

ABSTRACT

Title of dissertation: NOVEL INTEGRO-DIFFERENTIAL
SCHEMES FOR MULTISCALE
IMAGE REPRESENTATION
Prashant Athavale, Doctor of Philosophy, 2009

Dissertation directed by: Professor Eitan Tadmor
Department of Applied Mathematics
& Statistics and Scientific Computation

Multiscale representation of a given image is the problem of constructing a family of images, where each image in this family represents a scaled version of the given image. This finds its motivation from biological vision studies. Using the hierarchical multiscale image representation proposed by Tadmor et. al. [32], an image is decomposed into sums of simpler ‘slices’, which extract more refined information from the previous scales. This approach motivates us to propose a novel *integro-differential equation* (IDE), for a multiscale image representation. We examine various properties of this IDE.

The advantage of formulating the IDE this way is that, although this IDE is motivated by variational approach, we no longer need to be associated with any minimization problem and can modify the IDE, suitable to our image processing needs. For example, we may need to find different scales in the image, while retaining or enhancing prominent edges, which may define boundaries of objects. We propose some edge preserving modifications to our IDE.

One of the important problems in image processing is deblurring a blurred image. Images get blurred due to various reasons, such as unfocused camera lens, relative motion between the camera and the object pictured, etc. The blurring can be modeled with a continuous, linear operator. Recovering a clean image from a blurry image, is an ill-posed problem, which is solved using Tikhonov-like regularization. We propose a different IDE to solve the deblurring problem.

We propose hierarchical multiscale scheme based on (BV, L^1) decomposition, proposed by Chan, Esedoğlu, Nikolova and Alliney [12, 25, 3]. We finally propose another hierarchical multiscale representation based on a novel weighted (BV, L^1) decomposition.

NOVEL INTEGRO-DIFFERENTIAL SCHEMES
FOR MULTISCALE IMAGE REPRESENTATION

by

Prashant Athavale

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Professor Eitan Tadmor, Chair/Advisor
Professor Radu Balan
Professor Joel Cohen
Professor Kayo Ide
Professor David Jacobs
Professor Doron Levy

© Copyright by
Prashant Athavale
2009

Dedication

To my mother, Nirmala.

Acknowledgments

I owe my gratitude to all the people who have made this dissertation possible. There are many people I must thank in this acknowledgement.

First and foremost, I'd like to thank my advisor, Professor Eitan Tadmor for giving me an invaluable opportunity to work on this extremely interesting project. Despite being extremely busy, he spent enormous amount of time with me. It was a wonderful experience to work with him, and get to know his approach towards research. As a matter of fact, no amount of words will express my gratitude towards my advisor. I also thank my committee members: Prof. Radu Balan, Prof. Joel Cohen, Prof. Kayo Ide, Prof. David Jacobs and Prof. Doron Levy, for taking their precious time out during the Summer and being on my committee. I extend my gratitude towards Prof. Georg Dolzmann, whom I worked with for some time, during my first few years in the University of Maryland.

I had a great fortune to know many people, who enriched my educational experience. Unfortunately, I do not have enough space to list all my teachers who did their job beyond the call of duty. I must thank Prof. Ezzatollah Salari, from University of Toledo, who deepened my interest in image analysis and of course, Prof. Henry Wente, who has always served as one of my inspiration.

Table of Contents

List of Figures	vi
List of Abbreviations	vii
1 Introduction	1
1.1 What are images ?	1
1.2 Main problems in image processing	2
1.3 Multiscale representation	3
1.4 Multiscale representation with PDE and variational methods	5
1.4.1 Heat equation	5
1.4.2 Perona-Malik model	6
1.4.3 Catté, Lions, Coll, Morel model	7
1.4.4 Nordström model	7
1.4.5 Rudin, Osher, Fatemi model	8
1.5 Some properties of ROF decomposition	9
1.6 BV spaces	12
1.6.1 Definitions and general properties of BV space	12
1.6.2 Two types of convergence processes	14
1.6.3 Structure of BV functions	17
2 The Tadmor-Nezzar-Vese hierarchical multiscale representation	19
2.1 Introduction	19
2.2 The hierarchical (BV, L^2) decomposition	20
2.3 Initialization	21
2.4 Convergence results for the hierarchical multiscale expansion	22
2.5 General hierarchical decompositions	28
2.6 Euler-Lagrange equations	30
3 Integro-differential equation for multiscale image representation	32
3.1 Introduction	32
3.2 Motivation for the integro-differential equation	33
3.3 On the scaling function $\lambda(t)$	37
3.4 Modified IDE models	40
3.4.1 Perona-Malik models revisited	41
3.4.2 Modified IDE with filtered diffusion	44
3.4.3 IDE with tangential smoothing modification	48
4 Deblurring integro-differential equation model	50
4.1 Introduction	50
4.2 Deblurring based on hierarchical (BV, L^2) decomposition	50
4.3 A novel deblurring integro-differential equation	52

5	Hierarchical (BV, L^1) and <i>weighted</i> (BV, L^1) multiscale schemes	55
5.1	Introduction	55
5.2	Basic properties of (BV, L^1) model	57
5.3	Hierarchical (BV, L^1) image decomposition	62
5.4	Euler-Lagrange equations	63
5.5	Scale-space generated by hierarchical (BV, L^1) decomposition	64
5.6	Edge enhancing modification to (BV, L^1) decomposition	64
5.7	Weighted (BV, L^1) scheme	65
5.8	Euler-Lagrange equations	66
6	Numerical scheme	72
6.1	Introduction	72
6.2	Numerical scheme for hierarchical (BV, L^2) multiscale representation	72
6.3	Numerical scheme for integro-differential equations	74
6.4	Numerical scheme for deblurring integro-differential equations	77
6.5	Numerical scheme for hierarchical (BV, L^1) scheme	78
6.6	Numerical scheme for hierarchical weighted (BV, L^1) decomposition .	80
	Bibliography	82

List of Figures

1.1	Examples of digital images.	2
1.2	The problem of segmentation	3
1.3	The denoising problem	4
1.4	The deblurring problem	5
1.5	Solutions of the heat equation.	6
3.1	Result of the proposed integro-differential model.	36
3.2	Result of the integro-differential model with filtered diffusion.	45
3.3	Result of the integro-differential model for an MRI image.	46
3.4	Result of the integro-differential model with filtered diffusion for an MRI image.	47
3.5	Result of the integro-differential model with only tangential diffusion	49
4.1	Result of the deblurring integro-differential model	54
5.1	Result of the proposed (BV, L^1) multiscale model.	67
5.2	Result of the hierarchical (BV, L^1) multiscale model with diffusion controlling function.	68
5.3	Result of the <i>weighted</i> (BV, L^1) multiscale model for $p = 1.2$	69
5.4	Result of the <i>weighted</i> (BV, L^1) multiscale model for $p = 2$	70
5.5	Result of the hierarchical (BV, L^2) multiscale model.	71

List of Abbreviations and Notations

α	an algorithm specific denoising parameter
β	scaling constant used in the diffusion controlling function
δ	small positive constant
$\hat{\delta}$	function with a constant value δ
ϵ, ε	small positive constants
η	additive noise
ν	outward pointing normal to the boundary of $\Omega \subset \mathbb{R}^N$
λ	scaling parameter
$\lambda(t)$	scaling function
e	Euler's number $\approx 2.71828182845904523536 \dots$
f	given image
f_{Ω}	mean value of f on Ω
g	diffusion controlling function
h	grid size of the mesh covering the region Ω
t	artificial time
u	large scale image extracted from given f
v	the residual image $v := f - u$
$B_R(0)$	open ball of radius R , centered at 0
Du	distributional derivative
G_{σ}	Gaussian kernel with standard deviation σ
T	a linear, continuous blurring operator
Φ	the fundamental solution of the heat equation
χ_E	characteristic function on set E , $\chi_E(x) = 1$ if $x \in E$, else $\chi_E(x) = 0$
$\ f\ _{L^p(\Omega)}$	$(\int_{\Omega} f ^p)^{\frac{1}{p}}$
\mathbb{R}^N	Euclidean N - dimensional space
$\Omega \subset \mathbb{R}^N$	open set of an N - dimensional Euclidean space
$\bar{\Omega}$	closure of Ω
$\partial\Omega$	boundary of the region Ω
$BV(\Omega)$	the space of functions of bounded variations in Ω
$C(\Omega)$	space of scalar continuous functions in an open set $\Omega \subset \mathbb{R}^N$
$C_c(\Omega)$	space of scalar continuous functions with compact support in Ω
$\mathcal{M}(\Omega, \mathbb{R}^N)$	the space of all \mathbb{R}^N - values Borel measures
$L^p(\Omega)$	space of real valued functions f with $\ f\ _{L^p(\Omega)} < \infty$
$\text{supp}(f)$	closure of $\{x : f(x) \neq 0\}$
$\text{Per}(E, \Omega)$	perimeter of E in Ω
IDE	integro-differential equation
PM	Perona, malik model
ROF	Rudin, Osher, Fatemi model
TNV	Tadmor, Nezzar, Vese model

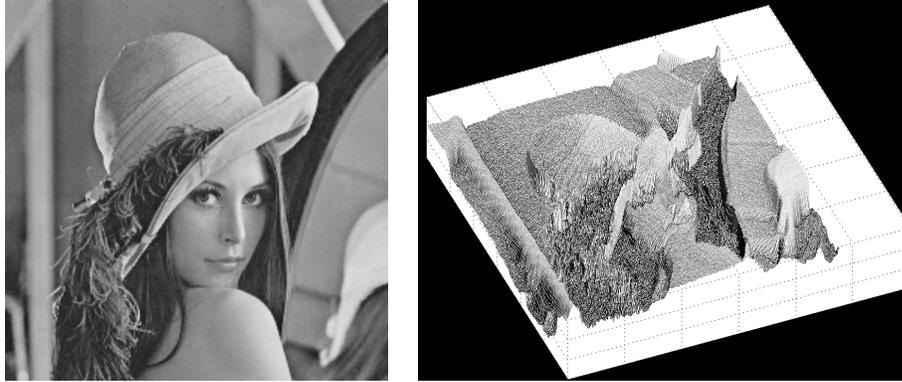
Chapter 1

Introduction

1.1 What are images ?

We are surrounded with images. Understanding images and interpreting them has been a crucial for survival. Mankind has been fascinated with the concept of storing images on paper, long before Ibn al-Haytham's 'Book of optics' [2] was written in 1021. Since then, Image capturing has been made very easy through the advent of digital cameras. Digital images are essentially the analogue images, sampled and quantized. If the images use b number of bits, then we have 2^b number of intensity levels available to digitally represent the image. The darkest spot in an image is assigned the value 'zero' and the brightest spot is assigned the value $2^b - 1$. For example, in case of an 8 bit image, there are 256 intensity levels. There could also be many channels used to describe an image. For example, in case of a typical color image there could be three channels, one for each colors: red, green, and blue.

For mathematical purposes, a black and white digital image is viewed as a function $f : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}$, sampled on a regular grid. Here, Ω is the region where the image is defined. The value of the function, $f(x)$, denotes the intensity of the image at the point $x \in \Omega$, see Figure 1.1.



(a)

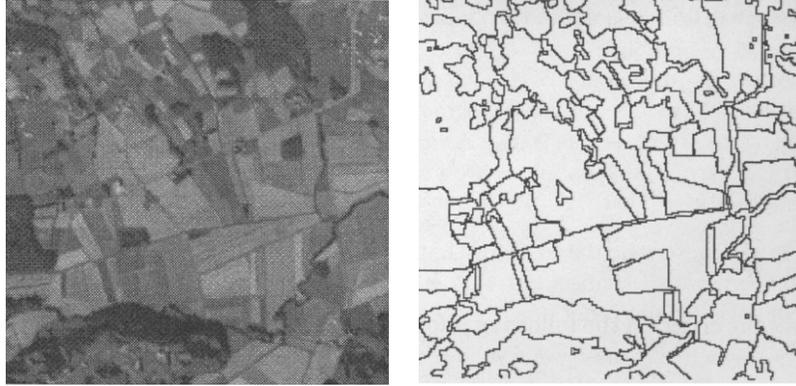
(b)

Figure 1.1: In (a) we can see the 24-bit image of Lenna, and in (b) we can see the same image at an angle, viewed as a graph of a function.

1.2 Main problems in image processing

Most problems in image processing fall under two broad categories: *image segmentation* and *image restoration*. *Image segmentation* is the problem of identifying constituent parts of the image. However, this definition is rather ambiguous. The meaning of segmentation mostly depend on the image at hand. For example, in Figure 1.2 we need to divide the image into its constituent parts depending on the difference in the texture.

Image restoration consists of denoising and deblurring. Additive noise, denoted by η , is inadvertently added to the image due to various reasons like limitations of the image capturing facilities or transmission losses, see Figure 1.3. Besides noise, the image could also be blurred due to reasons such as unfocused camera lens, relative motion between the camera and the object pictured. In Figure 1.4(a), we see a blurred image of Lenna and we seek to recover the Lenna, ideally as shown



(a)

(b)

Figure 1.2: The problem of segmentation: can we identify components in (a) and get a segmented image as in (b) ?

in Figure 1.4(b). This blurring is modeled by a linear, continuous blurring operator $T : L^2(\Omega) \mapsto L^2(\Omega)$, for example, convolution with a Gaussian kernel. Thus, we express the the observed image f as $f = TU + \eta$, where U is the clean image without blurring and noise. The recovery of this clean image U , given its blurred and noisy version f , is called *image restoration*. This is an ill-posed problem and can be solved by variational techniques using Tikhonov-like regularization [14, 36]. Several approaches like variational and PDE-based methods, filtering, stochastic modeling and wavelets were developed for solving these image processing problems [18, 17, 20, 21, 13].

1.3 Multiscale representation

In the absence of blurring, image-denoising naturally leads to *decomposition* of the given image f into a denoised part U_α and a noisy part $\eta_\alpha := f - U_\alpha$, where α



(a)

(b)

Figure 1.3: The denoising problem: can we go from a noisy image (a) to a restored image in (b) ?

is an algorithm-specific denoising parameter. For example, in the case of Gaussian smoothing, the variance of the Gaussian kernel acts as α . We get a larger scale version U_α of the given image f , as the small-scale features of f , categorized as noise, are forced to be in the noisy part η_α .

Small scale features, categorized as noise, are forced to be in η_α , resulting in a larger scale version U_α of the original image f . Thus, denoising of f generates a *multiscale representation*: $\{U_\alpha\}_{\alpha \in A}$ with α as a *scaling parameter*. In this exposition, we primarily deal with PDE-based methods, which as we shall see later, are intimately related to the variational methods. We now see some examples where denoising methods generate multiscale representation.



(a)

(b)

Figure 1.4: The deblurring problem: can we go from a blurred image (a) to a restored image in (b) ?

1.4 Multiscale representation with PDE and variational methods

1.4.1 Heat equation

One of the earliest PDE-based methods for denoising a given image $f := U(\cdot, 0)$ is the heat equation

$$\frac{\partial U}{\partial t} = \Delta U, \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \left. \frac{\partial U}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0. \quad (1.1a)$$

This yields a family of images, $\{U(\cdot, t) : \Omega \mapsto \mathbb{R}\}_{t \geq 0}$, which can be viewed as smoothed versions of f . In this linear set up, smoothing is a convolution with the two-dimensional Gaussian kernel, $G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x|^2}{2\sigma^2}\right)$ with standard deviation $\sigma = \sqrt{2t}$. Hence, details with a scale smaller than $\sqrt{2t}$ are smoothed out. We can say that $\{U(\cdot, t)\}_{t \geq 0}$ is a multiscale representation of f , as $U(\cdot, t)$ diffuses from the small scales in f into increasingly larger scales. Here, $\lambda(t) := \sqrt{2t}$ acts as a *scaling function*.

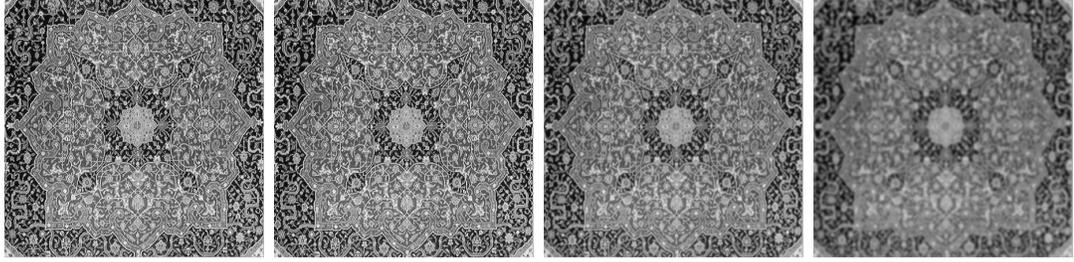


Figure 1.5: Solutions of the heat equation. As one can see the heat equation produces different scaled versions of the image, but at the same time diffusing the edges as well.

1.4.2 Perona-Malik model

Image denoising by the heat equation is based on isotropic diffusion, and consequently blurs all edges, which contain useful information about the image. Diffusion of the edges is observed in Figure 1.5 which depicts the results of the heat equation (1.1a). This drawback was removed by Perona-Malik (PM) model [28], which is based on *nonlinear diffusion*

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|\nabla U|)\nabla U), \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0, \quad (1.1b)$$

with an initial condition $U(\cdot, 0) := f$. Here, the diffusion controlling function, g , is a real valued function that vanishes at infinity, so that the amount of diffusion decreases as the gradient $|\nabla U|$ increases. Thus, g is responsible for the anisotropic nature of the PM model. Typical choices of g include $\frac{1}{\sqrt{1+\frac{s}{\beta}}}$, $\frac{1}{1+(\frac{s}{\beta})^2}$, $e^{-\frac{s}{\beta}}$ for some constant β .

1.4.3 Catté, Lions, Coll, Morel model

The family of PM models are not well-posed. They also pose a problem for noisy images. Since noise produces high gradients, it can be confused with relevant edges. These shortcomings were removed by Catté et. al. [10] by replacing $g(|\nabla U|)$ with $g(|G_\sigma \star \nabla U|)$, where $G_\sigma \star \nabla U$ denotes convolution of the two-dimensional Gaussian kernel G_σ with the gradient ∇U i.e.

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|G_\sigma \star \nabla U|)\nabla U), \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \left. \frac{\partial U}{\partial \mathbf{n}} \right|_{\partial\Omega} = 0, \quad (1.1c)$$

subject to $U(\cdot, 0) := f$.

1.4.4 Nordström model

The models (1.1) still suffer from a major drawback: the solution $U(t)$ diffuses to the average value \bar{f} as $t \rightarrow \infty$. Thus, a stopping criteria $t = t_c$ must be sought, so that the desired denoised image $U_c := U(t_c)$ is obtained. This raises the question of an appropriate stopping time t_c . The necessity of finding a stopping time is removed by Nordström's biased anisotropic model [26]

$$\frac{\partial U}{\partial t} = f - U + \operatorname{div}(g(|\nabla U|)\nabla U), \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \left. \frac{\partial U}{\partial \mathbf{n}} \right|_{\partial\Omega} = 0. \quad (1.2)$$

In this case, the solution $U(\cdot, t)$ varies from the initial condition $U(\cdot, 0) := 0$ to a desired denoised image U_c , as $t \rightarrow \infty$. Thus, the family $\{U(\cdot, t)\}_{t \geq 0}$ is an inverse scale representation of U_c , with t acting as an *inverse scale* parameter.

1.4.5 Rudin, Osher, Fatemi model

Variational approaches for image processing like Mumford-Shah segmentation [23, 24], Rudin-Osher-Fatemi (ROF) decomposition [30, 29] etc. fall under a general category of Tikhonov regularization [36], where one solves the ill-posed problem of finding u , given $f = Tu + \eta$. The discussion of deblurring problem is postponed to Chapter 3. We will first restrict our attention to the problem of pure denoising with no blur i.e. $T = I$. Here, we seek a faithful, noise free approximation $u \in X$ of $f \in L^2$, where $X \subsetneq L^2$ is an appropriate space adapted to measure edges and textures in images. This leads to the following minimization problem [8]

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_X + \lambda \|v\|_{L^2}^2 \}.$$

The term $\|u\|_X$ is a regularizing term and λ is a positive scaling parameter. In the case of ROF model, $X = BV(\Omega)$, the space of bounded variations [5], yields (BV, L^2) decomposition of f :

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \}, \quad (1.3)$$

where $\|u\|_{BV} := \int_\Omega |\nabla u|$. For more details on the BV space refer to section 1.6. The Euler-Lagrange equation characterizing the minimizer, u_λ for the minimization problem (1.3) is

$$u_\lambda = f + \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right), \quad (1.4a)$$

which can be obtained as a steady state solution of the following nonlinear, parabolic differential equation

$$\frac{\partial u}{\partial t} = f - u + \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right), \quad u \equiv u(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0. \quad (1.4b)$$

Subject to $u(\cdot, 0) := f$, this PDE gives rise to a multiscale family $\{u(\cdot, t)\}_{t \geq 0}$, where t is a *forward* scaling parameter. Also note that the variational ROF model (1.3) relates to the PDE based Nordström model (1.2) with $g(s) := \frac{1}{2\lambda s}$.

