UNIVERSITY OF CALIFORNIA

Los Angeles

Contribution to Problems in Image Restoration, Decomposition, and Segmentation by Variational Methods and Partial Differential Equations

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Mathematics

by

Linh Hue Lieu

2006

The dissertation of Linh Hue Lieu is approved.

Andrea Bertozzi

Tony Chan

Stanley Osher

Paul Thompson

Luminita Vese, Committee Chair

University of California, Los Angeles

2006

I would like to dedicate this manuscript to my mother whose unconditional love and unfaltering spirit have been the main source of strength that helped me complete graduate school.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank my Ph.D. thesis advisor, Luminita Vese, for her dedication in teaching and her deep knowledge in the field of mathematical image processing. My success in this area is virtually impossible without her guidance.

I am very thankful to my beloved Kim Yon-Seo, whose presence in my life has been a great blessing. I thank him for his love and for the revival of my spirit, without which I may still be as asleep and unaware as I had been before meeting him.

I am also very thankful to my family for their mental support. I thank my brother and sisters for their love and understanding. In particular, I thank my mother for the sacrifices she has made through her entire life for her family. Without her I certainly wouldn't be given life and be present in America today.

I would like to thank my doctoral committee members Prof. Andrea Bertozzi, Dean Tony Chan, Prof. Stanley Osher, Prof. Paul Thompson, and Prof. Luminita Vese. I am grateful to have received their support and service.

I would like to thank Larry Hill of Los Alamos National Lab and Prof. Andrea Bertozzi of UCLA Mathematics Department for financial support and coordination for the streak-camera image segmentation project in Chapter 4 of this manuscript.

Last but not least, I acknowledge the help of my friend and colleague Triet Le. Much of the work in Chapter 3 of this manuscript is done by Triet. I thank him for his talent in computing and his generosity when he insisted that I be part of that project.

VITA

1978	Born, Soc-Trang, Vietnam
1999	B.S. in Mathematics, $magna\ cum\ laude$ University of the Pacific, Stockton, California
2000-2003	Teaching Assistant Department of Mathematics University of California, Los Angeles
2003	M.A. in Mathematics University of California, Los Angeles
2004-2006	Teaching Associate, Research Assistant Department of Mathematics University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

A. Bertozzi, Y-H. Huang, L.G. Hill, L. Lieu, and M-Q. Zhu (October, 2005). *Edge detection in streak-camera images*. Paper presented at Workshop on Image Processing for Streak Photography, Los Alamos National Lab, Los Alamos, New Mexico.

L. Lieu and L. Vese (October, 2005). *Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces*. Paper presented at American Mathematical Society (AMS) Sectional Meetings, Special Session on Mathematical Methods for the Analysis of Images and High-Dimensional Data, Annandale-on-Hudson, New York.

G. Chung, T. Le, L. Lieu, N. Tanushev and L. Vese (2006). *Computational methods for image restoration, image segmentation, and texture modeling*. Proceedings of SPIE Vol. 6065, Computational Imaging IV, (C.A. Bouman, E.L. Miller and I. Pollak, Editors).

L. Lieu and L. Vese (May, 2006). *Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces*. Paper presented at Society for Industrial and Applied Mathematics (SIAM) Conference on Imaging Science, Minisymposium 44: Variational and PDE Models for Image Decomposition, Minneapolis, Minnesota.

ABSTRACT OF THE DISSERTATION

Contribution to Problems in Image Restoration, Decomposition, and Segmentation by Variational Methods and Partial Differential Equations

by

Linh Hue Lieu

Doctor of Philosophy in Mathematics, 2006

University of California, Los Angeles

Professor Luminita Vese, Chair

In the first part of this dissertation, we present two models for image reconstruction and decomposition, based on ideas of Rudin-Osher-Fatemi bounded total variation regularization [57], and on Y. Meyer's ideas of using distributional normed spaces for modeling oscillatory functions [46]. Given a degraded image $f = Ku + n$, where K is a blurring operator and n represents additive noise, to extract a clean image u, we consider variational problems of the form $\inf\{E(u) := \lambda F_1(u) + F_2(f - Ku)\}$, where F_1 and F_2 are functionals acting on some normed spaces. Inspired by Meyer and motivated by Mumford-Gidas [49] and Osher-Solé-Vese [52], we propose for the first model $F_1(u) = \int |Du|$, the total variation of u, and $F_2(f - Ku) = ||f - Ku||_{H^{-s}}$, the norm on the Hilbert-Sobolev space of negative exponent. For the second model, we consider for F_1 a general regularization penalty $\Phi(u) = \int \phi(|Du|)dx < \infty$, where ϕ is positive, increasing, and grows at most linearly at infinity, while for F_2, we propose the dual penalty $F_2(f - Ku) = \Phi^*(f - Ku)$. We present algorithms for computing the dual Φ^* and for solving the minimization problem

$\inf\{E(u) := \lambda\Phi(u) + \Phi^*(f - Ku)\}$. We also present theoretical and numerical results. In the particular case when Φ is the total variation of u, we show that the proposed model recovers the (BV, BV^*) decomposition, as in Y. Meyer's model.

In the second part of the dissertation, we present a solution to the problem of edge detection in streak-camera images, which arises from scientific experiments. We present an adaptation of the Chan-Vese segmentation model [18]: a novel approach in which we segment two dimensional images by a 1D model.

Numerical results from each model will be presented in each part.

Notations and Symbols

$\bar{\mathbb{R}}$	$\mathbb{R} \cup \{-\infty, +\infty\}$
$\mathbb{R}^+$	the set of non-negative real numbers, i.e. $[0, \infty)$
$\mathbb{Z}^+$	the set of positive integers, i.e. $1, 2, 3, ...$
$\mathbb{N}$	the set of natural numbers, i.e. $0, 1, 2, 3, ...$

Let $n \in \mathbb{Z}^+,\ \alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n,\ x = (x_1, ..., x_n) \in \mathbb{R}^n$:

$\lvert\alpha\rvert$	the modulus of α, $\lvert\alpha\rvert = \alpha_1 + \cdots + \alpha_n$
$\binom{m}{\alpha}$	the binomial coefficient $\binom{m}{\alpha} = m!/(\alpha_1! \cdots \alpha_n!)$
∂_i	the partial derivative with respect to x_i, $\partial_i = \partial/\partial x_i$
D^α	$D^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$
∇	the gradient operator, $\nabla f = (\partial_1 f, \ldots, \partial_n f)$
$\mathrm{div},\ \nabla\cdot$	the divergence operator, $\mathrm{div}\vec{p} = \nabla \cdot \vec{p} = \partial_1 p_1 + \cdots + \partial_n p_n$

Ω	a bounded domain in $\mathbb{R}^n$, with Lipschitz boundary
$\bar{\Omega}$	the closure of Ω in $\mathbb{R}^n$
Γ	the boundary of $\Omega,\ \Gamma = \partial\Omega$
ν	the outward unit normal to Γ,

$\mathcal{F}f$ or $\hat{f}$	the Fourier Transform of f (cf definition 2.1.2)
χ_Ω	the characteristic function of Ω (cf definition 1.1.3)

$C^\infty(\mathbb{R}^n)$, $C^\infty(\Omega)$	the space of infinitely differentiable functions
$C^k(\mathbb{R}^n)$, $C^k(\Omega)$	the space of differentiable functions whose partial derivatives up to order $k\mathrm{th}$ are continuous

$C_c^k(\mathbb{R}^n)$, $C_c^k(\Omega)$	the space of functions in $C^k(\mathbb{R}^n)$, $C^k(\Omega)$, with compact support
$\mathcal{D}(\mathbb{R}^n)$, $\mathcal{D}(\Omega)$	the subspace of functions in $C^\infty(\mathbb{R}^n)$, $C^\infty(\Omega)$, with compact support in $\mathbb{R}^n$, Ω; also denoted $C_c^\infty(\mathbb{R}^n)$, $C_c^\infty(\Omega)$
$\mathcal{S}(\mathbb{R}^n)$	the space of rapidly decreasing functions of $C^\infty(\mathbb{R}^n)$ such that all derivatives satisfy: $\lvert x\rvert^k \lvert D^\alpha f(x)\rvert \to 0$ as $\lvert x\rvert \to \infty$, $\forall k \in \mathbb{N},\ \alpha \in \mathbb{N}^n$, (also called the Schwartz space)
$\mathcal{D}(\mathbb{R}^n)^n$	the space of n-tuples $(\varphi_1, \ldots, \varphi_n)$ where $\varphi_i \in \mathcal{D}(\mathbb{R}^n)$ for $i = 1, ..., n$
$\mathcal{D}(\bar{\Omega})$	the space consisting of $\varphi\vert_\Omega$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$
$\mathcal{D}(\bar{\Omega})^n$	the space consisting of $\varphi\vert_\Omega$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)^n$
$BV(\Omega), BV(\mathbb{R}^n)$	the space of functions of bounded variations on Ω, $\mathbb{R}^n$
$SBV(\Omega), SBV(\mathbb{R}^n)$	the space of special functions of bounded variations on Ω, $\mathbb{R}^n$

X', or X^*	the dual of the space X
$\mathcal{S}'(\mathbb{R}^n)$	the space of tempered distributions on $\mathbb{R}^n$, i.e. the set of continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$
$\mathcal{D}'(\Omega)$	the space of distributions on Ω, i.e. the set of continuous linear functionals on $\mathcal{D}(\Omega)$
$\mathcal{M}(\Omega)$	the space of Radon measures on Ω

Let $m \in \mathbb{N}$, $p \geq 1$ and $s \in \mathbb{R}$:

$L^p_{\text{loc}} = L^p_{\text{loc}}(\mathbb{R}^n)$, or $L^p_{\text{loc}}(\Omega)$	the space of classes of measurable functions on $\mathbb{R}^n$, or Ω, such that $\lvert f(x)\rvert^p$ is locally integrable

$L^p = L^p(\mathbb{R}^n)$,	the space of classes of measurable functions on $\mathbb{R}^n$, or Ω,
or $L^p(\Omega)$	such that $\lvert f(x)\rvert^p$ is integrable
$L^\infty = L^\infty(\mathbb{R}^n)$,	the space of classes of measurable functions on $\mathbb{R}^n$, or Ω,
or $L^\infty(\Omega)$	such that $\lvert f(x)\rvert$ is essentially bounded
$W^{m,p} = W^{m,p}(\mathbb{R}^n)$,	Sobolev space consisting of all $f \in L^p$ (resp. $f \in L^p(\Omega)$)
or $W^{m,p}(\Omega)$	such that $D^\alpha f \in L^p$ (resp. $L^p(\Omega)$), $\forall \alpha \in \mathbb{N}^n, \lvert\alpha\rvert \leq m$
$W_0^{m,p}(\Omega)$	the closure of $\mathcal{D}(\Omega)$ in $W^{m,p}(\Omega)$
$H^m = H^m(\mathbb{R}^n)$,	$W^{m,2}$, (resp. $W^{m,2}(\Omega)$)
or $H^m(\Omega)$	
$H_0^m(\Omega)$	the closure of $\mathcal{D}(\Omega)$ in $H^m(\Omega)$
$H^{-m}(\Omega)$	the dual space of $H_0^m(\Omega)$
$H^s = H^s(\mathbb{R}^n)$	the Sobolev space of functions or distributions on $\mathbb{R}^n$
	(cf definition 2.1.3)
$H^s_{\text{loc}}(\mathbb{R}^n)$	the space $\{f \in \mathcal{S}'(\mathbb{R}^n) \text{ s.t. } f\varphi \in H^s,\ \forall \varphi \in \mathcal{S}(\mathbb{R}^n)\}$

$\lvert\cdot\rvert_X$	the semi-norm in the space X
$\lVert\cdot\rVert_X$	the norm in the space X
$\lVert\cdot\rVert_p$	the norm on L^p
$\lVert\cdot\rVert_\infty$	the norm on L^∞, i.e. essential sup norm
$\lVert\cdot\rVert_{-s}$	the norm in the space H^{-s}

$\rightharpoonup$	weakly convergent (convergent in the weak topology)
$\overset{*}{\rightharpoonup}$	weak* convergent (convergent in the weak* topology)
$\mathcal{L}^n$	the n-dimensional Lebesgue measure
$\lvert\Omega\rvert$	the volume of Ω in $\mathbb{R}^n$, $\lvert\Omega\rvert = \mathcal{L}^n(\Omega)$
$\mathcal{H}^n$	the n-dimensional Hausdorff measure

PART I

IMAGE RESTORATION AND DECOMPOSITION

This part of the dissertation is devoted to the study of image reconstruction and decomposition via the calculus of variations and partial differential equations.

In Chapter 1, we introduce the problem and outline the mathematical framework of our study. We also review previous works in image reconstruction that have motivated the works in this dissertation.

In Chapter 2, we present a new variational model for image restoration and decomposition. The model involves bounded total variation for regularization (in the framework of Rudin, Osher, Fatemi) and Negative-Hilbert-Sobelev spaces for fitting (following the ideas of Yves Meyer [46], Mumford-Gidas [49]). We present theoretical results on the existence and uniqueness of solutions and properties of the proposed model, as well as numerical algorithm for computing discretely the solutions. We end the chapter with numerical results from our proposed model, and comparisions.

In Chapter 3, we propose a new model for image restoration and decomposition using dual functionals and dual norms. We impose a standard regularization penalty $\Phi(u) = \int \phi(|\nabla u|)dx < \infty$ on u, where ϕ is positive, increasing and has at most linear growth at

infinity. On the residual $f - Ku$ we impose a dual penalty $\Phi^*(f - Ku) < \infty$. In particular, when ϕ is convex, homogeneous of degree one, and with linear growth (for instance the total variation of u), we recover the (BV, BV^*) decomposition of the data f, as in Y. Meyer [46]. We present a minimization algorithm for solving our proposed model. We also present experimental results and comparisons to validate our proposed model.

Chapter 1

Introduction And Motivations

In this dissertation, we will consider a gray-scale image as a function defined on a bounded subset Ω in $\mathbb{R}^2$. Generally in practice, Ω is a rectangle. In this dissertation we focus on studying variational problems, in the framework of functions of bounded variation, arising from image reconstruction.

During the formation, transmission or recording process, images may be degraded by noise and blur. Thus, an important problem in image processing is the reconstruction of the original true image u from an observed image u_0. The degradation in u_0 is generally the result of two phenomena. The first phenomenon is deterministic and is related to the mode of image acquisition (e.g. the computation of integral projections in tomography) or to defects in the imaging system (e.g. blur created by motion, lens adjustment, or atmospheric turbulence). The second phenomenon is the noise introduced by random distortions in the signal transmission. We will only consider additive noise. Then, the degradation model in consideration is:

$$u_0 = Ku + \eta,$$

where K is a linear operator representing the blur, and η is the random noise.

The recovery of u from u_0 can be considered as an inverse problem which is ill-posed in general. The information provided by u_0 is not sufficient to ensure existence, uniqueness and stability of a solution u. It is therefore necessary to regularize the problem by adding an a-prior constraint on the solution. One common regularization technique, which we shall consider extensively in this dissertation, is the bounded total variation regularization.

The process of removing noise from a noisy image u_0 can also be viewed as the decomposition of u_0 into a sum of two parts: $u_0 = u + v$, where u is the clean restored image and v is the residual. When total variation regularization is applied, the u component will be piecewise smooth. This corresponds to a 'cartoon' image. Then, the v component will necessarily contain all oscillatory patterns (such as noise and texture). We shall refer to this process as a cartoon + texture decomposition.

Our purpose in this dissertation is to study known models and to develop new models for image restoration and decomposition. In a general setting, the models in which we shall explore are energy minimization problems of the form

$$\inf_{(u,v)\in(X_1,X_2)} \left\{F(u,v) = F_1(u) + \lambda F_2(v), u_0 = Ku + v\right\}, \tag{1.1}$$

where F_1, F_2 are functionals acting on normed spaces X_1, X_2, K is a continuous linear operator from X_1 to X_2, u_0 is the initial data, and $\lambda > 0$ is a parameter. We shall apply mathematical tools from the calculus of variations and partial differential equations to our study of these problems.

1.1 Mathematical Preliminaries

In this section we outline some mathematical background essential to this manuscript. A significant part of this section covers the theories of the space of functions of bounded

variation (BV). We will also review briefly the theories on convex optimization and the calculus of variations.

1.1.1 Convexity and Lower Semicontinuity in Minimization Problems

Definitions 1.1.1 *Let X be a Banach Space with the norm $||x||$, and $F : X \to \mathbb{R}$.*

- **coercivity:** *F is X-coercive if*

$$\lim_{||x|| \to +\infty} F(x) = +\infty.$$

- **convexity:** *F is convex on X if*

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y), \quad \forall x, y \in X, \quad \lambda \in [0,1].$$

- **lower semicontinuity:** *F is called lower semicontinuous (l.s.c) if for all sequence $x_n \to x_0$, then*

$$\liminf_{x_n \to x_0} F(x_n) \geq F(x_0).$$

The same definition can be given with a weak topology, in which case we say F is weakly *l.s.c.*

Proposition 1.1.1 *Assume $F : X \to \mathbb{R}$ is convex. Then, F is weakly l.s.c. if and only if F is l.s.c.*

Consider the following minimization problem

$$\inf_{x \in X} F(x).$$

To prove existence of a minimizer, we follow the following steps:

(1) Construct a minimizing sequence $x_n \in X$, i.e.

$$\lim_{n\to\infty} F(x_n) = \inf_{x\in X} F(x).$$

(2) Obtain a weakly convergent subsequence in some way.

For example: If F is X-coercive, we obtain a uniform bound $||x_n|| \leq C$. If X is reflexive, then x_n has a weakly convergent subsequence, $x_{n_j} \rightharpoonup x_0$, for some $x_0 \in X$.

(3) To prove that x_0 is a minimizer, it suffices to show

$$\liminf_{x_{n_j}\rightharpoonup x_0} F(x_{n_j}) \geq F(x_0), \tag{1.2}$$

which implies that $F(x_0) = \inf_{x\in X} F(x)$.

Notice that if F is weakly l.s.c, then (1.2) is automatically satisfied.

1.1.2 Definitions and Properties of the Space BV

In this section we outline some of the main points on the theory of functions of bounded variation. For further details, we refer the readers to [30, 8, 11].

Let Ω denotes a domain (open, connected) of $\mathbb{R}^n$ ($n > 1$). When Ω is bounded in $\mathbb{R}^n$, we also assume the boundary $\Gamma = \partial\Omega$ of Ω is Lipschitz.

Let $\mathcal{D}(\Omega) := C_c^\infty(\Omega)$ denote the space of infinitely differentiable functions with compact support in Ω, and let $\mathcal{D}'(\Omega)$ denote the space of distributions on Ω.

Definition 1.1.1 (Functions of bounded variation) *Let $u \in L^1(\Omega)$. We say that u is a function of bounded variation in Ω if the distributional derivative of u is representable by*

a finite Radon measure, i.e.

$$\int_\Omega u\frac{\partial\varphi}{\partial x_j} = -\int_\Omega \varphi d(D_j u), \quad \forall\varphi \in C_c^1(\Omega)^n, \quad j = 1, ..., n, \tag{1.3}$$

for some $\mathbb{R}^n$-valued measure $Du = (D_1u, ..., D_nu)$ in Ω.

The space of all functions of bounded variation in Ω is denoted by $BV(\Omega)$, and the total mass of Ω in the Du-measure is denoted by $||Du||(\Omega)$.

An equivalent definition of the space $BV(\Omega)$ is obtained via the definition of the dual space as follows:

Definition 1.1.2 *Let $u \in L^1(\Omega)$. The total variation of u in Ω is defined by*

$$TV(u) := \sup\left\{\int_\Omega u\mathrm{div}(\varphi)dx,\ \varphi \in C_c^1(\Omega)^n, |\varphi| \leq 1\right\}. \tag{1.4}$$

Proposition 1.1.2 *Let $u \in L^1(\Omega)$, $u \in BV(\Omega)$ if and only if $TV(u) < \infty$. Moreover, $TV(u) = ||Du||(\Omega)$ for all $u \in BV(\Omega)$.*

Notation: As a result of Proposition 1.1.2, it is valid to refer to $||Du||(\Omega)$ as the total variation of u in Ω. We will also use the notation $\int_\Omega |Du|$ for $||Du||(\Omega)$.

For $u, v \in BV(\Omega), \alpha \in \mathbb{R}$, we clearly have

a) $||D(u + v)||(\Omega) \leq ||Du||(\Omega) + ||Dv||(\Omega)$, and

b) $||\alpha u||(\Omega) = |\alpha|||Du||(\Omega)$.

That is, the variation measure $||Du||(\Omega)$ is a seminorm on $BV(\Omega)$. It is a norm on the quotient space $BV(\Omega)/P_0(\Omega)$ where $P_0(\Omega)$ is the space of constant functions on Ω. When refering to the seminorm, we will also make use of the notation

$$|u|_{BV(\Omega)} \stackrel{\text{def}}{=} \int_\Omega |Du|.$$

The space $BV(\Omega)$ is a Banach space endowed with the norm

$$||u||_{BV(\Omega)} \stackrel{\text{def}}{=} ||u||_{L^1(\Omega)} + |u|_{BV(\Omega)}, \quad u \in BV(\Omega).$$

Example 1.1.1 *Suppose that Ω is bounded in $\mathbb{R}^n$. Recall $W^{1,1}(\Omega) \subset L^1(\Omega)$. For $v \in W^{1,1}(\Omega)$, the weak first derivative Dv exists and belongs to $L^1(\Omega)$. And, for any $\varphi \in C_c^1(\Omega)^n$, with $|\varphi| \leq 1$, we have*

$$\int_\Omega v \mathrm{div}\varphi \, dx := -\int_\Omega Dv \cdot \varphi \, dx \leq \int_\Omega |Dv| dx < \infty.$$

Hence,

$$W^{1,1}(\Omega) \subset BV(\Omega).$$

Since Ω is bounded, $L^p(\Omega) \subset L^1(\Omega)$ for $1 \leq p \leq \infty$, hence

$$W^{1,p}(\Omega) \subset BV(\Omega), \text{ for all } 1 \leq p \leq \infty.$$

Definition 1.1.3 *An $\mathcal{L}^n$-measurable subset $E \in \mathbb{R}^n$ has* **finite perimeter** *in Ω if*

$$\chi_E \in BV(\Omega),$$

where χ_E is the characteristic function of E defined as

$$\chi_E(x) = \begin{cases} 1, & x \in E \\ 0, & x \notin E. \end{cases}$$

We shall denote the total variation of χ_E in Ω, called the **perimeter of** E *in* Ω, *by*

$$||\partial E||(\Omega).$$

For E with finite perimeter in Ω, $||\partial E||$ is the *perimeter measure* of E, whose value on any open set $U \subset\subset \Omega$ is

$$||\partial E||(U) = \sup\left\{\int_E \mathrm{div}\varphi \, dx, \ \varphi \in C_c^1(U)^n, |\varphi(x)| \leq 1 \text{ for } x \in U\right\}.$$

By the Structure Theorem ([30], Ch. 5), there exists a $||\partial E||$-measurable function $\nu_E : \Omega \to \mathbb{R}^n$ such that

$$\int_E \mathrm{div}\varphi \, dx = \int_\Omega \varphi \cdot \nu_E \, d||\partial E||$$

for all $\varphi \in C_c^1(\Omega)^n$.

Example 1.1.2 *Assume E is a domain with smooth boundary in $\mathbb{R}^n$ such that $\mathcal{H}^{n-1}(\partial E \cap K) < \infty$ for each compact set $K \subset \Omega$. Then for any $\varphi \in C_c^1(\Omega)$, $|\varphi| \leq 1$,*

$$\int_E \mathrm{div}\varphi \, dx = \int_{\partial E} \varphi \cdot \nu \, d\mathcal{H}^{n-1},$$

where ν denotes the outward unit normal along ∂E. Hence, for any $V \subset\subset \Omega$,

$$\int_E \mathrm{div}\varphi \, dx = \int_{\partial E \cap V} \varphi \cdot \nu \, d\mathcal{H}^{n-1} \leq \mathcal{H}^{n-1}(\partial E \cap V) < \infty.$$

Therefore E has locally finite perimeter in Ω. Moreover,

$$||\partial E||(\Omega) = \mathcal{H}^{n-1}(\partial E \cap \Omega)$$

and

$$\nu_E = \nu \quad \mathcal{H}^{n-1}\text{-a.e. on } \partial E \cup \Omega.$$

Thus $||\partial E||(\Omega)$ *measures the "size" of* ∂E *in* Ω. *Since* $\chi_E \notin W^{1,1}_{\text{loc}}(\Omega)$ *(according to, for instance, Theorem 2, Section 4.9.2 in [30]),*

$$W^{1,1}_{\text{loc}}(\Omega) \subsetneq BV_{\text{loc}}(\Omega),$$
$$W^{1,1}(\Omega) \subsetneq BV(\Omega).$$

Theorem 1.1.1 (lower semicontinuity of variation measure) *Suppose* $u_k \in BV(\Omega)$ *(*$k = 1, ...$*) and* $u_k \to u$ *in* $L^1_{\text{loc}}(\Omega)$. *Then*

$$||Du||(\Omega) \leq \liminf_{k\to\infty} ||Du_k||(\Omega).$$

Approximation by Smooth Functions:

Definition 1.1.4 *Let* $u_k, u \in BV(\Omega),\ (k = 1, ...)$ *We say that* u_k strictly converges *to* u *in* $BV(\Omega)$ *if*

(i) $u_k \to u$ *in* $L^1(\Omega)$,

(ii) $||Du_k||(\Omega) \to ||Du||(\Omega)$ *as* $k \to \infty$.

