. If images are acquired from ARMD patients, lesions(e.g. drusen) are present in the image as well(Right, Figure 7.3).

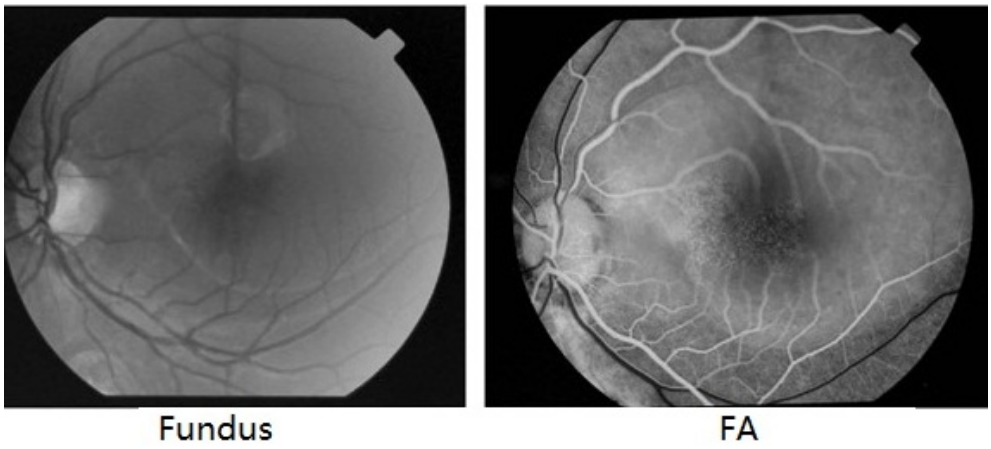


Figure 7.2: Left: Fundus image; Right: Fundus autofluorescence

Key lesions of ARMD are drusen, and most people over the age of 40 have at least one druse. Small, hard drusen are referred to simply as drusen, soft drusen over 63

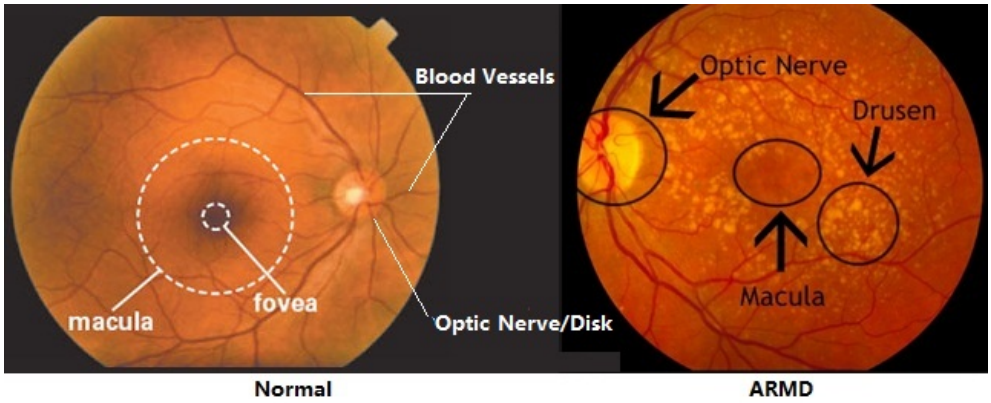


Figure 7.3: Structures within Normal v.s. ARMD retina fundus image

microns in diameter are statistically associated with visual pathology and are termed early ARMD. For a screening program, the severity of ARMD is decided by a ARMD grading system(e.g. [1, 73, 74]). However, the grading systems for ARMD disease are quite different across continents including Europe, U.S. and Japan, resulting in different conclusions in portions of ARMD patients etc.

7.2 Related Work

7.2.1 Drusen Detection

Automatic detection of drusen is not a easy image analysis problem, and the main reasons can be summarized into three points: (1) The similar visual appearance between drusen and other lesions, optic disk. The similarity makes simply thresholding the image intensity fail; (2) Boundaries of drusen especially soft drusen are vague and it makes the edge detection ineffective; (3) Within a single retina fundus image, size

and shape of drusen vary making template matching ineffective. Besides, the nonuniform illumination caused in the image acquisition process can introduce additional artifacts in the images.

Sbeh et.al. [60] approach the problem using a morphology operator to detect the brightest points (peaks) within individual drusen. Rapantzikos et. al.[75] used an adaptive local histogram to identify an appropriate local threshold to segment each druse. Thaibaoui [76] used a fuzzy logic approach to classify pixels as drusen or non-drusen based on image intensity.

In retinal image analysis, multi-level, multi-scale analysis becomes natural due to the limitation of single level segmentation and the existence of multiple sizes of drusen. The idea of multi-scale method is to represent the image at varying scales and segment image structures at an appropriate scale. For example, wavelet based filtering has already been used in problems such as blood vessel detection in retinal images[77] successfully. Zhang et. al. [29] introduced an neural net approach to multi-level segmentation of medical images. A top level neural net provides the coarsest segmentation, and each top level segmented element is passed to lower level neural nets which perform successively finer segmentations.

7.2.2 Blood vessel detection

Introduction

Retinal vessel segmentation and delineation of structural features of retinal blood vessels, including length, width, tortuosity, branching pattern are used for diagnosis, screening, treatment and evaluation of various opthalmologic diseases such as diabetes, hypertension, arteriosclerosis and chorodial neovascularization [3, 4] etc.

Automatic detection of vasculature structures can assist in the implementation of screening programs for diabetic retinopathy, computer assisted laser surgery, and study of the relation between vessel diameter and hypertension. The extraction of retinal branch points can be used for retinal image registration, optic disk localization; The extraction of retinal vascular tree which is found to be unique for each person can be used for biometric identification.

In general, vessel structures are complex, and it makes the manual tracking of retinal blood vessels a long and tedious task requiring training and skill. In addition, manual segmentation results vary a lot inter- and intra-observers. The automatic quantification of retinal vessels is the first step in the development of a computer-assisted diagnosis, and has been a subject of intensive research interest recently.

Retinal vessel segmentation

The retinal vascular structure is composed of arteries and veins appearing as elongated shapes. The vessel width varies from one pixel to 20 pixels, depending on the

width of the vessel and the image resolution. The vessel cross-section intensity profile approximates a Gaussian function [78] or a mixture of Gaussians in the case where a central vessel reflex is present. The orientation and intensity level of a vessel does not change abruptly. They are locally linear and gradually change in intensity along their lengths. The shape, size and local gray level of blood vessels can vary hugely and some background features may have similar attributes to vessels. Vessel crossing and branching can complicate the task. In addition, as in the processing of most medical images, signal noise, lack of image contrast pose significant challenges to accurate extraction of blood vessels.

Classification of Retinal vessel segmentation methods

Vessel segmentation in retinal fundus images can be grouped into seven main categories: 1) Pattern recognition techniques; 2) Matched filtering; 3) Vessel Tracking/tracing; 4) Mathematical morphology; 5) Multi-scale approaches; 6) Model based approaches; 7) Parallel hardware based approaches. Methods of different categories can overlap. In our study, we are most interested in category 3(vessel tracking) and category 5(multi-scale analysis).

Vessel tracking

A vessel tracking algorithm segments a vessel between two points using local information and works at the level of a single vessel rather than the entire vasculature in the image. The center of the longitudinal cross-section of a vessel is determined

with various properties of the vessel including average width, gray level intensity and tortuosity.

Tracking follows vessel center line guides by local information, usually trying to find the path which best matches a vessel profile model. The main advantage of vessel tracking is that they can provide highly accurate vessel widths, and can provide information about individual vessels usually unavailable using other methods. In addition, as vessels are connected within the retina image, the tracking is limited within the vessel region instead of wasting time examining vast majority of non-vessel regions. Vessel tracking can also give information on vessel structure such as branching and connectivity. The underlying disadvantages of this method include: 1) Unable to detect vessels or vessel segments which have no seed points; 2) missing branching points can result in undetected sub-trees.

Blood vessel in multi-scale space

The width of a vessel decreases as it travels radially outward from the optic disk and such a change in vessel caliber is a gradual one. Therefore, a vessel is defined as a contrasted pattern with a gaussian shape cross-section profile, piecewise connected, locally linear, and with a gradually decreasing vessel width. The variety of vessel widths makes the scale-space representation natural. Among these studies, the idea of Frangi [26] have been used widely by others and it has broad applications in vessel structure enhancement and detection.