1.5 Some properties of ROF decomposition

Meyer introduced [22] following Banach space to analyze properties of ROF model.

Definition 1.1. *Let $G(\mathbb{R}^2)$ denote the Banach space consisting of all generalized functions $v(x)$ which can be written as*

$$v(x) = \partial_1 g_1 + \partial_2 g_2 = \operatorname{div} \mathbf{g}, \quad g_1, g_2 \in L^\infty(\mathbb{R}^2)$$

The space $G(\mathbb{R}^2)$ is

$$\|v\|_* := \inf_{v=\operatorname{div} \mathbf{g}} \left\{ \|\mathbf{g}\|_{L^\infty} = \operatorname{ess\,sup}_{x \in \mathbb{R}^2} |\sqrt{(|g_1|^2 + |g_2|^2)}(x)| : g_1, g_2 \in L^\infty(\mathbb{R}^2) \right\}$$

With this definition and the density of $C^\infty \cap BV(\Omega)$ in $BV(\Omega)$ (see Theorem 1.3 for details), we have the following lemma [22].

Lemma 1.1. *If $v \in L^2(\mathbb{R}^2)$, then*

$$|(u, v)_2| \leq \|u\|_{BV} \|v\|_* \tag{1.5}$$

Proof. We observe that the assertion (1.5) is true by definition 1.1 for $u = \varphi$, a test function. Thus, it must be true for all Sobolev functions $u \in W_0^{1,1}$ by the density of the test functions. The extension to a general BV image u with zero trace is then achieved due to Theorem 1.3 by standard mollification procedures. \square

With these tools, we now have the following properties of the ROF decomposition (1.3) [22].

Theorem 1.1. *If $\|f\|_* > \frac{1}{2\lambda}$, then the ROF decomposition, $f = u + v$, is characterized by the following conditions.*

$$\|v\|_* = \frac{1}{2\lambda} \quad \text{and} \quad (u, v)_2 = \frac{1}{2\lambda} \|u\|_{BV} \quad (1.6)$$

If $\|f\|_ \leq \frac{1}{2\lambda}$, then $u = 0$ and $v = f$.*

Proof. Since u minimizes the functional $J(u) = \|u\|_{BV} + \lambda\|v\|_{L^2}$, for any $h \in BV$ and a scalar ϵ we have $J(u) \leq J(u + \epsilon h)$, i.e.

$$\begin{aligned} \|u\|_{BV} + \lambda\|v\|_{L^2}^2 &\leq \|u + \epsilon h\|_{BV} + \lambda\|v - \epsilon h\|_{L^2}^2 \\ &\leq \|u\|_{BV} + |\epsilon|\|h\|_{BV} + \lambda\|v\|_{L^2}^2 - 2\epsilon\lambda(v, h)_2 + \epsilon^2\lambda\|h\|_{L^2}^2. \end{aligned} \quad (1.7)$$

Thus, we now have

$$\epsilon(v, h)_2 \leq |\epsilon|\frac{1}{2\lambda}\|h\|_{BV} + \frac{\epsilon^2}{2}\|h\|_{L^2}^2.$$

Letting $\epsilon \rightarrow 0$, we obtain

$$|(v, h)_2| \leq \frac{1}{2\lambda}\|h\|_{BV}, \quad (1.8)$$

which implies that the star-norm of v does not exceed $\frac{1}{2\lambda}$ i.e. $\|v\|_* \leq \frac{1}{2\lambda}$. Now we take $h = u$ in (1.7). If $\epsilon > 0$, we obtain $(u, v)_2 \leq \frac{1}{2\lambda}\|u\|_{BV}$ and if $\epsilon < 0$ we get $(u, v)_2 \geq \frac{1}{2\lambda}\|u\|_{BV}$. When combined, we conclude that

$$(u, v)_2 = \frac{1}{2\lambda}\|u\|_{BV}.$$

We now proceed to prove the converse implication. We assume that the conditions (1.6) hold and we show that $f = u + v$ is the ROF decomposition. From (1.8) we have $\|u + \epsilon h\|_{BV} \geq 2\lambda(u + \epsilon h, v)_2$, and we write

$$\begin{aligned}
\|u + \epsilon h\|_{BV} + \lambda\|v - \epsilon h\|_{L^2}^2 &\geq 2\lambda(u + \epsilon h, v)_2 + \lambda\|v\|_{L^2}^2 - 2\lambda\epsilon(v, h)_2 + \lambda\epsilon^2\|h\|_{L^2}^2 \\
&= 2\lambda(u, v)_2 + \lambda\|v\|_{L^2}^2 + \lambda\epsilon^2\|h\|_{L^2}^2 \\
&= \|u\|_{BV} + \lambda\|v\|_{L^2}^2 + \lambda\epsilon^2\|h\|_{L^2}^2 \quad (\text{from(1.6)}) \\
&\geq \|u\|_{BV} + \lambda\|v\|_{L^2}^2.
\end{aligned}$$

Thus, u minimizes the ROF functional.

Finally we prove that $u = 0$ and $v = f$ is the minimizing pair for the variational problem (1.3). To this end, observe that $u = 0$ and $v = f$ minimizes the ROF functional if and only if for any $h \in BV$ and any scalar ϵ we have the following

$$\|0\|_{BV} + \lambda\|f\|_{L^2}^2 \leq \|\epsilon h\|_{BV} + \lambda\|f - \epsilon h\|_{L^2}^2.$$

Expanding the L^2 - norm we get

$$2\lambda\epsilon(f, h)_2 \leq |\epsilon|\|h\|_{BV} + \lambda\epsilon^2\|h\|_{L^2}^2.$$

Letting $\epsilon \rightarrow 0$ we obtain

$$|(f, h)_2| \leq \frac{1}{2\lambda}\|h\|_{BV}.$$

This implies from (1.8) that $\|f\|_* < \frac{1}{2\lambda}$, which completes the proof. \square

Tadmor, Nezzar and Vese's hierarchical multiscale image representation (TNV model) makes use of the above theorem. The TNV scheme consists of iterative use of

the ROF decomposition for a sequence of consecutively increasing scaling parameters $\{\lambda_k\}_{k=0}^\infty$. In the next chapter, we discuss the TNV model, as it serves as a motivation for the integro-differential model that we will propose later in Chapter 3, but before that we present some important results regarding the BV spaces.

1.6 BV spaces

In this section, we review some basic definitions and theorems regarding the BV space where we follow [5, 6, 22]. Throughout this Chapter Ω denotes a generic open set in \mathbb{R}^N and $\mathcal{M}(\Omega, \mathbb{R}^N)$ denotes the space of all \mathbb{R}^N -valued Borel measures, which we know is the dual of the space $C_0(\Omega, \mathbb{R}^N)$ of all continuous functions φ , vanishing at infinity, equipped with the uniform norm.

1.6.1 Definitions and general properties of BV space

We begin with the definition.

Definition 1.2 (Function of bounded variation). *We say that a function $u : \Omega \mapsto \mathbb{R}$ is a function of bounded variation if and only if it belongs to $L^1(\Omega)$ and its gradient Du in the distributional sense belongs to $\mathcal{M}(\Omega, \mathbb{R}^N)$. We denote the set of all functions of bounded variations by $BV(\Omega)$. The following four assertions are equivalent:*

1. $u \in BV(\Omega)$.
2. $u \in L^1(\Omega)$ and $\frac{\partial u}{\partial x_i} \in \mathcal{M}(\Omega)$ for all $i = 1, \dots, N$.

3. $u \in L^1(\Omega)$ and $\|Du\| := \sup\{\langle Du, \varphi \rangle : \varphi \in C_c(\Omega, \mathbb{R}^N), \|\varphi\|_\infty \leq 1\} < \infty$.

4. $u \in L^1(\Omega)$ and $\|Du\| := \sup\{\int_\Omega u \operatorname{div} \varphi : \varphi \in C_c(\Omega, \mathbb{R}^N), \|\varphi\|_\infty \leq 1\} < \infty$.

The product $\langle \cdot, \cdot \rangle$ in 3 is defined by

$$\langle Du, \varphi \rangle := \sum_{i=1}^N \int_\Omega \varphi_i \frac{\partial u}{\partial x_i}.$$

Equivalence between 2 and 3 is a consequence of the density of the space $C_c(\Omega, \mathbb{R}^N)$ in $C_0(\Omega, \mathbb{R}^N)$ equipped with the uniform norm. Equivalence between 3 and 4 can be easily established by the density of $C_c^\infty(\Omega, \mathbb{R}^N)$ in $C_c(\Omega, \mathbb{R}^N)$ and $C_c^1(\Omega, \mathbb{R}^N)$. Here we recall Riesz-Alexandroff representation theorem.

Theorem 1.2 (Riesz-Alexandroff representation theorem). *The topological dual of $C_0(\Omega)$ can be isometrically identified with the space of bounded Borel measures. more precisely, to each bounded linear functional Φ on $C_0(\Omega)$ there is a unique Borel measure μ on Ω such that for all $f \in C_0(\Omega)$,*

$$\Phi(f) = \int_\Omega f d\mu$$

Moreover, $\|\Phi\| = |\mu|(\Omega)$.

Thus, we have in the present context $\|Du\|$ is also the total mass $|Du|(\Omega) = \int_\Omega |Du|$ of the total variation Du of the measure Du .

Remark 1.1. *According to the Radon-Nikodym theorem, there exists $\nabla u \in L^1(\Omega, \mathbb{R}^N)$ and a measure $D_s u$, singular with respect to the N -dimensional Lebesgue measure $\mathcal{L}^N|_\Omega$, restricted to Ω , such that $Du = \nabla u \mathcal{L}^N|_\Omega + D_s u$. Consequently, $W^{1,1}(\Omega)$ is a subspace of the vectorial space $BV(\Omega)$ and $u \in W^{1,1}(\Omega)$ if and only if $Du = \nabla u \mathcal{L}^N|_\Omega$.*

The space $BV(\Omega)$ is equipped with the following norm, which extends to the classical norm in $W^{1,1}(\Omega)$:

$$\|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + \|Du\|.$$

1.6.2 Two types of convergence processes

We define two of convergence processes in $BV(\Omega)$. The first is *weak convergence* which provides compactness of bounded sequences. The second is an *intermediate convergence* between the weak and strong convergence associated with the norm.

Definition 1.3 (Weak convergence). *A sequence $\{u_n\}_{n \in \mathbb{N}}$ in $BV(\Omega)$ weakly converges to some u in $BV(\Omega)$, and we write $u_n \rightharpoonup u$ if and only if following hold:*

1. $u_n \rightarrow u$ strongly in $L^1(\Omega)$.
2. $Du_n \rightharpoonup Du$ weakly in $\mathcal{M}(\Omega, \mathbb{R}^N)$.

The following proposition establishes a compactness result related to the weak convergence, and the lower semicontinuity of the total mass.

Proposition 1.1. *Let $\{u_n\}_{n \in \mathbb{N}}$ be a sequence in $BV(\Omega)$ strongly converging to some u in $L^1(\Omega)$ and satisfying $\sup_{n \in \mathbb{N}} \int_{\Omega} |Du_n| < \infty$. Then*

1. *The limit function $u \in BV(\Omega)$ and $\int_{\Omega} |Du| \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |Du_n|$ and*
2. *The sequence u_n weakly converges to u in $BV(\Omega)$.*

Proof. For all $\varphi \in C_c^1(\Omega, \mathbb{R}^N)$ such that $\|\varphi\|_{\infty} \leq 1$, we have

$$\int_{\Omega} u \operatorname{div} \varphi = \lim_{n \rightarrow \infty} \int_{\Omega} u_n \operatorname{div} \varphi \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |Du_n|,$$

and the assertion 1 is proved by taking the supremum in the first member, over all the functions $\varphi \in C_c^1(\Omega, \mathbb{R}^N)$, with $\|\varphi\|_\infty \leq 1$. We now establish assertion (2). Since u_n strongly converges to u in $L^1(\Omega)$, for all $\varphi \in C_c^\infty(\Omega, \mathbb{R}^N)$ we have

$$\langle Du_n, \varphi \rangle = - \int_{\Omega} u_n \operatorname{div} \varphi \rightarrow - \int_{\Omega} u \operatorname{div} \varphi = \langle Du, \varphi \rangle.$$

By using the density of $C_c^\infty(\Omega, \mathbb{R}^N)$ in $C_0^\infty(\Omega, \mathbb{R}^N)$ for the uniform norm and the boundedness of $\{Du_n\}_{n \in \mathbb{N}}$, we conclude that the sequence $\{Du_n\}_{n \in \mathbb{N}}$ weakly converges to Du . \square

Remark 1.2. *As a consequence of the semicontinuity property 1, $BV(\Omega)$ is a complete normed space.*

We introduce a different type of convergence called *intermediate convergence*.

Definition 1.4 (Intermediate convergence). *Let $\{u_n\}_{n \in \mathbb{N}}$ be a sequence in $BV(\Omega)$ and $u \in BV(\Omega)$. We say that u_n converges to u in the sense of intermediate convergence if and only if*

1. $u_n \rightarrow u$ strongly in $L^1(\Omega)$.
2. $\int_{\Omega} |Du_n| \rightarrow \int_{\Omega} |Du|$.

Remark 1.3. *The term ‘intermediate convergence’ is due to Temam [35] and is also called strict convergence.*

Now we elaborate on the reason for not using the strong BV norm 1.1. The space $C^\infty(\bar{\Omega})$ is not dense in $BV(\Omega)$ when $BV(\Omega)$ is equipped with its strong norm

1.1. Indeed, its closure is the space $W^{1,1}$. Nevertheless, one can approximate every element of $BV(\Omega)$ by a function of $C^\infty(\bar{\Omega})$ in the sense of the intermediate convergence. More precisely, we have the following theorem.

Theorem 1.3. *The space $C^\infty(\Omega) \cap BV(\Omega)$ is dense in $BV(\Omega)$ equipped with the intermediate convergence. Consequently, $C^\infty(\bar{\Omega})$ is also dense in BV for the intermediate convergence.*

We refer the reader to [6] for a detailed proof of 1.3. This theorem essentially states that for any $u \in BV(\Omega)$, one can find a sequence of approximations $\{u_n\}_{n=0}^\infty$ such that

1. $u_n \in C^\infty(\Omega)$ for $n = 0, 1, 2, \dots$,
2. $u_n \rightarrow u$ in $L^1(\Omega)$ as $n \rightarrow \infty$,
3. $\int_\Omega |Du_n| \rightarrow \int_\Omega |Du|$ as $n \rightarrow \infty$.

The following theorem gives the embedding $BV(\Omega) \hookrightarrow L^p(\Omega)$.

Theorem 1.4. *Let Ω be a regular open subset of R^N . For all p such that $1 \leq p \leq \frac{N}{N-1}$, the embedding*

$$BV(\Omega) \hookrightarrow L^p(\Omega)$$

is continuous. More precisely, there exists a constant C which depends only on Ω, p , and N , such that for all u in $BV(\Omega)$,

$$\left(\int_\Omega |u|^p \right)^{\frac{1}{p}} \leq C \|u\|_{BV(\Omega)}$$

Proof. First notice that $C^\infty(\Omega) \cap BV(\Omega) = C^\infty(\Omega) \cap W^{1,1}(\Omega)$. Let $\{u_n\}_{n \in \mathbb{N}}$ be a sequence of functions in $C^\infty(\Omega) \cap BV(\Omega)$ which converges to some u in $BV(\Omega)$ for the intermediate convergence. Since the embedding $W^{1,1}(\Omega) \hookrightarrow L^p(\Omega)$ is continuous for $1 \leq p \leq \frac{N}{N-1}$, there exists a constant C , which depends only on Ω, p and N such that

$$\left(\int_{\Omega} |u_n|^p \right)^{\frac{1}{p}} \leq C \left(\|u_n\|_{L^1(\Omega)} + \int_{\Omega} |Du_n| \right) < \infty.$$

We deduce that $u_n \rightharpoonup u$ in $L^p(\Omega)$ and according to the weak lower semicontinuity of $L^p(\Omega)$ norm,

$$\begin{aligned} \left(\int_{\Omega} |u|^p \right)^{\frac{1}{p}} &\leq \liminf_{n \rightarrow \infty} \left(\int_{\Omega} |u_n|^p \right)^{\frac{1}{p}} \\ &\leq \liminf_{n \rightarrow \infty} C \left(\|u_n\|_{L^1(\Omega)} + \int_{\Omega} |Du_n| \right) \\ &= C \|u\|_{BV(\Omega)} \end{aligned}$$

where the last equality follows from the intermediate convergence. □

Remark 1.4. *For image processing problems we deal with real-valued functions on $\Omega \in \mathbb{R}^2$. Thus, Theorem 1.4 implies that $BV(\Omega) \hookrightarrow L^2(\Omega)$.*

1.6.3 Structure of BV functions

The functions in $BV(\Omega)$ inherit their properties from their level sets $[u > t] := \{x \in \Omega : u(x) > t\}$, where t varies in \mathbb{R} . The following property generalizes the coarea formula to BV functions [16].

Theorem 1.5 (Coarea formula). *Let u be a given function in $BV(\Omega)$. Then, for almost every t in \mathbb{R} , the level set $E_t = \{x \in \Omega : u(x) > t\}$ of u is a set of finite*

perimeter in Ω and

$$Du(\Omega) = \int_{-\infty}^{\infty} D_{\chi_{E_t}} dt,$$
$$|Du|(\Omega) = \int_{-\infty}^{\infty} |D_{\chi_{E_t}}| dt.$$

This theorem essentially states that for almost every $t \in \mathbb{R}$, the level set $[u > t]$ of each BV function u has a finite perimeter in Ω . i.e.

$$|Du|(\Omega) = \int_{-\infty}^{\infty} \text{Per}(\{x \in \Omega : u(x) > t\}) dt$$

where the perimeter of a set $\Sigma \subset \mathbb{R}^N$, $\text{Per}(\Sigma)$ is defined as

$$\text{Per}(\Sigma) := \int_{\mathbb{R}^N} |D_{\chi_{\Sigma}}|,$$

We refer the reader to [5, 6] for detailed proof.

Chapter 2

The Tadmor-Nezzar-Vese hierarchical multiscale representation

2.1 Introduction

Tadmor, Nezzar and Vese propose in [32] a new, hierarchical and multiscale representation particularly adapted for images lying in intermediate spaces. The standard tool for studying intermediate spaces is forming scales of interpolation spaces $(X, Y)_\theta$, $\theta \in [0, 1]$, where $X \subsetneq Y$, see [8, 9, 34]. The canonical example involves the so-called K -functional

$$K(f, \eta) \equiv K(f, \zeta; X, Y) := \inf_{f=u+v} \{ \|u\|_Y + \zeta \|v\|_X \}.$$

This leads to another variant with a similar scale of intermediate spaces. In the context of image processing, one seeks the representation in the intermediate spaces between $X = BV(\Omega)$, the space of bounded variations and $Y = L^2(\Omega)$, defined over two-dimensional domain $\Omega \in \mathbb{R}^2$ and quantified in terms of the J -functional

$$J(f, \lambda) \equiv J_2(f, \lambda; BV, L^2) := \inf_{f=u+v} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \} \quad (2.1)$$

The functional $J(f, \lambda)$ measures how well an L^2 - object can be approximated by its BV features. As we have seen in the previous Chapter, the functional $J(f, \lambda)$ was introduced by Rudin, Osher and Fatemi [30]. They suggested extracting the main features of contour discontinuities u_λ , which are separated from the noisy part v_λ , by realizing the minimizing pair, $[u_\lambda, v_\lambda]$, of $J(f, \lambda)$. For $f \in L^2(\Omega)$ the problem

admits a unique minimizer [11, 1, 37], which decomposes an $L^2(\Omega)$ image, f , into two distinct components,

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \}.$$

The BV part, u_λ , captures the large-scale features of f while the small-scales, interpreted as noise is captured in the residual $v_\lambda := f - u_\lambda$. The ROF model requires a-priori information of the noise scaling λ . For a small values of λ , only a cartoon representation of f consisting of only large scales in f and for large values of λ , the image u_λ contains too many details of f . The parameter λ can be estimated if some statistical information on the noise is known [30, 11].