Remark: Note that the second condition does not imply $||D(u_k - u)||(\Omega) \to 0$.

Theorem 1.1.2 (local approximation by smooth functions) *Assume* $u \in BV(\Omega)$. *There exists a sequence* $u_k \in C^\infty(\Omega) \cap BV(\Omega)$ *such that* $u_k \to u$ *strictly in* $BV(\Omega)$.

Theorem 1.1.3 *Assume* $u \in BV(\Omega)$. *There exists a sequence* $u_k \in C^\infty(\Omega) \cap BV(\Omega)$ *such that*

(i) $u_k \to u$ *in* $L^1(\Omega)$,

(ii) if $U \subset\subset \Omega$ *is such that* $||Du||(\partial U) = 0$, *then* $||Du_k||(U) \to ||Du||(U)$ *as* $k \to \infty$.

Theorem 1.1.4 (compactness) *Assume* Ω *is a bounded domain in* $\mathbb{R}^n$ *with Lipschitz boundary* Γ. *Suppose* $u_k \in BV(\Omega)$ $(k = 1, ...)$ *is a sequence satisfying*

$$\sup_k ||u_k||_{BV(\Omega)} < \infty.$$

Then there exists a subsequence u_{k_j} *and a function* $u_\infty \in BV(\Omega)$ *such that*

$$u_{k_j} \to u_\infty \text{ in } L^1(\Omega).$$

Remark: For any bounded sequence $u_k \in BV(\Omega)$, the Compactness theorem together with the Lower Semicontinuity of the variation measure imply the existence of subsequence u_{k_j} and a function $u_\infty \in BV(\Omega)$ satisfying

$$u_{k_j} \to u_\infty \text{ in } L^1(\Omega),$$

$$||Du||(\Omega) \leq \liminf_{j\to\infty} ||Du_{k_j}||(\Omega).$$

This shall be applied frequently to prove existence of solutions to the variational problems to be presented in this manuscript.

Traces and Extensions:

Throughout this section, assume Ω is a bounded domain with Lipschitz boundary Γ. Then the outward unit normal $\nu(x)$ exists for $\mathcal{H}^{n-1}$-a.e. on Γ. The following theorem guarantees the existence of the trace for functions of bounded variations.

Theorem 1.1.5 (Trace theorem) *There exists a bounded (continuous) linear mapping*

$$T : BV(\Omega) \to L^1(\Gamma, \mathcal{H}^{n-1})$$

such that

$$\int_\Omega u \mathrm{div}(\varphi) dx = -\int_\Omega \varphi \cdot dDu + \int_\Gamma \varphi \cdot \nu T u d\mathcal{H}^{n-1}$$

for all $u \in BV(\Omega)$ *and* $\varphi \in C^1(\mathbb{R}^n)^n$. *Moreover, for any* $u \in BV(\Omega)$ *and* $\mathcal{H}^{n-1}$*-a.e.* $x \in \Gamma$, *we have*

$$\lim_{r \to 0+} \frac{1}{r^n} \int_{B(x,r) \cap \Omega} |u(y) - Tu(x)| dy = 0,$$

where $B(x, r)$ *is the ball of radius* r *centered at* x.

Theorem 1.1.6 (Extension theorem) *Let* $u_1 \in BV(\Omega), u_2 \in BV(\mathbb{R}^n \backslash \bar{\Omega})$. *Define*

$$v(x) = \begin{cases} u_1(x) & \text{if } x \in \Omega \\ u_2(x) & \text{if } x \in \mathbb{R}^n \backslash \bar{\Omega}. \end{cases}$$

Then $v \in BV(\mathbb{R}^n)$ *and*

$$||Dv||(\mathbb{R}^n) = ||Du_1||(\Omega) + ||Du_2||(\mathbb{R}^n \backslash \bar{\Omega}) + \int_\Gamma |Tu_1 - Tu_2| d\mathcal{H}^{n-1}.$$

Since Ω is compact with Lipschitz boundary $\Gamma = \partial\Omega$, it has finite perimeter, i.e. $\mathcal{H}^{n-1}(\Gamma) < \infty$. Then Extension Theorem 1.1.6 implies that the extension by zeros

$$Eu(x) = \begin{cases} u(x) & \text{if } x \in \Omega \\ 0 & \text{if } x \in \mathbb{R}^n \backslash \bar{\Omega}. \end{cases}$$

belongs to $BV(\mathbb{R}^n)$ whenever $u \in BV(\Omega)$.

Isoperimetric Inequalities:

Theorem 1.1.7 (Sobolev inequality) *There is a constant $C > 0$ such that*

$$||u||_{L^{n/(n-1)}(\mathbb{R}^n)} \leq C||Du||(\mathbb{R}^n) \quad \textit{for all } u \in BV(\mathbb{R}^n).$$

Again, assume Ω is bounded in $\mathbb{R}^n$. For $u \in L^1(\Omega)$, let

$$u_\Omega \stackrel{\text{def}}{=} \left(\frac{1}{|\Omega|} \int_\Omega u(x)dx \right) \chi_\Omega, \tag{1.5}$$

where $|\Omega| = \mathcal{L}^n(\Omega)$ is the volume of Ω in $\mathbb{R}^n$.

Theorem 1.1.8 (Poincaré's inequality) *For Ω a bounded domain in $\mathbb{R}^n$ with Lipschitz boundary, there exists a constant $C_\Omega > 0$ depending only on Ω such that*

$$\int_\Omega |u - u_\Omega| dx \leq C_\Omega ||Du||(\Omega) \quad \forall u \in BV(\Omega).$$

The Sobolev inequality implies $BV(\mathbb{R}^n) \subset L^{n/(n-1)}(\mathbb{R}^n)$. The following theorem says that this is also true for bounded domains Ω with Lipschitz boundary. By Hölder's inequality, we also have $BV(\Omega) \subset L^p(\Omega)$ for any $1 \leq p \leq n/(n-1)$.

Theorem 1.1.9 (Embedding theorem) *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary. Then the embedding $BV(\Omega) \hookrightarrow L^{n/(n-1)}(\Omega)$ is continuous and $BV(\Omega) \hookrightarrow L^p(\Omega)$ is compact for all $1 \leq p < n/(n-1)$.*

Theorem 1.1.10 (Sobolev-Poincaré inequality) *The Embedding theorem together with the Poincaré inequality imply the following* Sobolev-Poincaré *inequality (which is also known*

as Poincaré-Wirtinger *inequality):*

$$||u - u_\Omega||_p \leq C_\Omega ||Du||(\Omega), \quad \forall u \in BV(\Omega), 1 \leq p \leq \frac{n}{n-1},$$

for some $C_\Omega > 0$ *depending only on* Ω.

Decomposition of the Measure Du:

For $u \in BV(\Omega)$, Du decomposes into three mutually singular measures

$$Du = \nabla u \cdot \mathcal{L}^n + C_u + J_u,$$

where $\nabla u \in L^1(\Omega)^n$ is the Radon-Nikodym derivative of Du with respect to the Lebesgue measure $\mathcal{L}^n$, C_u is the Cantor part, and J_u is the *Hausdorff* or *jump* part. We also have

$$J_u = (u^+ - u^-)\nu \cdot \mathcal{H}^{n-1}_{|S_u},$$

where S_u is the set of jump discontinuities of u whose Hausdorff dimension is at most $(n-1)$. For $\mathcal{H}^{n-1}$-almost all $x \in S_u$, $u^+(x) - u^-(x) > 0$ is the height of the jump at x, and $\nu(x) \in S^{n-1}$.

Thus, $||Du||(\Omega) = \int_\Omega |\nabla u| dx + \int_{\Omega \setminus S_u} |C_u| + \int_{S_u} (u^+ - u^-) d\mathcal{H}^{n-1}$.

The Space of Special Functions of Bounded Variation (SBV):

Definition 1.1.5 *The space of* **special functions of bounded variation** *on* Ω*, denoted* $SBV(\Omega)$ *is defined as the space of those functions* $u \in BV(\Omega)$ *for which* $C_u = 0$.

Theorem 1.1.11 *Let* $u_k \in SBV(\Omega)$ *be a sequence such that there exists a constant* $C > 0$ *with* $|u_k(x)| \leq C$ *a.e.* $x \in \Omega$ *and* $\int_\Omega |\nabla u_k|^2 dx + \mathcal{H}^{n-1}(S_{u_k}) \leq C$*. Then there exists a*

subsequence u_{k_j} *converging a.e.* x *to a function* $u \in SBV(\Omega)$. *Moreover,* ∇u_{k_j} *converges weakly in* $L^2(\Omega)^n$ *to* ∇u, *and* $\underline{\lim}\mathcal{H}^{n-1}(S_{u_{k_j}}) \geq \mathcal{H}^{n-1}(S_u)$.
(See [4] for proof).

Pairings between measures and BV functions:

Integrability conditions on the divergence of a vector field $\vec{p}$ in Ω yield trace properties for the normal component of $\vec{p}$ on $\Gamma = \partial\Omega$, as well as bounded measures that naturally arise from pairing $\vec{p}$ with the total variation measure. We will outline here several results from [60, 8].

Let Ω be a bounded domain in $\mathbb{R}^n$ with Lipschitz boundary $\Gamma = \partial\Omega$, and let $\nu(x)$ denote the outward unit normal to Γ.

We will consider the following spaces:

- $BV(\Omega)_c \stackrel{\text{def}}{=} BV(\Omega) \cap L^\infty(\Omega) \cap C(\Omega)$,
- $X(\Omega)_p \stackrel{\text{def}}{=} \{\vec{p} \in L^\infty(\Omega)^n, \operatorname{div}(\vec{p}) \in L^p(\Omega)\}$,
- $X(\Omega)_\mu \stackrel{\text{def}}{=} \{\vec{p} \in L^\infty(\Omega)^n, \operatorname{div}(\vec{p})$ is a bounded measure in $\Omega\}$.

Lemma 1 *There exists a bilinear map* $\langle \vec{p}, u\rangle_\Gamma : X(\Omega)_\mu \times BV(\Omega)_c \to \mathbb{R}$ *such that*

$\langle \vec{p}, u\rangle_\Gamma = \int_\Gamma u(x)\vec{p}(x) \cdot \nu(x) d\mathcal{H}^{n-1}$ *if* $\vec{p} \in C^1(\Omega)^n$)

$|\langle \vec{p}, u\rangle_\Gamma| \leq ||\vec{p}||_\infty \int_\Gamma |u(x)| d\mathcal{H}^{n-1} \quad \forall (\vec{p}, u) \in X(\Omega)_\mu \times BV(\Omega)_c.$

Furthermore, there is a linear operator $\gamma : X(\Omega)_n \to L^\infty(\Gamma)$ *such that*

$||\gamma(\vec{p})||_\infty \leq ||\vec{p}||_\infty,$

$\langle \vec{p}, u\rangle_\Gamma = \int_\Gamma \gamma(\vec{p})(x)u(x) d\mathcal{H}^{n-1} \quad \forall u \in BV(\Omega)_c$

$\gamma(\vec{p})(x) = \vec{p}(x) \cdot \nu(x)$ *for all* $x \in \Gamma$ *when* $\vec{p} \in C^1(\bar{\Omega})^n$.

(Proof of these results can be found in ([8], pp. 311-317))

The function $\gamma(\vec{p})$ is a weakly defined trace on Γ of the normal component of $\vec{p}$. We shall denote $\gamma(\vec{p})$ by $\vec{p} \cdot \nu$.

Clearly $X(\Omega)_p \subset X(\Omega)_\mu$ for all $p \geq 1$. Thus the trace $\vec{p} \cdot \nu$ is defined for any $\vec{p} \in X(\Omega)_p$. We also have the following Green's Formula:

Lemma 2 (Green's Formula) *Assume* $1 \leq p \leq \infty,\ p' = p/(p-1)$. *Then for all* $\vec{p} \in X(\Omega)_p$ *and* $u \in W^{1,1}(\Omega) \cap L^{p'}(\Omega)$, *we have*

$$\int_\Omega u \mathrm{div}(\vec{p})\, dx + \int_\Omega \vec{p} \cdot \nabla u dx = \int_\Gamma (\vec{p} \cdot \nu)(x) u(x) d\mathcal{H}^{n-1}.$$

Lemma 3 *For every* $u \in L^{n/(n-1)}(\Omega)$ *and for every* $\vec{p} \in L^n(\Omega)^n$ *with* $\mathrm{div}\vec{p} \in L^n(\Omega)$, *we can define a distribution on* Ω, *denoted* $\vec{p} \cdot Du$, *via the formula*

$$\langle \vec{p} \cdot Du, \varphi \rangle = -\int_\Omega (u \mathrm{div}\vec{p}) \varphi dx - \int_\Omega u(\vec{p} \cdot \nabla \varphi) dx, \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{1.6}$$

Furthermore, the mapping $(u, \vec{p}) \rightarrow \vec{p} \cdot Du$ *is bilinear weakly continuous from* $L^{n/(n-1)}(\Omega) \times X$ *into* $\mathcal{D}'(\Omega)$, *the space of distributions on* Ω, *where*

$$X = \{\vec{p} \in L^n(\Omega)^n, \mathrm{div}\vec{p} \in L^n(\Omega)\}, \tag{1.7}$$

which is a Banach Space with the natural norm

$$||\vec{p}||_X = ||\vec{p}||_n + ||\mathrm{div}\vec{p}||_n. \tag{1.8}$$

(A proof is in the Appendix).

Lemma 4 *Under the assumptions of Lemma 1 and if moreover* $u \in BV(\Omega)$ *and* $\vec{p} \in L^\infty(\Omega)^n$, *then* $\vec{p} \cdot Du$ *is a bounded signed measure with*

$$\int_\Omega |\vec{p} \cdot Du| \leq ||\vec{p}||_\infty \cdot \int_\Omega |Du|, \tag{1.9}$$

and the following generalized Green's formula holds

$$\int_\Omega \vec{p} \cdot Du = \int_\Gamma u(\vec{p} \cdot \nu) d\mathcal{H}^{n-1} - \int_\Omega u \mathrm{div}(\vec{p}) dx, \tag{1.10}$$

(A proof is in the Appendix).

1.2 Rudin-Osher-Fatemi (ROF) and Rudin-Osher Models

Assume Ω is an open and bounded subset of R^2, with $\partial\Omega$ Lipschitz. For two dimensional gray-scale images, Ω is in general a rectangle in the plane.

Assume the observed image $u_0 : \Omega \to \mathbb{R}$ is degraded by an additive noise $\eta : \Omega \to \mathbb{R}$ of zero mean and variance σ^2, i.e.

$$u_0(x, y) = u(x, y) + \eta(x, y),$$

where $u : \Omega \to R$ is the clean unknown image.

The reconstruction of u from u_0 may be attempted by utilizing the known statistics of the noise. We may try to find u such that

$$\int_\Omega u = \int_\Omega u_0, \quad \text{and} \quad \int_\Omega (u - u_0)^2 = \sigma^2.$$

The first condition corresponds to the noise having zero mean, and the second to that the

variance of the noise is σ^2.

Unfortunately, this problem is ill-posed. Existence, uniqueness, and stability of the solution u is not guaranteed due to insufficient information provided by u_0. Therefore regularization by adding a-priori constraint on the solution is necessary. It is desirable that the choice of regularization be made so that sharp edges in u_0 are maintained in u.

L. Rudin, S. Osher, and E. Fatemi propose in [57] to regularize u by requiring it to be of bounded variation. They propose the following constrained minimization problem, refered to as the ROF model, for reconstructing u from the observed image u_0 :

$$\begin{aligned} &\text{Minimize } \textstyle\int_\Omega |Du| \\ &\text{subject to } \textstyle\int_\Omega u = \int_\Omega u_0, \text{ and } \frac{1}{2}\int_\Omega |u-u_0|^2 = \sigma^2|\Omega|, \end{aligned} \tag{1.11}$$

where $\int_\Omega |Du|$ is the total variation of u.

Remark: This ROF model can be applied to decompose a function $u_0 \in L^2(\Omega)$ into a sum $u_0 = u + v$ where $u \in BV(\Omega)$ is piecewise smooth and $v \in L^2(\Omega)$ is oscillatory.

Assume further that blurring is present in the degradation process, i.e.

$$u_0(x,y) = Ku(x,y) + \eta(x,y),$$

where $K : L^2(\Omega) \to L^2(\Omega)$ is a linear and continuous operator. Rudin and Osher propose in [56] the following generalized version of (1.11) for recovering u :

$$\begin{aligned} &\text{Minimize } \textstyle\int_\Omega |Du| \\ &\text{subject to } \textstyle\int_\Omega Ku = \int_\Omega u_0, \text{ and } \frac{1}{2}\int_\Omega |Ku-u_0|^2 = \sigma^2|\Omega|. \end{aligned} \tag{1.12}$$

The space of functions of bounded variation (BV) consists of piecewise smooth func-

tions. Imposing that the reconstructed u belongs to the space (BV) allows u to have discontinuities, i.e. the reconstructed image is allowed to have sharp edges while having spurious oscillations such as noise removed. Then, the objects (which are defined by their edges) will appear sharply in the reconstructed image.

A. Chambolle and P.-L. Lions have shown in [17], via the use of the Lagrange multiplier $\lambda > 0$, that problem (1.12) is equivalent to the following unconstrained problem:

$$\inf_{u \in BV(\Omega)} \left\{ \int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega |u_0 - Ku|^2 \right\}, \tag{1.13}$$

under the following natural assumptions:

H1. K is a continuous linear operator of $L^2(\Omega)$,

H2. $K\chi_\Omega = \chi_\Omega$ (equivalently, $\int_\Omega K^* v = \int_\Omega v$ for all $v \in L^2(\Omega)$),

H3. $\eta(x)$ is an oscillatory function, representing an additive white noise,

H4. $\int_\Omega \eta = 0$, and $\sigma^2 = \int_\Omega |\eta|^2$ is known,

H5. $\int_\Omega Ku.\eta = 0$, i.e. the noise η and Ku are totally uncorrelated signals, (this and the degradation model would imply that $||u_0 - u_{0,\Omega}||_2 \geq \sigma$, where $u_{0,\Omega} := (\frac{1}{|\Omega|} \int_\Omega u_0)\chi_\Omega$.

For fixed u_0 and Ω, the Lagrange multiplier $\lambda > 0$ corresponds uniquely to the given σ, provided $0 < \sigma < ||u_0 - u_0, \Omega||_2$, (see Lemma 2.3 in [17]).

The unconstrained problem (1.13) can be viewed as a penalty method approach to solving the constrained problem (1.12). The Lagrange multiplier $\lambda > 0$ is the penalty (or regularization) parameter. It controls the tradeoff between the goodness of the fit to the data, as measured by $||u_0 - Ku||_2^2$, and the variability of the solution u, as measured by $|u|_{BV}$.

Remark: In [57], the authors already used (1.13) as a numerical device. Chambolle and Lions [17] gave theoretical analysis and proved the equivalence of the two problems.

Notations: We denote

$$J_0(u) := \int_\Omega |Du|, \quad \text{and} \quad F(u) := J_0(u) + \frac{\lambda}{2}||u_0 - Ku||_2^2.$$

Remark: F is strictly convex, i.e. $F(\frac{1}{2}u+\frac{1}{2}v) < \frac{1}{2}F(u)+\frac{1}{2}F(v)$ for all $u \neq v$. This is true since $J(u)$ is convex and $||\cdot||_2^2$ is strictly convex: $J(\frac{1}{2}u+\frac{1}{2}v) \leq \frac{1}{2}J(u)+\frac{1}{2}J(v)$, and $||\frac{1}{2}Ku+\frac{1}{2}Kv||_2^2 = \frac{1}{4}||Ku||_2^2 + \frac{1}{4}||Kv||_2^2 + \frac{1}{2}\langle u, v\rangle \leq \frac{1}{4}||Ku||_2^2 + \frac{1}{4}||Kv||_2^2 + \frac{1}{2}||Ku||_2||Kv||_2 < \frac{1}{2}||Ku||_2^2 + \frac{1}{2}||Kv||_2^2$.

Existence of a solution of (1.13), for any $u_0 \in L^2(\Omega)$ and $\lambda > 0$, is obtained as follows (see also [1, 17, 64]):

1. Since $F(0) = (\lambda/2)||u_0||_2^2 < \infty$, $\inf_{u\in BV(\Omega)} F(u) < \infty$. Therefore, we can find a minimizing sequence $\{u_k\} \in BV(\Omega)$, such that $J_0(u_k) + ||u_0 - Ku_k|| < M$, for all k.

2. Sobolev-Poincaré inequalities imply $||u_k - u_{k,\Omega}||_p \leq CJ_0(u_k) < M$, for all $k = 1, ...,$ $1 \leq p \leq 2$, where $C > 0$ depends only on Ω, and $u_{k,\Omega} = (\frac{1}{|\Omega|}\int_\Omega u_k)\chi_\Omega$.

3. $M > ||u_0 - K(u_k - u_{k,\Omega}) - Ku_{k,\Omega}||_2^2 \geq (||u_0 - K(u_k - u_{k,\Omega})||_2 - ||Ku_{k,\Omega}||_2)^2$
   $\geq ||Ku_{k,\Omega}||_2 \left(||Ku_{k,\Omega}||_2 - 2||u_0 - K(u_k - u_{k,\Omega})||_2\right)$
   $\geq ||Ku_{k,\Omega}||_2(||Ku_{k,\Omega}||_2 - 2(||u_0||_2 + ||K|| \cdot ||(u_k - u_{k,\Omega})||_2))$.
   K continuous and $||(u_k - u_{k,\Omega})||_2 < M \Rightarrow \ ||u_0||_2 + ||K|| \cdot ||(u_k - u_{k,\Omega})||_2 < M'$.
   Hence, $\frac{||K\chi_\Omega||_2}{|\Omega|}\left|\int_\Omega u_k\right| = ||Ku_{k,\Omega}||_2 < M''$. Thus, $\left|\int_\Omega u_k\right|$ is uniformly bounded.

4. For $1 \leq p \leq 2$, we have $||u_k||_p = ||u_k - u_{k,\Omega} + u_{k,\Omega}||_p \leq ||u_k - u_{k,\Omega}||_p + |\int_\Omega u_k| <$

$M + M''$, i.e. $||u_k||_p$ is uniformly bounded.

5. $||u_k||_{BV(\Omega)} := J_0(u_k) + ||u_k||_1$ is uniformly bounded. Then, Compactness Theorem implies existence of a subsequence, still denote u_k, and $u_\infty \in BV(\Omega)$ such that $u_k \longrightarrow u_\infty$ in $L^1(\Omega)$, and $u_k \rightharpoonup u_\infty$ weakly in $L^2(\Omega)$.

7. K is continuous $\Rightarrow \quad Ku_k \rightharpoonup Ku_\infty$ weakly in $L^2(\Omega)$. Then, $||u_0 - Ku_\infty||_2^2 \leq \liminf ||u_0 - Ku_k||_2^2$.

8. Lower-semicontinuity of the variation measure implies $J_0(u_\infty) \leq \liminf J_0(u_k)$.

9. Therefore, $F(u_\infty) \leq \liminf F(u_k)$, hence u_∞ is a minimizer.

Uniqueness of the minimizer is obtained when K is further assumed to be injective. Suppose that $u, v \in BV(\Omega)$ are two minimizers. F is strictly convex; therefore, if $Ku \neq Kv$, then $F(\frac{1}{2}u + \frac{1}{2}v) < \frac{1}{2}F(u) + \frac{1}{2}F(v) = \inf F$, which contradicts with u and v being minimizers. Therefore, $Ku = Kv$. Thus, we obtain uniqueness if K is injective. $\square$

Remark: We deduce from the existence proof that the functional F is: (1) lower semi-continuous with respect to the L^2 topology ($\equiv$ to the BV-w^* topology); (2) BV-coercive; to see this more clearly (see also [1]), assume $u \in BV(\Omega)$, and let $w = \left(\frac{\int_\Omega u dx}{|\Omega|}\right)\chi_\Omega$, $v = u - w$. Then,

(a) $\int_\Omega v dx = 0$, so $||v||_p \leq C_1 J_0(v)$, where $C_1 > 0$, $1 \leq p \leq 2$, (by Poincaré-Wirtinger inequality).

(b) $||Kw||_2 = C_2||w||_1$, for some $C_2 > 0$, and $||u_0 - Kv||_2 \leq ||K|| \cdot ||v||_2 + ||u_0||_2 \leq ||K|| \cdot C_1 J_0(v) + ||u_0||_2$.

(c) $||u||_{BV} = ||v + w||_1 + J_0(v + w) \leq ||w||_1 + (C_1 + 1)J_0(v)$.

(d) $F(u) = J_0(u) + \frac{\lambda}{2}||u_0 - Ku||_2^2$

$= J_0(v) + \frac{\lambda}{2}||(u_0 - Kv) - Kw||_2^2$

$\geq J_0(v) + \frac{\lambda}{2}\left(||(u_0 - Kv)||_2 - ||Kw||_2\right)^2$

$\geq J_0(v) + \frac{\lambda}{2}||Kw||_2\left(||Kw||_2 - 2||(u_0 - Kv)||_2\right)$

$\geq J_0(v) + \frac{\lambda}{2}C_2||w||_1\left(C_2||w||_1 - 2(||K|| \cdot C_1 J_0(v) + ||u_0||_2)\right).$

(e) If $C_2||w||_1 - 2(||K|| \cdot C_1 J_0(v) + ||u_0||_2) \geq 1$, then $F(u) \geq \frac{\lambda}{2}C_2||w||_1$, so $||u||_{BV} \leq (C_1 + 1)J_0(v) + ||w||_1 \leq (C_1 + 1 + \frac{2}{\lambda C_2})F(u)$.