In Frangi's work [26], a vesselness measure is obtained on the basis of eigenvalue

analysis of multi-scale second order local structures of a given image. Two gray-level invariant geometric ratios are defined on the basis of eigenvalues. The final vesselness measure is defined using one geometric ratio, the eigenvalues and the Frobenius matrix norm. This vessel measure can be used to enhance vessel structures for segmentation. It was shown with a 2D DSA image that the noise and background can suppressed and vessel can be enhanced at the same time.

Based on the analysis of total variation based multi-scale model and that TV/L_1 model has advantage in preserving geometric structures of images, one example utilizing the multi-scale properties of TV/L_1 model in retinal fundus image analysis will be shown and discussed.

7.3 TV/L_1: Blood Vessel and Drusen Detection

7.3.1 Method

Preprocess

For color fundus images, the green channel is used for analysis. Studies in [79] show that structures are often better discerned on monochromatic red-free(i.e. green) images when compared with color fundus images. The non-uniform illumination is corrected based on the previous work on bias correction[80]. A binary mask for the region of interest (ROI) i.e. the region of the fundus image, is generated using simple threshold and shape approximation.

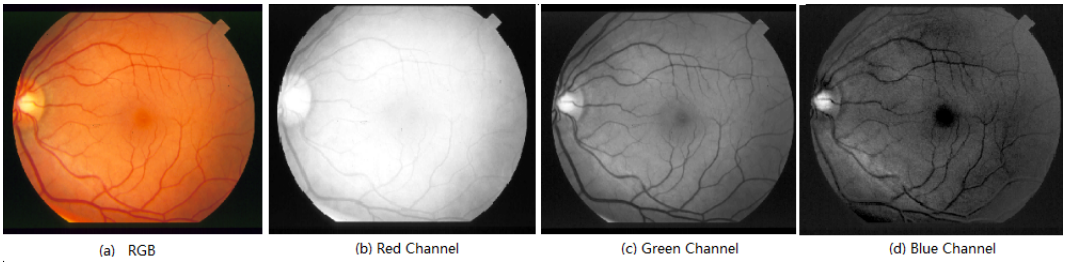


Figure 7.4: RGB channels of color retina fundus image. Data source: STARE dataset

Separate eye structures from background

Shown in Figure 7.3, pixels in a fundus image can be categorized into two groups: 1, eye structures i.e. retinal blood vessels, lesions(e.g. drusen or exudates), optic disk and fovea-macula region; 2, background i.e. pixels not within eye structures. In general, the optic disk and lesions(especially drusen, exudates etc.) are brighter and blood vessels are darker than the surroundings.

TV/L_1 multi-scale space is constructed using the method in [81], the parameter λ in Equation 4.1.6 varies from 0 to 10, with step size equals to 0.1. $u(x, \lambda_{min} = 0)$ is the original image, $\lambda_{max} = 10$ is chosen such that the evolution of the image can generate structure changes in the image. The set $\{\lambda_i\}$ is used for images analyzed in the results section. In real applications, the proper choice of λ_{max} depends closely on image sizes.

In the multi-scale space $M(f, \lambda) = \{u(\lambda, x\}$, pixels in image background and eye structures behave differently as λ increases, a typical example is shown in Figure 7.5(c). To characterize the difference between pixels, total variation from λ_{min} to λ_{max} for each pixel is defined using a quantification function. It is worth mentioning that

the multi-scale profiles $\{u(x, \lambda_i)\}$ have rich structural information, as a preliminary step of the study, we propose the following definition with the intuitive meaning based on Figure 7.5(c).

$$F_{resp}(x) = TV((u(x, \lambda_{min}), \cdots, u(x, \lambda_{max}))), x \in \Omega \tag{7.3.1}$$

$$\text{where} \quad TV(y) = \sum_n |y_{n+1} - y_n|, y = (y_1, \cdots, y_n) \tag{7.3.2}$$

It can be seen in Figure 7.5(b) that F_{resp} is a map which is able to enhance eye structures from the background. As commonly used in image analysis, simply thresholding for F_{resp} can not yield satisfying results for an arbitrary image. Instead, the threshold T_{resp} is defined adaptively for each image using the histogram information of F_{resp}. The threshold is defined as the response level where the sharpest decrease of histogram bin occurs. Shown in Figure 7.5(d), the x-axis is the total variation level, y-axis is number of the pixel below the total variation level. The response level where the sharpest decrease occurs will be used as the threshold T^{ϵ}_{resp} for F_{resp}.

$$G(y) = |\{p : F_{resp}(P) \leq y, P \in \Omega\}|$$

$$T^{\epsilon}_{resp} = \arg \min_{\min F_{resp} < y < \max F_{resp}} \{y : G'(y) < \epsilon\}, \epsilon \sim 0$$

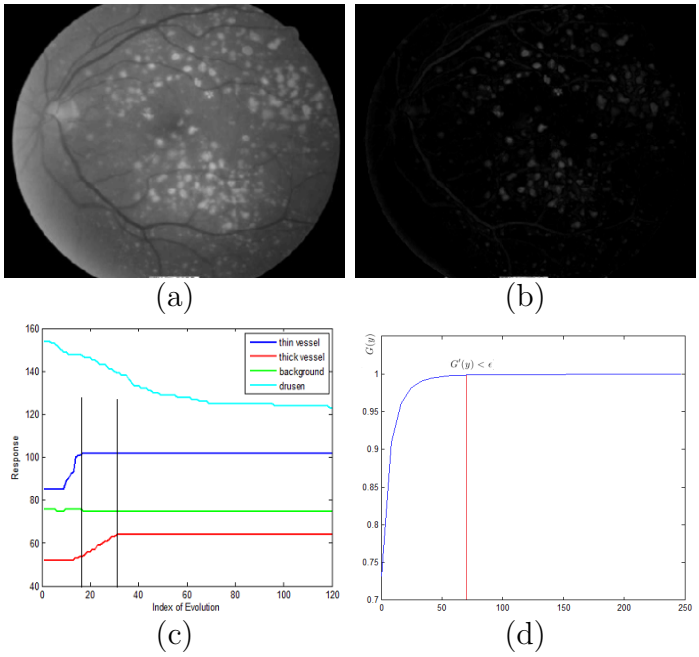


Figure 7.5: (a) Green channel of a color fundus image; (b) F_{resp}; (c) Multi-scale behavior of different structures, black vertical lines indicate sharp change point λ^*; x-axis is index i of λ_i, y-axis is the intensity level; (d) Cumulative distribution function $G(\cdot)$ of F_{resp}

Separate blood vessel

In Figure 7.5(b), it is shown that all eye structures are enhanced compared to the image background. In order to separate blood vessels from other structures, we use some prior knowledge of differences between these structures. Differences between eye structures include two aspects: (1) geometric shape: retinal vessels have elongated shapes, while optic disk, fovea and lesions such as drusen in ARMD are blob-like structures; (2) image intensity that optic disk, and lesions are in general brighter than the surroundings; fovea-macula region is darker than the surrounding and since blood vessels have lower reflectance than other retinal surfaces, they appear darker relative to the background. These structural differences provide useful prior knowledge for

vessel detection. Based on the difference in relative image intensity, it is expected that as λ increases to λ_{max}, the intensity value inside retinal vessels will increase as shown in Figure 7.5(c) while intensity values for the optic disk and drusen will decrease.

Blood vessel refinement: Connected component analysis

It can be observed after vessel separation that there can be certain connected components that are also classified as blood vessel, some are from fovea region, and some are random noisy points, which exist in the image background. In the binary mask for blood vessels, any connected component with area smaller than four pixels is removed. The value of component area is chosen manually here, under the assumption that the area of blood vessels should be larger than four pixels.

7.3.2 Vessel Detection: I, Global Performance

The method was firstly tested on a synthetic image which is a simplification of a real fundus image, as shown in Figure 7.6. Different levels of gaussian noise($\sigma = 0.01, 0.005, 0.01$) are added to the synthetic image, and the performance of our proposed methods are measured by True Positive Rate (TPR), False Positive Rate(FPR) and Accuracy(ACC). TPR is the ratio of the number of correctly classified vessel pixels to the number of total vessel pixels in the ground truth. FPR is the ratio of the number of incorrectly classified vessel pixels to the number of non-vessel pixels in the ground truth and ACC is the ratio of correctly classified pixels to the number of all pixels. Results are shown below in Figure 7.6.