2.2 The hierarchical (BV, L^2) decomposition

Tadmor, Nezzar and Vese observe in [32] that for a small value of the scaling parameter λ_0 , the residual image v_{λ_0} may still contain important details when viewed at a finer scale. Thus, v_{λ_0} can be further decomposed with a scaling parameter $\lambda_1 > \lambda_0$,

$$v_{\lambda_0} = u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \underset{v_{\lambda_0}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_1 \|v\|_{L^2}^2 \}.$$

We can continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \dots$

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^2}^2 \}.$$

Repeating this refinement step, we obtain the following *hierarchical decomposition* of f :

$$\left. \begin{aligned} f &= u_{\lambda_0} + v_{\lambda_0} \\ &= u_{\lambda_1} + u_{\lambda_1} + v_{\lambda_1} \\ &= \dots \\ &= u_{\lambda_0} + u_{\lambda_1} + \dots + u_{\lambda_N} + v_{\lambda_N}. \end{aligned} \right\} \quad (2.2)$$

This yields a *hierarchical multiscale image decomposition*, $f \sim u_{\lambda_0} + u_{\lambda_1} + \dots + u_{\lambda_N}$, with a residual v_{λ_N} . This construction of the hierarchical decomposition (2.2) is independent of a-priori parameters. The accumulated sum, $\sum_{k=0}^N u_{\lambda_k}$, provides a multilayered description of f , which lies in the intermediate scale of spaces, in between BV and L^2 . This multiscale representation of f is *essentially nonlinear* in the sense that its ‘slices’, $u_{\lambda_k} = u_{\lambda_k}(f)$, depend on the original image f .

2.3 Initialization

The question as to how to start the TNV algorithm is answered using Meyer’s Theorem [22, Theorem 3]. It asserts that if $\|f\|_* < \frac{1}{2\lambda_0}$ then the minimization of the functional $J(f, \lambda_0)$ is the trivial pair $[u_{\lambda_0}, v_{\lambda_0}] = [0, f]$. Thus, in order to initiate the TNV algorithm, one must choose the initial λ_0 so that $\lambda_0 \geq \frac{1}{2\|f\|_*}$. In general the information about the star-norm of f is not known. If the initial choice of λ_0 is smaller than $\frac{1}{2\|f\|_*}$ then we get a trivial pair as the minimizer, on the other hand, if λ_0 is too large then some larger scales may be missing. In this case, TNV propose in [32] a refinement procedure to capture a hierarchical representation of the missing scales.

2.4 Convergence results for the hierarchical multiscale expansion

The BV dual norm, called the star-norm,

$$\|v\|_* := \sup_{\|\varphi\|_{BV} \neq 0} \frac{(v, \varphi)_2}{\|\varphi\|_{BV}},$$

plays important role in understanding the ROF decomposition and TNV multiscale representation. Meyer [22] interprets the minimization of the functional $J(f, \lambda)$ in (2.1) as a decomposition, of $f = u_\lambda + v_\lambda$, so that u_λ extracts the main features with smooth parts with prominent edges and v_λ captures ‘textures’. As we have seen in Theorem 1.1 if $\|f\|_* \leq \frac{1}{2\lambda}$, then $u_\lambda = 0$ and $v_\lambda = f$, whereas if $\|f\|_* > \frac{1}{2\lambda}$ then the ROF decomposition is characterized by the following two conditions

$$\|v_\lambda\|_* = \frac{1}{2\lambda}, \quad (u_\lambda, v_\lambda)_2 = \frac{1}{2\lambda} \|u_\lambda\|_{BV}. \quad (2.3)$$

Tadmor et. al. [32] provide the following result quantifying the convergence $\sum_{k=0}^N u_{\lambda_k} \rightarrow f$ as $N \rightarrow \infty$.

Proposition 2.1. *For (BV, L^2) hierarchical multiscale decomposition (3.2) we have the following inequality*

$$\sum_{k \geq 0} \frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} \leq \|f\|_{L^2}^2. \quad (2.4)$$

Proof. Compare the decomposition $v_{\lambda_k} = u_{\lambda_{k+1}} + v_{\lambda_{k+1}}$ furnished by the minimizer of $J(v_{\lambda_k}, \lambda_{k+1})$, vs. the trivial pair $[0, v_{\lambda_k}]$, to find

$$\|u_{\lambda_{k+1}}\|_{BV} + \lambda_{k+1} \|v_{\lambda_{k+1}}\|_{L^2}^2 \leq \lambda_{k+1} \|v_{\lambda_k}\|_{L^2}^2.$$

It follows that

$$\begin{aligned}
\sum_{k \geq 0} \frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} &= \frac{1}{\lambda_0} \|u_{\lambda_0}\|_{BV} + \sum_{k=0} \frac{1}{\lambda_{k+1}} \|u_{\lambda_{k+1}}\|_{BV} \\
&\leq \|f\|_{L^2}^2 - \|v_0\|_{L^2}^2 + \sum_{k=0} [\|v_{\lambda_k}\|_{L^2}^2 - \|v_{\lambda_{k+1}}\|_{L^2}^2] \\
&\leq \|f\|_{L^2}^2.
\end{aligned}$$

□

A more precise (BV, L^2) hierarchical statement is provided in the following theorem.

Theorem 2.1. *Consider $f \in L^2$. Then f admits the following hierarchical decomposition.*

$$f = \sum_{k=0}^{\infty} u_{\lambda_k}, \quad \left\| f - \sum_{k=0}^{\infty} u_{\lambda_k} \right\|_* = \frac{1}{\lambda_{k+1}}, \quad (2.5)$$

and the following energy estimate holds:

$$\sum_{k=0}^{\infty} \left[\frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} + \|u_{\lambda_k}\|_{L^2}^2 \right] \leq \|f\|_{L^2}^2 \quad (2.6)$$

Proof. Recall Theorem 1.1, which states that for (BV, L^2) decomposition of f , if $\|f\|_* \leq \frac{1}{2\lambda}$, then $u_\lambda = 0$ and $v_\lambda = f$; on the other hand, if $\|f\|_* > \frac{1}{2\lambda}$, then $\|v_\lambda\|_* = \frac{1}{2\lambda}$ and $(u_\lambda, v_\lambda)_2 = \frac{1}{2\lambda} \|u_\lambda\|_{BV}$.

The first statement (2.5) follows from the basic hierarchical expansion, $f = \sum_{k=0}^N u_{\lambda_k} + v_{\lambda_N}$, while noting that $\|v_{\lambda_N}\|_* = \frac{1}{2\lambda_N}$. For the second statement, (2.6), we begin by squaring the basic refinement step, $u_{\lambda_{k+1}} + v_{\lambda_{k+1}} = v_{\lambda_k}$, in the L^2 -inner product:

$$\|v_{\lambda_{k+1}}\|_{L^2}^2 + \|u_{\lambda_{k+1}}\|_{L^2}^2 + 2(u_{\lambda_{k+1}}, v_{\lambda_{k+1}})_2 = \|v_{\lambda_k}\|_{L^2}^2, \quad j = -1, 0, 1, \dots$$

Observe that the last equality holds for $j = -1$ with $v_{\lambda_{-1}}$ interpreted as $v_{\lambda_{-1}} := f$.

We recall that $[u_{\lambda_{k+1}}, v_{\lambda_{k+1}}]$ is a minimizing pair for $J(v_k, \lambda_{k+1})$, and hence, by (2.3),

$$2(u_{\lambda_{k+1}}, v_{\lambda_{k+1}})_2 = \frac{1}{\lambda_{k+1}} \|u_{k+1}\|_{BV}$$

yielding a precise refinement of (2.4)

$$\frac{1}{\lambda_{k+1}} \|u_{\lambda_{k+1}}\|_{BV} + \|u_{\lambda_{k+1}}\|_{L^2}^2 = \|v_{\lambda_k}\|_{L^2}^2 - \|v_{\lambda_{k+1}}\|_{L^2}^2.$$

We sum up, obtaining

$$\begin{aligned} \sum_{k=0}^N \left[\frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} + \|u_{\lambda_k}\|_{L^2}^2 \right] &= \sum_{k=-1}^N \left[\frac{1}{\lambda_{k+1}} \|u_{\lambda_{k+1}}\|_{BV} + \|u_{\lambda_{k+1}}\|_{L^2}^2 \right] \\ &= \|f\|_{L^2}^2 - \|v_{\lambda_N}\|_{L^2}^2. \end{aligned} \tag{2.7}$$

□

Note that according to (2.7), equality holds in (2.6) if and only if we have a string L^2 -convergence, $\|f - \sum_{k=0}^N u_{\lambda_k}\|_{L^2} = \|v_{\lambda_N}\|_{L^2} \rightarrow 0$. The situation is reminiscent of the passage in a linear setup, from Bessel-energy inequality into the Parseval equality. Since the present setup is nonlinear, the linear sense of completeness of $\{u_{\lambda_k}\}_{k \geq 0}$ does not apply. We prove that the equality in (2.6) holds [32] by adding minimal amount of smoothness. To this effect, we prove the following lemma.

Lemma 2.1. *Consider the (BV, L^2) hierarchical decomposition of $f \in BV$, $f = \sum_{k=0}^{\infty} u_{\lambda_k}$, for an increasing sequence of scaling parameters $\{\lambda_k\}_{k=0}^{\infty}$ such that $\frac{\lambda_k}{\lambda_{2k}} \downarrow 0$. Then the residuals, $v_N := f - \sum_{k=0}^N u_{\lambda_k}$, converge strongly to 0 in L^2 ,*

$$\lim_{N \rightarrow \infty} \|v_{\lambda_N}\|_{L^2} = 0.$$

Proof. The starting point is the decomposition $v_{\lambda_{2N}} = \sum_{k=N+1}^{2N} u_{\lambda_k} + v_{\lambda_N}$. Multiplication against $v_{\lambda_{2N}}$ yields

$$\|v_{\lambda_{2N}}\|_{L^2}^2 = - \left(v_{\lambda_{2N}}, \sum_{k=N+1}^{2N} u_{\lambda_k} \right)_2 + (v_{\lambda_{2N}}, v_{\lambda_N}) =: I + II. \quad (2.8)$$

From (2.3) we know that $\|v_{\lambda_{2N}}\|_* = \frac{1}{2\lambda_{2N}}$. Thus, $|(v_{\lambda_{2N}}, h)| \leq \frac{\|h\|_{BV}}{2\lambda_{2N}}$. We get the following limit on the first term in (2.8), $I := - \left(v_{\lambda_{2N}}, \sum_{k=N+1}^{2N} u_{\lambda_k} \right)_2$

$$|I| \leq \frac{1}{2\lambda_{2N}} \sum_{k=N+1}^{2N} \|u_{\lambda_k}\|_{BV} \leq \sum_{k=N+1}^{2N} \frac{1}{2\lambda_k} \|u_{\lambda_k}\|_{BV}.$$

From (2.4), we note that the term $\sum_{k=N+1}^{2N} \frac{1}{2\lambda_k} \|u_{\lambda_k}\|_{BV}$ is a Cauchy subsequence of the bounded series $\sum \frac{1}{2\lambda_k} \|u_{\lambda_k}\|_{BV} \leq \|f\|_{L^2}^2$, and thus, it decays to zero as $N \rightarrow \infty$.

To show that the second term in (2.8), $II := (v_{\lambda_{2N}}, v_{\lambda_N})$ also goes to zero. To this end we note that the BV norm of v_{λ_N} does not grow faster than λ_N ; indeed we get the following upper bound on the BV norm of $v_N := f - \sum_{k=0}^N u_{\lambda_k}$.

$$\begin{aligned} \|v_{\lambda_k}\|_{BV} &\leq \|f\|_{BV} + \sum_{k=0}^N \|u_{\lambda_k}\|_{BV} \\ &\leq \|f\|_{BV} + \lambda_N \sum_{k=0}^N \frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} \\ &\leq \|f\|_{BV} + \lambda_N \|f\|_{L^2}^2. \end{aligned}$$

Here we made use of (2.4) in the last inequality. Using the fact $II := (v_{\lambda_{2N}}, v_{\lambda_N})_2 \leq \frac{1}{2\lambda_{2N}} \|v_{\lambda_N}\|_{BV}$, we conclude that the term II vanishes as $k \rightarrow \infty$:

$$|II| \leq \frac{1}{2\lambda_{2N}} \|v_{\lambda_N}\|_{BV} \leq \frac{1}{2\lambda_{2N}} [\|f\|_{BV} + \lambda_N \|f\|_{L^2}^2] \downarrow 0.$$

□

Using the Lemma 2.1 in (2.7) we get the following result:

Theorem 2.2. Consider the (BV, L^2) hierarchical decomposition of $f \in BV$, $f = \sum_{k=0}^{\infty} u_{\lambda_k}$, with an increasing sequence of scaling parameters $\{\lambda_k\}_{k=0}^{\infty}$ such that $\frac{\lambda_k}{\lambda_{2k}} \downarrow 0$. Then the energy of f is given by

$$\sum_{k=0}^{\infty} \left[\frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} + \|u_{\lambda_k}\|_{L^2}^2 \right] = \|f\|_{L^2}^2.$$

This result can be extended to f s beyond the BV space. For example, we can prove a similar result for f in the interpolation space $X_{\theta} := (L^2, BV)_{\theta}$, $\theta > 0$. Characterization of this scale of space can be found in [15].

Lemma 2.2. Consider the (BV, L^2) hierarchical decomposition of $f \in X_{\theta} := (L^2, BV)_{\theta}$, $\theta > 0$, $f = \sum_{k=0}^{\infty} u_{\lambda_k}$, for an increasing sequence of scaling parameters $\{\lambda_k\}_{k=0}^{\infty}$ such that $\frac{\lambda_k}{\lambda_{2k}} \downarrow 0$. Then the residuals, $v_N := f - \sum_{k=0}^N u_{\lambda_k}$, converge strongly to 0 in L^2 ,

$$\lim_{N \rightarrow \infty} \|v_{\lambda_N}\|_{L^2} = 0.$$

Proof. We notice that the first term I in (2.8) vanishes for arbitrary f , while we have an upper bound for the second term $II := (v_{\lambda_{2N}}, v_{\lambda_N})_2$ using (2.3) and (2.4):

$$\begin{aligned} |(v_{\lambda_{2N}}, v_{\lambda_N})_2| &= \left| (v_{\lambda_{2N}}, f)_2 - \sum_{k=0}^N (v_{\lambda_{2N}}, u_{\lambda_k})_2 \right| \\ &\leq |(v_{\lambda_{2N}}, f)_2| + \frac{1}{\lambda_{2N}} \sum_{k=0}^N \|u_{\lambda_k}\|_{BV} \\ &\leq |(v_{\lambda_{2N}}, f)_2| + \frac{\lambda_N}{\lambda_{2N}} \sum_{k=0}^N \frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} \\ &\leq |(v_{\lambda_{2N}}, f)_2| + \frac{\lambda_N}{\lambda_{2N}} \|f\|_{L^2}^2. \end{aligned}$$

Thus, we only need to show that the moments $|(v_{\lambda_{2N}}, f)_2| \rightarrow 0$. Let $X_{-\theta}$ denote the dual space, the collection of all functions v such that $\|v\|_{X_{-\theta}} := \sup_{\varphi} \frac{(v, \varphi)_2}{\|\varphi\|_{X_{\theta}}} < \infty$. We

recall that $v_{\lambda_{2N}}$ is in L^2 with $\|v_{\lambda_{2N}}\|_{L^2} < \|f\|_{L^2}$ and its star-norm is $\|v_{\lambda_{2N}}\|_* < \frac{1}{2\lambda_{2N}}$.

by convexity argument of Riesz we have

$$\|v_{\lambda_{2N}}\|_{X_{-\theta}} \lesssim \|v_{\lambda_{2N}}\|_{L^2}^{1-\theta} \|v_{\lambda_{2N}}\|_*^\theta \lesssim \lambda_{2N}^{-\theta} \|f\|_2^{1-\theta}$$

Thus, we have that

$$|(v_{\lambda_{2N}}, f)_2| \leq \|v_{\lambda_{2N}}\|_{X_{-\theta}} \|f\|_{X_\theta} \lesssim \lambda_{2N}^{-\theta} \|f\|_2^{1-\theta} \|f\|_{X_\theta}.$$

This implies the strong L^2 -convergence of texture terms, $\|v_{\lambda_N}\|_{L^2} \rightarrow 0$. □

This lemma gives us the following extension of the Theorem 2.2 when $f \in X_\theta := (L^2, BV)_\theta$, $\theta > 0$.

Corollary 2.1. *Consider the (BV, L^2) hierarchical decomposition of $f \in f \in X_\theta := (L^2, BV)_\theta$, $\theta > 0$, $f = \sum_{k=0}^{\infty} u_{\lambda_k}$, with an increasing sequence of scaling parameters $\{\lambda_k\}_{k=0}^{\infty}$ such that $\frac{\lambda_k}{\lambda_{2k}} \downarrow 0$. Then the energy of f is given by*

$$\sum_{k=0}^{\infty} \left[\frac{1}{\lambda_k} \|u_{\lambda_k}\|_{BV} + \|u_{\lambda_k}\|_{L^2}^2 \right] = \|f\|_{L^2}^2.$$

We observe that the condition of strong L^2 -convergence of the texture terms, $\|v_{\lambda_N}\|_{L^2} \rightarrow 0$ is essential to obtain the Parseval-type equality (2.1) in this non-linear set-up. In fact, we see in the next section that similar energy equality follows in more general setting, if the corresponding texture terms decay to zero in L^2 -norm.

2.5 General hierarchical decompositions

In this section we examine properties of a general hierarchical decomposition:

$$\left. \begin{aligned} f &= u_0 + v_0 \\ v_0 &= u_1 + v_1 \\ \dots &= \dots \\ v_{k-1} &= u_k + v_k \\ \dots &= \dots \end{aligned} \right\} \quad (2.9)$$

This procedure produces a hierarchical family of pairs $f \sim \{u_k, v_k\}_{k=0}^{\infty}$. In the context of a general decomposition of type (2.9) we derive the following result.

Lemma 2.3. *Let $f \sim \{u_k, v_k\}_{k=0}^{\infty}$ and $g \sim \{\hat{u}_k, \hat{v}_k\}_{k=0}^{\infty}$ be hierarchical decompositions of type (2.9) so that u_k, v_k, \hat{u}_k and \hat{v}_k belong to an inner product space X with an inner product $(\cdot, \cdot)_X$. We define a new inner product $\langle\langle \cdot, \cdot \rangle\rangle$*

$$\langle\langle f, g \rangle\rangle := \sum_{k=0}^{\infty} (u_k, \hat{u}_k)_X + (u_k, \hat{v}_k)_X + (v_k, \hat{u}_k)_X, \quad (2.10)$$

where $(\cdot, \cdot)_X$ is the inner product. Then $\langle\langle f, g \rangle\rangle = (f, g)_X$ if and only if the inner products of the residuals, $(v_k, \hat{v}_k)_X$, converge to 0.