(f) If $C_2||w||_1 - 2(||K|| \cdot C_1 J_0(v) + ||u_0||_2) < 1$, then $||w||_1 < \frac{1}{C_2}(1 + 2(||K|| \cdot C_1 J_0(v) + ||u_0||_2))$, so $||u||_{BV} - \frac{1+2||u_0||_2}{C_2} \leq \left(\frac{2||K||C_1}{C_2} + C_1 + 1\right)F(u)$.

Thus we can conclude from (e) and (f) that $\lim_{||u||_{BV}\to\infty} F(u) = \infty$.

We can formally solve problem (1.13) via the calculus of variations, in which case we obtain the following Euler-Lagrange equations:

$$\operatorname{div}\left(\frac{Du}{|Du|}\right) + \lambda K^*(u_0 - Ku) = 0 \text{ in } \Omega, \tag{1.14}$$

$$\frac{Du}{|Du|} \cdot \nu = 0 \text{ on } \partial\Omega. \tag{1.15}$$

where K^* is the adjoint of K.

Non-differentiability of the functional J_0 in (1.13) when $Du = 0$ presents a computational disadvantage. Indeed, equation (1.14) makes sense only when $|Du| \neq 0$. In [1], R. Acar and R.C. Vogel consider the following approximation for J_0:

$$J_\beta(u) := \int_\Omega \sqrt{\beta + |Du|^2}\, dx, \tag{1.16}$$

where $\beta \geq 0$. When $\beta = 0$, this reduces to the original functional J_0. The authors have

also established the following results:

R1. $J_0(u) \leq J_\beta(u) \leq J_0(u) + \sqrt{\beta}|\Omega|$, for any $u \in L^1(\Omega)$. Hence, when $u \in BV(\Omega)$,

$$\lim_{\beta \to 0} J_\beta(u) = J_0(u).$$

R2. J_β is convex and weakly lower semicontinuous with respect to the L^p topology for $1 \leq p < \infty$, for any $\beta \geq 0$.

R3. Define $F_\beta(u) := J_\beta(u) + \frac{\lambda}{2}||u_0 - Ku||_2^2,\ \beta \geq 0$. Existence and uniqueness of minimizer of F_β can be shown in the same manner as for problem (1.13).

Remark: Similar to F in problem (1.13), F_β is strictly convex, BV-coercive, and lower semicontinuous with respect to L^2 topology. Notice also that (R1) implies $0 \leq F_\beta(u) - F(u) = J_\beta(u) - J_0(u) \leq \beta|\Omega| \longrightarrow 0$ as $\beta \to 0$.

R4. (Stability of Minimizers) Assume that $1 \leq p < 2$ and that the functionals $F_\infty,\ F_k$ are BV-coercive, lower semicontinuous with respect to L^p-topology and have unique minimizers $\bar{u},\ \bar{u}_k$, respectively. Assume in addition

(i) (Uniform BV-coercivity) For any $v_j \in BV(\Omega)$,

$$\lim_{||v_j||_{BV} \to \infty} F_k(v_j) = +\infty, \quad \text{for all } k.$$

(ii) (Consistency) $F_k \to F_\infty$ uniformly on BV-bounded sets, i.e. given $B > 0$ and $\varepsilon > 0$, there exists N such that

$$|F_k(u) - F_\infty(u)| < \varepsilon \quad \text{whenever} \quad k > N,\ ||u||_{BV} \leq B.$$

Then,

$$||\bar{u}_k - \bar{u}||_p \longrightarrow 0.$$

For $p = 2$, we replace lower semicontinuity assumption by weak lower semicontinuity assumption, then the convergence is weak:

$$\bar{u}_k - \bar{u} \rightharpoonup 0 \quad \text{in } L^2(\Omega).$$

Example: We will now illustrate an application of the results above.

Assume $\lambda_k, \lambda_0 > 0$, with $\lambda_k \to \lambda_0$, $(k = 1, ...)$. Fix $\beta \geq 0$. Let

$$F_k(u) = J_\beta(u) + \frac{\lambda_k}{2}||u_0 - Ku||_2^2, \quad \text{for } k = 0, 1, ...$$

Let $\bar{u}_k$ denote the unique minimizers of F_k. Clearly the hypotheses in (R4) are satisfied. So we have $\bar{u}_k \to \bar{u}$ strongly in $L^p(\Omega)$ for all $1 \leq p < 2$, and $\bar{u}_k \rightharpoonup \bar{u}$ weakly in $L^2(\Omega)$.

Thus far we have seen that given an observed image u_0, a blurring operator K, and a regularization parameter λ, problem (1.13) possesses a unique solution. Yves Meyer [46] have shown that minimizers of (1.13) are characterized by certain properties. We now study characterizations of minimizers in a similar manner as Theorem 4 in [46].

Denote by $\langle , \rangle$ the scalar product in $L^2(\Omega)$. For $f \in L^2(\Omega)$ of zero mean, let

$$\|f\|_* \stackrel{\text{def}}{=} \sup_{h \in BV(\Omega),\ |h|_{BV} \neq 0} \frac{\langle f, h \rangle}{|h|_{BV}} \leq \infty. \tag{1.17}$$

Remark: If $f \in L^2(\Omega)$ is such that $\|f\|_* < \infty$, then f has zero mean. Indeed, let $h \in BV(\Omega)$ be such that $|h|_{BV} \neq 0$, or $h \neq$ constant functions. Let c be an arbitrary constant. Then $|h + c|_{BV} = |h|_{BV}$ and $\frac{\langle f, h+c \rangle}{|h+c|_{BV}} = \frac{\langle f, h \rangle + \langle f, c\chi_\Omega \rangle}{|h|_{BV}} = \frac{\langle f, h \rangle}{|h|_{BV}} + c\frac{\langle f, \chi_\Omega \rangle}{|h|_{BV}}$. This has

to be bounded from above for any constant c having the same sign with $\langle f, \chi_\Omega \rangle$. This can hold only if $\langle f, \chi_\Omega \rangle = 0$, or if $\int_\Omega f(x)dx = 0$. We also deduce that if $\|f\|_* \leq C < \infty$ according to the definition (1.17), then $\langle f, h \rangle \leq C|h|_{BV}$ for any $h \in BV(\Omega)$.

Theorem 1.2.1 *$u = 0$ is a minimizer of (1.13) if and only if $\|K^* f\|_* \leq \frac{1}{\lambda}$.*

Proof: Assume that $u = 0$ is a minimizer of (1.13). Then for any $h \in BV(\Omega)$, we have

$$\frac{\lambda}{2}\|f\|_2^2 \leq \frac{\lambda}{2}\|Kh - f\|_2^2 + |h|_{BV},$$

or

$$\lambda\langle f, Kh \rangle \leq \frac{\lambda}{2}\|Kh\|_2^2 + |h|_{BV}.$$

Changing h into ϵh, with $\epsilon \rangle 0$, we obtain

$$\lambda\epsilon\langle f, Kh \rangle \leq \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 + \epsilon|h|_{BV},$$

then first dividing by ϵ, and then taking the limit as $\epsilon \to 0$, gives

$$\lambda\langle f, Kh \rangle \leq |h|_{BV}, \text{ or } \lambda\langle K^* f, h \rangle \leq |h|_{BV},$$

for any $h \in BV(\Omega)$. Therefore $\|K^* f\|_* \leq \frac{1}{\lambda}$.

Conversely, assume that $\|K^* f\|_* \leq \frac{1}{\lambda}$, i.e. $\sup_{h \in BV(\Omega),\ |h|_{BV} \neq 0} \frac{\langle K^* f, h \rangle}{|h|_{BV}} \leq \frac{1}{\lambda}$. Therefore, $\lambda\langle K^* f, h \rangle \leq |h|_{BV}$, for any $h \in BV(\Omega)$ with $|h|_{BV} \neq 0$. Let $u = 0$. Then, for any $h \in BV(\Omega)$, with $|h|_{BV} \neq 0$, we have:

$$
\begin{aligned}
|h|_{BV} + \frac{\lambda}{2}\|f - Kh\|_2^2 &= |h|_{BV} + \frac{\lambda}{2}\|f\|_2^2 + \frac{\lambda}{2}\|Kh\|_2^2 - \lambda\langle f, Kh\rangle \\
&= |h|_{BV} + \frac{\lambda}{2}\|f\|_2^2 + \frac{\lambda}{2}\|Kh\|_2^2 - \lambda\langle K^* f, h\rangle \\
&\geq |h|_{BV} + \frac{\lambda}{2}\|f\|_2^2 + \frac{\lambda}{2}\|Kh\|_2^2 - |h|_{BV} \\
&= \frac{\lambda}{2}\|f\|_2^2 + \frac{\lambda}{2}\|Kh\|_2^2 \\
&\geq \frac{\lambda}{2}\|f\|_2^2 = \frac{\lambda}{2}\|f - Ku\|_2^2 + |u|_{BV}.
\end{aligned}
$$

Due to the assumption that $\frac{\int_\Omega f(x)dx}{|\Omega|} = 0$, it is easy to show that if a minimizer is a constant c, then c must be zero. Therefore $u = 0$ is a minimizer of (1.13). □

Theorem 1.2.2 *If $\|K^* f\|_* > \frac{1}{\lambda}$, then u is a minimizer of (1.13) if and only if*

$$
\|K^*(f - Ku)\|_* = \frac{1}{\lambda} \ \textit{and} \ \ < u, K^*(f - Ku) >= \frac{1}{\lambda}|u|_{BV}. \tag{1.18}
$$

Proof: Assume u is a minimizer. Then, for any $h \in BV(\Omega)$, we have

$$
\frac{\lambda}{2}\|f - K(u + \epsilon h)\|_2^2 + |u + \epsilon h|_{BV} \geq \frac{\lambda}{2}\|f - Ku\|_2^2 + |u|_{BV}, \text{ or}
$$

$$
\frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 - \lambda\epsilon\langle Kh, f - Ku\rangle + |u + \epsilon h|_{BV} \geq |u|_{BV}. \tag{1.19}
$$

By triangle inequality,

$$
\frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 - 2\lambda\epsilon\langle Kh, f - Ku\rangle + |u|_{BV} + |\epsilon||h|_{BV} \geq |u|_{BV},
$$

or, for $\epsilon > 0$,

$$
|u|_{BV} + \frac{\lambda}{2}\epsilon\|Kh\|_2^2 \geq \lambda\langle Kh, f - Ku\rangle.
$$

As ϵ decreases to 0^+, we obtain that for any $h \in BV(\Omega)$,

$$|h|_{BV} \geq \lambda \langle Kh, f - Ku \rangle,$$

and in particular

$$\frac{1}{\lambda} \geq \frac{\langle h, K^*(f - Ku) \rangle}{|h|_{BV}}$$

if $|h|_{BV} \neq 0$. Therefore

$$\|K^*(f - Ku)\|_* \leq \frac{1}{\lambda}. \tag{1.20}$$

Now, letting $h = u$ in (1.19), with $-1 < \epsilon < 0$, we have

$$\frac{\lambda}{2}\epsilon^2 \|Ku\|_2^2 + (1 + \epsilon)|u|_{BV} \geq |u|_{BV} + \lambda\epsilon \langle u, K^*(f - Ku) \rangle, \text{ or}$$

$$\frac{\lambda}{2}\epsilon^2 \|Ku\|_2^2 + \epsilon|u|_{BV} \geq \lambda\epsilon \langle u, K^*(f - Ku) \rangle.$$

For $-1 < \epsilon < 0$, after division by ϵ and letting $\epsilon \to 0^-$, we obtain that

$$\frac{1}{\lambda}|u|_{BV} \leq \langle u, K^*(f - Ku) \rangle.$$

Now, since u is a minimizer, we have that $|u|_{BV} \neq 0$. Indeed, if by contradiction u with $|u|_{BV} = 0$ is minimizer, then u must be a constant c; however, c must be zero because f has zero mean, therefore $u = 0$ would be a minimizer, contradiction with the hypothesis and the previous theorem. Therefore,

$$\frac{1}{\lambda} \leq \frac{\langle u, K^*(f - Ku) \rangle}{|u|_{BV}}. \tag{1.21}$$

Combining (1.20) and (1.21), we obtain (1.18).

Conversely, assume (1.18) for some $u \in BV(\Omega)$. Then $\|K^*(f - Ku)\|_* = \frac{1}{\lambda}$ implies

that $\lambda\langle K^*(f-Ku), u+\epsilon h\rangle \leq |u+\epsilon h|_{BV}$, for any real parameter ϵ and any $h \in BV(\Omega)$. We then have, for any $h \in BV(\Omega)$:

$$
\begin{aligned}
&|u+\epsilon h|_{BV} + \frac{\lambda}{2}\|f-K(u+\epsilon h)\|_2^2 \\
&= |u+\epsilon h|_{BV} + \frac{\lambda}{2}\|f-Ku\|_2^2 + \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 - \lambda\epsilon\langle f-Ku, Kh\rangle \\
&\geq \lambda\langle K^*(f-Ku), u+\epsilon h\rangle + \frac{\lambda}{2}\|f-Ku\|_2^2 + \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 - \lambda\epsilon\langle f-Ku, Kh\rangle \\
&= \lambda\langle K^*(f-Ku), u\rangle + \lambda\epsilon\langle K^*(f-Ku), h\rangle + \frac{\lambda}{2}\|f-Ku\|_2^2 \\
&+ \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 - \lambda\epsilon\langle K^*(f-Ku), h\rangle \\
&= \lambda\langle K^*(f-Ku), u\rangle + \frac{\lambda}{2}\|f-Ku\|_2^2 + \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 \\
&= |u|_{BV} + \frac{\lambda}{2}\|f-Ku\|_2^2 + \frac{\lambda}{2}\epsilon^2\|Kh\|_2^2 \geq |u|_{BV} + \frac{\lambda}{2}\|f-Ku\|_2^2,
\end{aligned}
$$

therefore u satisfying (1.18) is a minimizer. □

To motivate the problems we are studying in this dissertation, we give here some examples in which exact solutions to the Rudin-Osher-Fatemi model (1.13) are computed.

Example 1.2.1 *We first examine a simple example considered in [59]. Assume Ω is a compact domain in $\mathbb{R}^2$. Let $\Omega_1 \subset\subset \Omega$ with smooth boundary ($\partial\Omega_1 \subset \Omega$), and define the function u_0 as*

$$
u_0(x) = \begin{cases} 1 & \text{for } x \in \Omega_1 \\ 0 & \text{for } x \in \Omega_2 := \Omega\backslash\Omega_1. \end{cases}
$$

Let u_λ denote the solution of the ROF model (1.13) with initial data u_0 and a fixed $\lambda > 0$. Since u_0 is piecewise constant and noise-free, we assume that

$$
u(x) = \begin{cases} 1+\delta_1 & \text{for } x \in \Omega_1 \\ \delta_2 & \text{for } x \in \Omega_2, \end{cases}
$$

for some δ_1, δ_2 to be determined. Applying Example 1.1.2 to u, we obtain the total variation of u in Ω to be

$$J_0(u) = |1 + \delta_1 - \delta_2||\partial\Omega_1|.$$

Since $|1 + \delta_1 - \delta_2|$ is the height of the jumps of the function u at points on $\partial\Omega_1$, we may assume that $1 + \delta_1 - \delta_2 \geq 0$. The fitting error between u and u_0 is

$$\int_\Omega (u_0 - u)^2 dx = \delta_1^2|\Omega_1| + \delta_2^2|\Omega_2|.$$

To find δ_1, δ_2, we solve the following minimization problem

$$\min_{\delta_1,\delta_2} \left\{ (1 + \delta_1 - \delta_2)|\partial\Omega_1| + \frac{\lambda}{2}(\delta_1^2|\Omega_1| + \delta_2^2|\Omega_2|) \right\}.$$

Differentiating with respect to each δ_i yields

$$\begin{aligned} \lambda\delta_1|\Omega_1| + |\partial\Omega_1| = 0 \implies \delta_1 = -\tfrac{|\partial\Omega_1|}{\lambda|\Omega_1|} \\ \lambda\delta_2|\Omega_2| - |\partial\Omega_1| = 0 \implies \delta_2 = \tfrac{|\partial\Omega_1|}{\lambda|\Omega_2|} \end{aligned}$$

This Example 1.2.1 shows that the ROF model decreases the height of the jumps discontinuities. Then, the residual $v = u_0 - u$ will contain this lost height. Consequently, edges will appear in v.

Next we study an example considered in [40] which shows that not all oscillatory functions have small L^2 norm. Hence, the space L^2 is not a well-suited space for modeling oscillatory patterns.

Example 1.2.2 [40, 46] *Fix $a > 0,\ n > 1$, and let φ be a smooth function defined on $\mathbb{R}$ such that*

$$\varphi(x) = \begin{cases} 1 & \textit{for } |x| < n, \\ 0 & \textit{for } |x| > n + 1, \end{cases}$$

and φ is smoothly increasing on $[-n-1,-n]$ and smoothly decreasing on $[n,n+1]$. Let $m>1$ and define

$$f(x)=\frac{1}{m}\varphi'(x)sin(mx)+\varphi(x)cos(mx).$$

Then, $||f||_2^2 \geq 2a^2\int_0^n |cos(mx)|^2dx = a^2(n+\frac{1}{2m}sin(2mn)) \geq a^2(n-\frac{1}{2})>0$. Therefore, the L^2 norm of f can be quite large.

We conclude from these examples that the ROF model is not very well suited for decomposition of an image u_0 into a sum of a piecewise-smooth component u and an oscillatory component v : $u_0 = u+v$.

1.3 Yves Meyer's Models of Oscillatory Patterns

We are interested in an energy minimization model of the form (1.1) that is suitable for cartoon + texture decomposition. The fact that BV regularization allows discontinuities makes it an acceptable space for modeling cartoon images. However, as we have seen in the previous section, L^2 space is not appropriate for modeling oscillatory patterns. The idea is to use a norm that has small weights on oscillatory functions, i.e. weaker norm than the L^2 norm. Motivated by these observations, Yves Meyer [46] suggests the use of the dual to the space $BV(\Omega)$, which possesses the inclusions $BV(\Omega) \subset L^2(\Omega) \subset (BV(\Omega))'$ for any bounded domain Ω in $\mathbb{R}^2$, to capture oscillatory functions. However, there is no known integral representation for continuous linear functionals on $BV(\Omega)$. T. De Pauw has presented a description of the dual of $SBV(\Omega)$ in [25], but it leads to a complicated representation. To overcome this, Y. Meyer [46] suggests to approximate $(BV(\Omega))'$ by the following slightly larger space:

Definition 1.3.1 (The space $\mathbf{G}(\mathbb{R}^2)$) [46] *Let G denote the Banach space consisting of all generalized functions $f(x)$ which can be written as*

$$f(x) = \mathrm{div}(\vec{g}(x)) := \partial_1 g_1(x) + \partial_2 g_2(x), \quad \vec{g}(x) \in L^\infty(\mathbb{R}^2)^2.$$

Define the norm $||f||_G$ *by*

$$\| f \|_G \stackrel{\text{def}}{=} \inf \left\{ \| \sqrt{g_1(x)^2 + g_2(x)^2} \|_\infty, \ f = \mathrm{div}(\vec{g}), \ \vec{g} \in L^\infty(\mathbb{R}^2)^2 \right\}. \tag{1.22}$$

Recall that the homogeneous Sobolev space $W_0^{1,1} := W_0^{1,1}(\mathbb{R}^2)$ is defined as the closure in $W^{1,1} := W^{1,1}(\mathbb{R}^2)$ of the *Schwartz* class $\mathcal{S}(\mathbb{R}^2)$, (see def. 2.1.1). This space G is identical to $W^{-1,\infty}(\mathbb{R}^2)$, the dual of $W_0^{1,1}$, according to the following lemma:

Lemma 1.3.1 [46] *Let* BV_0 *be the closure in* BV *(in the norm* $|\cdot|_{BV}$*) of the* Schwartz *class* $\mathcal{S}(\mathbb{R}^2)$. *Then the space* G *is the dual to the space* BV_0.

In addition to the space G, Meyer [46] has also suggested two additional spaces E and F for modeling oscillatory patterns. For completeness, we shall give the definition of these spaces here:

Definition 1.3.2 (The space $BMO(\mathbb{R}^2)$**)** [40] *Let* $f \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$. *We say that* f *belongs to the* John-Nirenberg space of bounded mean oscillation*, denoted* $BMO(\mathbb{R}^2)$, *if*

$$\frac{1}{Q} \int_Q |f - f_Q| \leq A,$$

for all squares Q *(it is sufficient to consider squares with sides parallel to the axes). Here,* $f_Q = \frac{1}{Q} \int_Q f(x)dx$. *The smallest of all upper bounds* A *is the norm of* f *in* $BMO(\mathbb{R}^2)$, *denoted* $||f||_{BMO(\mathbb{R}^2)}$, *i.e.*

$$||f||_{BMO(\mathbb{R}^2)} \stackrel{\text{def}}{=} \sup_{Q=square} \frac{1}{Q} \int_Q |f - f_Q|.$$

The spaces $F,\ E$ are defined as follows [40]:

Definition 1.3.3 *Let F be the space consisting of all generalized functions f which can be written as*

$$f(x) = \mathrm{div}(\vec{g}(x)), \quad \vec{g}(x) \in BMO(\mathbb{R}^2)^2.$$

Define the norm $||f||_F$ by

$$\| f \|_F \stackrel{\text{def}}{=} \inf \left\{ (||g_1||_{BMO(\mathbb{R}^2)} + ||g_2||_{BMO(\mathbb{R}^2)}),\ f = \mathrm{div}(\vec{g}),\ \vec{g} \in BMO(\mathbb{R}^2)^2 \right\}. \quad (1.23)$$

Definition 1.3.4 [32] *A generalized function f belongs to the space $E(\mathbb{R}^2)$ if it can be written as $f = \Delta g$, where g satisfies*

$$\sup_{|y|>0} \frac{||g(.+y) - 2g(.) + g(.-y)||_\infty}{|y|} < +\infty.$$

Equivalently, E is the Besov space $\dot{B}^{-1}_{\infty,\infty}$ which is dual to the homogeneous Besov space $\dot{B}^1_{1,1}$.

Remark: The following embeddings hold in $\mathbb{R}^2$,

$$W^{1,1} \subset BV \subset L^2 \subset G \subset F \subset E.$$

The space G possesses the following properties:

Lemma 1.3.2 *If $g \in L^2(\mathbb{R}^2)$, then*

$$\left| \int g(x) f(x) dx \right| \leq |f|_{BV} ||g||_G. \quad (1.24)$$

Lemma 1.3.3 *Let $f_k,\ k \geq 1$ be a sequence of functions in $L^2(\Omega)$ satisfying the following three conditions:*

(a) there exists a compact set K such that the supports of f_k are contained in K, for all $k \geq 1$,

(b) there exists an exponent $q > 2$ and a constant C such that $||f_k||_q \leq C$, and

(b) the sequence $f_k \longrightarrow 0$ in the distributional sense.

Then $||f_k||_G \longrightarrow 0$ as $n \to \infty$.

Example 1.3.1 *We now repeat an example considered in [40] to illustrate Lemma 1.3.3. Let f be as in Example 1.2.2 and let $g(x) = \frac{\varphi(x)}{m} sin(mx) + c$, for any $c = constant$. Then $||g(x)||_\infty = \frac{a}{m}$ and $f = g'$. Thus*

$$||f||_G := \inf \{||g||_\infty, f = g', g \in L^\infty\} \leq \frac{a}{m} \longrightarrow 0 \text{ as } m \to \infty,$$

i.e. the G norm of f tends to 0 as the oscillation increases.

Lemma 1.3.3 and Example 1.3.1 show that a function in the space G may have large oscillations but nevertheless small norm. Therefore the G norm is well suited for capturing the oscillations of a function in an energy minimization model. However, in order to apply the $||\cdot||_G$ norm to our energy minimization problem, we need a proper definition of $G(\Omega)$ when Ω is a compact domain in $\mathbb{R}^2$ with Lipschitz boundary. A natural definition is considered in [40] by restricting to Ω.