Noise Level(σ)	TPR	FPR	ACC
0	1	0.0086	0.9920
0.005	1	0.0031	0.9971
0.01	0.9983	0.0027	0.9974

Figure 7.6: Left: Synthetic image($\sigma = 0.01$); Right: Detection performance

Performance Measurement

Database: Performance of our algorithm is validated on two public databases for vessel detection: STARE and DRIVE. (1) The STARE database [3] contains 20 images for blood vessel segmentation, ten of these contain pathology. The digitized slides are captured by a TopCon TRV-50 fundus camera at 35 degree field of view. The slides were digitized to 605×700 pixels, 8 bits per color channel. The approximate diameter of the FOV is 650×500pixels. Two observers manually segmented all the images. The first observer segmented many more of the thinner vessels than the first one. Performance is computed using the segmentation of the second observer as the ground truth; (2) The DRIVE(Digital Retinal Images for Vessel Extraction)[82] consists of 40 color fundus photographs. The photographs were obtained from a diabetic retinopathy screening program in the Netherlands. Of the 40 images in the database, 7 contain pathology, namely exudates, hemorrhages and pigment epithelium changes. The images were acquired using a Canon CR5 non-mydriatic 3-CCD camera with a

45 degree field of view(FOV). Each image is captured using 8bits per color plane at 768×584pixels. The FOV of each image is circular with a diameter of approximately 540 pixels. The set of 40 images was divided into a test and training set both containing 20 images.

Property of database: To understand the vessel compositions in the two databases, we calculated the ratio of vessel for each fundus image based on the manual segmentation results provided by each database.

$$\text{Vessel Ratio} = \frac{\text{Area(Vessel)}}{\text{Area(ROI)}} \tag{7.3.3}$$

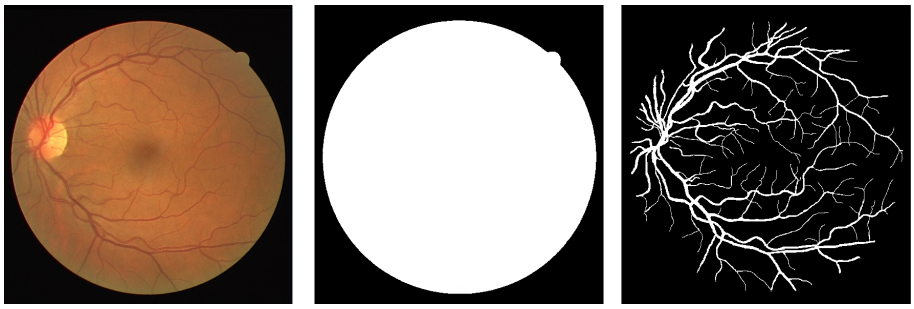


Figure 7.7: Left to right: Image from DRIVE, Region of Interest(ROI), Vessel Region

Table 7.1: Vessel Ratio(VR) distribution

VR	STARE-1	STARE-2	STARE-3	DRIVE-1	DRIVE-2	DRIVE-3
Mean	14.8%	10.4%	10.8%	12.7%	12.3%	12.5%
Std	4.2 %	1.9%	1.0%	1.2%	1.4%	1.8%

In Table 7.1, STARE-1, STARE-2, STARE-3 is based on the vessel segmentation

results from two manual segmentation results and one automatic method by Hoover [3]. DRIVE-1, DRIVE-2, DRIVE-3 is based on one manual segmentation for training set and two manual segmentation results for testing set. Each dataset has 20 images. The distributions of vessel ratio distribution within each database is shown in Figure 7.8. The first manual segmentation for STARE tends to segment more thinner vessels, and the first box-plot has higher mean and variance of blood vessel ratio.

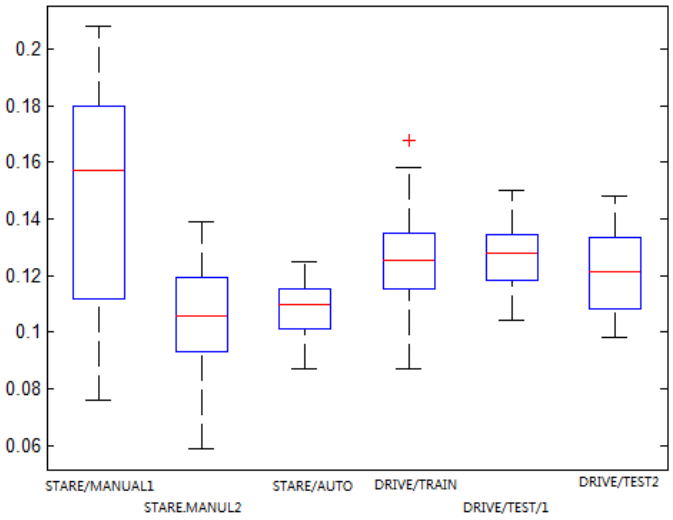


Figure 7.8: Boxplot of vessel ratio distribution within each dataset

Performance measurement: In the previous literature [3], comparisons of different methods are based on True Positive Rate (TPR), False Positive Rate(FPR) and Accuracy(ACC). Be more specific, denote the image region by Ω, $A := \{P \in \Omega | P \text{ is detected as vessel}\}$, $B := \{P \in \Omega | P \text{ is vessel in ground truth}\}$. $A^C = \Omega - A, B^C = \Omega - B$. $|S|$ denotes the cardinal of any given set S.

$$TPR = \frac{|A \cap B|}{|B|}, FPR = \frac{|A \cap B^C|}{|B|}, ACC = \frac{|A \cap B| + |A^C \cap B^C|}{|\Omega|}$$

$$TFR = \frac{|A^C \cap B^C|}{|B^C|}, FTR = \frac{|A \cap B^C|}{|B^C|}, Precision = \frac{|A \cap B|}{|A|}$$

Remark 7.3.1. TPR, FPR, ACC etc. are global measurements for segmentation performance. To interpret performance measures in a proper way, it is important to also consider the inherent properties of images. The reason can be explained as follows. From Table 7.1, the average blood vessel ratio across the two databases is around 12%. Given any retinal fundus image, denote $TPR : t$, $TFR : f$, and the ratio of blood vessel is p, here $p = |B|/|\Omega| \sim 12\%$, then

$$ACC = \frac{|A \cap B|}{|B|}\frac{|B|}{|\Omega|} + \frac{|A^C \cap B^C|}{|B^C|}\frac{|B^C|}{|\Omega|} = p \cdot f + (1-p) \cdot g$$

Here p and $(1-p)$ can be regarded as the weight for TPR and TFR respectively, and the ratio of the two weights is $(\frac{1}{p}-1) \sim 7.3$. It means that ACC will be weighted more on TFR, implying the accuracy detecting non-vessel point can have stronger effects on ACC value. Assume $p = 12\%$, if all pixels within ROI are labeled as non-vessel, *i.e.*$f = 0, g = 1$, then $ACC = 1 - P = 88\%$. In addition, $(f, g) = (70\%, 95\%)$ and $(f, g) = (66\%, 95.5\%)$ both have $ACC = 92\%$, however when TPR is decreased by 4%, it only requires the improvement of TFP by 0.5%. It is worthwhile to point out that in order to measure the performance of blood vessel detection, these measures especially index ACC should be used with caution to avoid bias introduced by the inherent properties of images.

Vessel detection: Global Performance

Results of our method are compared with several state of the art methods using the STARE[3] databases. In the comparison, the first manual mask with fewer narrow vessels is used as the ground truth. The performance measures are shown in Table 7.2 below. The results of other methods are from [3, 4, 83, 84, 85]. From the results, our

Table 7.2: Vessel Detection Results for STARE database

Method	TPR	FPR	ACC
2nd Human observer	0.8949	0.0610	0.9354
Hoover	0.6751	0.0433	0.9267
Soares	0.7165	0.0252	0.9480
Mendonca	0.6996	0.0270	0.9440
MF-FDOG	0.7177	0.0247	0.9484
Our method	0.7001	0.0212	0.9493

method has best FPR, ACC compared to previous methods, though according to the analysis in Section 7.3.2, it may not necessarily mean that our method is superior. Depending on different research purposes, we may need to balance between TPR and FPR. An example of our segmentation result is show in Fig 7.9.