Proof. Noting that $u_k + v_k = v_{k-1}$ and $\hat{u}_k + \hat{v}_k = \hat{v}_{k-1}$, we obtain

$$\begin{aligned}
\langle\langle f, g \rangle\rangle &= \sum_{k=0}^{\infty} (v_{k-1}, \hat{u}_k)_X + (u_k, \hat{v}_k)_X \\
&= \sum_{k=0}^{\infty} (v_{k-1}, \hat{u}_k)_X + (u_k, \hat{v}_k)_X + (v_k, \hat{v}_k)_X - (v_k, \hat{v}_k)_X \\
&= \sum_{k=0}^{\infty} (v_{k-1}, \hat{u}_k)_X + (u_k + v_k, \hat{v}_k)_X - (v_k, \hat{v}_k)_X \\
&= \sum_{k=0}^{\infty} (v_{k-1}, \hat{u}_k)_X + (v_{k-1}, \hat{v}_k)_X - (v_k, \hat{v}_k)_X \\
&= \sum_{k=0}^{\infty} (v_{k-1}, \hat{u}_k + \hat{v}_k)_X - (v_k, \hat{v}_k)_X \\
&= \sum_{k=0}^{\infty} (v_{k-1}, \hat{v}_{k-1})_X - (v_k, \hat{v}_k)_X.
\end{aligned}$$

Now, the n^{th} partial sum

$$\begin{aligned}
s_n &= \sum_{k=0}^n (v_{k-1}, \hat{v}_{k-1})_X - (v_k, \hat{v}_k)_X \\
&= (v_{-1}, \hat{v}_{-1})_X - (v_n, \hat{v}_n)_X \\
&= (f, g)_X - (v_n, \hat{v}_n)_X.
\end{aligned}$$

Thus, we conclude that $\langle\langle f, g \rangle\rangle = (f, g)_X$ if and only if $\lim_{k \rightarrow \infty} (v_k, \hat{v}_k)_X = 0$. \square

Remark 2.1. *Note this is a weak Lemma, in the sense that we only assume that the inner products $(u_k, v_k)_X$ are defined.*

Using the above Lemma we get the following energy estimate.

Corollary 2.2. *For a general hierarchical decomposition $f \sim \{u_k, v_k\}_{k=0}^{\infty}$ of the type (2.9) the following energy estimate holds*

$$\sum_{k=0}^{\infty} [\|u_k\|_X^2 + 2(u_k, v_k)_X] = \|f\|_{L^2}^2$$

if and only if the residuals, v_k , converge to 0 in X - norm.

Proof. This follows directly by letting $g = f$ in the Lemma 2.3. □

Remark 2.2. We notice in Lemma 2.1 that in the case of (BV, L^2) hierarchical decomposition, $f \sim \{u_{\lambda_k}, v_{\lambda_k}\}$, of a function $f \in L^2$, we have

$$\sum_{k=0}^{\infty} [\|u_k\|_{L^2}^2 + 2(u_k, v_k)_2] = \|f\|_{L^2}^2$$

if and only if $\|v_{\lambda_N}\|_{L^2} \rightarrow 0$. By (2.3) we have $(u_k, v_k)_2 = \frac{1}{2\lambda_k}$. We immediately conclude that

$$\sum_{k=0}^{\infty} \left[\|u_k\|_{L^2}^2 + \frac{1}{2\lambda_k} \|u_k\|_{BV} \right] = \|f\|_{L^2}^2.$$

Similarly, from Lemma 2.2 we get the same energy equality for f in the interpolation space $X_\theta := (L^2, BV)_\theta$, $\theta > 0$

2.6 Euler-Lagrange equations

The minimizer of functional $J(f, \lambda)$ are characterized by the Euler-Lagrange differential equation

$$u_\lambda - \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right) = f.$$

When restricted to a bounded domain Ω , the Euler-Lagrange equations are augmented by the Neumann boundary condition

$$\frac{\partial u_\lambda}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0.$$

To construct the hierarchical multiscale decomposition of f , $f \sim \sum_{k=0}^N u_{\lambda_k}$, the slices, u_{λ_k} , are constructed as approximate solutions of the recursive relation gov-

erned by the elliptic PDE

$$u_{\lambda_{k+1}} - \frac{1}{2\lambda_{k+1}} \operatorname{div} \left(\frac{\nabla u_{\lambda_{k+1}}}{|\nabla u_{\lambda_{k+1}}|} \right) = -\frac{1}{2\lambda_k} \operatorname{div} \left(\frac{\nabla u_{\lambda_k}}{|\nabla u_{\lambda_k}|} \right).$$

We propose a *novel integro-differential equation* (IDE) in the following chapter, motivated by a TNV scheme and the associated Euler-Lagrange differential equations.

Chapter 3

Integro-differential equation for multiscale image representation

3.1 Introduction

In this Chapter we introduce a novel integro-differential equation (IDE) for multiscale representation of f

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0, \quad (3.1)$$

subject to $u(\cdot, 0) := 0$. The integral $\int_0^t u(\cdot, s) ds =: U(t)$ gives a scaled version of the image f for a given t . The image $U(t)$ evolves with t , from a coarse, larger scale images, to consecutively smaller scale images with finer details. Thus, this is an *inverse scale* method, as opposed to the *forward scale* methods such as heat equation or PM models (1.1).

The motivation behind this IDE comes from the hierarchical (BV, L^2) multiscale image decomposition [32], [33] of Tadmor et. al., which we will elaborate upon in the next section. We will derive an important relationship between the star-norm (see [22] for more details) of the residual image $V(t) := f - \int_0^t u(\cdot, s) ds$ and the function $\lambda(t)$ in section 3.3. In sections 3.4.1 and 3.4.3, we will propose PDE-based modifications for our IDE.

3.2 Motivation for the integro-differential equation

Rudin, Osher and Fatemi introduced a BV-based minimization functional for image denoising in [30], which in turn led to the unconstrained (BV, L^2) decomposition (1.3) in [11]. The minimizer of (1.3), u_λ , is a coarse representation of the image f , containing smooth parts and prominent edges, whereas the residual v_λ contains texture and finer details, declared as “noise” of f . The parameter λ is the *inverse* scale parameter of u_λ , i.e. a small value of λ corresponds to more details in v_λ and thus, the image u_λ is more coarse and vice versa.

As a first step, we realize that the intensity of images is quantized. If we let τ denote the small intensity quanta, then we rescale the coarse representation u_λ in τ -units. The corresponding (BV, L^2) image decomposition (1.3) takes the form

$$f = \tau u_{\lambda_0} + v_{\lambda_0}, \quad [u_{\lambda_0}, v_{\lambda_0}] := \operatorname{arginf}_{f=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_0}{\tau} \|v\|_{L^2}^2 \right\}.$$

Tadmor, Nezzar and Vese observed in [32] that for a small value of the scaling parameter λ_0 , the residual image v_{λ_0} may still contain important details when viewed at a finer scale. Thus, v_{λ_0} can be further decomposed using a refined scaling parameter $\lambda_1 > \lambda_0$,

$$v_{\lambda_0} = \tau u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \operatorname{arginf}_{v_{\lambda_0}=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_1}{\tau} \|v\|_{L^2}^2 \right\}.$$

We can continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \dots$

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} + v_{\lambda_j}, \quad [u_{\lambda_j}, v_{\lambda_j}] := \operatorname{arginf}_{v_{\lambda_{j-1}}=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \right\}. \quad (3.2)$$

Repeating this refinement step, we obtain the following *hierarchical* multiscale rep-

resentation of f , [32]

$$\begin{aligned}
f &= \tau u_{\lambda_0} + v_{\lambda_0} \\
&= \tau u_{\lambda_1} + \tau u_{\lambda_1} + v_{\lambda_1} \\
&= \dots\dots \\
&= \tau u_{\lambda_0} + \tau u_{\lambda_1} + \dots \tau u_{\lambda_N} + v_{\lambda_N}.
\end{aligned}$$

Thus, we have

$$\sum_{j=0}^N u_{\lambda_j} \tau = f - v_{\lambda_N}. \tag{3.3}$$

The Euler-Lagrange equations characterizing minimizers of (3.2) are

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} - \frac{1}{2\lambda_j} \operatorname{div} \left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right). \tag{3.4}$$

From (3.4) and (3.2) we get

$$v_{\lambda_j} = -\frac{1}{2\lambda_j} \operatorname{div} \left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right),$$

and inserting this into (3.3) yields the hierarchical decomposition of f as

$$\sum_{j=0}^N u_{\lambda_j} \tau = f + \frac{1}{2\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right). \tag{3.5}$$

We consider a multiscale scaling, continuous in time, $u(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ such that $u_{\lambda_j}(x) \mapsto u(x, t^j := j\tau)$. Observe that the right hand side of (3.5) is homogeneous of degree zero. Letting $\tau \rightarrow 0$, the hierarchical description (3.5) motivates a multiscale representation $u(x, \cdot)$ which is sought as a solution to our IDE (3.1),

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0. \tag{3.6}$$

So far, we have not specified the parameter $\lambda(t)$, which can be chosen as *any* positive, increasing function, to serve as an *inverse scaling function*: as $\lambda(t) \rightarrow \infty$, the image computed in (3.6)

$$U(t) := \int_0^t u(x, s) ds,$$

extracts consecutively smaller scale slices of the original image f . Here, $u(x, t)$ denotes the *speed* at which the image $U(t)$ changes with time. The residual, $V(t) := f - U(t)$ contains texture and noisy parts of f . An example for a multiscale representation of an image f ,

$$\left\{ U(\cdot, t) := \int_0^t u(\cdot, s) ds \right\}_{t \geq 0},$$

is depicted in Figure 3.1. The numerical scheme for its evolution using the IDE (3.1) is prescribed in section 6.3.

Remark 3.1. *It is instructive to compare our IDE model (3.6) with the time dependent PDE used in solving the ROF minimization, (1.4). In contrast to the forward scale PDE realization of (1.4b), where the solution evolves from $u(\cdot, 0) := f$ to a bigger scale image u_λ , our IDE model (3.6) is an ‘inverse scale’ model, whose solution evolves from $u(\cdot, 0) \equiv 0$ to f as $\lambda(t) \rightarrow \infty$.*

Our IDE model is motivated by a variational formulation. An important advantage of the IDE model, however, is that it is no longer limited to a variational formulation and we can therefore extend it using PDE-based modifications similar to (1.1b) and (1.1c). We will discuss these modifications in sections 3.4.1 and 3.4.3.



Figure 3.1: The above images show $\int_0^t u(\cdot, s) ds$ for the integro-differential scheme in (3.1) for $t = 1, 2, \dots, 9$. From top to bottom, left to right the images where the function $\lambda(t)$ is 0.002×2^t .

3.3 On the scaling function $\lambda(t)$

It is argued in [22] that the dual norm,

$$\|w\|_* := \sup_{\|\varphi\|_{BV} \neq 0} \frac{(w, \varphi)}{\|\varphi\|_{BV}},$$

is a proper norm to measure texture. The critical role of the scaling function $\lambda(t)$ in the ODE model (3.6) and its relationship with the star-norm is outlined in the following theorem.

Theorem 3.1. *Consider the IDE model (3.6)*

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

and let $V(\cdot, t)$ be the residual

$$V(\cdot, t) := f - U(\cdot, t).$$

Then size of the residual is dictated by the scaling function $\lambda(t)$,

$$\|V(\cdot, t)\|_* = \frac{1}{2\lambda(t)}. \quad (3.7)$$

Proof. For $\varphi \in BV(\Omega)$ we have the following

$$|(V(\cdot, t), \varphi)| = \left| \left(\frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(\cdot, t)}{|\nabla u(\cdot, t)|} \right), \varphi \right) \right| \leq \frac{1}{2\lambda(t)} \|\varphi\|_{BV}. \quad (3.8)$$

Thus, we have $\|V(\cdot, t)\|_* \leq \frac{1}{2\lambda(t)}$. Letting $\varphi = u(\cdot, t)$, we get

$$\left| \left(\frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(\cdot, t)}{|\nabla u(\cdot, t)|} \right), u(\cdot, t) \right) \right| = \frac{1}{2\lambda(t)} \|u(\cdot, t)\|_{BV}. \quad (3.9)$$

From (3.8) and (3.9) we get the desired result (3.7). \square

The importance of Theorem 3.1 lies in the fact that it enables us to dictate the star-norm of the residual. For small values of $\lambda(t)$, we get a significant amount of texture in the residual and thus, the image $U(t) := \int_0^t u(\cdot, s) ds$ will contain only features with big scale. On the other hand, as $\lambda(t)$ increases, more and more details will appear in $U(t)$. Hence, the function $\lambda(t)$ can be viewed as an ‘*inverse scale function*’ for $U(t)$. In particular, if we choose the scaling function $\lambda(t)$, such that $\lim_{t \rightarrow \infty} \lambda(t) = c$ with a prescribed constant c , then $\lim_{t \rightarrow \infty} \|V(t)\|_* = \frac{1}{2c}$. Thus, Theorem 3.1 enables us to denoise images to any pre-determined level in the BV^* sense.

The previous theorem establishes the weak convergence, $U(t) \xrightarrow{*} f$ for all L^2 -images. In fact, a stronger L^2 -convergence holds for slightly more regular images $f \in BV$. To this end we first prove the following energy decomposition, interesting in its own sake, along the lines of [32, Theorem 2.2].

Theorem 3.2. *Consider the IDE model (3.6)*

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

associated with an L^2 -image f , and let $V(\cdot, t)$ be the residual, $V(t) = f - U(t)$.

Then the following energy decomposition holds

$$\int_{s=0}^t \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV} ds + \|V(\cdot, t)\|_{L^2}^2 = \|f\|_{L^2}^2. \quad (3.10)$$

Proof. To verify (3.10), integrate (3.6) against $u(\cdot, t)$ in space and time to find

$$\int_{s=0}^t \left(U(\cdot, s), U_s(\cdot, s) \right) ds - (f, U(\cdot, t)) = - \int_{s=0}^t \frac{1}{2\lambda(s)} \|u(\cdot, s)\|_{BV} ds.$$

The expression on the left is then rewritten as

$$\begin{aligned} \int_{s=0}^t \left(U(\cdot, s), U_s(\cdot, s) \right) ds - \left(f, U(\cdot, t) \right) &= \frac{1}{2} \|U(\cdot, t)\|_{L^2}^2 - \left(f, U(\cdot, t) \right) \\ &\equiv \frac{1}{2} \left[\left(U(\cdot, t) - f, U(\cdot, t) - f \right) \right] - \frac{1}{2} \|f\|_{L^2}^2, \end{aligned}$$

and (3.10) follows from the last two equalities. \square

Remark 3.2. *A different, equivalent way of stating Theorem 3.2 is that $(u(t), V(t))$ form a maximal pair in the sense that they turn the inequality $(w, \varphi) \leq \|w\|_{BV} \|\varphi\|_*$ into equality:*

$$(u(\cdot, t), V(\cdot, t)) = \|u(\cdot, t)\|_{BV} \|V(\cdot, t)\|_*. \quad (3.11)$$

Indeed, differentiating (3.10) with respect to time we find

$$\frac{1}{\lambda(t)} \|u(\cdot, t)\|_{BV} + 2(V(\cdot, t), -u(\cdot, t)) = 0,$$

and (3.11) follows in view of (3.7), $\|V(\cdot, t)\|_* = 1/2\lambda(t)$.

We now turn to upper-bound the L^2 -size of the residual. Using the usual duality estimate together with (3.7) to find

$$\|V(\cdot, t)\|_{L^2}^2 \leq \|V(\cdot, t)\|_* \|V(\cdot, t)\|_{BV} = \frac{1}{2\lambda(t)} \|V(\cdot, t)\|_{BV}, \quad (3.12)$$

and it remains to study how fast $\|V(\cdot, t)\|_{BV}$ grows. To this end we write

$$V(x, t) = f(x) - \int_{s=0}^{t/2} u(x, s) ds - \int_{s=t/2}^t u(x, s) ds,$$

which implies

$$\|V(\cdot, t)\|_{BV} \leq \|f\|_{BV} + \lambda(t/2) \int_{s=0}^{t/2} \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV} ds + \lambda(t) \int_{s=t/2}^t \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV} ds.$$

Inserting this into (3.12) we end up with the desired upper bound,

$$\|V(\cdot, t)\|_{L^2}^2 \leq \frac{1}{2\lambda(t)} \|f\|_{BV} + \frac{\lambda(t/2)}{2\lambda(t)} \|f\|_{L^2}^2 + \int_{s=t/2}^t \frac{1}{2\lambda(s)} \|u(\cdot, s)\|_{BV} ds.$$

Now, the first term on the right vanishes for $f \in BV$ at the $t = \infty$ -limit as $\lambda(t) \uparrow \infty$; the second term vanishes if $\lambda(t)$ increases fast enough to form a Hadamard sequence so that $\lambda(t)/\lambda(t/2) \uparrow \infty$ (e.g., $\lambda(t) \sim 2^t$); and the third term vanishes at $t \uparrow \infty$ as the tail of the uniformly bounded time integral in the energy bound (3.10). We summarize by stating the following.

Theorem 3.3. *Given an image $f \in BV$, we consider the IDE model (3.6)*

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

with rapidly increasing scaling function $\lambda(t)$ so that

$$\frac{\lambda(t/2)}{\lambda(t)} \xrightarrow{t \rightarrow \infty} 0.$$

Then, f admits the multiscale representation (where equality is interpreted in L^2 -sense)

$$f(x) = \int_{s=0}^{\infty} u(x, s) ds,$$

with energy decomposition

$$\|f\|_{L^2}^2 = \int_{s=0}^{\infty} \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV} ds.$$

3.4 Modified IDE models

One of the important advantages of formulating this IDE model is that even though it is motivated by a variational formulation, it no longer needs to be asso-

ciated with any minimization problem, and we can incorporate PDE-based modifications similar to (1.1b) and (1.1c), based on our image processing needs. We will discuss these modifications in sections 3.4.1 and 3.4.3.

3.4.1 Perona-Malik models revisited

We examine the PM models and propose PM model-like modifications to IDE (3.1). Recall that one of the drawbacks of using the heat equation:

$$\frac{\partial U}{\partial t} = \Delta U,$$

for denoising is the fact that the Laplacian operator Δ results in an isotropic diffusion. Indeed, the fundamental solution of the heat equation is

$$\Phi(x, t) := \begin{cases} \frac{1}{4\pi t} e^{-\frac{|x|^2}{4t}} & (x \in \mathbb{R}^2, t > 0) \\ 0 & (x \in \mathbb{R}^2, t < 0) \end{cases}$$

which is essentially a Gaussian kernel with a standard deviation $\sigma = \sqrt{2t}$. Thus, the heat equation results in smoothing of the original image f , in all directions. Objects in a digital image are separated by prominent edges. Thus, edges are important for the interpretation of a digital image. Heat equation smooths out these important edges, along with noise.

Another way of interpreting the isotropic nature of the heat equation is to examine the decomposition of the diffusion operator. For each point $x \equiv (x_1, x_2) \in \mathbb{R}^2$, where $|\nabla U(x)| \neq 0$, we can define the vectors $\mathbf{N} = \frac{\nabla U}{|\nabla U|}$ and \mathbf{T} an orthogonal unit vector to \mathbf{N} . With the usual notation of $U_{x_1}, U_{x_2}, U_{x_1 x_2}, \dots$ for the first and partial derivatives of U , we can formally write U_{TT} and U_{NN} , the second derivatives

of U in the \mathbf{T} - direction and \mathbf{N} - direction, respectively:

$$U_{TT} = (\mathbf{T}, \nabla^2 U \mathbf{T})_2 = \frac{1}{|\nabla U|^2} (U_{x_1}^2 U_{x_2 x_2} + U_{x_2}^2 U_{x_1 x_1} - 2U_{x_1} U_{x_2} U_{x_1 x_2}), \quad (3.13)$$

$$U_{NN} = (\mathbf{N}, \nabla^2 U \mathbf{N})_2 = \frac{1}{|\nabla U|^2} (U_{x_1}^2 U_{x_1 x_1} + U_{x_2}^2 U_{x_2 x_2} - 2U_{x_1} U_{x_2} U_{x_1 x_2}).$$

Using U_{TT} and U_{NN} we can write the Laplacian operator

$$\Delta U = U_{TT} + U_{NN}, \quad (3.14)$$

which implies that the tangential component U_{TT} and the normal component U_{NN} are *equally weighted* in Laplacian diffusion. In fact, decomposition of the divergence term as a weighted sum of the two directional derivatives along \mathbf{T} and \mathbf{N} can be achieved for most classical diffusion operators [19].