Definition 1.3.5 (The space $\mathbf{G^V(\Omega)}$) *Let $G^V(\Omega)$ be the space of all distributions $T \in D'(\Omega)$ which can be written as*

$$T = \operatorname{div}(\vec{g}), \quad \vec{g} \in L^\infty(\Omega)^2,$$

i.e. $T(\varphi) = -\int_\Omega (g_1 \frac{\partial \varphi}{\partial x_1} + g_2 \frac{\partial \varphi}{\partial x_2}) dx$, *for all* $\varphi \in \mathcal{D}(\Omega)$. *Define the* $||\cdot||_{G^V}$ *norm by*

$$\| T \|_{G^V} \stackrel{\text{def}}{=} \inf \left\{ \| \sqrt{g_1(x)^2 + g_2(x)^2} \|_\infty,\ T = \operatorname{div}(\vec{g}),\ \vec{g} \in L^\infty(\Omega)^2 \right\}. \tag{1.25}$$

We again recall that the homogeneous Sobolev space $W_0^{1,1}(\Omega)$ is the closure of $\mathcal{D}(\Omega) := C_c^\infty(\Omega)$ in $W^{1,1}(\Omega)$. Therefore, functions in $W_0^{1,1}(\Omega)$ have zero trace on $\partial\Omega$, and is identical to the quotient space $W^{1,1}(\Omega)/P_0(\Omega)$, where $P_0(\Omega)$ is the set of constant functions on Ω. (If $u \in W_0^{1,1}(\Omega)$ and $u = constant$ then $u = 0$ a.e.) Thus, $||u||_{W_0^{1,1}} := \int_\Omega |\nabla u| dx$ is a norm on $W_0^{1,1}(\Omega)$, and it is equivalent to the usual $W^{1,1}$ norm (by Sobolev inequality ([67], pp. 56)). Hence, $W_0^{1,1}(\Omega)$ is a Banach space under this norm $||u||_{W_0^{1,1}}$. Furthermore, the dual to $W_0^{1,1}(\Omega)$ with respect to this norm and that with respect to the usual $W^{1,1}$ norm coincide as spaces of distributions, and the dual norms are equivalent. We shall denote by $W^{-1,\infty}(\Omega)$ the dual to $W_0^{1,1}(\Omega)$.

Theorem 1.3.1 *The space* $G^V(\Omega)$ *under the* $||\cdot||_{G^V}$ *norm is isometrically isomorphic to* $W^{-1,\infty}(\Omega)$ *under the norm* $||\cdot||_{W^{-1,\infty}}$ *dual to the norm* $||\cdot||_{W_0^{1,1}} = \int_\Omega |\nabla u| dx$. A proof is given in the Appendix.

In [10], the authors considered a slightly smaller space for a definition of $G(\Omega)$:

Definition 1.3.6 $G^A(\Omega)$ *is the subspace of* $W^{-1,\infty}(\Omega)$ *given by*

$$G^A(\Omega) = \left\{ v \in L^2(\Omega), v = \operatorname{div}(\vec{g}), \vec{g} \in L^\infty(\Omega)^2, \vec{g} \cdot \nu = 0 \text{ on } \partial\Omega \right\}$$

The norm $||\cdot||_{G^A}$ *on* $G^A(\Omega)$ *is given by*

$$||v||_{G^A} = \inf \left\{ ||\sqrt{g_1(x)^2 + g_2(x)^2}||_\infty,\ v = \operatorname{div}(\vec{g}),\ \vec{g} \cdot \nu = 0 \text{ on } \partial\Omega \right\}.$$

Under definition 1.3.6, $G^A(\Omega) = \left\{v \in L^2(\Omega), \int_\Omega v = 0\right\}$. (A proof of this identity is given in [10]). With this, the inequality from Lemma 1.3.2 is also valid:

$$\left| \int u(x)v(x)dx \right| \leq |u|_{BV}||v||_{G^A}, \quad \forall v \in G(\Omega).$$

Remark: We would like to emphasize that the difference between the space $G^V(\Omega)$ and $G^A(\Omega)$ is that $G^V(\Omega)$ contains generalized functions (since it is identified with $W^{-1,\infty}(\Omega)$) while $G^A(\Omega)$ is only the intersection of $W^{-1,\infty}(\Omega)$ with the space $\left\{v \in L^2(\Omega), \int_\Omega v = 0\right\}$.

With the definition and properties of the space G given, Meyer proposes for cartoon + texture decomposition the following minimization problem [46]:

$$\inf_{u \in BV(\Omega)} \left\{ |u|_{BV(\Omega)} + \lambda \parallel u_0 - u \parallel_G \right\}. \tag{1.26}$$

We note that upon replacing the $||\cdot||_G$ norm in (1.26) with the norms $||\cdot||_F$ and $||\cdot||_E$, we obtain two additional energy minimization models.

1.4 Vese-Osher and Osher-Solé-Vese Models

Some numerical models [66, 52, 10, 13, 14] have been proposed in literature to solve Meyer's minimization problem (1.26), and the one most relevant to this dissertation is the Osher-Solé-Vese model [52], which in turn is a particular case of an earlier work by Vese-Osher [66].

The Vese-Osher [66] model is inspired by Meyer's problem (1.26) and is based on the following observations:

- $G(\Omega) = W^{-1,\infty}(\Omega)$.

- The spaces $W^{-1,p}(\Omega)$ are dual to $W_0^{1,p'}(\Omega)$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

- On a bounded set, $||\sqrt{g_1^2 + g_2^2}||_\infty = \lim_{p\to\infty} ||\sqrt{g_1^2 + g_2^2}||_p$, therefore, the spaces $W^{-1,p}(\Omega)$ approximates $W^{-1,\infty}(\Omega)$ as $p \to \infty$.

- Finally, for any $1 \leq p < \infty$, $G(\Omega) \subset W^{-1,p}(\Omega)$, hence they allow for different choices of weaker norms for the oscillatory component v.

Hence, Vese and Osher propose in [66] the following energy minimization problem:

$$\inf_{u\in BV(\Omega),\vec{g}=(g_1,g_2)\in L^p(\Omega)^2} \left\{ G_p(u,\vec{g}) = \int_\Omega |Du| + \lambda \int_\Omega |u_0 - (u + \mathrm{div}\vec{g})|^2 dx \right. \\ \left. + \mu\Big[\int_\Omega \Big(\sqrt{g_1^2 + g_2^2}\Big)^p dxdy \Big]^{\frac{1}{p}} \right\}, \tag{1.27}$$

where $\lambda, \mu > 0$ are tuning parameters, and $p \geq 1$.

In (1.27), the unknowns are u, g_1, g_2. The first term ensures that $u \in BV(\Omega)$, the second that $v = u_0 - u \approx \mathrm{div}(\vec{g})$, and the last that v is approximately in the space $W^{-1,p}(\Omega)$.

To see that (1.27) is an approximation to (1.26), take $\lambda \to \infty$ and $p \to \infty$, then in the minimization, $u_0 - u = \mathrm{div}(\vec{g})$ a.e. x, for all those $\vec{g}$ with smallest $L^\infty(\Omega)$ norm; the middle term disappears and the last term becomes $\| u_0 - u \|_G$.

In the particular case when $p = 2$, the residual $v = \mathrm{div}(\vec{g})$ in (1.27) belongs to $H^{-1}(\Omega)$, the dual to the space $H_0^1(\Omega)$, with the following dual norm:

$$||v||^2_{H^{-1}(\Omega)} := \inf_{\substack{\vec{g}\in L^2(\Omega)^2 \\ v=\mathrm{div}\vec{g}}} \left\{ \int_\Omega (g_1^2 + g_2^2) dx \right\}.$$

Assume $\vec{g}$ possesses a unique Hodge decomposition $\vec{g} = \nabla P + \vec{Q}$, where $P \in H^1(\Omega)$ and $\mathrm{div}(\vec{Q}) = 0$. Then, $v = \mathrm{div}(\nabla P) = \triangle P$ (or equivalently $P = \triangle^{-1} v$).

In fact, if $v \in L^2(\Omega)$ and $\int_\Omega v dx = 0$, then there is a unique $P \in H^1(\Omega)$, $\int_\Omega P dx = 0$,

such that $-\triangle P = v$, $(\partial P)/(\partial \nu)|_{\partial\Omega} = 0$, ([22], pp. 348-380). The uniqueness of P gives sense to the expression $P = \triangle^{-1}v$. Then, $\vec{g} = \nabla(\triangle^{-1}v) + \vec{Q}$. Thus, on the subspace $\{v \in L^2(\Omega), \int_\Omega v dx = 0\}$ of $H^{-1}(\Omega)$, we can write the norm $||\cdot||_{H^{-1}}$ in the form

$$||v||^2_{H^{-1}(\Omega)} := \int_\Omega |\nabla P|^2 = \int_\Omega |\nabla(\triangle^{-1}v)|^2. \tag{1.28}$$

With these, Osher, Solé, Vese propose in [52] the following energy minimization problem:

$$\inf_{u\in BV(\Omega)} \Big\{ \int_\Omega |\nabla u| + \lambda \int_\Omega |\nabla(\triangle^{-1}(u_0 - u))|^2 \Big\}. \tag{1.29}$$

The solution to problem (1.29) is obtained by solving for steady state solution of the following fourth-order nonlinear PDE:

$$\begin{aligned} &u_t = -\tfrac{1}{2\lambda}\triangle\left[\operatorname{div}\left(\tfrac{\nabla u}{|\nabla u|}\right)\right] - (u - f), \\ &u(0,x,y) = f(x,y), \quad \tfrac{\partial u}{\partial \nu}|_{\partial\Omega} = 0, \quad \tfrac{\partial \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right)}{\partial \nu}|_{\partial\Omega} = 0. \end{aligned} \tag{1.30}$$

1.5 Mumford-Gidas and Other Related Works

In [49] D. Mumford and B. Gidas apply techniques in statistical inference to analyze images in various categories, including vegetation, manmade, road surfaces, sky (with clouds), to name a few. In this work, the authors observe that the statistics of natural images are scale-invariant. Since there are no scale invariance probability measures supported on a space of functions, the authors inferred that natural images should be modeled by generalize functions (distributions in the sense of Schwartz). In particular, the authors mention that Gaussian white noise, measured by Gaussian probability measure, is supported in

$$\cap_{\epsilon>0} H^{-n/2-\epsilon}_{\text{loc}},$$

where H^{-s} is the Hilbert-Sobolev spaces of negative exponent (see definition 2.1.3). In short, the Mumford-Gidas advocate, along the lines of Y. Meyer's, that we search in the space of distributions for models appropriate for oscillatory functions.

J-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle is a group of authors who, inspired by Y. Meyer's [46], have work extensively in image decomposition models using total variation and dual norms, capturing texture with norms on distributional spaces, and numerical algorithms for computing these norms. We refer the readers to [10, 13, 14, 15] for all details on their work.

T. Le and L. Vese [40], and also with J. Garnett [32], inspired by Y. Meyer's, have proposed various numerical models for image decomposition and modeling texture with Meyer's G-norm and F-norm, and also with Besov norms.

For some works in image decomposition via wavelets approach, we refer the readers to the works by I. Daubechies and G. Teske [20, 21], J.-L. Starck, M. Elad, and D.L. Donoho [63].

Related to Chapter 2 of this dissertation is a work inspired by the Osher-Soé-Vese and using the Fourier Transform for computation done by S. Roudenko [55].

For additional recent works in image reconstruction and decomposition using variational methods and partial differential equations, we refer the readers to the works by E. Tadmor, S. Nezzar, and L. Vese [62], S. Esedoglu and S.J. Osher [28], S. Kindermann, S.J Osher, and J. Xu [38], S.E. Levine [41], A. Marquina and S. Osher [44, 45].

Chapter 2

(BV, H^{-s}) Model

In this chapter we present a new model for image restoration and decomposition. We follow Yves Meyer's suggestion of modeling oscillatory patterns with generalized functions and Rudin-Osher-Fatemi's idea of regularization by total variation minimization.

The proposed model decomposes a given (degraded or textured) image u_0 into a sum $u+v$, where $u \in BV$ is a function of bounded variation (the cartoon component of u_0), and the noisy (or textured) component v is modeled by tempered distributions belonging to the negative Hilbert-Sobolev space H^{-s}. It can be seen as generalization of Osher-Solé-Vese's model and have been motivated also by Mumford-Gidas [49].

We present proofs for existence and uniqueness of solution to the proposed model, as well as two characterizations of the solution. We also give a numerical algorithm for solving the minimization problem. And finally, we present numerical results on denoising, deblurring, and decompositions of both real and synthetic images.

One related work has been done by Daubechies and Teschke [20, 21] via wavelets approach. In [20, 21], the authors modified the Vese-Osher and Osher-Solé-Vese energies in (1.27) and (1.29) by replacing $BV(\Omega)$ in the regularizing term with the space $B_1^1(L^1(\Omega))$, and limiting themselves to the case $p = 2$, whereby they arrived to a new minimization

problem which gives a decomposition $u_0 \approx u + v$, with $u \in B_1^1(L^1(\Omega))$ and $v \in H^{-1}(\Omega)$.

Other interesting related work for image decomposition and cartoon and texture separation using generalized functions and dual spaces are by Aujol and collaborators [13, 14, 10, 15], Starck, Elad and Donoho [63], Tadmor, Nezzar and Vese [62], Esedoglu and Osher [28], among others. A related preliminary work inspired from the OSV model[52] and also using the Fourier Transform for computations and variants is by Roudenko [55].

2.1 Definitions and Assumptions

Definition 2.1.1 (Schwartz Space and Space of Tempered Distributions) [31, 22]

- *The Schwartz Space is defined as:*

$$\mathcal{S}(\mathbb{R}^n) \stackrel{\text{def}}{=} \left\{ u \in C^\infty(\mathbb{R}^n) \text{ s.t. } \forall \alpha \in \mathbb{N}, \beta \in \mathbb{N}^n, |x|^\alpha |D^\beta u| \to 0 \text{ as } |x| \to +\infty \right\}.$$

- *The dual to $\mathcal{S}(\mathbb{R}^n)$, denoted $\mathcal{S}'(\mathbb{R}^n)$, is called the space of tempered distributions.*

Definition 2.1.2 (Fourier Transforms) [22]

- *The Fourier Transform of a function $f \in \mathcal{S}(\mathbb{R}^n)$, denoted $\hat{f}$, is defined by*

$$\hat{f}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f(x) e^{-2\pi i\, x\cdot\xi} dx,\ \xi \in \mathbb{R}^n, \quad (i = \sqrt{-1}).$$

  The Fourier transform $\hat{f}$ is also well defined for integrable functions f on $\mathbb{R}^n$.

- *Let the brackets $\langle , \rangle$ denote the duality pairing between $\mathcal{S}'(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$. The Fourier Transform of a tempered distribution g is the tempered distribution $\hat{g}$ defined by:*

$$\langle \hat{g}, \varphi \rangle = \langle g, \hat{\varphi} \rangle \text{ for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Definition 2.1.3 (The space $H^s(\mathbb{R}^n)$) *For any $s \in \mathbb{R}$,*

$$H^s(\mathbb{R}^n) \stackrel{\text{def}}{=} \left\{ g \in \mathcal{S}'(\mathbb{R}^n) \text{ s.t. } (1+|\xi|^2)^{s/2} \cdot \hat{g} \in L^2(\mathbb{R}^n) \right\},$$

where $\hat{g}$ is the Fourier transform of g.

$H^s(\mathbb{R}^n)$ is a Hilbert space equipped with the inner product

$$\langle f, g \rangle_s = \int_{\mathbb{R}^n} (1+|\xi|^2)^s \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi,$$

and the associated norm $\| f \|_s = \sqrt{\langle f, f \rangle_s} = \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{f}(\xi)|^2 d\xi \right)^{1/2}$.

Remarks 2.1.1

- *When $s = m$ is an integer, then $H^s(\mathbb{R}^n)$ is the same as the Sobolev Space $H^m(\mathbb{R}^n)$ with equivalent norms. The dual to $H^{-s}(\mathbb{R}^n)$ is the space $H^s(\mathbb{R}^n)$.*
- *If $s_1 > s_2 \geq 0$, then $\| f \|_{s_1} > \| f \|_{s_2}$. Thus we have the following continuous embeddings (injections) of spaces*

$$\mathcal{S}(\mathbb{R}^n) \subset H^{s_1}(\mathbb{R}^n) \subset H^{s_2}(\mathbb{R}^n) \subset \cdots \subset H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n) \subset \cdots$$
$$\cdots \subset H^{-s_2}(\mathbb{R}^n) \subset H^{-s_1}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n).$$

2.1.1 Extension from Ω to $\mathbb{R}^2$

Often when working with images, we work with functions defined on a bounded domain $\Omega \subset \mathbb{R}^2$. On the other hand, the H^s norm is given via the Fourier Transform which is defined only for functions with values given on $\mathbb{R}^2$. To resolve this, whenever necessary, we shall extend a function to $\mathbb{R}^2$ by setting it to zero at points outside of Ω. It is clear that extension by zeros is a continuous embedding of $L^2(\Omega)$ into $L^2(\mathbb{R}^2)$. Moreover, it is

an embedding of $BV(\Omega)$ into $BV(\mathbb{R}^2)$ [30]. In addition, since $\partial\Omega$ is Lipschitz, Poincaré-Wirtinger Inequality [30] implies that $BV(\Omega)$ is continuously embedded into $L^2(\Omega)$. We thus have the following continuous embeddings: $BV(\Omega) \subset L^2(\Omega) \subset L^2(\mathbb{R}^2) \subset H^{-s}(\mathbb{R}^2)$, $s \geq 0$.

Henceforth, for analysis in the continuous setting, we will apply extension by zeros when necessary. We shall assume that the initial data image u_0 belongs to $H^{-s}(\mathbb{R}^2)$. For functions $u \in BV(\Omega)$, we will first apply the Trace Theorem (1.1.5) to extend u to $\bar{\Omega}$. Then we extend $u = 0$ to $\mathbb{R}^2 \backslash \overline{\Omega}$

In practice, we actually work in the discrete setting (for example see Section 2.6). In this case, we apply the Discrete Fourier Transform (DFT) and the Inverse Discrete Fourier Transform (IDFT). Then, the explicit extension by zeros of discrete functions outside $\overline{\Omega}$ is no longer needed. Indeed, observe that when u is defined only on $\bar{\Omega}$, and we apply extension by zeros to $\mathbb{R}^2$, then the Fourier Transform

$$\hat{u}(\xi) := \int_{\mathbb{R}^2} u(x) e^{-2\pi i \xi \cdot x} dx = \int_{\Omega} u(x) e^{-2\pi i \xi \cdot x} dx.$$

Therefore, in the case when Ω is a rectangular domain, we obtain an implicit equivalence between the Fourier Transform of the extended function and the Fourier Transform on a torus. Moreover, $\hat{u}(\xi) = 0$ for all non-integer ξ (i.e. $\xi \notin \mathbb{Z}$). That is, the Fourier Transform of the extended function is nonzero only at integer frequencies.

We are now ready to describe a new model, which we propose for image reconstruction and decomposition, using bounded total variation regularization and negative-exponent Hilbert Sobolev fitting.

2.2 Description of the (BV, H^{-s}) Model

From hereforth, $s \geq 0$, $L^2 = L^2(\mathbb{R}^2)$, $H^{-s} = H^{-s}(\mathbb{R}^2)$, Ω is an open, bounded and connected subset of $\mathbb{R}^2$ with Lipschitz boundary $\Gamma = \partial\Omega$, and the blurring operator K is an injective continuous linear operator from L^2 into L^2 such that $K1_\Omega \neq 0$.

We propose the following new variational model for image restoration and decomposition:

$$\inf_{u \in BV(\Omega)} F(u) \ := \ \lambda |u|_{BV(\Omega)} + \ \| \ u_0 - Ku \ \|^2_{-s}, \tag{2.1}$$

where $s \geq 0$ is kept as a parameter for the fidelity norm, $\lambda > 0$ is a regularization parameter, $|u|_{BV(\Omega)} := \int_\Omega |Du|$ is the regularization (total variation) term, $\| \ u_0 - Ku \ \|^2_{-s} = \int_{\mathbb{R}^2} (1 + |\xi|^2)^{-s} |\hat{u}_0 - \widehat{Ku}|^2 \, d\xi$ is the fidelity term.

Observe that the regularizing term is a convex functional on the space $BV(\Omega)$ which consists of functions defined only on Ω. The fitting term is a strictly convex functional defined on the space H^{-s} consisting of generalized functions given on the whole frequency domain $\mathbb{R}^2$. Clearly, the energy $F(u)$ is convex. Secondly, observe that since the H^{-s}-norm is bounded by the L^2-norm, the operator K is also continuous in the H^{-s}-norm on $L^2 \subset H^{-s}$. Hence, by Hahn-Banach Theorem, K can be extended to a continuous linear operator $K : H^{-s} \to H^{-s}$.

Remarks 2.2.1 *(i) Taking K to be the identity operator and $s = 0$ and applying Parseval Identity $\int |g|^2 dx = \int |\hat{g}|^2 d\xi$, (2.1) becomes exactly (1.13), i.e. our proposed model recovers the ROF model.*

(ii) We can obtain an equivalent form of (1.29) from (2.1) by setting $s = 1$ and writing $\| \ u_0 - u \ \|_{H_0^{-s}} := \int |\xi|^{-2s} |\hat{u}_0 - \hat{u}|^2 d\xi$, that is, we change the H^{-s} norm into a semi-norm by removing 1 from the first term in the integrand. Thus, we obtain an equivalent formulation of the OSV model [52].

Example 2.2.1 *In this example, we show that the spaces L^2 and L^1 do not always model oscillatory functions very well, whereas the spaces H^{-s} for any $s > 0$ are better candidates for modeling oscillatory functions. (We paraphrase Example 1.2.2 for the sake of completeness.)*

Let p be a fixed positive integer, and denote by χ_p the characteristic function of the interval $[-p, p]$, that is

$$\chi_p(x) = \begin{cases} 1, & \text{if } |x| < p, \\ 0, & \text{if } |x| \geq p. \end{cases}$$

Let $f : [-p, p] \to \mathbb{R}$ be given by $f(x) = cos(2\pi qx) \cdot \chi_p(x)$, where q is a positive integer. We have

(i) $||f||_{L^2}^2 = 2\int_0^p |cos(2\pi qx)|^2 dx = p > 0.$

(ii) $||f||_{L^1} = 2\int_0^p |cos(2\pi qx)| dx > 2\int_0^p |cos(2\pi qx)|^2 dx = p > 0.$

(iii) Since the Fourier Transform of a product is the convolution of the Fourier Transforms,

$$\begin{aligned} \hat{f}(\xi) &= \mathcal{F}\left(cos(2\pi qx)\right) * \mathcal{F}\left(\chi_p\right) \\ &= \left(\frac{1}{2}\left[\delta(y+q) + \delta(y-q)\right] * \frac{sin(2\pi py)}{\pi y}\right)(\xi) \\ &= \frac{1}{2\pi}\left(\frac{sin(2\pi p(\xi+q))}{\xi+q} + \frac{sin(2\pi p(\xi-q)}{\xi-q}\right) \\ &= \frac{1}{2\pi q}\left(\frac{sin(2\pi p\xi)}{\xi/q+1} + \frac{sin(2\pi p\xi)}{\xi/q-1}\right). \end{aligned}$$

So, for any $s > 0$,

$$
\begin{aligned}
||f||_{-s}^2 &= \frac{1}{(2\pi q)^2}\left[\int_{\mathbb{R}} \frac{|sin(2\pi p\xi)|^2}{(\xi/q+1)^2(1+\xi^2)^s}d\xi\right. \\
&\quad + \int_{\mathbb{R}} \frac{|sin(2\pi p\xi)|^2}{(\xi/q-1)^2(1+\xi^2)^s}d\xi \\
&\quad \left. + \int_{\mathbb{R}} \frac{2|sin(2\pi p\xi)|^2}{(\xi/q+1)(\xi/q-1)(1+\xi^2)^s}d\xi\right] \\
&\leq \frac{C}{(2\pi q)^2},
\end{aligned}
$$

where C is a constant independent of p and q. So, as $q \to +\infty$, $||f||_{-s} \to 0$.

Therefore, as the oscillation increases, the H^{-s} norm goes to zero, whereas the L^2 and L^1 norms stay bounded below by $p > 0$. In other words, in a minimization model, the oscillatory components will be better captured by a H^{-s} norm with $s > 0$, rather than by an L^p norm with $p \geq 1$.

2.3 Existence and Uniqueness of Solutions

We now prove existence and uniqueness of minimizers for the proposed model, adapting the techniques from [64, 17, 1] to the (BV, H^{-s}) case.

Theorem 2.3.1 *Given $\Omega \subset \mathbb{R}^2$, open, bounded and connected, with Lipschitz boundary, $u_0 \in H^{-s}$, $\lambda > 0$, and $K : L^2 \to L^2$ is an injective continuous linear operator such that $K1_\Omega \neq 0$, then the minimization problem*

$$
\inf_{u \in BV(\Omega)} F(u) = \lambda |u|_{BV(\Omega)} + \parallel u_0 - Ku \parallel_{-s}^2, \quad s > 0,
$$

has a unique solution in $BV(\Omega)$.

Proof: Let $u_n \in BV(\Omega)$ be a minimizing sequence. Then there exists a constant $M > 0$ for which $|u_n|_{BV(\Omega)} \leq M$ and $\| u_0 - Ku_n \|_{-s}^2 \leq M$ for all n. By the Poincaré-Wirtinger inequality, we have a constant $C > 0$ which depends only on Ω such that

$$\| u_n - \frac{1}{|\Omega|} \int_\Omega u_n \|_{L^2(\Omega)} \leq C|u_n|_{BV(\Omega)} \leq CM, \text{ for all } n. \tag{2.2}$$

Claim 2.3.1 *There exists a constant $C' > 0$ such that $\frac{1}{|\Omega|} \left| \int_\Omega u_n \right| \leq C', \ \forall n.$*

(We postpone the proof to this claim until later).