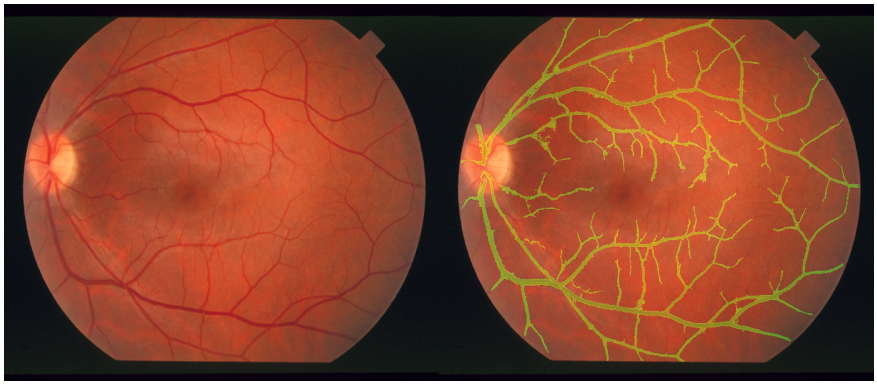


Figure 7.9: Left: image in STARE database; right: with vessel mask

7.3.3 Vessel Detection: II, Local Performance

Choice of subregions

Necessity of local comparison: To measure blood vessel segmentation performances, people used global indices (e.g. TPR, FPR, ACC etc.) as shown before. In fact, each individual segmentation method has its own advantages and disadvantages. In real computer aided diagnosis, instead of achieving better global performance, it can be important to achieve high accuracy in each local region for the assistance of diagnosis. For this purpose, we conducted a study evaluating performance of blood vessel detection within subregions of retinal images.

Subregion: For each retinal fundus image, a region of size 256×256(pixel×pixel) is cropped randomly. *Randomly* means that the region is cropped without knowing the vessel detection performance of each method. It is cropped in such way that the bias of method preference will not be introduced in comparisons. Regions are selected from 5 groups: 1, contain optic disk; 2, contain fovea; 3, contain lesion; 4, contain blood vessel only; 5, low contrast between background and blood vessel.

As stated in Section 7.2.2, optic disk, fovea, lesions are structures that exist commonly within a retinal image. Low contrast and inhomogeneous illuminations are conditions that exist in real retinal fundus images. How the method performs in these situations can be important and useful to evaluate the robustness of the specific method. In our preliminaries studies, we found that different methods can have different performance under different conditions. In my point of view, visual comparisons are important,

as it can evaluate performance locally in a more comprehensive way. Examples of categories of subregions is shown in Figure 7.10.

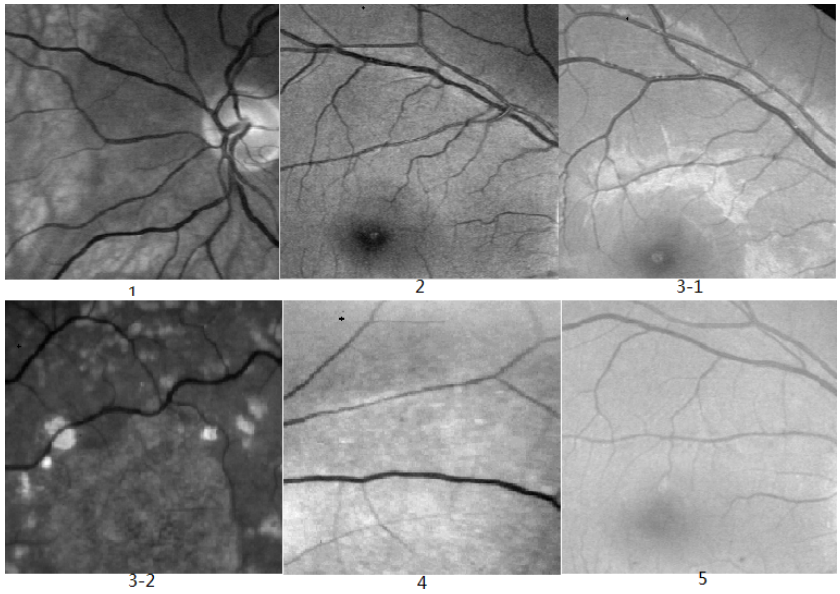


Figure 7.10: (1)Optic Disk;(2) Fovea;(3-1,3-2) Lesions;(4) Vessel Only;(5) Low Contrast

Comparisons study based on Database STARE

Methods for comparison: For STARE database, methods used for comparisons include 1) Manual Segmentation by Adam Hoover[3]; 2) Our Algorithm; 3)Frangi's vessleness measurement [26]; 4) Local entropy Thresholding[86]; 5) Matched Filter Response[3]. 1), 5) are chosen since the segmentation results are provided on the website of STARE database [3]; 3) Frangi's method of vesselness is a widely used method especially for multi-scale analysis of blood vessel. In addition, the source code of Frangi's method is available [87]; The source code of 4) is also available [86] and claimed to be superior to many current methods. Comparisons of detection performance are analyzed below.

1. **Subregion containing Optic Disk**

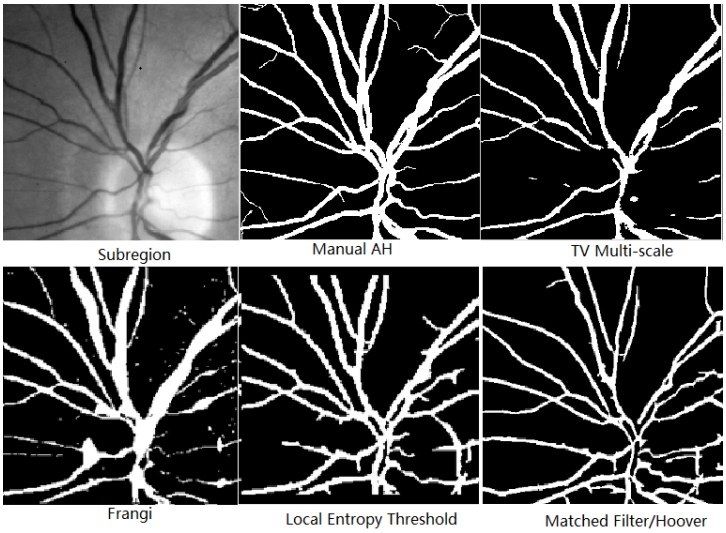


Figure 7.11: From left to right: Example of subregion, Segmentation using Method 1-5

Remark 7.3.2. Evaluations of methods are focused on the last four figures in Figure 7.11. *Method 2* is the only method that can rule out optic disk pixels successfully. The relative width of vessels are also maintained in the segmentation. However since the optic disk region is removed, the blood vessels that cross the OD are not detected; *Method 3*, the segmentation is generated by thresholding on the vesselness proposed in [26]. The threshold is chosen to be low(e.g. $> 0.01 \max(vesselness)$) such that most of vessels are detected. It can be observed that there are two obvious disadvantages. A, Frangi's vesselness map is scale-dependent. Considering the variations of vessel width within retinal images, the performance of vessel detection can be distorted if the scale parameter i.e. σ (used in gaussian filtering) is not chosen properly. It can be seen that the vessels detected are wider than the actual vessel width, due to the

effects of gaussian blurring. On the right of the result, two branches of vessels are merged together as a wider vessel. *B*, some of pixels on the boundary of optic disk are labeled as vessel; *Method 4*, the detection of blood vessel by local entropy thresholding is visually acceptable. The disadvantages include *A*, part of OD boundary is detected as vessel; *B*, the vessel width in the segmented image seems uniform, i.e. fails to capture the variations of vessel width; *Method 5* performs similar with *Method 4*.

2. **Subregion containing fovea**

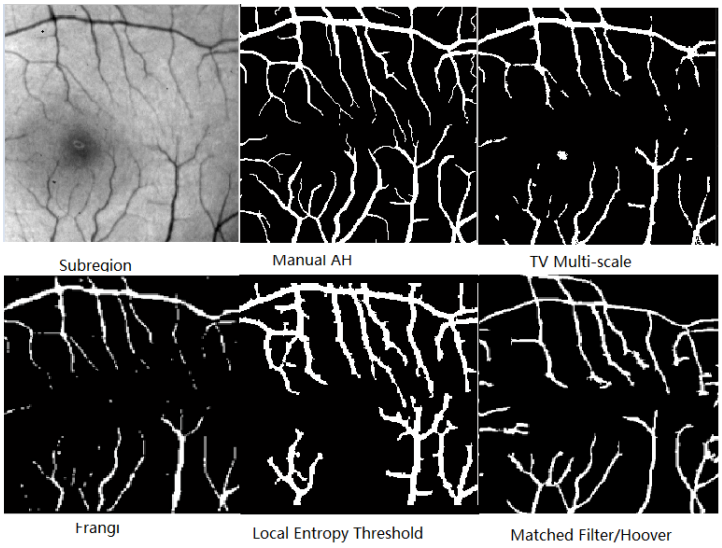


Figure 7.12: From left to right: Example of subregion, Segmentation using Method 1-5

Remark 7.3.3. In general, *Method 2,3,5* have similar performances in vessel detection. In *Method 2*, one small spots around the fovea region is detected as fovea region is considered also relatively darker than the surroundings; In *Method 3*, the detected vessels are blurred, especially the widest vessels in the

subregion; In *Method 5*, the main vessel structures are detected, however with relatively uniform vessel width. For *Method 4*, there is one main branch at the bottom that is not detected. The missed branch has similar intensity properties with others, and it is unclear the reason the branch is not detected. In addition, the vessel widths in the results barely vary.