Perona Malik [28] introduced a nonlinear diffusion

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|\nabla U|)\nabla U), \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0, \quad (3.15)$$

The diffusion uses $U(\cdot, 0) := f$ as the initial condition. Before going further we note that if we choose $g = 1$, we recover the heat equation. We choose the function g to vanish at ∞ to control the diffusion near edges. Hence, in smooth regions i.e when $|\nabla U|$ is relatively small, $g \approx 0$, and the PM model behaves like a heat equation, resulting in isotropic diffusion. The gradient $|\nabla U|$ takes a large value when an edge is encountered, and the function g is close to zero. Thus, at places where the edges are prominent the diffusion is controlled and the edges are not smoothed. More precisely, we can use the decomposition (3.13) to get

$$\begin{aligned} & \operatorname{div}(g(|\nabla U|)\nabla U) \\ &= 2(U_{x_1}^2 U_{x_1 x_1} + U_{x_2}^2 U_{x_2 x_2} + 2U_{x_1} U_{x_2} U_{x_1 x_2})g'(|\nabla U|) + (U_{x_1 x_1} + U_{x_2 x_2})g(|\nabla U|). \end{aligned}$$

If we define $h(s) := g(s) + 2sg'(s)$, then the PM model reads

$$\frac{\partial U}{\partial t} = g(|\nabla U|)U_{TT} + h(|\nabla U|)U_{NN}.$$

Therefore, the diffusion model now is a weighted sum of a diffusion in the \mathbf{T} - direction and a diffusion in the \mathbf{N} - direction [7]. Perona, Malik recommend the functions $e^{-s}, \frac{1}{1+s^2}$. The PM-models are not well-posed for these functions. Also, as we have discussed in Chapter 1, the gradient can get larger even for smooth regions in the presence of noise. Thus, this noise is confused as edges. These shortcomings were removed by Catteé et. al. [10]. They propose some regularization, for example, convolution with the Gaussian kernel, G_σ , to obtain a well posed problem (1.1c):

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|G_\sigma \star \nabla U|)\nabla U), \quad U \equiv U(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \left. \frac{\partial U}{\partial \mathbf{n}} \right|_{\partial\Omega} = 0,$$

where G_σ is defined as

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{|x|^2}{2\sigma^2}}.$$

We can also understand the the action of the diffusion function, g , using concepts in filtering. The function $g \approx 1$ in the absence of prominent edges, resulting in isotropic diffusion. The function g , being very small near edges, stops the diffusion and *passes* the edges, the high frequency, in the image $U(t)$. Thus, the diffusion controlling function acts as a *high-pass filter*. The practical usefulness of this filtered diffusion model can be demonstrated in case when certain edges are required in the scale-space for smaller values of t . For example, in Figure 3.3, the edges are blurred for smaller values of t with the standard IDE (3.1), but with the filtered diffusion IDE (3.16) we retain relevant edges, as shown in Figure 3.4.

3.4.2 Modified IDE with filtered diffusion

Based on the discussion of the PM models, we propose the following modification to the IDE (3.1)

$$\int_0^t u(x, s) ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x, t)|)}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad (3.16)$$

$$u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0,$$

subject to $u(\cdot, 0) := 0$. The function g is a diffusion controlling function with same properties as in the Perona-Malik model (3.15), i.e. $g(0) = 1$ and it vanishes at infinity. This approach controls the smoothing of edges as expected. The numerical experiments for (3.16) are shown in Figure 3.2.

Figure 3.2 displays the results of the modified IDE (3.16) with

$$\frac{1}{1 + (s/\beta)^2},$$

The constant β determines how the edges are preserved. If we choose β as a small number, the function g becomes zero very quickly, resulting in many edges being preserved. On the other hand, large values of β makes the function g remain closer to 1, which results in smoothing of relevant edges. Hence, one needs to compromise between preserving the edges vs. isotropic diffusion. As choices for such a g -filter,

Detailed discussion of the numerical scheme for the filtered diffusion model (3.16) are given in section 6.3.



Figure 3.2: The above images show $\int_0^t u(\cdot, s) ds$ for the modified integro-differential equation (3.16) for $t = 1, 2, \dots, 9$. From top to bottom, left to right the images where the function $\lambda(t)$ is 0.002×2^t . The diffusion controlling function used is $g(s) = \frac{1}{1+(s/\beta)^2}$ with $\beta = 5$.

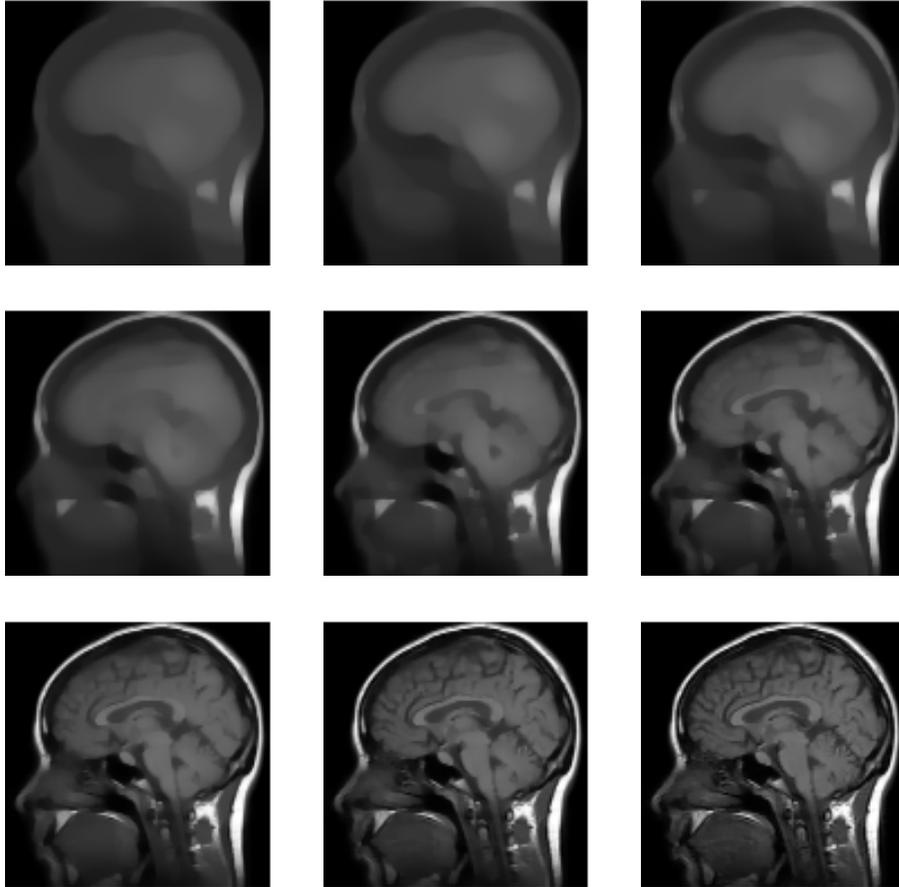


Figure 3.3: The images, $U(t) = \int_0^t u(\cdot, s) ds$, of the IDE (3.1) at $t = 1, \dots, 9$. Here, $\lambda(t) = 0.002 \times 2^t$. We see that $U(t)$ approaches the original image f as t increases.

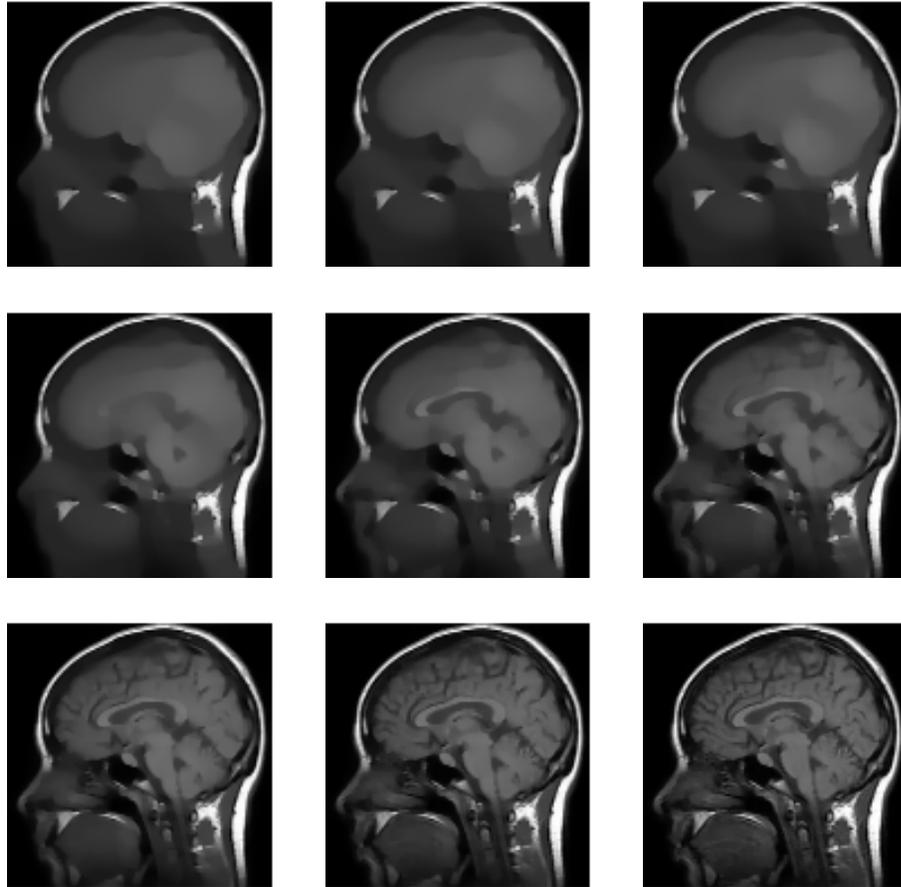


Figure 3.4: The images, $U(t) = \int_0^t u(\cdot, s) ds$, of the IDE (3.16) at $t = 1, \dots, 9$. Here, $\lambda(t) = 0.002 \times 2^t$. The diffusion controlling function used is $g(s) = \frac{1}{1+(s/\beta)^2}$ with $\beta = 5$. As in Figure 3.3 $U(t)$ approaches the original image f as t increases.

3.4.3 IDE with tangential smoothing modification

The approach of using the diffusion controlling function works well with natural images. In case of images, in which the boundaries of objects in a given image are marked with high gradients (for example, characteristic function χ_D , where D is a disk in Ω), we can choose to smooth only in the tangential direction to the boundaries of the objects. Indeed, as we see in (3.14), the Laplacian ΔU can be decomposed as a sum of the tangential and normal components of the second derivative $\nabla^2 U$. Thus, as in [4], we can only diffuse in the tangential direction. We look at the following integro-differential scheme

$$\int_0^t u(x, s) ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x, t)|) |\nabla u(x, t)|}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad (3.17)$$

$$u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0,$$

with $u(\cdot, 0) := 0$. As expected this formulation does not diffuse edges when they are well defined. The results of the numerical experiments are shown in Figure 3.5.

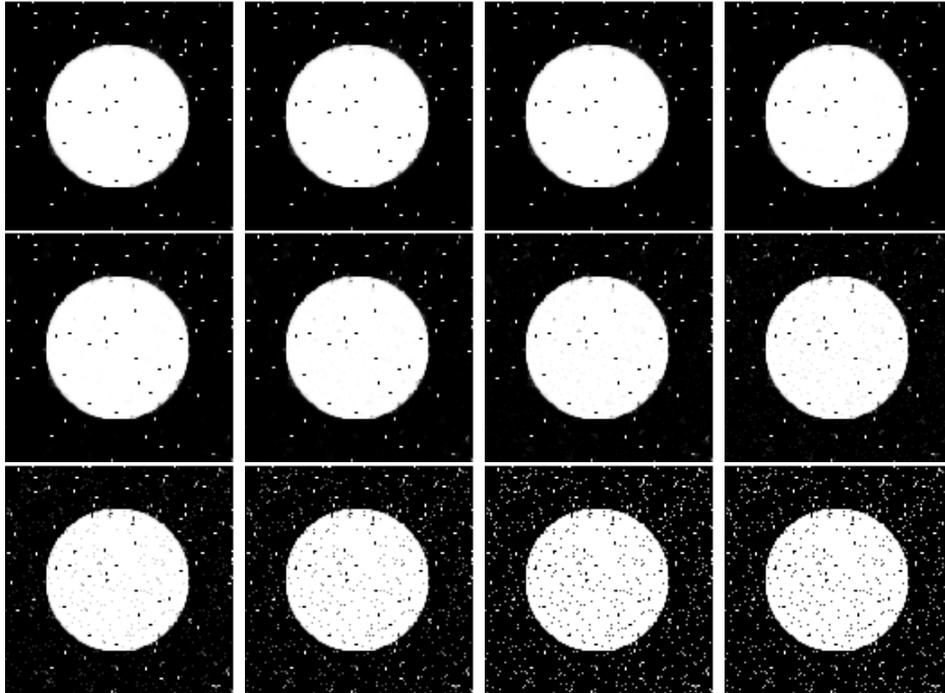


Figure 3.5: The first image above is the given blurred image f . The other images show $\int_0^t u(\cdot, s) ds$ for $t = 1, 2, \dots, 12$, with a deblurring integro-differential equation (3.17). The function $\lambda(t)$ is 0.002×2^t .

Chapter 4

Deblurring integro-differential equation model

4.1 Introduction

Blurring of an image is caused due to many reasons, for example, unfocused lens, relative motion between the camera and an object in the image, etc. Thus, instead of a clean image U , we capture its blurred version $f = TU$. Recovery of the unblurred image U , from a given blurred image f is called *image deblurring*. Blurring is modeled by a linear, continuous blurring operator, $T : L^2(\Omega) \rightarrow L^2(\Omega)$, such as a convolution with a Gaussian kernel, directional averaging etc. In section 4.2 we review a multiscale hierarchical decomposition method proposed by Tadmor, Nezzar Vese in [33]. In section 4.3 we propose a novel integro-differential equation based on the TNV deblurring decomposition, and examine its properties.

4.2 Deblurring based on hierarchical (BV, L^2) decomposition

In [33] Tadmor et. al. incorporate blurring in (BV, L^2) decomposition

$$f = Tu_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \operatorname{arginf}_{f=Tu+v} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \}. \quad (4.1)$$

The Euler-Lagrange equation for (4.1) is

$$T^*f = T^*Tu_\lambda - \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right),$$

where T^* is the adjoint of the operator T . As discussed in [33], starting with $v_{-1} = f$ and for a sequence of increasing positive real numbers $\lambda_0 < \lambda_1 < \lambda_2 \dots$ we get the following iteration

$$v_{\lambda_{k-1}} = Tu_{\lambda_k} + v_{\lambda_k}, \quad \operatorname{arginf}_{v_{\lambda_{k-1}}=Tu+v} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^2}^2 \}.$$

This gives us a *hierarchical multiscale representation* of the blurred image f

$$\left. \begin{aligned} f &= Tu_{\lambda_0} + v_{\lambda_0} \\ &= Tu_{\lambda_1} + Tu_{\lambda_1} + v_{\lambda_1} \\ &= \dots\dots\dots \\ &= Tu_{\lambda_0} + Tu_{\lambda_1} + \dots Tu_{\lambda_N} + v_{\lambda_N}. \end{aligned} \right\} \quad (4.2)$$

Thus,

$$\sum_{k=0}^N Tu_{\lambda_k} = f - v_{\lambda_N}.$$

Tadmor et. al. [33] consider the following variational problem:

$$f = Tu_{\lambda} + v_{\lambda}, \quad [u_{\lambda}, v_{\lambda}] := \operatorname{arginf}_{f=Tu+v} \{ \|u\|_{\phi} + \lambda \|v\|_{L^2}^2 \}. \quad (4.3)$$

where the *regularization functional* $\|\cdot\| : X \subset L^2(\Omega) \rightarrow [0, \infty]$ is a semi-norm which takes a general form,

$$\|u\|_{\phi} := \int_{\Omega} \phi(D^p u), \quad p \geq 1.$$

If $\|u\|_{\phi} = \|u\|_{BV} = \int_{\Omega} |\nabla u|$ is the BV -semi-norm of u , then the variational problem (4.3) is the same as (4.1). We have the following theorem [33] characterizing the minimizer u_{λ} of (4.3).

Theorem 4.1. *Let $T : L^2(\Omega) \rightarrow L^2(\Omega)$ be a linear continuous blurring operator with adjoint T^* then*

1. The variational problem (4.3) admits a minimizer u . Moreover, if $\|u\|_\phi$ is strictly convex, then a minimizer u_λ with $\|u_\lambda\|_\phi \neq 0$ is unique.

2. u_λ is a minimizer of (4.3) if and only if

$$(Tu_\lambda, v_\lambda)_2 = \|u_\lambda\|_\phi \|v_\lambda\|_* = \frac{1}{2\lambda} \|u_\lambda\|_\phi. \quad (4.4)$$

Note that the hierarchical decomposition (4.2) is of the type (2.9) described in Chapter 2 and thus, the Lemma 2.3 applies. Combining this fact with (4.4) we have,

$$\sum_{k=0}^{\infty} \left[\|u_k\|_2^2 + \frac{1}{2\lambda_k} \|u_k\|_\phi \right] = \|f\|_2^2$$

if and only if the residuals, v_k , strongly converge to 0 in L^2 .

For $f \in BV$ we have $\|v_k\|_2 \rightarrow 0$, which yields the following energy decomposition

$$\sum_{k=0}^{\infty} \left[\|u_k\|_2^2 + \frac{1}{2\lambda_k} \|u_k\|_{BV} \right] = \|f\|_2^2.$$

4.3 A novel deblurring integro-differential equation

We now extend our IDE model to deblur of images. Blurring is modeled by a continuous, linear operator $T : L^2(\Omega) \rightarrow L^2(\Omega)$. Examples of a blurring operator include convolution with a Gaussian kernel, directional averaging etc. Thus, a observed image is expressed as $f = TU$, where U is the “clean” unblurred image which we aim to recover. Hierarchical decomposition of blurred images was discussed in [33]. To this end, one sets a sequence of increasing scaling parameters

$\lambda_0 < \lambda_1 < \lambda_2 \dots$. Starting with $v_{-1} = f$, we get the following iteration

$$v_{\lambda_{j-1}} = \tau T u_{\lambda_j} + v_{\lambda_j}, \quad \underset{v_{\lambda_{j-1}} = \tau T u + v}{\operatorname{arginf}} \left\{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \right\}.$$

This gives us a *hierarchical multiscale representation* of the blurred image f presented in [33],

$$\begin{aligned} f &= \tau T u_{\lambda_0} + v_{\lambda_0} \\ &= \tau T u_{\lambda_1} + \tau T u_{\lambda_1} + v_{\lambda_1} \\ &= \dots \\ &= \tau T u_{\lambda_0} + \tau T u_{\lambda_1} + \dots \tau T u_{\lambda_N} + v_{\lambda_N}. \end{aligned}$$

Thus, after applying the conjugate T^* to the above equation we obtain,

$$\tau \sum_{j=0}^N T^* T u_{\lambda_j} = T^* f - T^* v_{\lambda_N}. \quad (4.5)$$

Using the Euler-Lagrange characterization of the minimizer in (4.3),

$$T^* v_{\lambda_{j-1}} = \tau T^* T u_{\lambda_j} - \frac{1}{2\lambda_j} \operatorname{div} \left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right),$$

which, in view of $T^* v_{\lambda_{j-1}} = \tau T^* T u_{\lambda_j} + T^* v_{\lambda_j}$ implies

$$T^* v_{\lambda_j} = -\frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right).$$

Using the above expression we can rewrite (4.5) as

$$\sum_{j=0}^N T^* T u_{\lambda_j} \tau = T^* f + \frac{1}{2\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right). \quad (4.6)$$

As $\tau \rightarrow 0$, the expression (4.6) motivates the following integro-differential equation

(IDE) for deblurring, where $u(x, t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ is sought such that

$$\int_0^t T^* T u(x, s) ds = T^* f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right); \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0. \quad (4.7)$$



(a)

(b)

Figure 4.1: Image (a) shows a blurred image of Lenna blurred using a Gaussian kernel with $\sigma = 1$. Image (b) shows the result of the deblurring integro-differential equation (4.7), as $t \rightarrow \infty$.

In this IDE, $\int_0^t u(\cdot, s) ds$ provides a multiscale representation of the *unblurred, clean image* $U(x, t) := \int_0^t u(x, s) ds$. Note that the blurring operator T is in general non-invertible for general L^2 images, but it is assumed to be invertible on the restricted set of multiscale representations $\int_0^t T^*Tu(x, s) ds$. Thus, the deblurring IDE (4.7) gives us a recipe to extract the unblurred image U from its blurred version f .