By the Claim, $\| u_n \|_{L^2(\Omega)} \leq M', \ \forall n$, implying $\| u_n \|_{L^1(\Omega)} \leq M', \ \forall n$. Hence

$$\| u_n \|_{BV(\Omega)} := \| u_n \|_{L^1(\Omega)} + |u_n|_{BV(\Omega)} \leq M'', \ \forall n.$$

Therefore there exist a subsequence, still denoted u_n, and a $u \in BV(\Omega)$ such that $u_n \to u$ in $L^1(\Omega)$ and

$$|u|_{BV(\Omega)} \leq \liminf_{n \to \infty} |u_n|_{BV(\Omega)}. \tag{2.3}$$

Moreover, by passing to a subsequence if necessary, $u_n \rightharpoonup u$ weakly in $L^2(\Omega)$. After extending u_n, u by zeros to $\mathbb{R}^2$, we still have $u_n \rightharpoonup u$ weakly in L^2. Since K is a continuous linear operator from L^2 to L^2, $\int_{\mathbb{R}^2} Ku_n(x)\varphi(x)dx = \int_{\mathbb{R}^2} u_n(x)K^*\varphi(x)dx, \quad \forall \varphi \in L^2$. Therefore, $Ku_n \rightharpoonup Ku$ weakly in L^2.

To show $Ku_n \rightharpoonup Ku$ weakly in H^{-s}, we recall that for any $\varphi \in H^{-s}$, $\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^{s/2}} \in L^2(\mathbb{R}^2; \mathbb{C})$, therefore $\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^s} \in L^2(\mathbb{R}^2; \mathbb{C})$. So $\mathcal{F}\Big(\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^s}\Big) \in L^2(\mathbb{R}^2; \mathbb{C})$. Hence (by subsequently taking the real part and then the imaginary part and combining them afterwards)

$$\int_{\mathbb{R}^2} \overline{\mathcal{F}\Big(\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^s}\Big)}(x) Ku_n(x)dx \to \int_{\mathbb{R}^2} \overline{\mathcal{F}\Big(\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^s}\Big)}(x) Ku(x)dx \quad \text{as } n \to \infty.$$

Applying $\int v\bar{\hat{w}} = \int \bar{\hat{v}} w$, we thus have for any $\varphi \in H^{-s}$,

$$\int_{\mathbb{R}^2} \frac{\hat{\varphi}\overline{\widehat{Ku_n}}}{(1+|\xi|^2)^s} d\xi \to \int_{\mathbb{R}^2} \frac{\hat{\varphi}\overline{\widehat{Ku}}}{(1+|\xi|^2)^s} d\xi \quad \text{as } n \to \infty.$$

Therefore, $u_0 - Ku_n \rightharpoonup u_0 - Ku$ weakly in H^{-s} and by the lower semi-continuity of the H^{-s}-norm, we get

$$\| u_0 - Ku \|_{-s}^2 \leq \liminf_{n\to\infty} \| u_0 - Ku_n \|_{-s}^2 .$$

Hence,

$$F(u) := \lambda |u|_{BV(\Omega)} + \| u_0 - Ku \|_{-s}^2 \leq \liminf_{n\to\infty} \Big(\lambda |u_n|_{BV(\Omega)} + \| u_0 - Ku_n \|_{-s}^2 \Big).$$

So u is a minimizer.

For uniqueness, suppose $u, v \in BV(\Omega)$ are two minimizers, i.e. $F(u) = F(v) = \inf_{w\in BV(\Omega)} F(w)$, and suppose also $Ku \neq Kv$. Then we would have

$$\begin{aligned} \| u_0 - \frac{1}{2}Ku - \frac{1}{2}Kv \|_{-s}^2 = \frac{1}{4} \| u_0 - Ku \|_{-s}^2 & \\ + \frac{1}{2}\mathbf{Re}\langle u_0 - Ku, u_0 - Kv\rangle_{-s} & \\ + \frac{1}{4} \| u_0 - Kv \|_{-s}^2 & \\ \langle \frac{1}{2} \| u_0 - Ku \|_{-s}^2 + \frac{1}{2} \| u_0 - Kv \|_{-s}^2 . & \end{aligned}$$

Then $\| u_0 - \frac{1}{2}Ku - \frac{1}{2}Kv \|_{-s}^2 + \lambda |\frac{1}{2}(u+v)|_{BV(\Omega)} \leq \| u_0 - \frac{1}{2}Ku - \frac{1}{2}Kv \|_{-s}^2 + \lambda \frac{1}{2}(|u|_{BV(\Omega)} + |v|_{BV(\Omega)}) < \frac{1}{2} \| u_0 - Ku \|_{-s}^2 + \frac{1}{2} \| u_0 - Kv \|_{-s}^2 + \lambda \frac{1}{2}(|u|_{BV(\Omega)} + |v|_{BV(\Omega)}) = inf_{w\in BV(\Omega)} F(w)$. This implies $F(\frac{1}{2}(u+v)) < F(u) = F(v)$, which cannot be so if u and v are minimizers. Thus, $Ku = Kv$. Since K is injective, $u = v$. $\square$

Proof: (of Claim 2.3.1) [1, 64]

Denote by $w_n = \left(\frac{1}{|\Omega|}\left|\int_\Omega u_n\right|\right)\chi_\Omega$ and $v_n = u_n - w_n$, where u_n is a minimizing sequence as above. Then clearly $w_n, v_n \in BV(\Omega)$. Moreover, we have from (2.2)

$$\| v_n \|_{-s} \leq \| v_n \|_{L^2} = \| u_n - \frac{1}{|\Omega|}\int_\Omega u_n \|_{L^2(\Omega)} \leq C,$$

for some constant $C > 0$ independent of n, for all n. Thus,

$$\begin{aligned}
M \geq & \| u_0 - Ku_n \|_{-s}^2 \\
= & \| u_0 - Kv_n - Kw_n \|_{-s}^2 \\
= & \| u_0 - Kv_n \|_{-s}^2 + \| Kw_n \|_{-s}^2 - 2\mathbf{Re}\langle u_0 - Kv_n, Kw_n\rangle_{-s} \\
\geq & \| Kw_n \|_{-s}^2 - 2 \| u_0 - Kv_n \|_{-s} \| Kw_n \|_{-s} \\
= & \| Kw_n \|_{-s} \big(\| Kw_n \|_{-s} - 2 \| u_0 - Kv_n \|_{-s} \big). \\
\geq & \| Kw_n \|_{-s} \Big(\| Kw_n \|_{-s} - 2\big(\| u_0 \|_{-s} + \| K \| \| v_n \|_{-s} \big)\Big).
\end{aligned}$$

Denote by $x_n = \| Kw_n \|_{-s}$ and $a_n = \| u_0 \|_{-s} + \| K \| \| v_n \|_{-s}$. Then

$$x_n(x_n - 2a_n) \leq M, \text{ and } 0 \leq a_n \leq \| u_0 \|_{-s} + \| K \| \cdot C = M', \ \forall n.$$

Hence $0 \leq x_n \leq a_n + \sqrt{a_n^2 + M} \leq M''$. This implies

$$\frac{1}{|\Omega|}\left|\int_\Omega u_n\right| \| K\chi_\Omega \|_{-s} = \| Kw_n \|_{-s} \leq M'', \ \forall n.$$

Therefore, $\frac{1}{|\Omega|}\left|\int_\Omega u_n\right| \leq C', \ \forall n.$

Remark: The above existence and uniqueness result also holds for other regularizing functionals on $BV(\Omega)$ instead of the total variation. For example, $|u|_{BV(\Omega)}$ can be

substituted by $\int_\Omega \phi(Du)$ defined in the sense of convex functions of measures (see [26], [64]), where $\phi : \mathbb{R}^2 \to [0, \infty)$ is continuous, even, convex, and satisfying $\phi(0) = 0$, $a|x| - b \leq \phi(x) \leq a|x| + b$ for some constants $a > 0$ and $b \geq 0$, and any $x \in \mathbb{R}^2$. Examples are $\phi(Du) = |u_{x_1}| + |u_{x_2}|$, $\phi(Du) = \sqrt{\alpha + |Du|^2}$, $\phi(Du) = \log\cosh(\alpha + |Du|^2)$, with $\alpha > 0$.

2.4 Characterization of Minimizers

In this section we present two approaches for characterization of the minimizer of our proposed model. One approach is to characterize the subdifferential of the functional F in (2.1) by studying the duality of convex optimization problems. This is a rigorous way to derive the optimality conditions which validate that the solution to problem (2.1) is indeed the solution to its associated Euler-Lagrange equation to be derived in Section 2.5. The second approach is to utilize the H^{-s} inner product to define a 'texture' norm that is dual to $|\cdot|_{BV}$, as in Meyer's [46]. We will show that this 'texture' norm provides a way for determining how our proposed model discriminate cartoon versus textures.

2.4.1 Characterization of minimizers via duality

We first review the theoretical results in convex analysis which allow us to make an association between a minimization problem $(\mathcal{P})$ and a maximization problem $(\mathcal{P}^*)$, called the dual problem of $(\mathcal{P})$. We shall apply these results to obtain characterizations of the solution to problem (2.1), via the properties of the dual problem. We refer the readers to [27, 64, 26, 61, 54] for all details covered in this section.

Duality in convex optimization:

Assume X is a Banach space, X^* is its dual endowed with the weak-star topology, and

$F : X \to \bar{\mathbb{R}}$ is a functional defined on X.

Definition 2.4.1 (the sub-differential of F) *For $u \in X$, we say that $\phi \in X^*$ is in the* sub-differential ∂F *of F at u* **iff**

$$F(u) \in \mathbb{R} \quad \textit{and} \quad F(u) - \langle \phi, u \rangle_{X \times X^*} \leq F(v) - \langle \phi, v \rangle_{X \times X^*}, \ \forall v \in X.$$

Theorem 1 $F(u) = \inf_{v \in X} F(v)$ *if and only if* $0 \in \partial F(u)$.

Definition 2.4.2 (Legendre transform) *Assume $\Phi : X \to \bar{\mathbb{R}}$ is linear, the* Legendre transform *or* polar function *of Φ, is the function $\Phi^* : X^* \to \bar{\mathbb{R}}$ given by*

$$\Phi^*(v^*) = \sup_{v \in X} \left\{ \langle v, v^* \rangle - \Phi(v) \right\}.$$

Theorem 2.4.1 (Duality Theorem) *Let V, Y be two locally convex topological spaces, and V^*, Y^* their dual. Let $\Lambda : V \to Y$ be linear continuous operator (its adjoint $\Lambda^* : Y^* \to V^*$ is linear and continuous). Assume $\mathcal{F}$ (resp. $\mathcal{G}$) is a convex lower semi-continuous function from V (resp. Y) into $\mathbb{R} \cup \{+\infty\}$, not identically $+\infty$. Denote the primal problem by $(\mathcal{P})$, i.e.*

$$(\mathcal{P}) \qquad \inf_{v \in V} \left\{ \mathcal{F}(v) + \mathcal{G}(\Lambda v) \right\}.$$

Its dual $(\mathcal{P}^*)$ is:

$$(\mathcal{P}^*) \qquad \sup_{p^* \in Y^*} \left\{ -\mathcal{F}^*(\Lambda^* p^*) - \mathcal{G}^*(-p^*) \right\}.$$

Then,

(1) $-\infty \leq$ sup of $\mathcal{P}^* \leq$ inf of $\mathcal{P} \leq +\infty$.

(2) The two conditions (a) inf of $\mathcal{P}$ is in $\mathbb{R}$, (b) $\exists u_0 \in V$ s.t. $\mathcal{F}(u_0) < +\infty$ and $\mathcal{G}$ is

continuous at Λu_0, *imply that*

$$\textit{Inf}\, \mathcal{P} = \textit{Sup}\, \mathcal{P}^*.$$

(3) $\mathcal{P}^*$ *has at least one solution.*

(4) If $\mathcal{P}$ *and* $\mathcal{P}^*$ *have solutions* $\bar{u}, \bar{p}^*$, *and* $\textit{Inf}\, \mathcal{P} = \textit{Sup}\, \mathcal{P}^*$, *then*

$$\mathcal{F}(\bar{u}) + \mathcal{F}^*(\Lambda^* \bar{p}^*) = \langle \Lambda^* \bar{p}^*, \bar{u} \rangle, \tag{2.4}$$

$$\mathcal{G}(\Lambda \bar{u}) + \mathcal{G}^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda \bar{u} \rangle. \tag{2.5}$$

Characterization of minimizers via duality:

We now follow the theoretical framework set forth above to formulate the dual to our minimization problem (2.1). This will allow us to characterize the subdifferential of F from which we obtain characterizations of the minimizer u of (2.1) by recalling Theorem 1.

First, we extend the functional F in (2.1) to the ambient space $L^2(\Omega)$ by setting $F(u) = +\infty$ for $u \in L^2(\Omega) \backslash BV(\Omega)$.

Now, the Fourier Transform can be written as a sum of two linear continuous operators in the following way:

$$\mathcal{F}f = \mathcal{F}_C f + i\mathcal{F}_S f, \quad (i = \sqrt{-1}).$$

When f is a function, these are exactly

$$\mathcal{F}_C f = \int_{\mathbb{R}^2} f(x) cos(2\pi x \cdot \xi) dx, \quad \mathcal{F}_S f = \int_{\mathbb{R}^2} f(x) sin(2\pi x \cdot \xi) dx.$$

The definitions of $\mathcal{F}_C$ and $\mathcal{F}_S$ can be extended in a natural way to tempered distributions

$f \in H^{-s}$ by

$$\langle \mathcal{F}_C f, \varphi \rangle = \langle f, \mathcal{F}_C \varphi \rangle, \text{ for all } \varphi \in \mathcal{S}(\mathbb{R}^2).$$

Let $\mathcal{F}_1, \mathcal{F}_2 : L^2(\Omega) \to L^2(\mathbb{R}^2)$ be linear continuous operators given by

$$(\mathcal{F}_1 v)(y) = (1 + |y|^2)^{-s/2} (\mathcal{F}_C K v_{ext})(y),$$

$$(\mathcal{F}_2 v)(y) = (1 + |y|^2)^{-s/2} (\mathcal{F}_S K v_{ext})(y),$$

where $v \in L^2(\Omega)$, v_{ext} denotes the extension by zeros of v from Ω to $\mathbb{R}^2$, and K is the usual blurring operator.

For the data image $u_0 \in H^{-s}$, we denote by

$\tilde{f}_1(y) \stackrel{\text{def}}{=} (1 + |y|^2)^{-s/2} (\mathcal{F}_C u_0)(y)$ and

$\tilde{f}_2(y) \stackrel{\text{def}}{=} (1 + |y|^2)^{-s/2} (\mathcal{F}_S u_0)(y)$.

With these notations, the functional F can be re-written as

$$F(u) = \int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1 u)^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2 u)^2 dx + \lambda \int_\Omega |Du|, \ \forall u \in L^2(\Omega).$$

Now, fix $u \in BV(\Omega)$, and assume that $\phi \in L^2(\Omega)$ belongs to $\partial F(u)$. Therefore, u is the minimum on $BV(\Omega)$ of the following problem:

$$\inf_{v \in BV(\Omega)} F(v) - \int_\Omega \phi \cdot v dx. \tag{2.6}$$

The infimum on $BV(\Omega)$ is the same as the infimum on $W^{1,1}(\Omega)$ (see [64]), so we replace $BV(\Omega)$ with $W^{1,1}(\Omega)$, and (2.6) becomes

$$(\mathcal{P}) \qquad \inf_{v \in W^{1,1}(\Omega)} F(v) - \int_\Omega \phi \cdot v dx,$$

where

$$F(v) = \int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1 v)^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2 v)^2 dx + \lambda \int_\Omega |\nabla v| dx, \text{ for } v \in W^{1,1}(\Omega).$$

We now formulate the problem $(\mathcal{P}^*)$ dual to $(\mathcal{P})$. First, recall that the Legendre transform of a linear functional $\Phi : X \to \bar{\mathbb{R}}$ is $\Phi^* : X^* \to \bar{\mathbb{R}}$ given by

$$\Phi^*(v^*) = \sup_{v \in X} \{\langle v, v^* \rangle - \Phi(v)\}.$$

Let $\Lambda : W^{1,1}(\Omega) \to L^2(\mathbb{R}^2)^2 \times L^1(\Omega)^2$ be the linear continuous operator given by

$$\Lambda v = (\mathcal{F}_1 v, \mathcal{F}_2 v, D_1 v, D_2 v).$$

Let $\mathcal{F} : W^{1,1}(\Omega) \to \mathbb{R}$, $\mathcal{G}_1 : L^2(\mathbb{R}^2) \to \mathbb{R}$, $\mathcal{G}_2 : L^2(\mathbb{R}^2) \to \mathbb{R}$, $\mathcal{G}_3 : L^1(\Omega)^2 \to \mathbb{R}$, and $\mathcal{G} : L^2(\mathbb{R}^2)^2 \times L^1(\Omega)^2 \to \mathbb{R}$, such that

- $\mathcal{F}(v) = -\int \phi \cdot v dx$,
- $\mathcal{G}_1(w_1) = \int_{\mathbb{R}^2} (\tilde{f}_1 - w_1)^2 dx$, $\mathcal{G}_2(w_2) = \int_{\mathbb{R}^2} (\tilde{f}_2 - w_2)^2 dx$, $\mathcal{G}_3(\bar{w}) = \lambda \int_\Omega |\bar{w}| dx$,
- $\mathcal{G}(w) = \mathcal{G}_1(w_1) + \mathcal{G}_2(w_2) + \mathcal{G}_3(\bar{w})$, where $w = (w_1, w_2, \bar{w}) \in L^2(\mathbb{R}^2)^2 \times L^1(\Omega)^2$.

Then the problem $(\mathcal{P})$ is

$$(\mathcal{P}) \qquad \inf_{v \in W^{1,1}(\Omega)} \{\mathcal{F}(v) + \mathcal{G}(\Lambda v)\}.$$

Since $\mathcal{F}$ and $\mathcal{G}$ are proper (i.e. not identically $+\infty$) convex lower semi-continuous functions, the dual problem $(\mathcal{P}^*)$ is

$$(\mathcal{P}^*) \qquad \sup_{p^* \in L^2(\mathbb{R}^2)^2 \times L^\infty(\Omega)^2} \{-\mathcal{F}^*(\Lambda^* p^*) - \mathcal{G}^*(-p^*)\},$$

where $\mathcal{F}^*$ (resp. $\mathcal{G}^*$) is the Legendre transform of $\mathcal{F}$ (resp. $\mathcal{G}$), and Λ^* is the adjoint of Λ.

We now compute $\mathcal{F}^*(\Lambda^* p^*)$ and $\mathcal{G}^*(-p^*)$.

$$\begin{aligned}
\mathcal{F}^*(\Lambda^* p^*) &:= \sup_{v \in W^{1,1}(\Omega)} \langle \Lambda^* p^*, v \rangle_{W^{1,1}(\Omega) \times (W^{1,1}(\Omega))^*} + \int_\Omega \phi \cdot v dx \\
&= \sup_{v \in W^{1,1}(\Omega)} \langle \Lambda^* p^* + \phi, v \rangle_{W^{1,1}(\Omega) \times (W^{1,1}(\Omega))^*} \\
&= \begin{cases} 0, & \text{if } \Lambda^* p^* + \phi = 0 \text{ on } W^{1,1}(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}
\end{aligned}$$

It is easy to see that $\mathcal{G}^*(p^*) = \mathcal{G}_1^*(p_1^*) + \mathcal{G}_2^*(p_2^*) + \mathcal{G}_3^*(\bar{p}^*)$, where $p^* = (p_1^*, p_2^*, \bar{p}^*) \in L^2(\mathbb{R}^2)^2 \times L^\infty(\Omega)^2$. Applying the same steps above, we have

$$\begin{aligned}
\mathcal{G}_1^*(p_1^*) &= \int_{\mathbb{R}^2} \left(\frac{p_1^*(x)^2}{4} + \tilde{f}_1(x) p_1^*(x) \right) dx, \\
\mathcal{G}_2^*(p_2^*) &= \int_{\mathbb{R}^2} \left(\frac{p_2^*(x)^2}{4} + \tilde{f}_2(x) p_2^*(x) \right) dx, \\
\mathcal{G}_3^*(\bar{p}^*) &= \sup_{\bar{p} \in L^1(\Omega)^2} \int_\Omega \bar{p} \cdot \bar{p}^* - \lambda |\bar{p}| dx = \begin{cases} 0, \text{ if } |\bar{p}^*| \leq \lambda, \\ +\infty, \text{ otherwise.} \end{cases}
\end{aligned}$$

Let $\mathcal{K} = \{ p^* \in L^2(\mathbb{R}^2)^2 \times L^\infty(\Omega)^2 : \ |\bar{p}^*| \leq \lambda, \ \Lambda^* p^* + \phi = 0 \text{ on } W^{1,1}(\Omega) \}$, then problem $(\mathcal{P}^*)$ can be re-written as

$$(\mathcal{P}^*) \qquad \sup_{p^* \in \mathcal{K}} \left\{ - \int_{\mathbb{R}^2} \left(\frac{(p_1^*)^2}{4} - \tilde{f}_1 p_1^* \right) dx - \int_{\mathbb{R}^2} \left(\frac{(p_2^*)^2}{4} - \tilde{f}_2 p_2^* \right) dx \right\}.$$

The condition $\Lambda^* p^* + \phi = 0$ on $W^{1,1}(\Omega)$ implies

$$\begin{aligned}
0 &= \langle \Lambda^* p^*, w \rangle + \langle \phi, w \rangle \\
&= \langle p^*, \Lambda w \rangle + \langle \phi, w \rangle \\
&= \langle p_1^*, \mathcal{F}_1 w \rangle + \langle p_2^*, \mathcal{F}_2 w \rangle + \langle \bar{p}^*, Dw \rangle + \langle \phi, w \rangle \\
&= \langle \mathcal{F}_1^* p_1^* + \mathcal{F}_2^* p_2^* - \mathrm{div}\bar{p}^* + \phi, w \rangle, \ \ \forall w \in W^{1,1}(\Omega).
\end{aligned}$$

Therefore $\mathcal{F}_1^* p_1^* + \mathcal{F}_2^* p_2^* - \mathrm{div}\bar{p}^* + \phi = 0$ as a distribution in $\mathcal{D}'(\Omega)$. Now, $\mathcal{F}_1^* p_1^*$, $\mathcal{F}_2^* p_2^*$, $\phi \in L^2(\Omega)$, for $\bar{p}^*$ to satisfy this relation, we must have $\mathrm{div}\bar{p}^* \in L^2(\Omega)$. Therefore, following the trace theorem from Lemma 1, we can define the trace $\bar{p}^* \cdot \vec{n}$ on $\Gamma = \partial\Omega$, where $\vec{n}$ is the outward unit normal to Γ. Moreover, applying integration by parts we get for all $v \in W^{1,1}(\Omega)$,

$$\begin{aligned}
\int_\Gamma (\bar{p}^* \cdot \vec{n}) v d\Gamma &= \int_\Omega \sum_{j=1}^{2} (D_j \bar{p}_j^* v) dx + \int_\Omega \sum_{j=1}^{2} (\bar{p}_j^* D_j v) dx \\
&= \langle \mathcal{F}_1^* p_1^*, v \rangle + \langle \mathcal{F}_2^* p_2^*, v \rangle + \langle \phi, v \rangle \\
&\quad - \langle p_1^*, \mathcal{F}_1 v \rangle - \langle p_2^*, \mathcal{F}_2 v \rangle - \langle v, \phi \rangle \\
&= 0.
\end{aligned}$$

Thus we deduce that if $p^* = (p_1^*, p_2^*, \bar{p}^*) \in \mathcal{K}$ then $\bar{p}^* \cdot \vec{n} = 0 \ d\Gamma - a.e.$ on Γ. Hence, we can write $\mathcal{K}$ as

$$\begin{aligned}
\mathcal{K} = \Big\{ & p^* = (p_1^*, p_2^*, \bar{p}^*) \in L^2(\mathbb{R}^2)^2 \times L^\infty(\Omega)^2 : \\
& |\bar{p}^*| \leq \lambda, \mathcal{F}_1^* p_1^* + \mathcal{F}_2^* p_2^* - \mathrm{div}\bar{p}^* + \phi = 0 \text{ in } \mathcal{D}'(\Omega), \bar{p}^* \cdot \vec{n} = 0 \text{ on } \Gamma \Big\}.
\end{aligned}$$

Now, we had assumed that ϕ is in $\partial F(u)$, i.e. u is the solution to (2.6). And since

the infimum of (2.6) is same as the infimum of $\mathcal{P}$, therefore $\inf \mathcal{P}$ is finite. In addition, the functional in $(\mathcal{P})$ is convex and continuous with respect to Λv in $L^2(\mathbb{R}^2)^2 \times L^1(\Omega)^2$. Thus, by duality theorem, $(\mathcal{P}^*)$ has a solution $M = (M_1, M_2, \bar{M}) \in \mathcal{K}$ and $\inf \mathcal{P} = \sup \mathcal{P}^*$ (minimax relation). So we have

$$\begin{aligned}\int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1(u))^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2(u))^2 dx + \lambda \int_\Omega |Du| - \int_\Omega \phi \cdot u dx \\ = - \int_{\mathbb{R}^2} \left(\frac{M_1^2}{4} - \tilde{f}_1 M_1 \right) dx - \int_{\mathbb{R}^2} \left(\frac{M_2^2}{4} - \tilde{f}_2 M_2 \right) dx. \end{aligned} \tag{2.7}$$