3. **Subregion containing lesion**

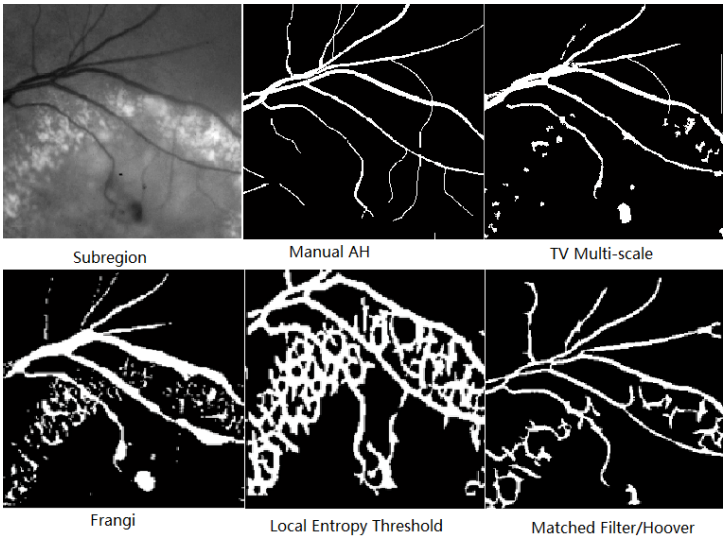


Figure 7.13: From left to right: Example of subregion, Segmentation using Method 1-5

Remark 7.3.4. In Figure 7.13, detection performances of *Method 2-5* are all affected by the existence of lesion structures. *Method 2* is least affected by the existence of lesion, as bright regions can be removed by tracking the multi-scale behavior of each pixel; *Method 3* is affected by the lesions in two aspects:A, the widths of vessel are highly distorted by the blurring effect and the existence of lesions; B, most of the lesion regions are detected as vessels. The possible

reason is that the lesion structures also reveal some vessel properties resulting in high vesselness values; *Method 4* performed worst compared with other four methods. It can be possible that the existence of lesions makes the choice of local entropy threshold harder; *Method 5*, the vessel structures within the lesions are also labeled vessel with uniform widths.

4. **Subregion with vessel only**

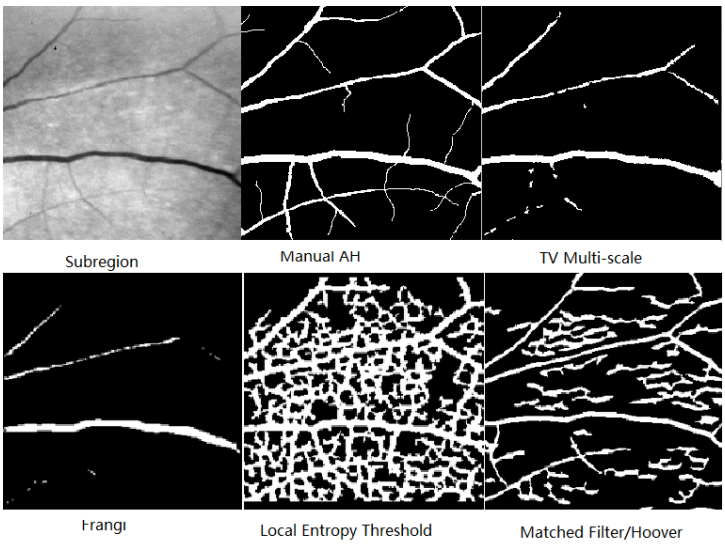


Figure 7.14: From left to right: Example of subregion, Segmentation using Method 1-5

Remark 7.3.5. By visual comparison, *Method 2,3* have better performance in blood vessel detection than *Method 4,5*. As pointed in previous examples e.g. Figure 7.11, 7.12 etc., *Method 3* blurred the main vessel structures, and missed the thin vessel branches. *Method 4,5* are affected by the inhomogeneities in the background of the image. For *Method 4*, it implies that the performance of vessel detection is highly dependent on the local image information. The

small inhomogeneity within a local region can affect the accuracy of vessel detection. In *Method 5*, although the inhomogeneity in the background is small, as all regions within the images are used for filter-matching, it resulted in vessel structures detected in the inhomogeneous background.

5. **Subregion with low image contrast**

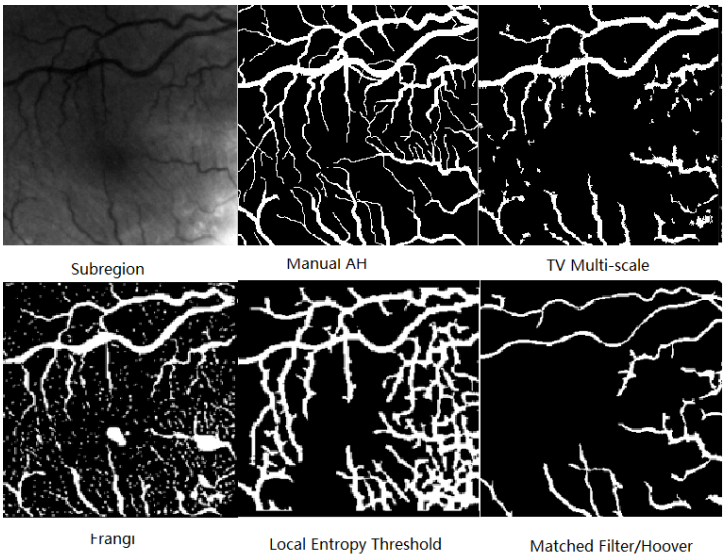


Figure 7.15: Left to right: Subregion, Results of Method 1-5

Remark 7.3.6. The case of low image contrast can be divided into two situations i.e. intensity of blood vessel and background are low(or high) and close. Our example is the first situation. *Method 2* captured the main vessel structures under the low contrast, partly missing some small vessel branches. For example the upper right of the figure; For *Method 3*, the noisy dots detected in the background can be removed by further processing(for example connected component analysis). The vessel structures are blurred, for example one vessel branch at

the bottom-right of image is blurred into blob; *Method 4*, since the bottom-right corner of the image is brighter than the surroundings, the detection of vessel at this corner labeled the vessel pixel false positively; *Method 5*, due to the low contrast of the image, large parts of the vessels on the left half of the image are missed.

In summary, based on visual comparisons for five categories of subregions, we have the following conclusions: *First*, by using the prior knowledge of retinal fundus image properties, our algorithm has consistently better performance in blood vessel detection compared with other four methods; *Second*, the vesselness, a scalar proposed is highly dependent on the proper choice of scale parameter σ. Considering the variations of blood vessel width even within in a retinal image, using the vesselness map to directly detect blood vessels can pose problems on vesselness thresholding. In fact, the improper choice of σ can distort the vessel profiles; *Third*, the method of local entropy thresholding is highly dependent on local image information, for example existence of local image inhomogeneity or existence of lesion structures can result in false positive pixels; *Fourth*, Hoover method is based on the idea of matched filter. The results have uniform vessel width within an image mainly due to the global choice of matched filter for a single image. This limitation can result in failure to capture vessel width variations. The performance of this method can also be affected by the low contrast of the image, resulting in missed blood vessels in low contrast regions.

7.3.4 Drusen Detection

The evaluation of drusen detection is based on a clinical dataset, which contains images obtained from ARMD patients. Cases with drusen presenting in the images are used to evaluate the drusen detection performance.

Method

The method of drusen detection follows the steps in blood vessel detection. The difference between vessel detection and drusen detection is that, drusen are considered brighter than the surroundings, while vessels are darker than the surroundings. Based on this property, after eye structures are extracted from the image background, vessel structures and drusen structures can be differentiated as shown in Figure 7.5.