We can see the deblurring result of (4.7) in Figure 4.1. Furthermore, we can modify the deblurring integro-differential equation using edge enhancing filtering, where a $U(x, t) = \int_0^t u(x, s) ds : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ is sought as a solution of

$$T^*TU(x, t) = T^*f(x) + \frac{g(|G_\sigma \star u(x, t)|)}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right); \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0. \quad (4.8)$$

In the next chapter, we will discuss the (BV, L^1) decomposition introduced by Tony Chan and Selim Esedoğlu [12]. We will propose a multiscale image decomposition and its variants base on (BV, L^1) decomposition.

Chapter 5

Hierarchical (BV, L^1) and *weighted* (BV, L^1) multiscale schemes

5.1 Introduction

In the previous chapters, we focused on minimization of the ROF functional

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \}. \quad (5.1)$$

This variational problem lead to (BV, L^2) multiscale hierarchical decomposition and the integro-differential, introduced in Chapter 3. Using the BV seminorm, $\|u\|_{BV}$, as a regularization term was the main contribution of the ROF model. This regularization term preserves prominent edges in the image, at the same time disfavors the small oscillations, thus, denoising the image f . Nevertheless, ROF algorithm has certain limitations. One of the issue with ROF model is the loss of contrast in solutions. This issue has been studied by Strong and Chan in [31], where they show that if the given image is a characteristic function of a disk of radius R , $f := \chi_{B_R(0)}$, then any minimizer of (5.1) is of the form cf , where $c \in [0, 1)$ is a constant. We never get $u_\lambda = \chi_{B_R(0)}$, irrespective of the value of the scaling parameter λ . Moreover, it is desirable for an image denoising method to have a large class of “noise-free” images that are left invariant. As we have observed in Chapter 2, if the scaling parameter $\lambda > \frac{1}{2\|f\|_*}$, then the star-norm of the residual image $\|f - u_\lambda\|_* = \frac{1}{2\lambda}$. From this property, Chan and Esedoğlu point out in [12] that for standard ROF model this

class consists of only the trivial image $f := 0$. The work of Meyer [22] has drawn attention to the role of the fidelity term in ROF model and has inspired Vese and Osher [38] and later Osher, Sole and Vese [27] to formulate variants of the ROF model that replace the fidelity term with weaker norm. In their exposition in [12], Chan and Esedoglu investigate the implications of using L^1 norm, $\|f - u\|_{L^1}$, as a fidelity term. They propose the following variational problem:

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda \|v\|_{L^1} \}. \quad (5.2)$$

The above variational model with L^1 fidelity term was introduced and studied in the context of image denoising and deblurring by Alliney and Nikolova in [3], [25]. We follow [12] and list the following properties of the variational formulation (5.2).

1. The functional in (5.2) is not strictly convex as opposed to the functional in (5.1). This leads to non-uniqueness of minimizers.
2. Unlike the standard model (5.1), the model (5.2) is contrast invariant in the following sense: if u_λ is a solution of the minimization problem (5.2) for a given image f , then cu_λ is a solution of the modified model for the observed image cf .
3. The regularization imposed on solution by the L^1 model is more geometric, in the sense that it has less dependence on the contrast of image features than their shapes.

In the next section we elaborate on these properties more.

5.2 Basic properties of (BV, L^1) model

The following proposition [13] asserts that the (BV, L^1) model almost decouples the level sets of the given image from each other. Thus, the minimization (5.2) becomes a geometry problem for each level set.

Proposition 5.1. *Consider the energy in the minimization problem (5.2),*

$$E(u, \lambda) := \|u\|_{BV} + \lambda \|f - u\|_{L^1}, \quad (5.3)$$

can be rewritten as follows:

$$E(u, \lambda) = \int_{-\infty}^{\infty} \text{Per}(\{x : u(x) > t\}) + \lambda |\{x : u(x) > t\} \Delta \{x : f(x) > t\}| dt$$

Proof. Recall the coarea formula for the BV functions [5] (also see section 1.6):

$$\|u\|_{BV} = \int_{-\infty}^{\infty} \text{Per}(\{x : u(x) > t\}) dt \quad (5.4)$$

We resolve the second term in (5.3) as follows:

$$\begin{aligned} & \int_{\Omega} |u(x) - f(x)| dx \\ &= \int_{\{x : u(x) > f(x)\}} u(x) - f(x) dx + \int_{\{x : f(x) > u(x)\}} f(x) - u(x) dx \\ &= \int_{\{x : u(x) > f(x)\}} \int_{f(x)}^{u(x)} dt dx + \int_{\{x : f(x) > u(x)\}} \int_{u(x)}^{f(x)} dt dx \\ &= \int_{\Omega} \int_{-\infty}^{\infty} \chi_{\{x : u(x) > f(x)\}} \chi_{[f(x), u(x))}(t) + \chi_{\{x : f(x) > u(x)\}} \chi_{[u(x), f(x))}(t) dt dx \\ &= \int_{-\infty}^{\infty} \int_{\Omega} \chi_{\{x : u(x) > f(x)\}} \chi_{[f(x), u(x))}(t) + \chi_{\{x : f(x) > u(x)\}} \chi_{[u(x), f(x))}(t) dt dx \end{aligned}$$

Notice that

$$\chi_{\{x : u(x) > f(x)\}} \chi_{[f(x), u(x))}(t) = 1 \quad ,$$

if and only if

$$x \in \{x : u(x) > f(x)\} \cap \{x : u(x) > t\} \cap \{x : f(x) > t\}^c$$

and 0 otherwise. Similarly,

$$\chi_{\{x:f(x)>u(x)\}}\chi_{[u(x),f(x)]}(t) = 1 \quad ,$$

if and only if

$$x \in \{x : f(x) > u(x)\} \cap \{x : f(x) > t\} \cap \{x : u(x) > t\}^c,$$

and 0 otherwise. This implies

$$\chi_{\{x:u(x)>f(x)\}}\chi_{[f(x),u(x)]}(t) + \chi_{\{x:f(x)>u(x)\}}\chi_{[u(x),f(x)]}(t) = \chi_{\{x:u(x)>t\}}\Delta\chi_{\{x:f(x)>t\}}(x).$$

Therefore,

$$\int_{\Omega} |u(x) - f(x)| dx = \int_{-\infty}^{\infty} |\{x : u(x) > t\} \Delta \{x : f(x) > t\}| dt \quad (5.5)$$

Putting together (5.4) and (5.5) we get the assertion of the proposition. \square

We noted that the class of images f which remain invariant, i.e. $u_{\lambda} = f$, under minimization of the standard ROF functional (5.1) consists of only the trivial function $f = 0$. Now we will examine [12] which functions remain unchanged under the minimization of the functional $E(u, \lambda)$ in (5.3).

Lemma 5.1. *Given an observed image $f \in BV(\mathbb{R}^2)$, if there exists a vector field \mathbf{g} with the following properties*

1. $\mathbf{g} \in C_c^1(\mathbb{R}^N, \mathbb{R}^N)$,

2. $|\mathbf{g}| \leq 1$ for all $x \in \mathbb{R}^N$,

3. $\|f\|_{BV} = \int f \operatorname{div} \mathbf{g}$,

then there exists a threshold $\lambda_H \geq 0$ such that $u_\lambda = f$, for all $\lambda \geq \lambda_H$.

Proof. We set $\lambda_H := \max_x |\operatorname{div} \mathbf{g}|$. Then for $\lambda \geq \lambda_H$ and any $u \in BV(\Omega)$ we have

$$\begin{aligned} E(u, \lambda) &= \|u\|_{BV} + \lambda \int |u - f| \\ &\geq \int u \operatorname{div} \mathbf{g} + \lambda \int |u - f| \\ &= \int f \operatorname{div} \mathbf{g} + \lambda \int |u - f| + \int (u - f) \operatorname{div} \mathbf{g} \\ &\geq E(f, \lambda) + \left(\lambda - \max_x |\operatorname{div} \mathbf{g}| \right) \int |u - f|. \end{aligned}$$

Since $\lambda > \lambda_H := \max_x |\operatorname{div} \mathbf{g}|$, the last inequality shows that $E(u, \lambda) > E(f, \lambda)$ unless $u = f$. Thus, if $u = u_\lambda$ is a minimizer of $E(u, \lambda)$ we must have $u_\lambda = f$. \square

This lemma can be applied to binary images to obtain important class of exact solutions.

Theorem 5.1. *Let $\Sigma \subset \mathbb{R}^2$ be a bounded domain with C^2 boundary. Let the observed images f be given by the characteristic function of Σ , $f = \chi_\Sigma$. Then there exists a threshold $\lambda_H \geq 0$ such that whenever $\lambda \geq \lambda_H$, the unique minimizer of $E(u, \lambda)$ is the observed image itself, i.e. $u_\lambda = \chi_\Sigma$.*

Proof. Since the boundary $\partial\Sigma$ of the bounded domain Σ is assumed to be C^2 , the outward unit normal vector field $\mathbf{n} : \partial\Sigma \rightarrow \mathbf{S}^{N-1}$ of $\partial\Sigma$ can be extended in a C^1 manner to a tubular neighborhood of $\partial\Sigma$, so one gets a vector field $\mathbf{g} \in C_c^1(\mathbb{R}^N, \mathbb{R}^N)$,

such that $\mathbf{g}|_{\partial\Sigma} = \mathbf{n}$ and $|\mathbf{g}| \leq 1$ for all $x \in \mathbb{R}^N$. This implies

$$\int f \operatorname{div} \mathbf{g} = \int_{\Sigma} \operatorname{div} \mathbf{g} = \int_{\partial\Sigma} \mathbf{g} \cdot \mathbf{n} = \int_{\partial\Sigma} 1 = \operatorname{Per}(\partial\Sigma) = \int |\nabla f| = \|f\|_{BV}$$

Thus, the vector field \mathbf{g} satisfies all the requirements of the Lemma (5.1). Hence, the claim of the theorem follows. \square

The above theorem can be extended to more general form. Indeed, if the level sets $\{x : f(x) = t\}$ of a given image f vary smoothly with respect to t , the same conclusion holds. Now we examine the case when the scale parameter λ is taken too small.

Proposition 5.2. *Let $R > 0$. Then there exists a threshold $\lambda_L = \lambda_L(R, N)$ such that, if $f \in L^1(\mathbb{R}^N)$ with $\operatorname{supp}(f) \subset B_R(0)$, then $u_\lambda = 0$ for any $\lambda \leq \lambda_L$.*

Proof. Let $C = C(N)$ be the isoperimetric constant

$$\|u\|_{BV(\mathbb{R}^N)} \geq C(N)\|u\|_{L^{N^*}(\mathbb{R}^N)}, \quad (5.6)$$

for all $u \in BV(\mathbb{R}^N)$, where $N^* := \frac{N}{N-1}$. Then set

$$\lambda_L := \frac{C(N)}{R\omega_N^{\frac{1}{N}}}.$$

where ω_N is the volume of the unit ball in \mathbb{R}^N . Take a $\lambda > \lambda_L$ and let u be one of the minimizers of the (BV, L^1) functional in (5.3) for the given scale parameter λ .

Using the isoperimetric inequality (5.6) with $E(u, \lambda) \leq E(0, \lambda)$ we get

$$C(N)\|u\|_{L^{N^*}(\mathbb{R}^N)} + \lambda\|u - f\|_{L^1(\mathbb{R}^N)} \leq \lambda\|f\|_{L^1(\mathbb{R}^N)} = \lambda\|f\|_{L^1(B_R(0))}. \quad (5.7)$$

We split the first term on the left hand side into integrations over $B_R(0)$ and $B_R^c(0)$,

$$\begin{aligned}\|u\|_{L^{N^*}(\mathbb{R}^N)} &= \left(\int_{B_R(0)} |u|^{N^*} \right)^{\frac{1}{N^*}} + \left(\int_{B_R^c(0)} |u|^{N^*} \right)^{\frac{1}{N^*}} \\ &= \|u\|_{L^{N^*}(B_R(0))} + \|u\|_{L^{N^*}(B_R^c(0))}\end{aligned}\quad (5.8)$$

By Hölder's inequality we have,

$$\begin{aligned}\int_{B_R(0)} |u| &\leq \left(\int_{B_R(0)} |u|^{N^*} \right)^{\frac{1}{N^*}} \left(\int_{B_R(0)} 1 \right)^{\frac{1}{N}}, \\ \|u\|_{L^1(B_R(0))} &\leq (R^N \omega_N)^{\frac{1}{N}} \|u\|_{L^{N^*}(B_R(0))} = R \omega_N^{\frac{1}{N}} \|u\|_{L^{N^*}(B_R(0))},\end{aligned}$$

Thus,

$$\frac{1}{R \omega_N^{\frac{1}{N}}} \|u\|_{L^1(B_R(0))} \leq \|u\|_{L^{N^*}(B_R(0))}. \quad (5.9)$$

Using (5.8) and (5.9) in (5.7) we obtain,

$$\begin{aligned}\frac{C(N)}{R \omega_N^{\frac{1}{N}}} \|u\|_{L^1(B_R(0))} + \lambda \|u - f\|_{L^1(B_R(0))} + C(N) \|u\|_{L^{N^*}(B_R^c(0))} &\leq \lambda \|f\|_{L^1(B_R(0))}, \\ (\lambda_L - \lambda) \|u\|_{L^1(B_R(0))} + \lambda \|u\|_{L^1(B_R(0))} + \lambda \|u - f\|_{L^1(B_R(0))} + C(N) \|u\|_{L^{N^*}(B_R^c(0))} \\ &\leq \lambda \|f\|_{L^1(B_R(0))} \leq \lambda \|u\|_{L^1(B_R(0))} + \lambda \|u - f\|_{L^1(B_R(0))}.\end{aligned}$$

Thus, we have

$$(\lambda_L - \lambda) \|u\|_{L^1(B_R(0))} + C(N) \|u\|_{L^{N^*}(B_R^c(0))} \leq 0. \quad (5.10)$$

From (5.10) we conclude that if $\lambda \leq \lambda_L$, then we must have

$$\|u\|_{L^1(B_R(0))} = \|u\|_{L^{N^*}(B_R^c(0))} = 0.$$

Thus, we conclude that if $\lambda \leq \lambda_L = \frac{C(N)}{R \omega_N^{\frac{1}{N}}}$, then the unique minimizer is the trivial minimizer $u = 0$. \square

5.3 Hierarchical (BV, L^1) image decomposition

From Theorem 5.2 and Proposition 5.1 we see that there exist two critical values of the scale parameter, λ_L and λ_H . If $\lambda \leq \lambda_L$ then we have a trivial minimizer, $u_\lambda = 0$. This lower critical value λ_L depends on the shape of the object in the image. This λ_L is in general not known a-priori.

Recall, in the case of the standard ROF model, a similar critical value of the scale parameter exists, but it depended on the star-norm of the function f . Moreover, in the present case of (BV, L^1) model, if $\lambda \geq \lambda_H$, then the minimizer u_λ is the given function f itself. Nevertheless, as we have noted in the Theorem 5.1, such a scaling parameter may not exist for all images, as this λ_H depends on the function itself. In [12], Chan and Esedoğlu demonstrate that the scale space generated by λ is essentially different than the scale space generated in case of the standard ROF model.

Using these properties as a basis we propose a hierarchical image decomposition as follows. Starting with a small value of λ_0 , we can decompose the given image f using the (BV, L^1) scheme as follows:

$$f = u_{\lambda_0} + v_{\lambda_0}, \quad [u_{\lambda_0}, v_{\lambda_0}] := \underset{f=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_0 \|v\|_{L^1} \}.$$

The image v_{λ_0} can further be decomposed into smaller scale with $\lambda_1 > \lambda_0$,

$$v_{\lambda_0} = u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \underset{v_{\lambda_0}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_1 \|v\|_{L^1} \}.$$

We can continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \dots$

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^1} \}. \quad (5.11)$$

Repeating this refinement step, we obtain the following *hierarchical* (BV, L^1) *decomposition* of f :

$$\left. \begin{aligned} f &= u_{\lambda_0} + v_{\lambda_0} \\ &= u_{\lambda_1} + u_{\lambda_1} + v_{\lambda_1} \\ &= \dots \\ &= u_{\lambda_0} + u_{\lambda_1} + \dots + u_{\lambda_N} + v_{\lambda_N}. \end{aligned} \right\}$$

This yields a *hierarchical* (BV, L^1) *multiscale image decomposition*,

$$f = \sum_{k=0}^N u_{\lambda_k},$$

with a residual v_{λ_N} .

5.4 Euler-Lagrange equations

The minimizer of the (BV, L^1) functional, $E(u, \lambda)$, are characterized by the Euler-Lagrange differential equation

$$\operatorname{sgn}(u_\lambda - f) - \frac{1}{\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right) = 0.$$

When restricted to a bounded domain Ω , the Euler-Lagrange equations are augmented by the Neumann boundary condition

$$\frac{\partial u_\lambda}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0.$$

To construct the hierarchical multiscale decomposition of f , $f \sim \sum_{k=0}^N u_{\lambda_k}$, the slices u_{λ_k} are constructed as approximate solutions of the recursive relation governed by the PDE

$$\operatorname{sgn}(u_{\lambda_{k+1}} - v_{\lambda_k}) - \frac{1}{2\lambda_{k+1}} \operatorname{div} \left(\frac{\nabla u_{\lambda_{k+1}}}{|\nabla u_{\lambda_{k+1}}|} \right) = 0.$$

starting with $v_{\lambda_{-1}} := f$.

5.5 Scale-space generated by hierarchical (BV, L^1) decomposition

The scale-space generated by hierarchical (BV, L^1) decomposition is fundamentally different. The numerical result of the (BV, L^1) are shown in Figure 5.1, where the images depict $\sum_{k=0}^N u_{\lambda_k}$, for $N = 0, 1, \dots$. As argued in Proposition 5.2, the images do not vary continuously; rather, certain details appear abruptly. Compare these results with Figure 5.5, where we see the images $\sum_{k=0}^N u_{\lambda_k}$ change rather smoothly, with no sudden changes.

5.6 Edge enhancing modification to (BV, L^1) decomposition

Recall, that the standard ROF variational problem leads to the IDE:

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

$$u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0,$$

We noted in Chapter 3 that the diffusion term can be modified by multiplication by a diffusion controlling function g which vanished at infinity, to get the following IDE:

$$\int_0^t u(x, s) ds = f(x) + \frac{g(G_\sigma \star \nabla u(x, t))}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

$$u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0,$$

This modification leads to controlling of the diffusion near the edges. We propose a similar modification for hierarchical (BV, L^1) decomposition, starting $v_{\lambda_{-1}} := f$:

$$\operatorname{sgn}(u_{\lambda_{k+1}} - v_{\lambda_k}) - \frac{g(G_\sigma \star \nabla u_{\lambda_{k+1}})}{2\lambda_{k+1}} \operatorname{div} \left(\frac{\nabla u_{\lambda_{k+1}}}{|\nabla u_{\lambda_{k+1}}|} \right) = 0. \quad (5.12)$$

Figure 5.2 depict results of (5.12). We observe the enhancement around the edges due to the term $g(G_\sigma \star \nabla u_{\lambda_{k+1}})$.

5.7 Weighted (BV, L^1) scheme

As we have seen that in section 5.2 the (BV, L^1) decomposition is contrast invariant. Nevertheless, since the (BV, L^1) functional (5.3)

$$E(u, \lambda) := \|u\|_{BV} + \lambda \|f - u\|_{L^1},$$

is it not strictly convex, we do not have a unique minimizer for $E(u, \lambda)$. On the other hand, for (BV, L^2) functional, we have uniqueness of the solution, but we lose the contrast invariance property. In this section, we introduce *weighted* (BV, L^1) functional:

$$E_p(f, \lambda) := \|u\|_{BV} + \lambda \|f - u\|_{L^1}^p, \quad p > 1.$$

Since this is a strictly convex functional, it has a unique minimizer. If p is chosen to be close to 1, and it behaves like (BV, L^1) functional. Thus, we propose a hierarchical image decomposition as follows. Starting with a small value of λ_0 ,

$$v_{\lambda_0} = u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \underset{v_{\lambda_0}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_1 \|v\|_{L^1}^p \} \quad p > 1.$$

We can continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \dots$

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^1}^p \} \quad p > 1. \quad (5.13)$$

Repeating this refinement step, we obtain the following *hierarchical* (BV, L^1) *decomposition* of f :

$$\begin{aligned}
f &= u_{\lambda_0} + v_{\lambda_0} \\
&= u_{\lambda_1} + u_{\lambda_1} + v_{\lambda_1} \\
&= \dots\dots \\
&= u_{\lambda_0} + u_{\lambda_1} + \dots + u_{\lambda_N} + v_{\lambda_N}.
\end{aligned}$$

This yields a *hierarchical weighted* (BV, L^1) *multiscale image decomposition*,

$$f = \sum_{k=0}^N u_{\lambda_k},$$

with a residual v_{λ_N} .