Since $M \in \mathcal{K}$, $\mathcal{F}_1^* M_1 + \mathcal{F}_2^* M_2 - \mathrm{div}\bar{M} + \phi = 0$ in $\mathcal{D}'(\Omega)$. So (2.7) becomes

$$\begin{aligned}&\int_{\mathbb{R}^2} \left(\frac{M_1^2}{4} - \tilde{f}_1 M_1 \right) dx + \int_{\mathbb{R}^2} \left(\frac{M_2^2}{4} - \tilde{f}_2 M_2 \right) dx \\ &+ \int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1 u)^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2 u)^2 dx + \lambda \int_\Omega |Du| \\ &+ \int_{\mathbb{R}^2} M_1 (\mathcal{F}_1 u) dx + \int_{\mathbb{R}^2} M_2 (\mathcal{F}_2 u) dx - \int_\Omega u \cdot \mathrm{div}\bar{M} dx = 0. \end{aligned} \tag{2.8}$$

Following Lemmas (3),(4) we can associate to u and $\bar{M}$ a Radon measure, denoted $Du \cdot \bar{M}$, defined as a distribution on Ω :

$$\langle Du \cdot \bar{M}, \psi \rangle = - \int_\Omega u (\mathrm{div}\bar{M}) \psi dx - \int_\Omega \bar{M} \cdot (\nabla \psi) u dx, \quad \forall \psi \in \mathcal{D}(\Omega).$$

By the generalized Green's formula,

$$\int_\Omega Du \cdot \bar{M} = - \int_\Omega u \cdot \mathrm{div}\bar{M} + \int_\Gamma u (\bar{M} \cdot \vec{n}) d\Gamma.$$

Since $\bar{M} \cdot \vec{n} = 0$ dΓ-a.e., (2.8) becomes

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left(\frac{M_1^2}{4} - \tilde{f}_1 M_1 \right) dx + \int_{\mathbb{R}^2} \left(\frac{M_2^2}{4} - \tilde{f}_2 M_2 \right) dx \\
&+ \int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1 u)^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2 u)^2 dx + \lambda \int_\Omega |Du| \\
&+ \int_{\mathbb{R}^2} M_1(\mathcal{F}_1 u) dx + \int_{\mathbb{R}^2} M_2(\mathcal{F}_2 u) dx + \int_\Omega Du \cdot \bar{M} = 0.
\end{aligned}$$

Applying the decomposition $Du = \nabla u dx + C_u + (u^+ - u^-)\vec{n} d\mathcal{H}^1|_{S_u}$, where $C_u(S_u) = 0$, $\nabla u dx$ is the Lebesgue part, S_u is the jump set, $J_u := (u^+ - u^-)\vec{n} d\mathcal{H}^1|_{S_u}$ denotes the jump part, and C_u is the Cantor part, we finally have

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left(\frac{M_1^2}{4} - \tilde{f}_1 M_1 \right) dx + \int_{\mathbb{R}^2} \left(\frac{M_2^2}{4} - \tilde{f}_2 M_2 \right) dx \\
&+ \int_{\mathbb{R}^2} (\tilde{f}_1 - \mathcal{F}_1 u)^2 dx + \int_{\mathbb{R}^2} (\tilde{f}_2 - \mathcal{F}_2 u)^2 dx \\
&+ \lambda \int_\Omega |\nabla u| dx + \lambda \int_{\Omega \setminus S_u} |C_u| + \lambda \int_{S_u} (u^+ - u^-) d\mathcal{H}^1 \\
&+ \int_{\mathbb{R}^2} M_1(\mathcal{F}_1 u) dx + \int_{\mathbb{R}^2} M_2(\mathcal{F}_2 u) dx \\
&+ \int_\Omega \nabla u \cdot \bar{M} dx + \int_{\Omega \setminus S_u} \bar{M} \cdot C_u + \int_{S_u} (u^+ - u^-) \bar{M} \cdot \vec{n} d\mathcal{H}^1 = 0.
\end{aligned} \tag{2.9}$$

By definition of $\mathcal{G}_1^*(-M_1)$, $\mathcal{G}_2^*(-M_2)$, and $\mathcal{G}_3^*(-\bar{M})$, we have

1°. $\left(\frac{M_1^2}{4} - \tilde{f}_1 M_1 \right) + (\tilde{f}_1 - \mathcal{F}_1 u)^2 + M_1(\mathcal{F}_1 u) \geq 0$ for dx $a.e.$

2°. $\left(\frac{M_2^2}{4} - \tilde{f}_2 M_2 \right) + (\tilde{f}_2 - \mathcal{F}_2 u)^2 + M_2(\mathcal{F}_2 u) \geq 0$ for dx $a.e.$

3°. $\nabla u \cdot \bar{M} + \lambda |\nabla u| \geq 0$ for dx $a.e.$ in Ω.

By the Radon-Nikodym derivative theorem, we have $C_u << |C_u|$ and $\exists h \in L^1(|C_u|)^2$ such that $|h| = 1$ and $C_u = h \cdot |C_u|$. Thus, $\bar{M} \cdot C_u = \bar{M} \cdot h |C_u|$, hence

4°. $(\bar{M} C_u + \lambda |C_u|) = (\lambda + \bar{M} \cdot h)|C_u| \geq 0$, since $|\bar{M}| \leq \lambda$.

Finally, when u^+ and u^- are defined on S_u, we have $u^+ - u^- \geq 0$ and $\lambda + \bar{M} \cdot \vec{n} \geq 0$ (since $|\bar{M}| \leq \lambda$), hence

5°. $(u^+ - u^-)(\lambda + \bar{M} \cdot \vec{n}) \geq 0$ for $d\mathcal{H}^1$ $a.e.$ in S_u.

In order for (2.9) to hold, we must have each of the quantity in 1°, 2°, 3°, 4°, and 5° to be exactly 0. The quantity in 3° equals 0 implies

$$\nabla u \cdot \bar{M} + \lambda |\nabla u| = 0, \ dx\text{-a.e. } x \in \Omega.$$

Hence

$$\bar{M} = -\lambda \frac{\nabla u}{|\nabla u|}, \quad \text{when } |\nabla u| \neq 0.$$

Therefore at points where $|\nabla u| \neq 0$, we have the Neumann boundary condition $\frac{\nabla u}{|\nabla u|} \cdot \vec{n} = 0$ $d\Gamma$-a.e., since $\bar{M} \cdot \vec{n} = 0$ $d\Gamma$-a.e. on Γ. A full characterization of the minimizer u can thus be obtained by setting all quantities in $1^o - 5^o$ to zero, with the above properties on M_1, M_2 and $\bar{M}$.

We can now state the characterization of the subdifferential $\partial F(u)$ of F at u as follows:

Theorem 2.4.2 *Let $\mathcal{F}_1, \mathcal{F}_2, \tilde{f}_1, \tilde{f}_1$ be defined as above. Let $\phi \in L^2(\Omega)$ and $u \in BV(\Omega)$, with $Du = \nabla u dx + C_u + J_u$ the decomposition of Du. Then $\phi \in \partial F(u)$ if and only if there exists $M(x) = (M_1(x), M_2(x), \bar{M}(x)) \in L^2(\mathbb{R}^2) \times L^2(\mathbb{R}^2) \times L^\infty(\Omega)^2$, $|\bar{M}(\cdot)| \leq \lambda$, such that*

$$\nabla u \cdot \bar{M} + \lambda |\nabla u| = 0, \ dx \ a.e. \ x \ \textit{in} \ \ \Omega, \tag{2.10}$$

$$\bar{M} \cdot \nu = 0 \ d\Gamma \ a.e. \ \textit{on} \ \Gamma = \partial\Omega, \tag{2.11}$$

$$\mathcal{F}_1^* M_1 + \mathcal{F}_2^* M_2 - \mathrm{div}\bar{M} + \phi = 0 \ \textit{in} \ \mathcal{D}'(\Omega), \tag{2.12}$$

$$\frac{M_1^2}{4} - \tilde{f}_1 M_1 + M_1(\mathcal{F}_1 u) = (\tilde{f}_1 - \mathcal{F}_1 u)^2 \ dx \ a.e. \ x \ \textit{in} \ \Omega, \tag{2.13}$$

$$\frac{M_2^2}{4} - \tilde{f}_2 M_2 + M_2(\mathcal{F}_2 u) = (\tilde{f}_2 - \mathcal{F}_2 u)^2 \ dx \ a.e. \ x \ \textit{in} \ \Omega, \tag{2.14}$$

$$\lambda + \bar{M} \cdot \nu = 0 \ \ (\textit{i.e.} \ |\bar{M}| = \lambda) \ \ d\mathcal{H}^1 \ a.e. \ x \in S_u, \tag{2.15}$$

$$\text{and finally, } \mathrm{supp}(|C_u|) \subset \mathcal{N}, \tag{2.16}$$

where $\mathcal{N} = \{x \in \Omega \backslash S_u, \ \lambda + \bar{M}(x) \cdot h(x) = 0$ *where* $h \in L^1(|C_u|)^2, |h| = 1, C_u = h|C_u|\}$.

Recall that u is a minimizer of F if and only if $0 \in \partial F(u)$. Setting ϕ in Theorem 2.4.2 above to 0 gives characterizations of the solution of problem (2.1).

2.4.2 Characterization of minimizers via "texture"-norm

In this section we present some more theoretical results that further characterize the solution to our proposed problem (2.1). Similar calculations for minimizers of the ROF model have be done by Y. Meyer [46] (see also [8] for related discussion).

We begin by defining a semi-norm $\| \cdot \|_*$ on H^{-s} dual to the $| \cdot |_{BV}$ norm as follows:

$$\| \ f \ \|_* = \sup_{u \in BV(\Omega), \ |u|_{BV(\Omega)} \neq 0} \frac{|\mathbf{Re}\langle f, u \rangle_{-s}|}{|u|_{BV(\Omega)}}, \quad f \in H^{-s}. \tag{2.17}$$

$\| \cdot \|_*$ is a semi-norm since for any $f, g \in H^{-s}$ and $u \in BV(\Omega)$, $\mathbf{Re}\langle f + g, u \rangle_{-s} = \mathbf{Re}\langle f, u \rangle_{-s} + \mathbf{Re}\langle g, u \rangle_{-s}$, so $|\mathbf{Re}\langle f+g, u \rangle_{-s}| \leq |\mathbf{Re}\langle f, u \rangle_{-s}| + |\mathbf{Re}\langle g, u \rangle_{-s}|$. This implies the triangle inequality. Moreover, it is clear that for any $\lambda \in \mathbb{R}$, $\| \ \lambda f \ \|_* = |\lambda| \ \| \ f \ \|_*$.

Lemma 2.4.1 *If $f \in H^{-s}$ is such that $\|f\|_* < \infty$, then $\mathbf{Re}\langle f, 1_\Omega\rangle_{-s} = 0$.*

Proof: Let $u \in BV(\Omega)$ be such that $|u|_{BV(\Omega)} \neq 0$. Then for any constant $c \in \mathbb{R}$

$$\begin{aligned}\frac{|\mathbf{Re}\langle f, u+c\rangle_{-s}|}{|u+c|_{BV(\Omega)}} &= \frac{|\mathbf{Re}\langle f, u+c\rangle_{-s}|}{|u|_{BV(\Omega)}} \\ &= \frac{|\mathbf{Re}(\langle f, u\rangle_{-s} + \langle f, c\rangle_{-s})|}{|u|_{BV(\Omega)}} \\ &\leq \|f\|_* < \infty.\end{aligned}$$

This implies that there is a constant C such that $|\mathbf{Re}\langle f, c\rangle_{-s}| \leq C < \infty$ for any $c \in \mathbb{R}$. Then $|c||\mathbf{Re}\langle f, 1_\Omega\rangle_{-s}| \leq C < \infty$ for any $c \in \mathbb{R}$, therefore $|\mathbf{Re}\langle f, 1_\Omega\rangle_{-s}|$ must be zero. $\square$

Remark 2.4.1 *The above Lemma implies that if $\|f\|_* < \infty$ defined by (2.17) is finite, then $|\mathbf{Re}\langle f, u\rangle_{-s}| \leq |u|_{BV(\Omega)}\|f\|_*$ for any $u \in BV(\Omega)$.*

Theorem 2.4.3 *Fix $\lambda > 0$. Let u be a minimizer of (2.1) and set $v := u_0 - Ku$. Then the following holds,*

(I) $\| K^*u_0 \|_* \leq \frac{\lambda}{2}$ *if and only if* $u = 0, v = u_0$.

(II) Suppose $\| K^*u_0 \|_* > \frac{\lambda}{2}$. *Then* $u \in BV(\Omega), v = u_0 - Ku$ *is minimizer if and only if* $\| K^*v \|_* \leq \frac{\lambda}{2}$ *and* $\mathbf{Re}\langle K^*v, u\rangle_{-s} = \frac{\lambda}{2}|u|_{BV}$. *In both cases, if in addition* $|u|_{BV(\Omega)} \neq 0$ *then* $\|K^*v\|_* = \frac{\lambda}{2}$.

Proof: (I) The minimizer is $u = 0, v = u_0$ if and only if $\forall h \in BV(\Omega)$ and $\forall \varepsilon \in \mathbb{R}, \varepsilon \neq 0$,

$$\lambda|\varepsilon h|_{BV(\Omega)} + \| u_0 - \varepsilon Kh \|^2_{-s} \geq \| u_0 \|^2_{-s} .$$

Expanding the term $\| u_0 - \varepsilon Kh \|^2_{-s}$, we get

$$\lambda|\varepsilon||h|_{BV(\Omega)} - 2\varepsilon\, \mathbf{Re}\langle u_0, Kh\rangle_{-s} + \varepsilon^2 \| Kh \|^2_{-s} \geq 0, \quad \text{for any } \varepsilon \neq 0.$$

If $\varepsilon > 0$:

$$\lambda|\varepsilon||h|_{BV(\Omega)} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq 2|\varepsilon|\mathbf{Re}\langle u_0, Kh\rangle_{-s}.$$

If $\varepsilon < 0$:

$$\lambda|\varepsilon||h|_{BV(\Omega)} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq -2|\varepsilon|\mathbf{Re}\langle u_0, Kh\rangle_{-s}.$$

Hence,

$$\lambda|\varepsilon||h|_{BV(\Omega)} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq 2|\varepsilon|\big|\mathbf{Re}\langle u_0, Kh\rangle_{-s}\big|.$$

Dividing both sides by $|\varepsilon|$ and taking $|\varepsilon| \to 0$, we obtain

$$\lambda|h|_{BV(\Omega)} \geq \big|\mathbf{Re}\langle u_0, Kh\rangle_{-s}\big| = \big|\mathbf{Re}\langle K^* u_0, h\rangle_{-s}\big|, \tag{2.18}$$

for all $h \in BV(\Omega)$. Hence, $\parallel K^* u_0 \parallel_* \leq \frac{\lambda}{2}$.

Conversely, if $\parallel K^* u_0 \parallel_* \leq \frac{\lambda}{2}$, then by the remark after the Lemma, the last inequality (2.18) holds for any $h \in BV(\Omega)$. From here, add $\parallel u_0 \parallel^2_{-s}$ to both sides of (2.18) and $\parallel Kh \parallel^2_{-s}$ to the left hand side, we obtain

$$\lambda|h|_{BV(\Omega)} + \parallel u_0 \parallel^2_{-s} - 2\,\mathbf{Re}\langle u_0, Kh\rangle_{-s} + \parallel Kh \parallel^2_{-s} \geq \parallel u_0 \parallel^2_{-s}.$$

That is,

$$\lambda|h|_{BV(\Omega)} + \parallel u_0 - Kh \parallel^2_{-s} \geq \parallel u_0 \parallel^2_{-s}.$$

This is true for all $h \in BV(\Omega)$, thus $u = 0$ is the minimizer.

(II) We have $u \in BV(\Omega), v = u_0 - Ku$ is minimizer if and only if $\forall h \in BV(\Omega), \forall \varepsilon \in \mathbb{R}$:

$$\lambda|u + \varepsilon h|_{BV(\Omega)} + \parallel v - \varepsilon Kh \parallel^2_{-s} \geq \lambda|u|_{BV(\Omega)} + \parallel v \parallel^2_{-s}.$$

Then,

$$\lambda|\varepsilon||h|_{BV(\Omega)} - 2\varepsilon\, \mathbf{Re}\langle v, Kh\rangle_{-s} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq 0,$$

for any $\varepsilon \in \mathbb{R}$, hence

$$\lambda|\varepsilon||h|_{BV(\Omega)} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq 2|\varepsilon||\mathbf{Re}\langle K^*v, h\rangle_{-s}|.$$

Divide both sides by $|\varepsilon|$ and taking $\varepsilon \to 0$, we get

$$\frac{\lambda}{2}|h|_{BV(\Omega)} \geq |\mathbf{Re}\langle K^*v, h\rangle_{-s}|, \quad \forall h \in BV(\Omega).$$

Hence, $\parallel K^*v \parallel_* \leq \frac{\lambda}{2}$.

Repeat the same calculation with specifically $h = u$ and $-1 < \varepsilon < 0$ (so that $1+\varepsilon > 0$), we get

$$\lambda(1+\varepsilon)|u|_{BV(\Omega)} + \parallel v \parallel^2_{-s} -2\varepsilon\, \mathbf{Re}\langle v, Ku\rangle_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq \lambda|u|_{BV(\Omega)} + \parallel v \parallel^2_{-s}.$$

$\Longrightarrow$ $\lambda\varepsilon|u|_{BV(\Omega)} - 2\varepsilon\, \mathbf{Re}\langle K^*v, u\rangle_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq 0$. Since $\varepsilon < 0$, this is the same as

$$-\lambda|\varepsilon||u|_{BV(\Omega)} + 2|\varepsilon|\mathbf{Re}\langle K^*v, u\rangle_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq 0.$$

Divide both sides by $|\varepsilon|$ and then let $\varepsilon \nearrow 0$, we obtain

$$2\, \mathbf{Re}\langle K^*v, u\rangle_{-s} \geq \lambda|u|_{BV(\Omega)}.$$

Since $|\mathbf{Re}\langle K^*v, u\rangle_{-s}| \leq \frac{\lambda}{2}|u|_{BV(\Omega)}$, we get $|\mathbf{Re}\langle K^*v, u\rangle_{-s}| = \frac{\lambda}{2}|u|_{BV(\Omega)}$. If $|u|_{BV(\Omega)} \neq 0$, then $\parallel K^*v \parallel_* = \frac{\lambda}{2}$.

Conversely, suppose $u \in BV(\Omega)$ and $v = u_0 - Ku$ satisfy $\parallel K^*v \parallel_* \leq \frac{\lambda}{2}$ (with equality

if $|u|_{BV(\Omega)} \neq 0$) and $\mathbf{Re}\langle K^*v, u\rangle_{-s} = \frac{\lambda}{2}|u|_{BV(\Omega)}$. Then $\forall h \in BV(\Omega)$ and $\varepsilon \in \mathbb{R}$,

$$\begin{aligned}
&\lambda|u+\varepsilon h|_{BV(\Omega)} + \parallel v - \varepsilon Kh \parallel^2_{-s} \\
&\geq 2\,\mathbf{Re}\langle K^*v, u+\varepsilon h\rangle_{-s} + \parallel v \parallel^2_{-s} - 2\varepsilon\,\mathbf{Re}\langle v, Kh\rangle_{-s} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \\
&= \parallel v \parallel^2_{-s} + 2\,\mathbf{Re}\langle K^*v, u\rangle_{-s} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \\
&\geq \parallel v \parallel^2_{-s} + \lambda|u|_{BV(\Omega)}.
\end{aligned}$$

Therefore, u is minimizer. □

Remark 2.4.2 *Note that if* $\mathbf{Re}\langle u_0, K1_\Omega\rangle_{-s} = 0$ *(can be simply obtained by subtracting a constant from* u_0*), then in part (II) above we always have* $|u|_{BV(\Omega)} \neq 0$ *(because if* u *is a constant minimizer, then* u *must be zero in this case, but this cannot hold in part (II)).*

Remark 2.4.3 *The above characterization of minimizers holds if the total variation* $|u|_{BV(\Omega)}$ *is substituted by another functional* Φ *on* $BV(\Omega)$ *that is convex, lower semi-continuous and positive homogeneous of degree 1.*

2.5 Derivation of the Euler-Lagrange Equation

In this section we will formally compute the Euler-Lagrange equation for the minimization problem (2.1). We have shown in Section 2.4.1, via the dual variable $\bar{M}$, that if u is the minimizer of (2.1), then u automatically satisfies the Neumann boundary condition $\frac{Du}{|Du|} \cdot \vec{n} = 0$ for $d\Gamma$-a.e. $x \in \partial\Omega$ where $|Du| \neq 0$. Recall also that $u = 0$ outside $\overline{\Omega}$ when $\hat{u}$ appears. The associated Euler-Lagrange equation for (2.1) is derived as follows:

Take any test function $\varphi \in C^\infty(\Omega)$. Define

$$
\begin{aligned}
g(\varepsilon) &= F(u+\varepsilon\varphi) \qquad (2.19)\\
&:= \lambda \int_\Omega \big|Du + \varepsilon D\varphi\big| + \int_{\mathbb{R}^2} \frac{(\hat{u}_0 - \widehat{Ku} - \varepsilon\widehat{K\varphi})(\bar{\hat{u}}_0 - \overline{\widehat{Ku}} - \varepsilon\overline{\widehat{K\varphi}})}{(1+|\xi|^2)^s} d\xi.
\end{aligned}
$$

Recall the two identities $\int v\hat{w}dx = \int \hat{v}wdx$ and $\int v\bar{\hat{w}}dx = \int \bar{\hat{v}}wdx$, for any $w(x) \in \mathbb{R}$. Applying these to (2.19), we obtain

$$
\begin{aligned}
g'(0) &= \lambda \int_\Omega \frac{Du}{|Du|}\cdot D\varphi - \int_{\mathbb{R}^2} \frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\widehat{K\varphi} - \int_{\mathbb{R}^2} \frac{\hat{u}_0 - \widehat{Ku}}{(1+|\xi|^2)^s}\overline{\widehat{K\varphi}}\\
&= \lambda \int_{\partial\Omega} \varphi \frac{Du}{|Du|}\cdot \vec{n}\, ds - \lambda \int_\Omega \operatorname{div}\big(\frac{Du}{|Du|}\big)\varphi\\
&\quad - \int_{\mathbb{R}^2} \mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\Big)K\varphi - \int_{\mathbb{R}^2} \overline{\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\Big)}K\varphi.
\end{aligned}
$$

The first integral vanishes by applying the implicit Neumann boundary condition. The last two integrals are conjugates of each other. Thus we obtain

$$
\begin{aligned}
g'(0) &= -\lambda \int_\Omega \operatorname{div}\big(\frac{Du}{|Du|}\big)\varphi - 2\int_{\mathbb{R}^2} \mathbf{Re}\Big\{\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\Big)\Big\}\, K\varphi\\
&= -\lambda \int_\Omega \operatorname{div}\big(\frac{Du}{|Du|}\big)\varphi - 2\int_{\mathbb{R}^2} K^*\Big(\mathbf{Re}\Big\{\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\Big)\Big\}\Big)\varphi
\end{aligned}
$$

for all $\varphi \in C^\infty(\Omega)$, where K^* is the adjoint of K. Thus, the Euler-Lagrange equation is

$$
\begin{cases}
\lambda \operatorname{div}\Big(\frac{Du}{|Du|}\Big) + K^*\Big(2\,\mathbf{Re}\Big\{\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \overline{\widehat{Ku}}}{(1+|\xi|^2)^s}\Big)\Big\}\Big) &= 0 \quad \text{in } \Omega\\
\frac{Du}{|Du|}\cdot\vec{n} &= 0 \quad \text{on } \partial\Omega,\\
u &= 0 \quad \text{outside } \overline{\Omega}
\end{cases} \qquad (2.20)
$$

Remarks 2.5.1 *(i) When $s = 0$ and K is the identity operator, the Euler-Lagrange*

equation in (2.20) above is the same as that of the ROF model (because $u_0 - u = \mathcal{F}(\bar{\hat{u}_0} - \bar{\hat{u}})$ when u_0 and u are real-valued). This proves that the obtained PDE is consistent with the energy (2.1), (see Remark 2.2.1(i) above).

(ii) When the operator K is convolution with a kernel k, (e.g. K is convolution with a Gaussian kernel), then owing to the identity $\widehat{K\varphi} = \hat{k}\hat{\varphi}$, equation (2.20) becomes

$$\lambda \mathrm{div}\big(\frac{Du}{|Du|}\big) + 2\,\mathbf{Re}\Big\{\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)\Big\} = 0.$$

In fact, $\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)$ is real-valued, hence we get

$$\lambda \mathrm{div}\Big(\frac{Du}{|Du|}\Big) + 2\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big) = 0. \tag{2.21}$$

(iii) If we integrate both sides of (2.21) over $\mathbb{R}^2$, imposing that $u = 0$ outside of $\overline{\Omega}$ and $\frac{Du}{|Du|} \cdot \vec{n} = 0$ on $\partial\Omega$($\vec{n}$ is the outward normal to $\partial\Omega$), then the first term vanishes, and we get

$$\int_{\mathbb{R}^2} \mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big) = 0, \text{ implying that } \mathcal{F}\left(\mathcal{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)\right)(0) = 0.$$

If we write $w(\xi) = \frac{\hat{u}_0(\xi) - \hat{k}(\xi)\hat{u}(\xi)}{(1+|\xi|^2)^s}\bar{\hat{k}}(\xi),$ then $\bar{w}(\xi) = w(-\xi)$. Then

$$\begin{aligned}
\hat{\bar{w}}(x) &= \int_{\mathbb{R}^2} w(-\xi) e^{-2\pi i \xi \cdot x} d\xi, \quad (\textit{make a change of variables } -\xi \to \xi), \\
&= \int_{\mathbb{R}^2} w(\xi) e^{2\pi i \xi \cdot x} d\xi \\
&= \check{w}(x)
\end{aligned}$$

Therefore, $\hat{\bar{w}}(0) = w(0) = \big(\hat{u}_0(0) - \hat{k}(0)\hat{u}(0)\big)\bar{\hat{k}}(0) = 0$. We impose that $\hat{k}(0) \neq 0$,

that is, k has nonzero mean, then $\hat{u}_0(0) - \hat{u}(0) = 0$, which means that the component $v = u_0 - u$ has zero mean.