The boundary of a druse is assumed to be closed and a druse is a compact region. A connected component with area smaller than 4 pixels and curve perimeter larger than half of the perimeter of Ω are removed. The two thresholds are used uniformly. The importance of region detection include, *First*, area of region can be used to remove regions that are impossible to be drusen; *second*, area of drusen provides a quantitative way to describe sizes of drusen, and can be used to grade the severity of ARMD. As mentioned in [78], drusen may have a similar appearance to other lesions, for example cotton wool spots, which appears as puffy white patches on the retina; some other drusen will also be confused with normal retina structures, for example, some background patterns caused by choroidal blood vessels. The complexity of

fundus images makes drusen detection challenging, prior knowledge of structures can help with the drusen detection. Three cases from this clinical dataset are shown in Figure 7.16, and binary masks for drusen are overlaid on original color images.

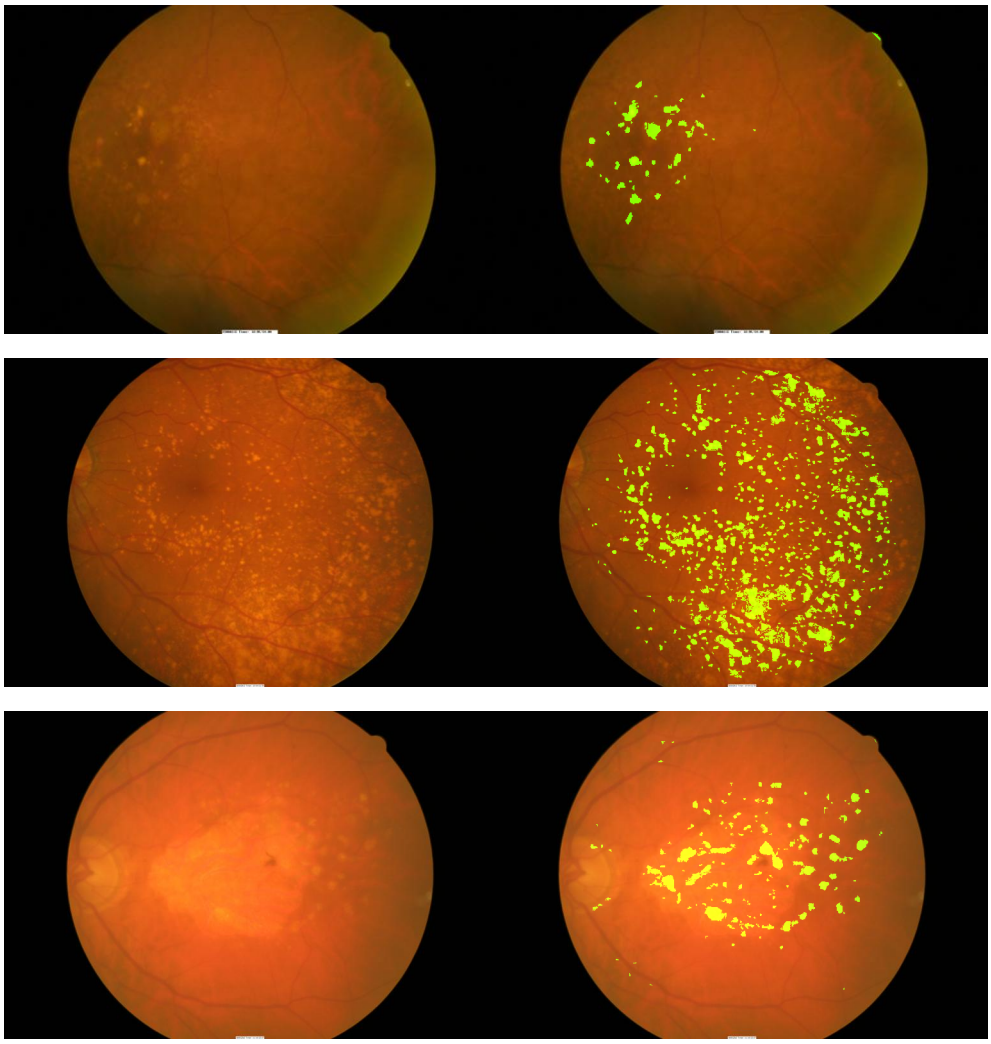


Figure 7.16: Left: Color fundus image; Right: Color image overlaid with drusen mask

Remark 7.3.7. Fundus images used in general clinical studies are complex, and the three cases in Figure 7.16 are typical ones. The first case has nonuniform illumination, the second case has relatively large region of lesions and the third case has hemor-

rhage as well as drusen. It shows that by using the preprocessing method to correct nonuniform illumination, drusen regions can still be detected correctly; considering the various sizes of drusen existing in the second case, drusen detection can still attain a fair performance. Since the multi-scale TV/L_1 model is based on local changes of structures, drusen overlaid on the hemorrhage region can also be detected, which for many other methods detect of the hemorrhage region will fail. Unfortunately, compared to blood vessel detection, manual segmentation results for drusen are not publicly available. Results shown are preliminary visual inspections.

In addition to whole region drusen detection, in ARMD grading system(e.g. [88]), the fovea center will be detected first, and three circular regions centered at the fovea center with increasing radius will be used as ROI. The number of drusen within each circle will be counted and used in the grading system shown as in Figure 7.17 in [88]. In our experiment, the algorithm is designed to work on rectangular regions in stead

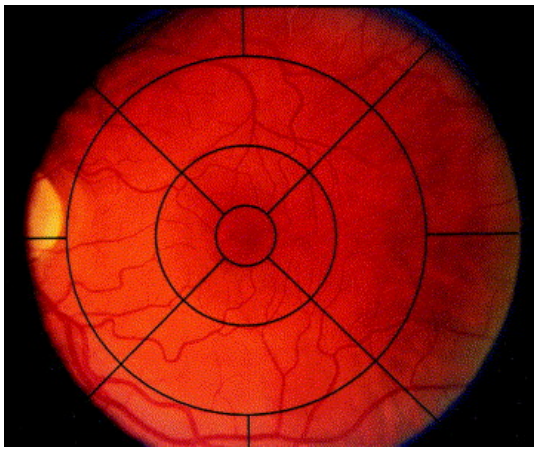


Figure 7.17: Example of ARMD Grading System

of circular regions. Instead of analysis within circular region, a square ROI centered

at fovea is cropped such that the side length equals to 1.2·diameter(second circle in Figure 7.17). Examples of drusen detection within the square are shown below.

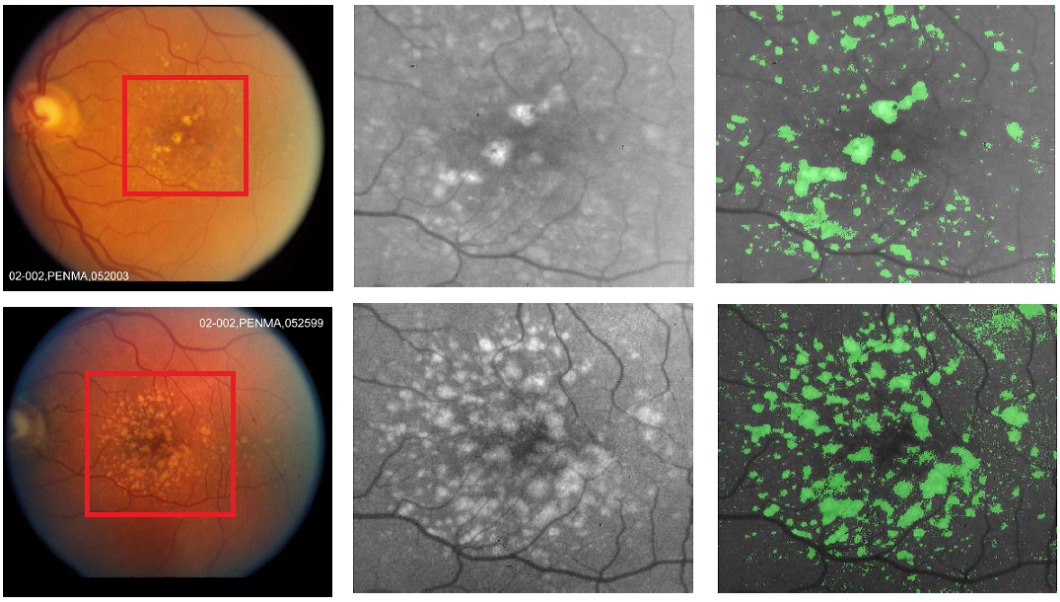


Figure 7.18: Drusen detection within fovea-macula region I.