5.8 Euler-Lagrange equations

The minimizer of the *weighted* (BV, L^1) functional, $E_p(f, \lambda)$, are characterized by the Euler-Lagrange differential equation

$$\operatorname{sgn}(u_\lambda - f) \|u_\lambda - f\|_{L^1}^{p-1} - \frac{1}{p\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right) = 0.$$

When restricted to a bounded domain Ω , the Euler-Lagrange equations are augmented by the Neumann boundary condition

$$\left. \frac{\partial u_\lambda}{\partial \mathbf{n}} \right|_{\partial\Omega} = 0.$$

To construct the hierarchical multiscale decomposition of f , $f \sim \sum_{k=0}^N u_{\lambda_k}$, the slices u_{λ_k} are constructed as approximate solutions of the recursive relation governed by the PDE

$$\operatorname{sgn}(u_{\lambda_{k+1}} - v_{\lambda_k}) \|u - v_{\lambda_k}\|_{L^1}^{p-1} - \frac{1}{p\lambda_{k+1}} \operatorname{div} \left(\frac{\nabla u_{\lambda_{k+1}}}{|\nabla u_{\lambda_{k+1}}|} \right) = 0.$$

starting with $v_{\lambda_{-1}} := f$.

Numerical results for the *weighted* (BV, L^1) scheme (5.13) for $p = 1.2$ are depicted in Figure 5.3. The numerical results with $p = 1$ are shown in Figure 5.4.

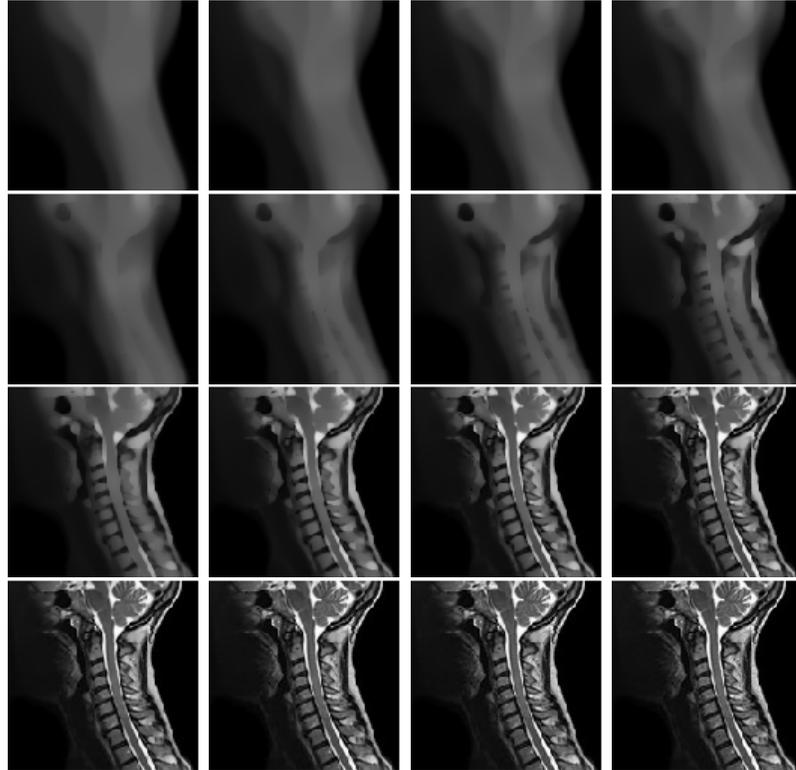


Figure 5.1: The above images show the accumulated images $\sum_{k=0}^N u_{\lambda_k}$, $N = 0, \dots, 16$, for the (BV, L^1) multiscale scheme in (5.11). From top to bottom, left to right the images the function λ_k is taken as 2^j and the grid-size, $h = 1/i_{max}$.

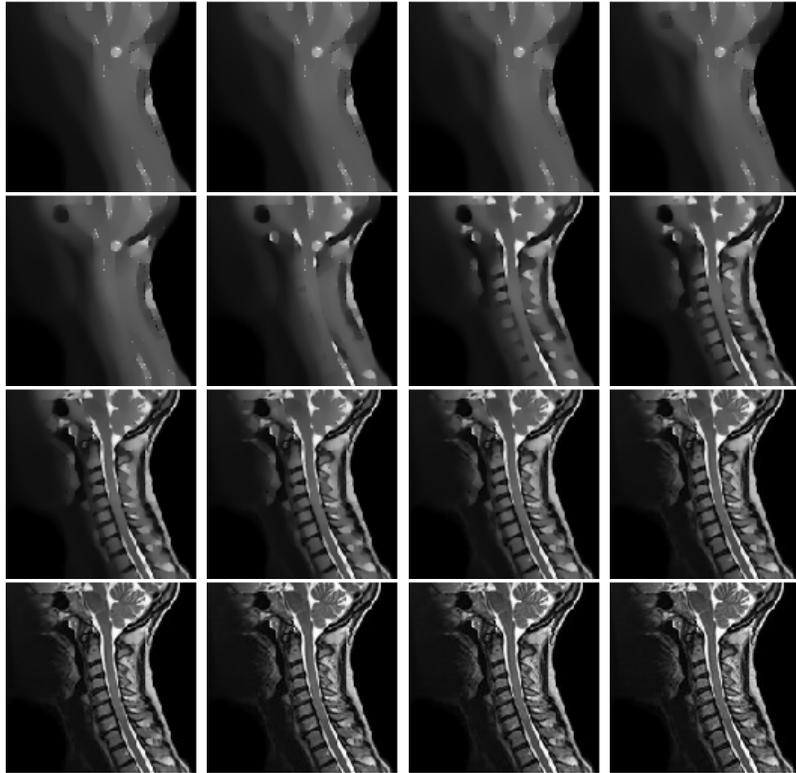


Figure 5.2: The above images show the accumulated images $\sum_{k=0}^N u_{\lambda_k}$, $N = 0, \dots, 16$, for the (BV, L^1) multiscale scheme with diffusion controlling function (5.12). From top to bottom, left to right the images the function λ_k is taken as 2^j and the grid-size, $h = 1/i_{max}$.

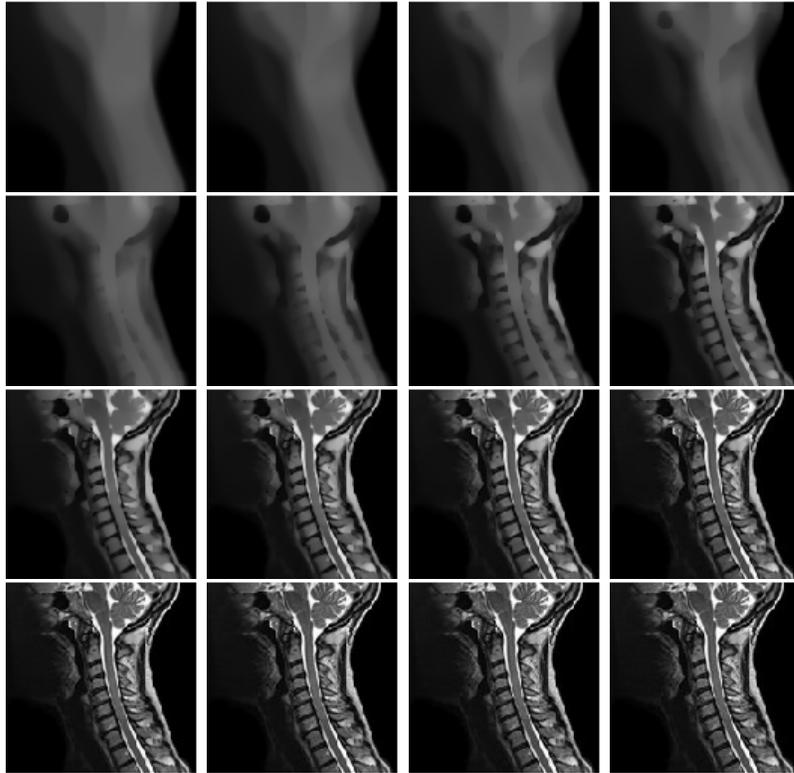


Figure 5.3: The above images show the accumulated images $\sum_{k=0}^N u_{\lambda_k}$, $N = 0, \dots, 16$, for the *weighted* (BV, L^1) multiscale scheme in (5.13) with $p = 1.2$. From top to bottom, left to right the images the function λ_k is taken as 2^j and the grid-size, $h = 1/i_{max}$.

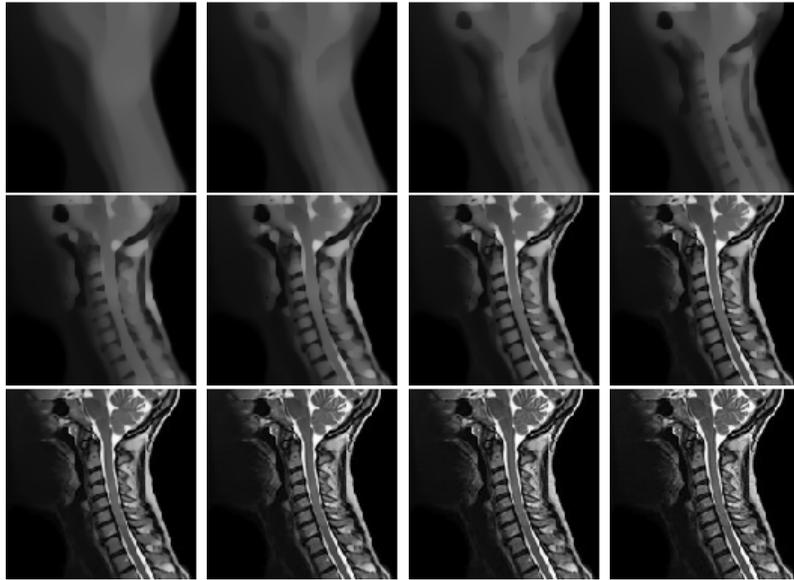


Figure 5.4: The above images show the accumulated images $\sum_{k=0}^N u_{\lambda_k}$, $N = 0, \dots, 8$, for the *weighted* (BV, L^1) multiscale scheme in (5.13) with $p = 2$. From top to bottom, left to right the images the function λ_k is taken as 0.1×2^j and the grid-size, $h = 1/i_{max}$.

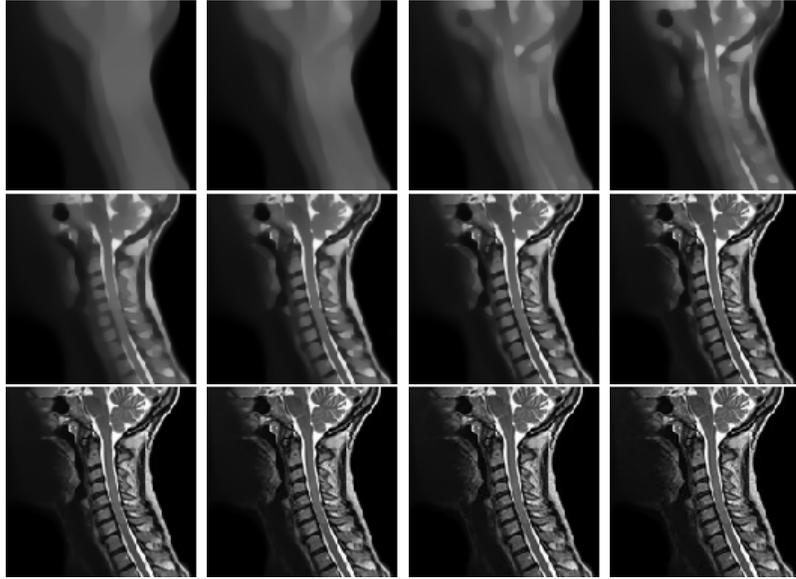


Figure 5.5: The above images show the accumulated images $\sum_{k=0}^N u_{\lambda_k}$, $N = 0, \dots, 12$, for the (BV, L^2) multiscale scheme in (5.11). From top to bottom, left to right the images the function λ_k is taken as 0.1×2^j and the grid-size, $h = 1/i_{max}$.

Chapter 6

Numerical scheme

6.1 Introduction

In Chapters 1 through Chapter 5 we proposed numerous algorithms and presented their results. In this Chapter we describe the numerical schemes used for the implementations of these models.

6.2 Numerical scheme for hierarchical (BV, L^2) multiscale representation

We begin with the hierarchical (BV, L^2) multiscale representation for a sequence of scale parameters $\{\lambda_k\}_{k \geq 0}$, starting with $v_{\lambda_{-1}} := f$,

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}}=u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^2}^2 \}. \quad (6.1)$$

We begin with regularization. To remove the singularity when $|\nabla u| = 0$, we replace $\|u\|_{BV} := \int_{\Omega} |\nabla u|$ with $\int_{\Omega} \sqrt{\varepsilon^2 + |\nabla u|^2}$. The minimizer u_{λ_0} for the *regularized* variational problem (6.1) with $v_{\lambda_{-1}} := f$ satisfies the following Euler-Lagrange equations in u :

$$u = f + \frac{1}{2\lambda_0} \operatorname{div} \left(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}} \right); \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0. \quad (6.2)$$

To compute the solution u_{λ_0} we follow [32]. The region Ω is covered with computational grid $(x_i, y_j) = (ih, jh)$ where h is a cell size. Let $D_+ = D_+(h)$, $D_- = D_-(h)$,

and $D_0 := (D_+ + D_-)/2$ denote the usual forward, backward, and centered divided difference. Thus, $D_{+x}u_{i,j} = (u_{i+1,j} - u_{i,j})/h$, $D_{-x}u_{i,j} = (u_{i,j} - u_{i-1,j})/h$, $D_{+y}u_{i,j} = (u_{i,j+1} - u_{i,j})/h$, $D_{-y}u_{i,j} = (u_{i,j} - u_{i,j-1})/h$, $D_{0x}u_{i,j} = (u_{i+1,j} - u_{i-1,j})/2h$ and $D_{0y}u_{i,j} = (u_{i,j+1} - u_{i,j-1})/2h$. With this notation (6.2) is discretized as follows:

$$\begin{aligned}
u_{i,j} &= f_{i,j} + \frac{1}{2\lambda_0} D_{-x} \left[\frac{D_{+x}u_{i,j}}{\sqrt{\varepsilon^2 + (D_{+x}u_{i,j})^2 + (D_{0y}u_{i,j})^2}} \right] \\
&\quad + \frac{1}{2\lambda_0} D_{-y} \left[\frac{D_{+y}u_{i,j}}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j})^2 + (D_{+y}u_{i,j})^2}} \right] \\
&= f_{i,j} + \frac{1}{2\lambda_0 h^2} \left[\frac{u_{i+1,j} - u_{i,j}}{\sqrt{\varepsilon^2 + (D_{+x}u_{i,j})^2 + (D_{0y}u_{i,j})^2}} - \frac{u_{i,j} - u_{i-1,j}}{\sqrt{\varepsilon^2 + (D_{-x}u_{i,j})^2 + (D_{0y}u_{i-1,j})^2}} \right] \\
&\quad + \frac{1}{2\lambda_0 h^2} \left[\frac{u_{i,j+1} - u_{i,j}}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j})^2 + (D_{+y}u_{i,j})^2}} - \frac{u_{i,j} - u_{i,j-1}}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j-1})^2 + (D_{-y}u_{i,j})^2}} \right].
\end{aligned} \tag{6.3}$$

One can use the following fixed point iteration to solve the above discrete regularized Euler-Lagrange equation (6.3):

$$\begin{aligned}
u_{i,j}^{n+1} &= f_{i,j} + \frac{1}{2\lambda_0 h^2} \left[\frac{u_{i+1,j}^n - u_{i,j}^{n+1}}{\sqrt{\varepsilon^2 + (D_{+x}u_{i,j}^n)^2 + (D_{0y}u_{i,j}^n)^2}} - \frac{u_{i,j}^{n+1} - u_{i-1,j}^n}{\sqrt{\varepsilon^2 + (D_{-x}u_{i,j}^n)^2 + (D_{0y}u_{i-1,j}^n)^2}} \right] \\
&\quad + \frac{1}{2\lambda_0 h^2} \left[\frac{u_{i,j+1}^n - u_{i,j}^{n+1}}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j}^n)^2 + (D_{+y}u_{i,j}^n)^2}} - \frac{u_{i,j}^{n+1} - u_{i,j-1}^n}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j-1}^n)^2 + (D_{-y}u_{i,j}^n)^2}} \right].
\end{aligned}$$

We introduce the following notations

$$\left. \begin{aligned}
c_E &\equiv c_E(u) := \frac{1}{\sqrt{\varepsilon^2 + (D_{+x}u_{i,j}^n)^2 + (D_{0y}u_{i,j}^n)^2}}, \\
c_W &\equiv c_W(u) := \frac{1}{\sqrt{\varepsilon^2 + (D_{-x}u_{i,j}^n)^2 + (D_{0y}u_{i-1,j}^n)^2}}, \\
c_S &\equiv c_S(u) := \frac{1}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j}^n)^2 + (D_{+y}u_{i,j}^n)^2}}, \\
c_N &\equiv c_N(u) := \frac{1}{\sqrt{\varepsilon^2 + (D_{0x}u_{i,j-1}^n)^2 + (D_{-y}u_{i,j}^n)^2}}.
\end{aligned} \right\} \tag{6.4}$$

Using these notations the fixed point iteration (6.4) reads

$$u_{i,j}^{n+1} = \frac{2\lambda_0 h^2 f_{i,j} + c_E u_{i+1,j}^n + c_W u_{i-1,j}^n + c_S u_{i,j+1}^n + c_N u_{i,j-1}^n}{2\lambda_0 h^2 + c_E + c_W + c_S + c_N}. \tag{6.5}$$

To ensure that the Neumann boundary condition $\frac{\partial u}{\partial \mathbf{n}}|_{\partial\Omega} = 0$ holds, we extend the image by reflecting it along the boundary $\partial\Omega$ at each step of (6.5). As the initial condition we set $u_{i,j}^0 = f_{i,j}$. In order to avoid grid effects, we rotate the starting point of the scheme (6.5) between the four corners of the grid, namely, $(1, 1)$, $(i_{max}, 1)$, (i_{max}, j_{max}) and $(1, j_{max})$, and alternate whether we run the algorithm row by row or column by column. To initialize the iteration we set $u^0 = f$. The scheme (6.5) iterated for $n = 0, 1, \dots, n_\infty$, until $\|u^{n_\infty} - u^{n_\infty-1}\|$ is reduced below a preassigned tolerance, so that $u_{i,j}^{n_\infty}$ produces an accurate approximation of the fixed point steady solution $u_{\lambda_0}(x_i, y_j)$. This completes the description of the Euler-Lagrange scheme for a fixed λ_0 , which is the first step in the hierarchical (BV, L^2) multiscale decomposition, with $v_{\lambda_{-1}} := f$. To continue the hierarchical scheme for a sequence of $\{\lambda_k\}_{k \geq 0}$, we reiterate this process, each time updating the value of $v_{\lambda_{k-1}}$ to $v_{\lambda_k} = v_{\lambda_{k-1}} - u_{\lambda_k}$.

With the notations developed in this scheme, we now describe numerical schemes for the integro-differential schemes in the next section.