2.6 Numerical Approximation of the Model

We present here a numerical algorithm for solving the proposed model (2.1). We shall work with the case when the blurring operator K is the convolution with a kernel k, i.e. the problem we are solving is

$$\inf_{u \in BV(\Omega)} F(u) = \lambda \int_\Omega |Du| dx + \int_{\mathbb{R}^2} \frac{|\hat{u}_0 - \hat{k}\hat{u}|^2}{(1+|\xi|^2)^s} d\xi. \tag{2.22}$$

Since the total variation is not differentiable at 0, we approximate (2.22) with the following:

$$\inf_{u_\varepsilon \in BV(\Omega)} F_\varepsilon(u_\varepsilon) = \lambda \int_\Omega \sqrt{\varepsilon^2 + |Du_\varepsilon|^2} dx + \int_{\mathbb{R}^2} \frac{|\hat{u}_0 - \hat{k}\hat{u}_\varepsilon|^2}{(1+|\xi|^2)^s} d\xi \tag{2.23}$$

Remark 2.6.1 *It can be shown that the regularized minimization problem (2.23) has a unique solution for each $\lambda > 0$ and $\varepsilon > 0$ by repeating the steps in Section 2.3. In addition, following the arguments in [17], it can be shown that as ε converges to zero, the unique solution of (2.23) converges to the unique solution of (2.22).*

Applying gradient descent method to solve (2.23), we obtain the following non-linear PDE:

$$\begin{cases} \frac{\partial u}{\partial t} = \lambda div\left(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}}\right) \; + \; 2\,\mathfrak{F}\left(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\right) & \text{in } \Omega \\ \frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}} \cdot \vec{n} \; = \; 0 & \text{on } \partial\Omega \\ u = 0 \text{ outside } \overline{\Omega}. \end{cases} \tag{2.24}$$

To proceed with the discretization of (2.24), let us assume that the initial discrete image u_0 is of $M \times M$ pixels, and that it has been sampled from its continuous version at uniformly

spaced points starting at $(0,0)$, i.e. $u_{0,j,l} = u_0(j\triangle x, l\triangle x)$ for $j,l = 0,1,\cdots,M-1$, where $\triangle x$ is to be determined.

2.6.1 The "force" term

Computing the force term $\mathfrak{F}\left(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\right)$ in (2.24) requires the Discrete Fourier Transform (DFT), which is defined for any $M \times M$ signal u as

$$\hat{u}_{m,n} = \frac{1}{M^2}\sum_{j,l=0}^{M-1} u_{j,l}e^{-2\pi i(jm+ln)/M}, \quad \text{for } m,n = 0,1,\cdots,M-1.$$

The Inverse Discrete Fourier Transform (IDFT) for u is defined as

$$\check{u}_{j,l} = \sum_{m,n=0}^{M-1} u_{m,n}e^{2\pi i(jm+ln)/M}, \quad \text{for } j,l = 0,1,\cdots,M-1.$$

The DFT array $(\hat{u})_{m,n}$ is, indeed, as taken from its continuous counterpart at frequencies $(m\triangle\xi, n\triangle\xi)$ $(m,n = 0,1,...,M-1)$. The inverse relation between the DFT and IDFT implies that $\triangle x$ and $\triangle\xi$ are inversely related by

$$\triangle\xi = \frac{1}{M\triangle x}.$$

Therefore, to give our numerical computations a balance weight between the spatial terms and the Fourier frequency terms, we shall choose

$$\triangle x = \frac{1}{\sqrt{M}} \quad \text{and} \quad \triangle\xi = \frac{1}{\sqrt{M}}.$$

A final note on computing the force term: before taking the DFT, we multiply $(-1)^{j+l}$ to the signal. This shifts the origin of the frequency domain to the center of the image. Thus,

for $0 \leq m, n < M$, the (m, n) entry corresponds to the Fourier coefficient at frequency $((m - \frac{M}{2})\triangle\xi, (n - \frac{M}{2})\triangle\xi)$. Therefore, we evaluate the weight function $\frac{1}{(1+|\xi|^2)^s}$ at points $\xi_1, \xi_2 = -\frac{M}{2}\triangle\xi, -\frac{(M-1)}{2}\triangle\xi, \cdots, \frac{(M-1)}{2}\triangle\xi$.

2.6.2 The curvature term

For the discrete gradient, we shall use the following usual notations:

$$\nabla^{+,+}u = (\nabla_x^+ u, \nabla_y^+ u), \qquad \nabla^{+,-}u = (\nabla_x^+ u, \nabla_y^- u)$$

$$\nabla^{-,+}u = (\nabla_x^- u, \nabla_y^+ u), \qquad \nabla^{-,-}u = (\nabla_x^- u, \nabla_y^- u)$$

where

$$\nabla_x^+ u = u_{j+1,l} - u_{j,l}, \qquad \nabla_x^- = u_{j,l} - u_{j-1,l}$$

$$\nabla_y^+ u = u_{j,l+1} - u_{j,l}, \qquad \nabla_y^- = u_{j,l} - u_{j,l-1}.$$

Since the dual operators to $\nabla^{+,+}$, $\nabla^{+,-}$, $\nabla^{-,+}$, $\nabla^{-,-}$ are, respectively, the operators $div^{-,-}$, $div^{-,+}$, $div^{+,-}$, $div^{+,+}$, we can write the regularized curvature term in one of four ways:

$$div(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}}) \approx div^{\alpha*,\beta*}\Big(\frac{\nabla^{\alpha,\beta} u}{\sqrt{\varepsilon^2 + |\nabla^{\alpha,\beta} u|^2}}\Big) \tag{2.25}$$

where $div^{\alpha*,\beta*}$ denotes the dual operator of $\nabla^{\alpha,\beta}$, with $\alpha, \beta = +, -$.

To make our numerical scheme rotationally invariant, we use all four approximations to the gradient operator by alternating them with each iteration [8]. For example, if u^n was computed using $\nabla^{+,+}$, then we use $\nabla^{+,-}$ to compute u^{n+1}, $\nabla^{-,+}$ to compute u^{n+2}, and $\nabla^{-,-}$ to compute u^{n+3}, and then repeat.

2.6.3 Numerical algorithm

We solve (2.24) with the following iterative semi-implicit scheme [64]:

1. u^0 is arbitrarily given (we can take $u^0 = u_0$)

2. Once u^n is calculated, compute the forcing term $F^n = \mathfrak{F}\Big(\frac{\bar{\hat{u}}_0 - \bar{\hat{k}}\bar{\hat{u}}^n}{(1+|\xi|^2)^s}\hat{k}\Big)$

3. Compute $u_{j,l}^{n+1},$ for $j, l = 1, 2, \cdots, M-2$ as the solution of the linear discrete equation:
   $u_{j,l}^{n+1} = u_{j,l}^n + \triangle t\Big(\frac{\lambda}{\triangle x} div^{\alpha*,\beta*}\Big(\frac{\nabla^{\alpha,\beta}u^{n+1}}{\sqrt{\varepsilon^2+|\nabla^{\alpha,\beta}u^n|^2}}\Big)_{j,l} + 2F_{j,l}^n\Big)$
   with $\varepsilon = \varepsilon'\triangle x$, some $\varepsilon' > 0$ small, and the boundary conditions
   $u_{0,l}^{n+1} = u_{1,l}^{n+1}, \quad u_{M-1,l}^{n+1} = u_{M-2,l}^{n+1}, \quad u_{j,0}^{n+1} = u_{j,1}^{n+1}, \quad u_{j,M-1}^{n+1} = u_{j,M-2}^{n+1}.$

To clarify the notations in step 3 of our algorithm, assume that we are solving at pixel (j, l) the equation

$$u_{j,l}^{n+1} = u_{j,l}^n + \triangle t\Big(\frac{\lambda}{(\triangle x)} div^{-,-}\Big(\frac{\nabla^{+,+}u^{n+1}}{\sqrt{\varepsilon^2 + |\nabla^{+,+}u^n|^2}}\Big)_{j,l} + 2F_{j,l}^n\Big).$$

Let $d_{j,l} = (\sqrt{\varepsilon^2 + |\nabla^{+,+}u^n|^2})_{j,l}$, which is known since u^n is already computed. Then

$$div^{-,-}\Big(\frac{\nabla^{+,+}u^{n+1}}{\sqrt{\varepsilon^2 + |\nabla^{+,+}u^n|^2}}\Big)_{j,l} = \frac{u_{l+1,l}^n - u_{j,l}^{n+1}}{d_{j,l}} - \frac{u_{j,l}^{n+1} - u_{j-1,l}^n}{d_{j-1,l}} + \frac{u_{j,l+1}^n - u_{j,l}^{n+1}}{d_{j,l}} - \frac{u_{j,l}^{n+1} - u_{j,l-1}^n}{d_{j,l-1}}$$

Basically, we set all terms at the current pixel (j, l) in the curvature term to be unknown. Setting $c_{j,l} = \frac{\lambda\triangle t}{d_{j,l}(\triangle x)}$, then we have

$$(1+c_{j-1,l}+2c_{j,l}+c_{j,l-1})u_{j,l}^{n+1} = u_{j,l}^n + c_{j-1,l}u_{j-1,l}^n + c_{j,l}u_{j+1,l}^n + c_{j,l}u_{j,l+1}^n + c_{j,l-1}u_{j,l-1}^n + 2\triangle t F_{j,l}^n.$$

Hence, $u_{j,l}^{n+1}$ is obtained by dividing the coefficient $(1 + c_{j-1,l} + 2c_{j,l} + c_{j,l-1})$ to both sides of the equation. We remark here that the semi-implicit scheme above approaches the steady

state equation much faster than an explicit scheme, and hence, an advantage.

2.6.4 The blurring kernel

In our implementation, we perform blurring directly in the Fourier frequency domain. Since the Fourier Transform of a Gaussian is again a Gaussian, to blur an image, we multiply (pixel-wise) a Gaussian kernel to the DFT of the image, then we take the inverse transform. For a Gaussian kernel we used the form

$$k(\xi_1, \xi_2) = exp\Big(- \frac{\xi_1^2 + \xi_2^2}{2/\alpha^2} \Big).$$

Thus, in our numerical computation, we never compute a discrete convolution.

2.7 Numerical results for image restoration

In this section we present numerical results obtained by applying our proposed new model to image denoising, deblurring and decomposition. For comparison, we also present results from application of the ROF model [57] to the same images. In our implementation of both models (our proposed model and the ROF model), we discretize the curvature term in the manner given in Section 2.6.2. In our proposed model, s is a parameter. We will show numerical results obtained with various values of s.

For image restoration, the parameter λ was chosen so that the best residual-mean-squared-error ($RMSE$) is obtained. For the $RMSE$ and SNR, we used the expressions

$$RMSE = \frac{\sqrt{\sum \left(u_{j,l} - u_{c,j,l}\right)^2}}{MN}, \qquad SNR = \frac{\sum \left(u_{j,l} - \frac{\sum u_{j,l}}{MN}\right)^2}{\sum \left(u_{j,l} - u_{c,j,l} - \frac{\sum (u_{j,l} - u_{c,j,l})}{MN}\right)^2},$$

where u_c is the clean image (which is known in all of our experiments), and M, N are the

size of the image.

We mention that alternative approaches for selecting λ is the method proposed by Gilboa-Sochen-Zeevi [37], where the authors choose λ in an optimal way, by maximizing an estimate of the SNR or by minimizing the correlation between u and $f - u$.

In Figure 2.2, we show the denoising results obtained from our proposed new model with H^{-1} and H^{-0} norms in the fidelity and the ROF model performed on a synthetic piece-wise constant image with additive Gaussian white noise of $\sigma = 30$ (shown in Fig. 2.1). In Figure 2.3, we show more results on the same image from our proposed model using $s = 0.5$ and $s = 2$, respectively, for the H^{-s} norm in the fidelity term. We also show the results from our model using H_0^{-1} semi-norm for the fidelity term (as discussed in Remark 2.2.1, equivalent case with the model from [52]). The $RMSE$ and SNR, together with the value of the parameter λ, for all results in this experiment are shown in Table 2.1. Comparing the results, our model with $H^{-0.5}$ norm gives results with the best $RMSE$, while the H_0^{-1} semi-norm gives results with the best SNR. Visually, the results from our model with H^{-1} norm preserves best the edges in the u component. Overall, our proposed model performs much better than the ROF model, as expected.

In Figure 2.5 we show results from another denoising application. We applied our model using $H^{-0.5}$ and H^{-1} and the ROF model to a noisy image of a woman. The noise is additive white noise with standard deviation $\sigma = 20$ (see Fig. 2.4). The $RMSE$ and SNR of the original noisy image is 0.0762599 and 8.01743, respectively. After denoising, the best results from the ROF, $H^{-0.5}$, and H^{-1} have the SNR of 30.64769, 32.97592, 32.82501, and the $RMSE$ of 0.03548729, 0.03461537, 0.03472017, respectively. The values for λ that yield these results are $14, 0.4, 0.25$ for the ROF,$H^{-0.5}$, H^{-1} model, respectively. Again, our model yields better results than the ROF model.

We also try texture removal with our H^{-1}-model and compare with ROF model. The results are shown in Figures 2.6, 2.7. In Figure 2.6, the texture image is synthetically

Restoration Model	λ	$RMSE$	SNR
Noisy Image		0.224578	4.03064
ROF	55	0.0526239	51.99021
H^0	9.3	0.05204916	53.11084
$H^{-0.5}$	3	0.04953237	58.99322
H^{-1}	1.75	0.04959289	59.41552
H^{-2}	0.9	0.05290355	52.67199
H_0^{-1}	3.8	0.04977601	59.905

Table 2.1: $RMSE$ and SNR for the denoising results on the synthetic piece-wise constant image shown in Figures 2.2,2.3.

created, and in Figure 2.7, the image is natural. We can see that the H^{-1}-norm models better the texture than the L^2-norm.

In Figures 2.8, 2.9, we show deblurring results on a synthetic image and a natural image of an office. The blur is done as described in Sect. 2.6.4 with $\alpha = 0.8$. For the result in Figure 2.8, the blurred image has $RMSE = 0.1016$. The improved image has $RMSE = 0.03618513$ using the parameter $\lambda = 0.0011$. For the result in Figure 2.9, the blurred image has $RMSE = 0.181955$, and the improved image has $RMSE = 0.10373$ using $\lambda = 0.0004$. Visually, we see a significant improvement in the recovered image as compared to the degraded one.

Our last experiment is recovering from a blurred and noisy image. In Figure 2.10, we add white noise of standard deviation $\sigma = 10$ to the blurry image in Figure 2.8. The degradation has $RMSE = 0.121289$. The improved image is obtained with $\lambda = 0.0675$, and the $RMSE = 0.0607262$.

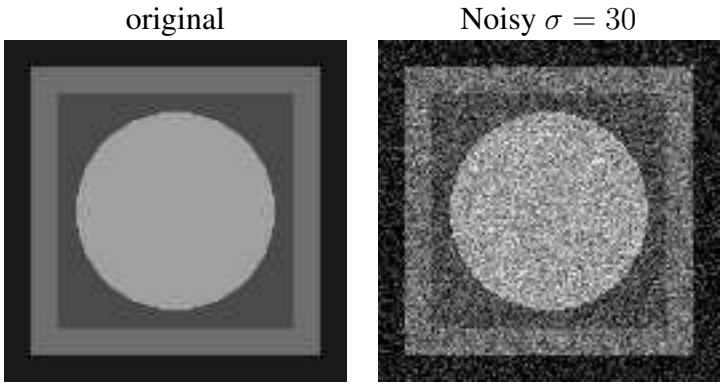


Figure 2.1: A synthetic image and its noisy version with additive Gaussian white noise with standard deviation 30 and zero mean.

2.8 Conclusion

We have proposed a new variational model for image restoration and decomposition, using the Hilbert-Sobolev spaces of negative degree of differentiability to capture oscillatory patterns. We have presented an algorithm to solve the proposed variational problem. We have also presented results of numerical experiments for our proposed model and for the classical ROF model. In each experiment, we have chosen the parameter λ so that the RMSE measurement is minimized. The results obtained from our proposed model are improved, visually and quantitatively (in terms of RMSE and SNR), over the ROF model, for values of $s \in [0.5, 1]$. Outside this interval, we have obtained results similar with those produced by the ROF model.

In each of our numerical experiments, we choose manually the value for s. The results from our numerical experiments with synthetic and real images suggest that Gaussian white noise is captured best by the H^{-s} norm when $0.5 \leq s \leq 1$. This agrees with Mumford and Gidas [49] whose work has motivated us to investigate the spaces H^{-s} for modeling oscillatory patterns.

We also conclude that the value $s = 1$ corresponding to H^{-1} can be used both for image

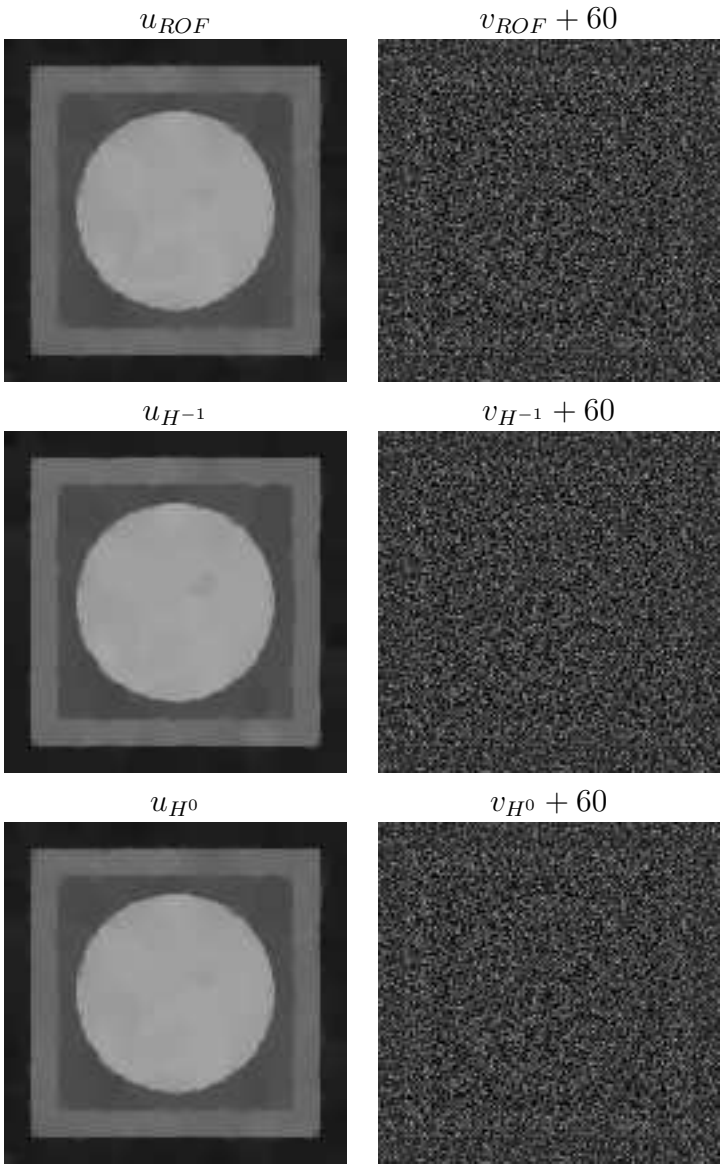


Figure 2.2: Comparison of results from our proposed model ($s = 0$ and $s = -1$) with the ROF model. Top: denoising results obtained from the ROF model. Middle: denoising results obtained with H^{-1}. Bottom: denoising results obtained with H^0.

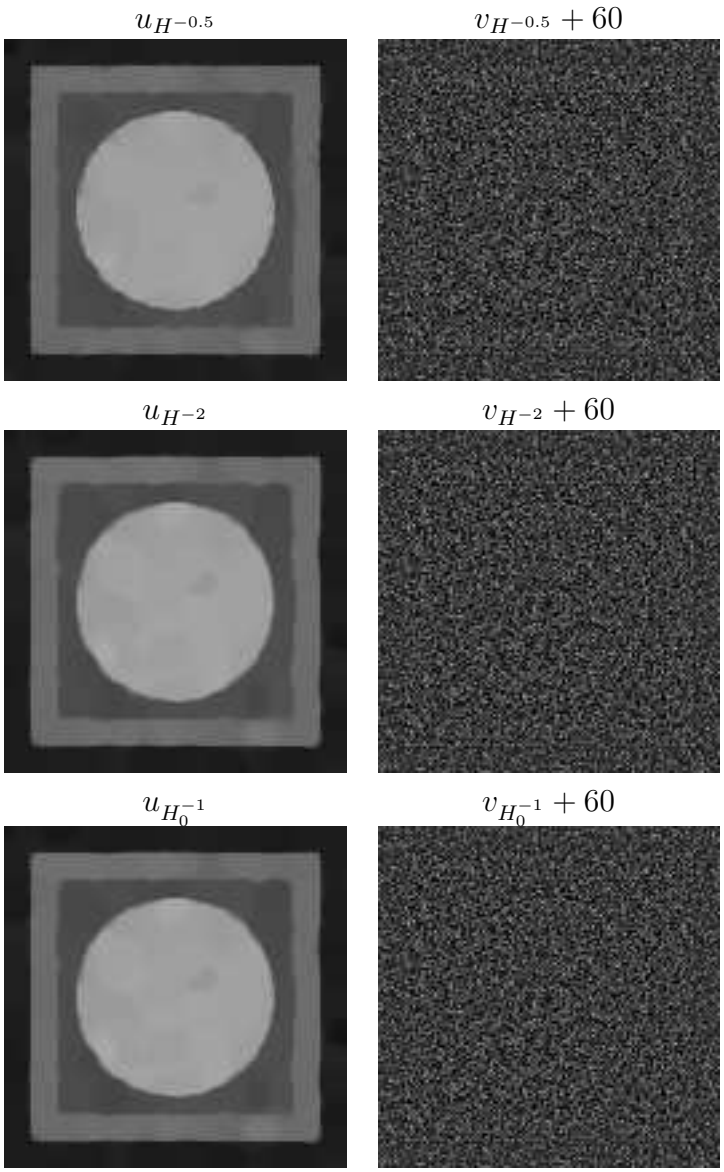


Figure 2.3: More results from our proposed model, using for the norm in the fidelity term $H^{-0.5}$ (row 1), H^{-2} (row 2), and H_0^{-1} semi-norm (row 3).

Figure 2.4: Lena image and its noisy version with additive white noise.

denoising as well as for cartoon and texture separation. However, we do not focus in this paper on $u + v + w$ decompositions, with u a cartoon component, v a texture component and w a noise component, and we leave the study of separating noise from texture in a future work; we refer the reader to Aujol-Chambolle [15] for a solution in this direction.

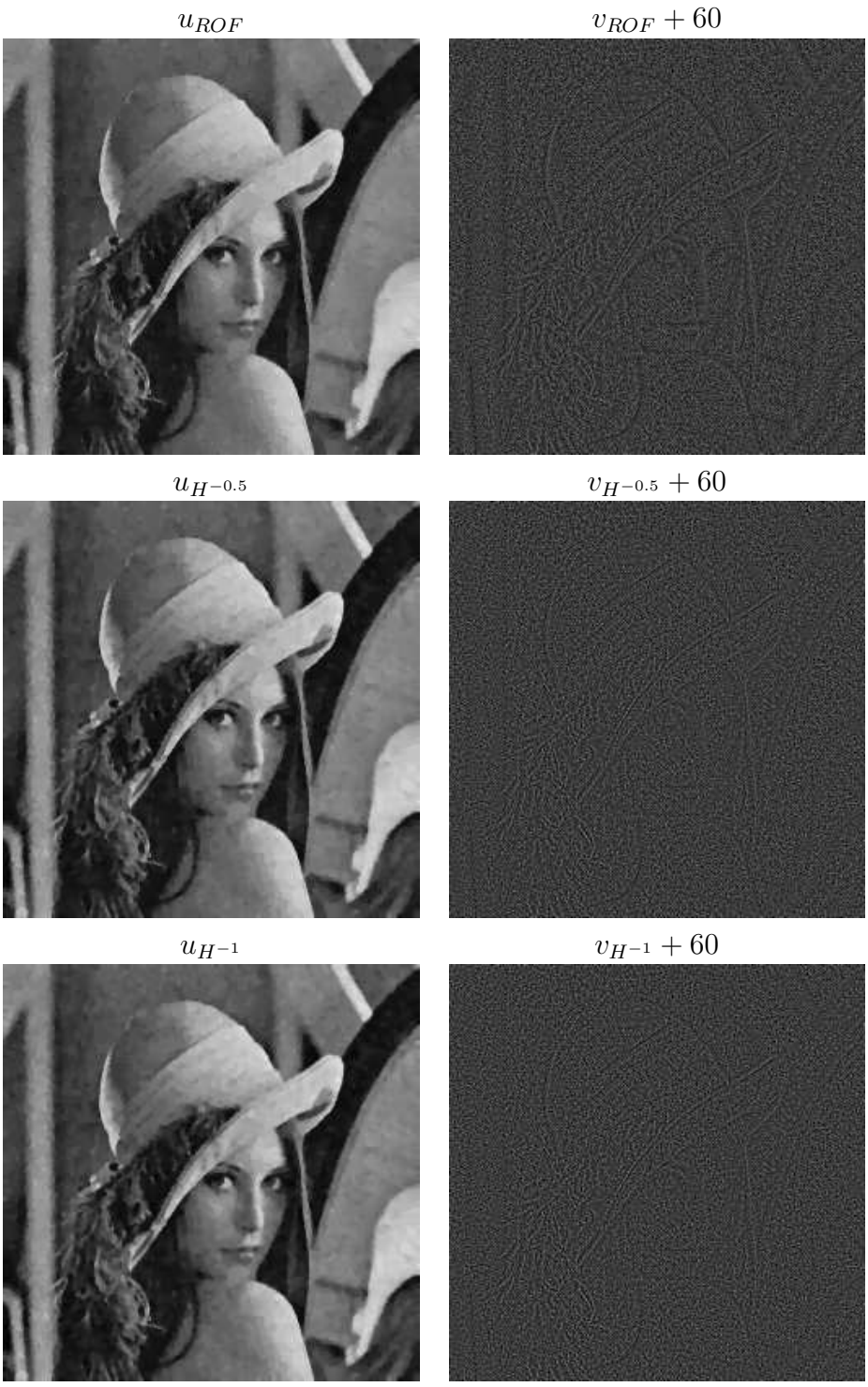


Figure 2.5: Denoising results on Lena image. Result of ROF model (top), result of our model with $H^{-0.5}$ (middle), result of our model with H^{-1} (bottom).