7.3.5 Discussion

Given the existence of different algorithms proposed for retina fundus image analysis, the examples shown in this chapter belong to the branch of multi-scale analysis. The advantages of this model lie in two facts: first, the multi-scale TV/L_1 model is known in theory to be better at capturing the geometric properties of structures. This multi-scale model is able to capture the local pattern of image structures better; second, prior knowledge of vessel-drusen, eye structure-background differences is utilized in the analysis of the fundus image. Instead of simply proposing a new method, we use this example so as to understand this mathematical model in a better sense and bridge

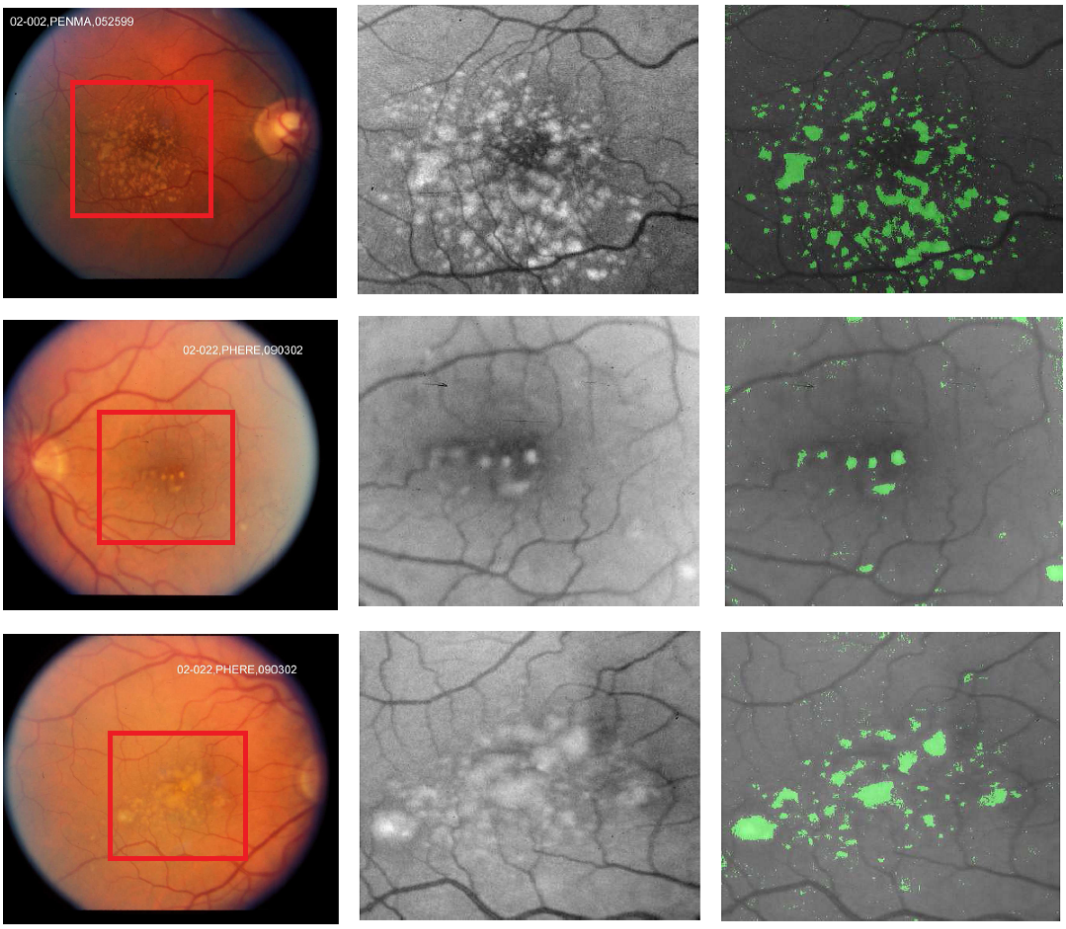


Figure 7.19: Drusen detection within fovea-macula region II.

the gap between the theoretical model and the applications in real image analysis.

Bibliography

[1] A. Bird, N. Bressler, S. Bressler, I. Chisholm, G. Coscas, M. Davis, P. De Jong, C. Klaver, B. Klein, R. Klein, *et al.*, "An international classification and grading system for age-related maculopathy and age-related macular degeneration," *Survey of ophthalmology*, vol. 39, no. 5, pp. 367–374, 1995.

[2] T. E. Clemons, L.-Y. Lee, E. Y. Chew, R. C. Milton, S. B. Bressler, and R. Klein, "A simplified severity scale for age-related macular degeneration," *Arch Ophthalmol*, vol. 123, pp. 1570–1574, 2005.

[3] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 203–210, 2000.

[4] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 9, pp. 1214–1222, 2006.

[5] W. Andrew, "Scale-space filtering," *Proceeding IJCAI'83 Proceedings of the Eighth international joint conference on Artificial intelligence*, vol. 2, 1983.

[6] T. Lindeberg, *Scale selection for differential operators*. Springer, 1994.

[7] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[8] A. Fick, "Ueber diffusion," *Annalen der Physik*, vol. 170, no. 1, pp. 59–86, 2006.

[9] M. A. Piech, "Decomposing the laplacian," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 8, pp. 830–831, 1990.

[10] J. J. Clark, "Authenticating edges produced by zero-crossing algorithms," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 1, pp. 43–57, 1989.

[11] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[12] B. Merriman, J. Bence, and S. Osher, *Diffusion Generated Motion by Mean Curvature*. Dept. of Mathematics, University of California, Los Angeles, 1992.

[13] J. Weickert, "Theoretical foundations of anisotropic diffusion in image processing," *Computing*, vol. 11, pp. 221–236, 1996.

[14] A. Chambolle, "An algorithm for mean curvature motion," *Interfaces and free Boundaries*, vol. 6, pp. 195–218, 2004.

[15] T. Brox and J. Weickert, "A TV flow based local scale measure for texture discrimination," *European Conf. on Computer Vision*, vol. 3022, pp. 578–590, 2004.

[16] J. Weickert, "A review of nonlinear diffusion filtering," *Scale-space theory in computer vision*, pp. 1–28, 1997.

[17] J. Weickert and B. Benhamouda, "A semidiscrete nonlinear scale-space theory and its relation to the perona-malik paradox," in *F. Solina (Ed.), Advances in computer vision*, Citeseer, 1997.

[18] S. Kichenassamy, "The perona–malik paradox," *SIAM Journal on Applied Mathematics*, vol. 57, no. 5, pp. 1328–1342, 1997.

[19] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel, "Axioms and fundamental equations of image processing," *Archive for rational mechanics and analysis*, vol. 123, no. 3, pp. 199–257, 1993.

[20] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[21] L. Rudin, S. Osher, C. Inc, and C. Santa Monica, "Total variation based image restoration with free local constraints," *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, vol. 1, 1994.

[22] T. F. Chan, G. H. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 20, no. 6, pp. 1964–1977, 1999.

[23] P. Blomgren, T. F. Chan, P. Mulet, and C.-K. Wong, "Total variation image restoration: numerical methods and extensions," in *Image Processing, 1997. Proceedings., International Conference on*, vol. 3, pp. 384–387, IEEE, 1997.

[24] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[25] J. Weickert, B. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 398–410, 1998.

[26] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted InterventationMICCAI98*, pp. 130–137, Springer, 1998.

[27] L. M. Bruce and R. R. Adhami, "Classifying mammographic mass shapes using the wavelet transform modulus-maxima method," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 12, pp. 1170–1177, 1999.

[28] G. Gimel, "What is texture@ONLINE," June 2009.

[29] M. Tuceryan and A. K. Jain, "Texture analysis," *Handbook of pattern recognition and computer vision*, vol. 276, 1993.

[30] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, no. 6, pp. 460–473, 1978.

[31] A. Witkin, "Scale-space filtering: A new approach to multi-scale description," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, vol. 9, 1984.

[32] A. Rosenfeld *et al.*, *Multiresolution image processing and analysis*. Springer-Verlag New York:, 1984.

[33] A. Klinger, "Patterns and search statistics," *Optimizing Methods in Statistics*, vol. 3, pp. 303–337, 1971.

[34] P. J. Burt, "Fast filter transform for image processing," *Computer graphics and image processing*, vol. 16, no. 1, pp. 20–51, 1981.

[35] A. Chehikian and J. Crowley, "Fast computation of optimal semi-octave pyramids," *7th SCIA, Aalborg*, 1991.

[36] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693, 1989.