6.3 Numerical scheme for integro-differential equations

In this section, we describe the numerical implementation of the proposed IDE, rewritten here for convenience:

$$\int_0^t u(x, s) ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}, \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = 0. \quad (6.6)$$

Let $\Delta\tau$ be the time interval step. Thus, after N steps, $t = N\Delta\tau$

$$\int_0^t u(x, s) ds = \int_0^{N\Delta\tau} u(x, s) ds = \sum_{j=0}^{N-1} \int_{j\Delta\tau}^{(j+1)\Delta\tau} u(x, s) ds.$$

With the notations:

$$U(t) = \int_0^t u(\cdot, s) ds, \quad U^N = \int_0^{N\Delta\tau} u(x, s) ds, \quad \omega^N := u^N \Delta\tau. \quad (6.7)$$

we have, the left hand side of (6.6) as

$$\int_0^{N\Delta\tau} u(x, s) ds = U^N \approx U^{N-1} + u^N \Delta\tau = U^{N-1} + \omega^N.$$

Hence, we get the following iteration to compute $\omega^N = u^N \Delta\tau$

$$\omega_{i,j}^{n+1} = \frac{2\lambda^N h^2 (f_{i,j} - U_{i,j}^{N-1}) + c_E \omega_{i+1,j}^n + c_W \omega_{i-1,j}^n + c_S \omega_{i,j+1}^n + c_N \omega_{i,j-1}^n}{2\lambda^N h^2 + c_E + c_W + c_S + c_N}, \quad (6.8)$$

where for $N = 0$ we take $U^0 \equiv 0$. The coefficients $c_E = c_E(\omega)$, $c_W = c_W(\omega)$, $c_S = c_S(\omega)$, $c_N = c_N(\omega)$ in (6.8) are as defined in (6.4), now functions of ω . We denote by λ^N a discretized version of the function $\lambda(t)$ at $t = N\Delta\tau$. As in the case of (BV, L^2) decomposition, the Neumann condition was ensured by reflecting the image ω along the boundary $\partial\Omega$ at each iterative step and the effect of the grid lines are minimized by rotation of the starting point and alternating between running the algorithm row-wise or column-wise. The scheme (6.8) iterated for $n = 0, 1, \dots, n_\infty$, until $\|\omega^{n_\infty} - \omega^{n_\infty-1}\|$ is reduced below a preassigned tolerance, so that ω^{n_∞} produces an accurate approximation of the fixed point steady solution ω^N . Thus, we get an accurate approximation for $\int_0^{N\Delta\tau} u(\cdot, s) ds \approx U^{N-1} + \omega^N$.

Now we discuss the diffusion controlling modification to the IDE

$$\int_0^t u(x, s) ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x, t)|)}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad (6.9)$$

$$x \in \Omega \subset \mathbb{R}^2, \quad t > 0; \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0.$$

Here, the only difference is the diffusion controlling function $g(|G_\sigma \star \nabla u(x, t)|)$, where G_σ is the two-dimensional Gaussian smoothing with a predetermined value of standard deviation, σ . We approximate $g(|G_\sigma \star \nabla u(x, t)|)$ with $g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right)$. The function $g(s) = \frac{1}{1+(s/\beta)^2}$ is used in these experiments, where β is a constant. Using the notations (6.7) and (6.4), we get the following fixed point iteration for the scheme (6.9)

$$\omega_{i,j}^{n+1} = \frac{2\lambda^N h^2 (f_{i,j} - U_{i,j}^{N-1}) + g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right) (c_E \omega_{i+1,j}^n + c_W \omega_{i-1,j}^n + c_S \omega_{i,j+1}^n + c_N \omega_{i,j-1}^n)}{2\lambda^N h^2 + g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right) (c_E + c_W + c_S + c_N)},$$

which converges to the steady state solution ω^N and we get an accurate approximation for $\int_0^{N\Delta\tau} u(\cdot, s) ds \approx U^{N-1} + \omega^N$.

We propose similar numerical scheme for IDE with tangential smoothing,

$$\int_0^t u(x, s) ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x, t)|) |\nabla u(x, t)|}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad (6.10)$$

$$u := u(x, t), \quad x \in \Omega \subset \mathbb{R}^2, \quad t > 0; \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0.$$

We propose the following fixed point iteration for (6.10)

$$\omega_{i,j}^{n+1} = \frac{2\lambda^N h^2 (f_{i,j} - U_{i,j}^{N-1}) + \xi^n (c_E \omega_{i+1,j}^n + c_W \omega_{i-1,j}^n + c_S \omega_{i,j+1}^n + c_N \omega_{i,j-1}^n)}{2\lambda^N h^2 + \xi^n (c_E + c_W + c_S + c_N)},$$

where $\xi^n \equiv g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right) \left|\frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|$. This iteration is carried out until $\omega^{n\infty}$ approaches a good approximation for ω^N and we obtain $\int_0^{N\Delta\tau} u(\cdot, s) ds \approx U^{N-1} + \omega^N$.

6.4 Numerical scheme for deblurring integro-differential equations

Now we will prescribe numerical schemes for the proposed models (4.7), rewritten here for convenience.

$$\int_0^t T^* T u(x, s) ds = T^* f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right),$$

$$x \in \Omega \subset \mathbb{R}^2, \quad t > 0; \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0.$$

As in the previous section, the right hand side of the above equation can be approximated as follows

$$\hat{U}(t = N\Delta\tau) = \int_0^t T^* T u(x, s) ds = \int_0^{N\Delta\tau} T^* T u(x, s) ds = \hat{U}^N,$$

$$\hat{U}^N = \hat{U}^{N-1} + T^* T u^N \Delta\tau = \hat{U}^{N-1} + T^* T \omega^N.$$

Hence, we get the following gradient descent scheme to compute ω^N .

$$\frac{\omega_{i,j}^{n+1} - \omega_{i,j}^n}{\delta t} = T f_{i,j} - \hat{U}_{i,j}^{N-1} + \frac{1}{2\lambda^N h^2} (c_E \omega_{i+1,j}^n + c_W \omega_{i-1,j}^n + c_S \omega_{i,j+1}^n + c_N \omega_{i,j-1}^n)$$

$$- \frac{1}{2\lambda^N h^2} \omega_{i,j}^{n+1} (c_E + c_W + c_S + c_N), \quad (6.11)$$

where c_E, c_W, c_S, c_N are as defined before in (6.4), but now functions of ω , and λ^N is the discretized version of the function $\lambda(t)$. The operator T denotes the blurring operator. In our experiments, T is modeled by Gaussian blur. The scheme (6.11) iterated for $n = 0, 1, \dots, n_\infty$, until $\|\omega^{n_\infty} - \omega^{n_\infty-1}\|$ is reduced below a preassigned tolerance, so that ω^{n_∞} produces an accurate approximation of the steady solution ω^N . Thus, we get $\int_0^{N\Delta\tau} u(s) ds \approx U^N = U^{N-1} + \omega^N$.

We modify this numerical scheme for deblurring IDE (4.8) with the diffusion

controlling function, rewritten here:

$$\int_0^t T^* T u(x, s) ds = T^* f(x) + \frac{g(|G_\sigma \star u(x, t)|)}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right), \quad (6.12)$$

$$u := u(x, t), \quad x \in \Omega \subset \mathbb{R}^2, \quad t > 0; \quad \frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial \Omega} = 0.$$

We approximate $g(|G_\sigma \star \nabla u(x, t)|)$ with $g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right)$. The function $g(s) = \frac{1}{1+(s/\beta)^2}$ is used in these experiments, where β is a constant. We get the following gradient descent scheme for (6.12):

$$\begin{aligned} \frac{\omega_{i,j}^{n+1} - \omega_{i,j}^n}{\delta t} &= T f_{i,j} - \hat{U}_{i,j}^{N-1} \\ &+ \frac{g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right)}{2\lambda^N h^2} (c_E \omega_{i+1,j}^n + c_W \omega_{i-1,j}^n + c_S \omega_{i,j+1}^n + c_N \omega_{i,j-1}^n) \\ &- \frac{g\left(\left|G_\sigma \star \frac{\nabla \omega_{i,j}^n}{\Delta \tau}\right|\right)}{2\lambda^N h^2} \omega_{i,j}^{n+1} (c_E + c_W + c_S + c_N). \end{aligned} \quad (6.13)$$

Rest of the scheme remains the same as that of (6.11). The coefficients c_E, c_W, c_S, c_N are as defined before in (6.4), but now functions of ω . The operator T denotes the blurring operator which is modeled by Gaussian blur in our experiments. The scheme (6.13) is iterated for $n = 0, 1, \dots, n_\infty$, until ω^{n_∞} produces an accurate approximation of the steady solution ω^N . Thus, we get $\int_0^{N\Delta\tau} u(s) ds \approx U^N = U^{N-1} + \omega^N$.

6.5 Numerical scheme for hierarchical (BV, L^1) scheme

Now we discuss numerical scheme for hierarchical (BV, L^1) scheme:

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \operatorname{arginf}_{v_{\lambda_{k-1}}=u+v} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^1} \}. \quad (6.14)$$

Recall, that $v_{\lambda_{-1}} := f$ and $\{\lambda_k\}_{k \geq 0}$ is an increasing sequence of positive scale parameters. The Euler-Lagrange equation associated with (6.14) for $v_{\lambda_{-1}} := f$ is as follows:

$$\operatorname{sgn}(u_{\lambda_0} - f) - \frac{1}{\lambda_0} \operatorname{div} \left(\frac{\nabla u_{\lambda_0}}{|\nabla u_{\lambda_0}|} \right) = 0.$$

We regularize $\operatorname{sgn}(u_{\lambda_0} - f)$ with $\frac{u_{\lambda_0} - f}{\sqrt{(u_{\lambda_0} - f)^2 + \delta}}$, where δ is a small positive scalar. Thus, we discretize the following

$$u = f - \frac{\sqrt{(u_{\lambda_0} - f)^2 + \delta}}{\lambda_0} \operatorname{div} \left(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}} \right),$$

where u approximates u_{λ_0} . To this effect we propose the following discretization:

$$\begin{aligned} u_{i,j} &= f_{i,j} \\ &+ \frac{\sqrt{(u_{i,j} - f_{i,j})^2 + \delta}}{\lambda_0} D_{-x} \left[\frac{D_{+x} u_{i,j}}{\sqrt{\varepsilon^2 + (D_{+x} u_{i,j})^2 + (D_{0y} u_{i,j})^2}} \right] \\ &+ \frac{\sqrt{(u_{i,j} - f_{i,j})^2 + \delta}}{\lambda_0} D_{-y} \left[\frac{D_{+y} u_{i,j}}{\sqrt{\varepsilon^2 + (D_{0x} u_{i,j})^2 + (D_{+y} u_{i,j})^2}} \right] \\ &= f_{i,j} \\ &+ \frac{\sqrt{(u_{i,j} - f_{i,j})^2 + \delta}}{\lambda_0 h^2} \left[\frac{u_{i+1,j} - u_{i,j}}{\sqrt{\varepsilon^2 + (D_{+x} u_{i,j})^2 + (D_{0y} u_{i,j})^2}} - \frac{u_{i,j} - u_{i-1,j}}{\sqrt{\varepsilon^2 + (D_{-x} u_{i,j})^2 + (D_{0y} u_{i-1,j})^2}} \right] \\ &+ \frac{\sqrt{(u_{i,j} - f_{i,j})^2 + \delta}}{\lambda_0 h^2} \left[\frac{u_{i,j+1} - u_{i,j}}{\sqrt{\varepsilon^2 + (D_{0x} u_{i,j})^2 + (D_{+y} u_{i,j})^2}} - \frac{u_{i,j} - u_{i,j-1}}{\sqrt{\varepsilon^2 + (D_{0x} u_{i,j-1})^2 + (D_{-y} u_{i,j})^2}} \right]. \end{aligned}$$

This leads us to the following fixed point iteration:

$$\begin{aligned} u_{i,j}^{n+1} &= f_{i,j} \\ &+ \frac{\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta}}{\lambda_0 h^2} \left[\frac{u_{i+1,j}^n - u_{i,j}^{n+1}}{\sqrt{\varepsilon^2 + (D_{+x} u_{i,j}^n)^2 + (D_{0y} u_{i,j}^n)^2}} - \frac{u_{i,j}^{n+1} - u_{i-1,j}^n}{\sqrt{\varepsilon^2 + (D_{-x} u_{i,j}^n)^2 + (D_{0y} u_{i-1,j}^n)^2}} \right] \\ &+ \frac{\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta}}{\lambda_0 h^2} \left[\frac{u_{i,j+1}^n - u_{i,j}^{n+1}}{\sqrt{\varepsilon^2 + (D_{0x} u_{i,j}^n)^2 + (D_{+y} u_{i,j}^n)^2}} - \frac{u_{i,j}^{n+1} - u_{i,j-1}^n}{\sqrt{\varepsilon^2 + (D_{0x} u_{i,j-1}^n)^2 + (D_{-y} u_{i,j}^n)^2}} \right]. \end{aligned}$$

With the notations (6.4) in the above scheme we get,

$$u_{i,j}^{n+1} = \frac{\lambda_0 h^2 f_{i,j} + \left(\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta} \right) (c_E u_{i+1,j}^n + c_W u_{i-1,j}^n + c_S u_{i,j+1}^n + c_N u_{i,j-1}^n)}{\lambda_0 h^2 + \left(\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta} \right) (c_E + c_W + c_S + c_N)}. \quad (6.15)$$

As in the case of the (BV, L^2) decomposition, we rotate the starting point of the iterations and ensure the Neumann boundary conditions by reflecting the image u^n along $\partial\Omega$. The iteration (6.15) is carried out until u^{n_∞} approaches an accurate approximation of the steady solution u_{λ_0} . To continue the hierarchical scheme for a sequence of $\{\lambda_k\}_{k \geq 0}$, we reiterate this process, each time updating the value of $v_{\lambda_{k-1}}$ to $v_{\lambda_k} = v_{\lambda_{k-1}} - u_{\lambda_k}$.

6.6 Numerical scheme for hierarchical weighted (BV, L^1) decomposition

We proposed a hierarchical weighted (BV, L^1) decomposition for a sequence of scaling parameters $\{\lambda_k\}_k$ in Chapter 5.

$$v_{\lambda_{k-1}} = u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}} = u + v}{\operatorname{arginf}} \{ \|u\|_{BV} + \lambda_k \|v\|_{L^1}^p \} \quad p > 1. \quad (6.16)$$

with $\lambda_{-1} := f$. The Euler-Lagrange equation associated with variational problem (6.16) with $k = 0$ is as follows:

$$\operatorname{sgn}(u_{\lambda_0} - f) \|u_{\lambda_0} - f\|_{L^1}^{p-1} - \frac{1}{p\lambda_0} \operatorname{div} \left(\frac{\nabla u_{\lambda_0}}{|\nabla u_{\lambda_0}|} \right) = 0.$$

which is augmented with Neumann boundary condition. The minimizer u_{λ_0} can be approximated with the solution of the following regularized Euler-Lagrange equa-

tion:

$$u = f + \frac{\sqrt{(u-f)^2 + \hat{\delta}}}{p\lambda_0 \|u-f + \hat{\delta}\|_{L^1}^{p-1}} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right). \quad (6.17)$$

where $\hat{\delta}$ denotes a constant function, $\hat{\delta} \equiv \delta$. The solution to (6.17) can be obtained by the fixed point iteration:

$$u_{i,j}^{n+1} = \frac{\Lambda^n h^2 f_{i,j} + \left(\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta} \right) (c_E u_{i+1,j}^n + c_W u_{i-1,j}^n + c_S u_{i,j+1}^n + c_N u_{i,j-1}^n)}{\Lambda^n h^2 + \left(\sqrt{(u_{i,j}^n - f_{i,j})^2 + \delta} \right) (c_E + c_W + c_S + c_N)} \quad (6.18)$$

where $\Lambda^n := p \left(\sum_{i,j} |u_{i,j}^n - f_{i,j}| h^2 \right)^{p-1} \lambda_0$. The iteration (6.18) is carried out until u^{n_∞} approaches an accurate approximation of the steady solution u_{λ_0} . To continue the hierarchical scheme for a sequence $\{\lambda_k\}_{k \geq 0}$, we reiterate this process, each time updating the value of $v_{\lambda_{k-1}}$ to $v_{\lambda_k} = v_{\lambda_{k-1}} - u_{\lambda_k}$.

Bibliography

- [1] R. Acar and C. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems **10** (1994), no. 6, 1217–1229.
- [2] Alhazen, *The optics of ibn al-haytham, books i-iii on direct vision / translated with introduction and commentary by a.i. sabra.*, Studies of the Warburg Institute, vol. 40, 1989.
- [3] S. Alliney, *Digital filters as absolute norm regularizers*, IEEE Transactions on Signal Processing **40** (1992), no. 6, 1548–1562.
- [4] L. Alvarez, P-L. Lions, and J-M. Morel, *Image selective smoothing and edge detection by nonlinear diffusion. II*, SIAM J. Numer. Anal. **29** (1992), no. 3, 845–866.
- [5] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.
- [6] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in sobolev and bv spaces, applications to pdes and optimization*, MPS-SIAM series on optimization, Society for Industrial and Applied Mathematics, Mathematical Programming Society, Philadelphia, 2005.
- [7] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: Partial differential equations and the calculus of variations (second edition)*, Applied Mathematical Sciences, vol. 147, Springer-Verlag, 2006.
- [8] C. Bennett and R. Sharpley, *Interpolation of spaces*, Academic Press, 1988.
- [9] J. Bergh and J. Lofstrom, *Interpolation spaces: an introduction*, Springer-Verlag, New York, 1976.
- [10] F. Catté, P-L. Lions, J-M. Morel, and T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal. **29** (1992), no. 1, 182–193.
- [11] A. Chambolle and P-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math. **76** (1997), no. 2, 167–188.
- [12] T. Chan and S. Esedoğlu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math. **65** (2005), no. 5, 1817–1837 (electronic).
- [13] T. Chan and J. Shen, *Image processing and analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005, Variational, PDE, wavelet, and stochastic methods.

- [14] T. Chan, J. Shen, and L. Vese, *Variational PDE models in image processing*, Notices Amer. Math. Soc. **50** (2003), no. 1, 14–26.
- [15] A. Cohen, W. Dahmen, I. Daubechies, and R. DeVore, *Harmonic analysis of the space BV*, Rev. Mat. Iberoamericana **19** (2003), no. 1, 235–263.
- [16] Wendell H. Fleming and Raymond Rishel, *An integral formula for total gradient variation*, Arch. Math. (Basel) **11** (1960), 218–222.
- [17] R. Gonzales and R. Woods, *Digital image processing*, 3 ed., Prentice Hall, 2008.
- [18] A. Jain, *Fundamentals of digital signal processing*, Prentice-Hall international editions, 1989.
- [19] P. Kornprobst and G. Deriche, R. aand Aubert, *Nonlinear operators in image restoration*, in Proceedings of the international conference on computer vision and pattern recognition, June 1997.
- [20] S. Li, *Markov random field modeling in computer vision*, Springer-Verlag, 1995.
- [21] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1998.
- [22] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, University Lecture Series, vol. 22, 2001.
- [23] D. Mumford and J. Shah, *Boundary detection by minimizing functionals*, in Proceedings of the IEEE Computer Vision Pattern Recognition Conference, San Francisco, CA, 1985, pp. 22–26.
- [24] ———, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math. **42** (1989), no. 5, 577–685.
- [25] M. Nikolova, *Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers*, SIAM J. Numer. Anal. **40** (2002), no. 3, 965–994 (electronic).
- [26] N. Nordström, *Biased anisotropic diffusion a unified regularization and diffusion approach to edge detection*, Image and Vision Computing **8** (1990), 318–327.
- [27] S. Osher, A. Solé, and L. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul. **1** (2003), no. 3, 349–370 (electronic).
- [28] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, Tech. report, Berkeley, CA, USA, 1988.

- [29] L. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, in Proceedings of IEEE International Conference on Image Processing, November 1994.
- [30] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D **60** (1992), 259–268.
- [31] David Strong and Tony Chan, *Edge-preserving and scale-dependent properties of total variation regularization*, Inverse Problems **19** (2003), no. 6, S165–S187, Special section on imaging.
- [32] E. Tadmor, S. Nezzar, and L. Vese, *A multiscale image representation using hierarchical (BV, L^2) decompositions*, Multiscale Model. Simul. **2** (2004), no. 4, 554–579.
- [33] ———, *Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation*, Commun. Math. Sci. **6** (2008), no. 2, 1–26.
- [34] L. Tartar, *An introduction to sobolev spaces and interpolation spaces*, Springer, New York, 2007.
- [35] R. Temam, *Problèmes mathématiques en plasticité*, Gauthier-Villars, Paris, 1983.
- [36] A. Tikhonov and V. Arsenin, *Solutions of ill-posed problems*, V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977, Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.
- [37] L. Vese, *A study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optim. **44** (2001), no. 2, 131–161.
- [38] L. Vese and S. Osher, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput. **19** (2003), no. 1-3, 553–572, Special issue in honor of the sixtieth birthday of Stanley Osher.