[37] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of l2 (r)," *Trans. Amer. Math. Soc*, vol. 315, no. 1, 1989.

[38] G. Sheikholeslami, S. Chatterjee, and A. Zhang, "Wavecluster: A multi-resolution clustering approach for very large spatial databases," in *Proceedings of the international conference on very large data bases*, pp. 428–439, INSTITUTE OF ELECTRICAL & ELECTRONICS ENGINEERS, 1998.

[39] B. Schiele and J. Crowley, "Object recognition using multidimensional receptive field histograms," *Computer Vision ECCV'96*, pp. 610–619, 1996.

[40] T. Lindeberg, "On scale selection for differential operators," in *Proceedings of the scandinavian conference on image anlaysis*, vol. 2, pp. 857–857, 1993.

[41] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150–1157, Ieee, 1999.

[42] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[43] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM*, vol. 9, no. 1, pp. 84–97, 1962.

[44] S. Twomey, "On the numerical solution of fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," *Journal of the ACM*, vol. 10, no. 1, pp. 97–101, 1963.

[45] B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," *JOSA*, vol. 62, no. 4, pp. 511–518, 1972.

[46] L. I. Rudin, "Images, numerical analysis of singularities and shock filters," 1987.

[47] T. Chan and E. S., "Aspects of total variation regularized L^1 function approximation," *Siam Journal of Applied Mathematics*, vol. 65, no. 5, pp. 1817–1837, 2005.

[48] F. Andreu-Vaillo, V. Caselles, and J. Mazón, *Parabolic quasilinear equations minimizing linear growth functionals*. Birkhauser, 2004.

[49] L. I. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *International Conference on Image Processing, 1994*, vol. 1, pp. 31–35, IEEE, 1994.

[50] M. Nikolova, "A variational approach to remove outliers and impulse noise," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 99–120, 2004.

[51] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Illumination normalization for face recognition and uneven background correction using total variation based image models," in *Computer Vision and Pattern Recognition, 2005*, vol. 2, pp. 532–539, IEEE, 2005.

[52] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1519–1524, 2006.

[53] W. Yin, D. Goldfarb, and S. Osher, "A comparison of three total variation based texture extraction models," *Journal of Visual Communication and Image Representation*, vol. 18, no. 3, pp. 240–252, 2007.

[54] W. Yin, T. Chen, S. X. Zhou, and A. Chakraborty, "Background correction for cdna microarray images using the tv+ l1 model," *Bioinformatics*, vol. 21, no. 10, pp. 2410–2416, 2005.

[55] T. Chen, T. S. Huang, W. Yin, and X. S. Zhou, "A new coarse-to-fine framework for 3d brain mr image registration," in *Computer Vision for Biomedical Image Applications*, pp. 114–124, Springer, 2005.

[56] W. Yin, D. Goldfarb, and S. Osher, "The Total Variation Regularized L^ 1 Model for Multiscale Decomposition," *Multiscale Modeling and Simulation*, vol. 6, no. 1, p. 190, 2007.

[57] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse problems*, vol. 19, pp. S165–S187, 2003.

[58] D. M. Strong, *Adaptive total variation minimizing image restoration.* Department of Mathematics, University of California, Los Angeles, 1997.

[59] P. Jones and T. Le, "Local scales and multiscale image decompositions," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 371–394, 2009.

[60] Z. Sbeh, L. Cohen, G. Mimoun, and G.Coscas, "A new approach of geodesic reconstruction for drusen segmentation in eye fundus images," *IEEE Transaction on medical imaging*, vol. 20, no. 12, pp. 1321–1333, 2001.

[61] F. Andreu, C. Ballester, V. Caselles, and J. Mazon, "Minimizing total variation flow," *Differential and Integral Equations*, vol. 14, no. 3, pp. 321–360, 2001.

[62] G. Bellettini, V. Caselles, and M. Novaga, "The total variation flow in N," *Journal of Differential Equations*, vol. 184, no. 2, pp. 475–525, 2002.

[63] F. I. Karahanoglu, I. Bayram, and D. Van De Ville, "A signal processing approach to generalized 1-d total variation," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5265–5274, 2011.

[64] X. Feng and A. Prohl, "Analysis of total variation flow and its finite element approximations," *ESAIM-Mathematical Modelling and Numerical Analysis*, vol. 37, no. 3, p. 533, 2003.

[65] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.

[66] J. Weickert, "Coherence-enhancing diffusion of colour images," *Image and Vision Computing*, vol. 17, no. 3, pp. 201–212, 1999.

[67] P. Mrázek and M. Navara, "Selection of optimal stopping time for nonlinear diffusion filtering," *International Journal of Computer Vision*, vol. 52, no. 2-3, pp. 189–203, 2003.

[68] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Estimation of optimal pde-based denoising in the snr sense," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2269–2280, 2006.

[69] A. Kuijper, "Geometrical pdes based on second-order derivatives of gauge coordinates in image processing," *Image and Vision Computing*, vol. 27, no. 8, pp. 1023–1034, 2009.

[70] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 679–698, 1986.

[71] P. S. Mettu, A. R. Wielgus, S. Ong, and S. W. Cousins, "Retinal pigment epithelium response to oxidant injury in the pathogenesis of early age-related macular degeneration," *Molecular Aspects of Medicine*, 2012.

[72] W. S. Grizzard, D. Arnett, and S. L. Haag, "Twin study of age-related macular degeneration," *Neuro-Ophthalmology*, vol. 10, no. 5, pp. 315–322, 2003.

[73] R. Klein, M. D. Davis, Y. L. Magli, P. Segal, B. Klein, L. Hubbard, *et al.*, "The wisconsin age-related maculopathy grading system.," *Ophthalmology*, vol. 98, no. 7, pp. 1128–1134, 1991.

[74] C. A. Curcio, N. E. Medeiros, and C. L. Millican, "The alabama age-related macular degeneration grading system for donor eyes.," *Investigative ophthalmology & visual science*, vol. 39, no. 7, pp. 1085–1096, 1998.

[75] B. Remeseiro, N. Barreira, D. Calvo, M. Ortega, and M. Penedo, "Automatic drusen detection from digital retinal images: Amd prevention," *Computer Aided Systems Theory-EUROCAST 2009*, pp. 187–194, 2009.

[76] A. Thaibaoui, A. Raji, and P. Bunel, "A fuzzy logic approach to drusen detection in retinal angiographic images," *International Conference on Pattern Recognition*, vol. 4, p. 4748, 2000.

[77] J. J. Leandro, R. M. Cesar Jr, and H. F. Jelinek, "Blood vessels segmentation in retina: Preliminary assessment of the mathematical morphology and of the wavelet transform techniques," in *Computer Graphics and Image Processing, 2001 Proceedings of XIV Brazilian Symposium on*, pp. 84–90, IEEE, 2001.

[78] L. Brandon and A. Hoover, "Drusen detection in a retinal image using multi-level analysis," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pp. 618–625, Springer, 2003.

[79] D. S.Shin, N. B.Javornik, and J. W.Berger, "Computer-assisted interactive fundus image processing for macular drusen quantitation," *Ophthalmology*, vol. 106, pp. 1119–1125, 1999.

[80] Y. Zheng, M. Grossman, S. P. Awate, and J. C. Gee, "Automatic correction of intensity nonuniformity from sparseness of gradient distribution in medical images," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pp. 852–859, Springer, 2009.

[81] D. Goldfarb and W. Yin, "Parametric maximum flow algorithms for fast total variation minimization," *Rice University CAAM Technical Report TR07-09*, 2007.

[82] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 4, pp. 501–509, 2004.

[83] A. M. Mendona and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Transactions on Medical Imaging*, vol. 25, pp. 1200–1213, 2006.

[84] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Transactions on Medical Imaging*, vol. 8, pp. 263–269, 1989.

[85] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of gaussian," *Comput. Biol. Med.*, vol. 40, pp. 438–445, April 2010.

[86] T. Chanwimaluang and G. Fan, "An efficient blood vessel detection algorithm for retinal images using local entropy thresholding," in *Circuits and Systems, 2003. ISCAS'03. Proceedings of the 2003 International Symposium on*, vol. 5, pp. V–21, IEEE, 2003.

[87] D.-J. Kroon, "Matlab code for frangi vesselness."

[88] "The age-related eye disease study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the age-related eye disease study report number 6," *The, Age-Related Eye Disease Study Research Group, American Journal of Ophthalmology*, vol. 132, no. 5, pp. 668–681, 2001.