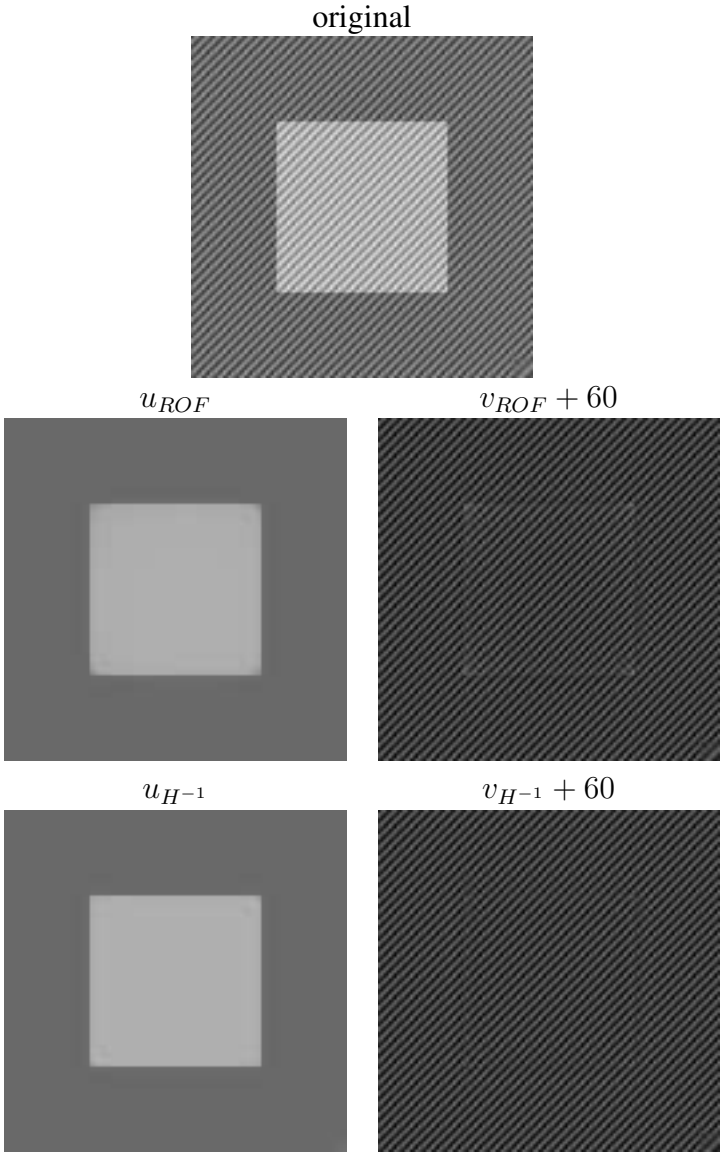


Figure 2.6: Decomposition of a synthetic textured image. Results from ROF model ($\lambda = 42$) and our model with H^{-1} ($\lambda = 0.5$).

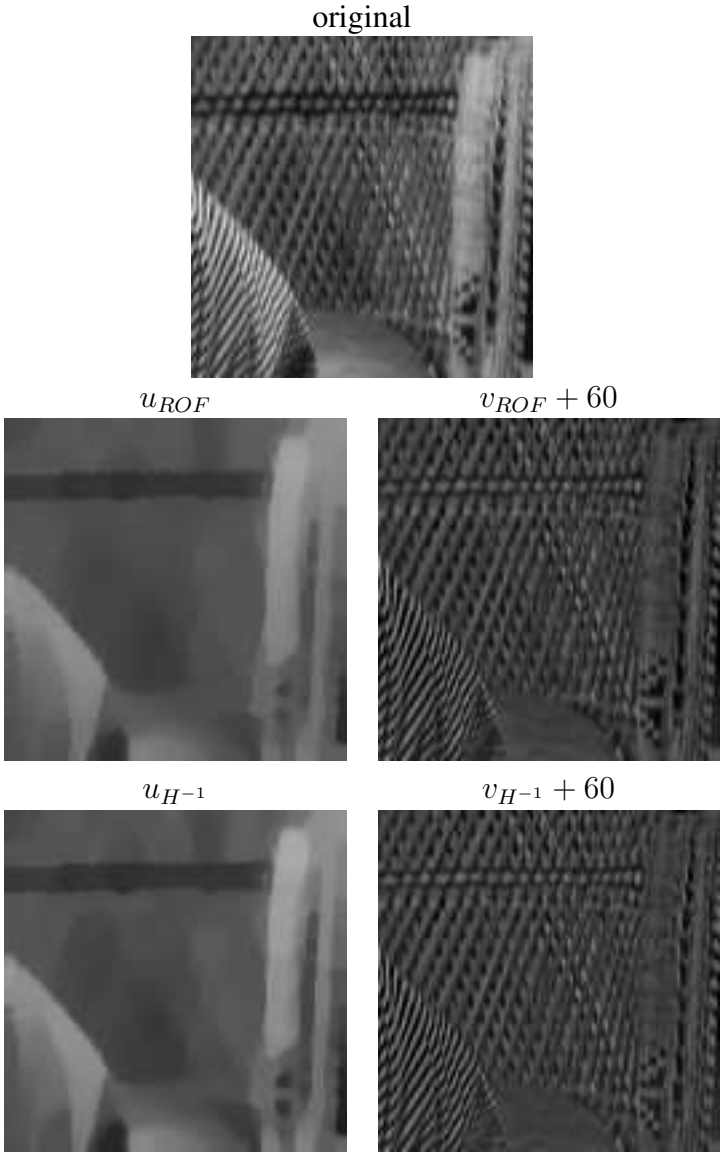


Figure 2.7: Decomposition of a natural textured image. Results from ROF model ($\lambda = 58$, middle) and our model with H^{-1} ($\lambda = 2.5$, bottom).

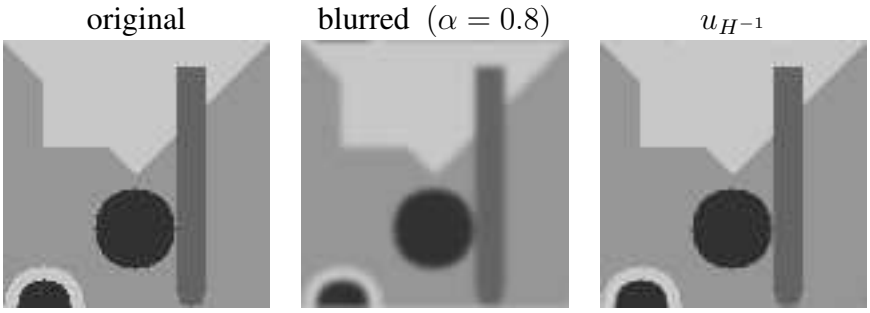


Figure 2.8: Deblurring on a synthetic image. Result is obtained from our model with H^{-1} ($\lambda = 0.0011$).

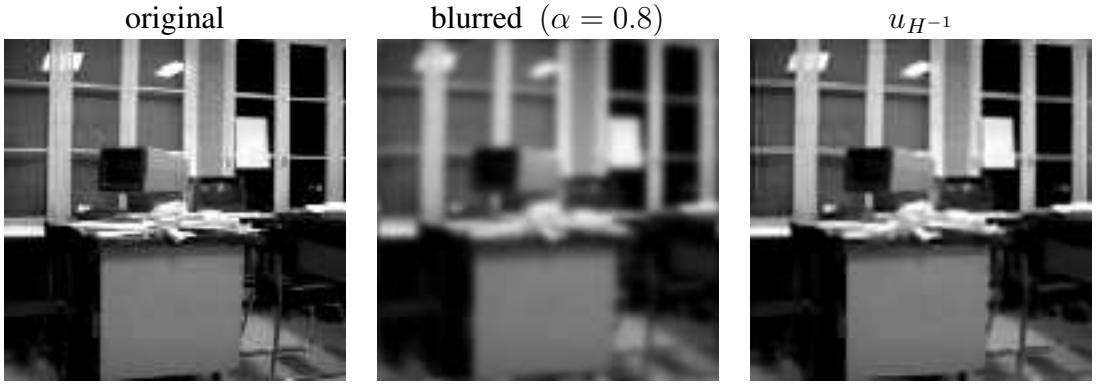


Figure 2.9: Deblurring on image of an office. Result is obtained from our model with H^{-1} ($\lambda = 0.0004$).

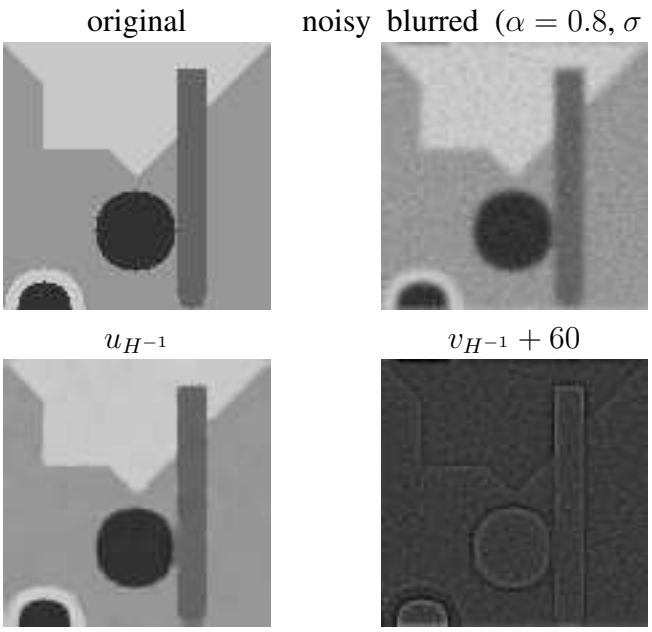


Figure 2.10: Denoising-deblurring result using our model with H^{-1} ($\lambda = 0.0675$).

Chapter 3

(Φ, Φ^*) Decomposition Model and Minimization Algorithms

In this chapter, we consider a decomposition model that imposes the standard regularization penalty $\Phi(u) = \int \phi(|\nabla u|)dx < \infty$ on u, where ϕ is positive, increasing and has at most linear growth at infinity. However, on the residual $f - Ku$, the model imposes the dual penalty $\Phi^*(f - Ku) < \infty$, instead of the usual $\|f - Ku\|_{L^2}^2$ fidelity term. In particular, when ϕ is convex, homogeneous of degree one, and with linear growth (for instance the total variation of u), we recover the (BV, BV^*) decomposition of the data f, as in Y. Meyer [46].

We present practical minimization algorithms, together with experimental results and comparisons to illustrate the validity of the proposed models.

The work in this chapter is jointly done with Triet Le. I thank him very much for his generous contribution.

We begin with some basic assumptions and motivational background:

Let Ω be an open and bounded subset of R^2, with $\partial\Omega$ Lipschitz. For two dimensional

images, Ω is in general a rectangle in the plane. Again, we assume the linear degradation model $f = Ku + n$, where $f, u : \Omega \to R$ are the degraded and the clean unknown images respectively, $K : L^2(\Omega) \to L^2(\Omega)$ is a linear and continuous operator, and n represents additive noise of zero mean. Recall our discussion in Chapter I that the problem of recovery of the unknown image u, given f and given this degradation model, is known to be an ill-posed problem. Therefore, regularization techniques as a-priori smoothness on the unknown u are usually imposed in a minimization approach, of the form

$$\inf_u E(u) = R(u) + \lambda F(f - Ku), \tag{3.1}$$

where the first term acts as a regularization term (usually depending on spatial derivatives of the unknown u), $F(f-Ku)$ acts as a fidelity term, and $\lambda \geq 0$ is a tunning parameter. The behavior of functionals R and F is chosen function of the a-priori smoothness assumptions on u and function of the statistics of the noise n. The standard case is when $R(u)$ depends on the gradient ∇u of u and on its discontinuity set S_u, and if additive Gaussian noise of zero mean is observed, then $F(u) = \|f - Ku\|_2^2$. These cases include the Perona-Malik model [53] ($\lambda = 0$), the Mumford and Shah model for image segmentation and piecewise-smooth regularization, the models of D. Geman, S. Geman and collaborators [33, 34, 35, 36] in the non-convex regularization, the total variation minimization of Rudin, Osher and Fatemi [57, 56]. Other related models and analysis in variational approach are by Acar-Vogel [1], A. Chambolle-P.L. Lions [17], Aubert-Vese [12], Vese [64]. In the PDE approach, we mention again the Perona-Malik equation [53], as well as the anisotropic smoothing [16, 9].

More recently, D. Mumford - B. Gidas [49] and Y. Meyer [46] advocated the use of generalized functions as distributions in dual spaces for modelling images with oscillations, such as natural images, noise, texture, oscillatory patterns; thus proposing spaces such as

$H^{-s}(\Omega)$ [49] and spaces that approximate the dual $BV^*(\Omega)$ of the space $BV(\Omega)$ [46]. Such oscillatory images are better modeled if weaker (dual) norms are considered as penalty or assumption, instead of the $\|\cdot\|_2^2$ fidelity penalty.

Here, we follow the approach suggested by Y. Meyer of using duality to obtain weaker norms to represent the oscillatory component $v = f - Ku$. When analyzing the Rudin-Osher-Fatemi model [57] in the book manuscripts by Y. Meyer [46] and Andreu-Vaillo, Caselles and Mazón [8], the dual functional $\Phi^*(v)$ of the total variation $\Phi(u) = |Du|(\Omega)$ already appears in the characterization of minimizers. We propose in this paper minimization models of the form

$$\inf_u \Phi(u) + \lambda\Phi^*(f - Ku).$$

We will consider in particular the penalty $u \in BV(\Omega)$, or more generally of the form $\phi(|Du|)(\Omega) < \infty$, with ϕ convex and of linear growth ($\phi(t) = |t|$, $\phi(t) = \sqrt{1+t^2}$, $\phi(t) = \log\cosh(|t|)$), as well as non-convex potentials. In the convex case these include the total variation minimization proposed by L. Rudin, S. Osher, and E. Fatemi [57]. Related recent work is proposed by J.-F. Aujol and A. Chambolle [15], and by S. Levine [41], where the authors use the duality given by the Legendre-Fenchel transform to solve cartoon and texture decomposition models. However, our approach proposed here is different.

In Vese-Osher [65, 66], Osher-Sole-Vese [52], Aujol and collaborators [13, 14], approximations to the (BV, BV^*) model of Y. Meyer have been previously proposed.

The use of the dual norm of the total variation has also appeared, independently and contemporaneously, in S. Kindermann, S. Osher, and J. Xu's work [38] in a different framework.

3.1 Description of the Model and Properties

Assume we have a normed space E and E^* its dual space. Let $\Phi : E \to [0, \infty]$ be any function. Let us define $\Phi^* : E^* \to [0, \infty]$, by

$$\Phi^*(v) = \sup \Big\{ \frac{\langle v, u \rangle}{\Phi(u)} : \ u \in E \},$$

with the convention that $\frac{0}{0} = 0$, $\frac{0}{\infty} = 0$. Here, $\langle v, u \rangle = v(u)$ denotes the duality pairing.

Note that $\Phi^*(v) \geq 0$, for any $v \in E^*$. Note also that the supremum is attained on the set of $u \in E$ such that $\langle v, u \rangle \geq 0$. Note also that we have the following Cauchy-Schwartz inequality

$$\langle v, u \rangle \leq \Phi^*(v)\Phi(u), \quad \text{if } \Phi(u) > 0.$$

We are interested in image decomposition models of the form

$$\inf \Big\{ \Phi(u) + \lambda \Phi^*(v), \ u \in E, \ v \in E^*, \ f = u + v \Big\} \tag{3.2}$$

where $f \in E^*$ is a given data. The component u models the cartoon part of f, while the component v models the rough, oscillatory, or noise part of f. We will choose Φ so that if $\Phi(u) < \infty$, then u is a piecewise-smooth function, with homogeneous regions and sharp boundaries.

3.2 (Φ, Φ^*) Decomposition Model

Consider the following general decomposition model:

$$\inf_u \Big\{ \int_\Omega \phi(|\nabla u|) \, dx + \lambda \sup_{w, \int_\Omega \phi(|\nabla w|) \neq 0} \frac{\int_\Omega (f - u) w \, dx}{\int_\Omega \phi(|\nabla w|) \, dx} \Big\}. \tag{3.3}$$

In the form of (3.2), $\Phi(u) = \int_\Omega \phi(|\nabla u|)\ dx$. We assume that ϕ is a differentiable, increasing function on $[0,\infty)$, possibly non-convex, with at most linear growth at infinity, and satisfying $\phi(0) > 0$ again to avoid division by zero. In the convex case, $\phi(|Du|)$ is well defined for $u \in BV(\Omega)$, convex and lower semi-continuous, as a convex function of measures (see Demengel-Temam [26]). In the non-convex case, we work with $u \in W^{1,1}(\Omega) \subset BV(\Omega) \subset L^2(\Omega)$ and the distributional gradient ∇u, as a function in $(L^1(\Omega))^2$.

In a time-dependent approach, the solution of (3.3) is obtained as follows: start with $(u^0, w^0) = (u(0,x), w(0,x))$, and solve for $t > 0$ the following evolutionary coupled system in the unknowns u, w

$$\frac{\partial w}{\partial t} = f - u + \frac{\int_\Omega (f-u)w dx}{\int_\Omega \phi(|\nabla w|)dx} \mathrm{div}\Big(\phi'(|\nabla w|)\frac{\nabla w}{|\nabla w|}\Big), \tag{3.4}$$

$$\frac{\partial u}{\partial t} = w + \frac{|w|_{BV(\Omega)}}{\lambda} \mathrm{div}\Big(\phi'(|\nabla u|)\frac{\nabla u}{|\nabla u|}\Big). \tag{3.5}$$

Details of this minimization algorithm is given in Section 3.3.1 for the particular case when $\Phi(u) = \int_\Omega |\nabla u| dx$.

3.3 (BV, BV^*) Image Decomposition

Recall that $u \in L^1(\Omega)$ has bounded variation in Ω if

$$\int_\Omega |Du| \stackrel{\text{def}}{=} \sup\Big\{ \int_\Omega u \mathrm{div}\phi dx : \ \phi \in C_c^1(\Omega, R^2),\ |\phi| \le 1 \Big\} < \infty.$$

The total variation of u in Ω, denoted $|u|_{BV(\Omega)} := \int_\Omega |Du|$, is a seminorm on the space $BV(\Omega)$ and vanishes only for constant functions. Therefore it is a norm on the quotient space $BV(\Omega)/P_0(\Omega) \stackrel{\text{def}}{=} \mathcal{BV}(\Omega)$ (i.e. functions in $BV(\Omega)$ that are different only by a constant are identified).

Recall also the Poincaré-Wirtinger inequality: in two dimensions, for $u \in BV(\Omega)$,

$$\|u - u_\Omega\|_2 \leq C \int_\Omega |Du|,$$

where $u_\Omega = \frac{\int_\Omega u(x)dx}{|\Omega|}$ denotes the mean of u in Ω.

Now, let us consider the particular case $E = L^2(\Omega)$, $E^* = E = L^2(\Omega)$, $\langle v, u \rangle = \int_\Omega uvdx$, for $u, v \in L^2(\Omega)$, and

$$\Phi(u) = \begin{cases} \int_\Omega |Du| \text{ if } u \in BV(\Omega), \\ +\infty \text{ if } u \in L^2(\Omega) \setminus BV(\Omega). \end{cases} \tag{3.6}$$

Let $X = \{v \in L^2(\Omega) : \int_\Omega v(x)dx = 0\}$. For $v \in X$, define

$$\|v\|_* := \Phi^*(v) = \sup_{\substack{w \in BV(\Omega) \\ |w|_{BV(\Omega)} \neq 0}} \frac{\int_\Omega vwdx}{|w|_{BV(\Omega)}} = \sup_{\substack{w \in BV(\Omega) \\ |w|_{BV(\Omega)} \neq 0}} \frac{|\int_\Omega vwdx|}{|w|_{BV(\Omega)}}. \tag{3.7}$$

Note that when $|w|_{BV(\Omega)} \neq 0$ and $v \in X$, then $\frac{|\int_\Omega vwdx|}{|w|_{BV(\Omega)}} = \frac{|\int_\Omega v(w+c)dx|}{|w+c|_{BV(\Omega)}}$, for any real constant c. Therefore, the supremum in the definition of $\|v\|_*$ can be computed over the quotient space $\mathcal{BV}(\Omega)$.

Note also that, if $v \in L^2(\Omega)$ and $\|v\|_* < \infty$, then $v \in X$. In other words v necessarily has zero mean. To see this, take any such v and any $w \in BV(\Omega)$ with $|w|_{BV(\Omega)} \neq 0$. Then for any constant c,

$$\frac{\int_\Omega v(w+c)dx}{|w+c|_{BV(\Omega)}} = \frac{\int_\Omega v(w+c)dx}{|w|_{BV(\Omega)}} = \frac{\int_\Omega vwdx + c\int_\Omega vdx}{|w|_{BV(\Omega)}} \leq \|v\|_* < \infty.$$

This can hold if and only if $\int_\Omega v(x)dx = 0$, therefore $v \in X$.

We make the following observations:

1. $||\cdot||_*$ is a norm on X. Furthermore, X is a complete space under this norm. This is clear, since if $v_n \in X$ is a sequence that converges to some v in this $||\cdot||_*$ norm, then for some index N, $||v_N - v||_* < \infty$ implying $v_N - v \in X$, so $v \in X$.

2. For any $v \in X$ and $w \in \mathcal{BV}$, $\left|\int_\Omega vwdx\right| \leq ||v||_*|w|_{BV}$.

3. Φ is a norm on $\mathcal{BV}$ and Φ^* defined via duality is a norm on X. We note that X is only a closed subspace of $F_v \in \mathcal{BV}^*(\Omega)$, the dual of $(\mathcal{BV}(\Omega), |\cdot|_{BV(\Omega)})$.

Let us provide more details to item 3. Denote by $w_\Omega := \frac{\int_\Omega w(x)dx}{|\Omega|}$ the mean on Ω of a function $w \in L^2(\Omega)$. To any $v \in X$, we associate a linear functional $F_v : \mathcal{BV}(\Omega) \to R$ defined by $F_v(w) = \int_\Omega vwdx$. If $v \in X$, then

$$|F_v(w)| = |\int_\Omega vwdx| =\leq \|v\|_*|w|_{BV(\Omega)} = C|w|_{BV(\Omega)}. \tag{3.8}$$

Therefore, F_v is linear and bounded (hence continuous) functional on $\mathcal{BV}(\Omega)$, i.e. $F_v \in \mathcal{BV}^*(\Omega)$. Note that not all elements of $\mathcal{BV}^*(\Omega)$ can be expressed in this way (see T. De Pauw [25] for a characterization of the dual of $SBV(\Omega)$, of special functions of bounded variation).

Finally, since $|F_v(w)| \leq \|v\|_*|w|_{BV(\Omega)}$ for all $w \in \mathcal{BV}(\Omega)$, we have $\|F_v\|_{\mathcal{BV}^*(\Omega)} = \|v\|_*$, recertifying $\|\cdot\|_*$ is a norm on a subspace X of $\mathcal{BV}^*$.

Recall the notation $X(\Omega)_p \stackrel{\text{def}}{=} \{\vec{p} \in L^\infty(\Omega)^2, \text{div}(\vec{p}) \in L^p(\Omega)\}$. It is shown in [8], Chapter 1, that

$$\Phi^*(v) = \inf\{\|\vec{g}\|_\infty : \vec{g} \in X(\Omega)_2,\ v = -\text{div}(\vec{g}) \text{ in } \mathcal{D}'(\Omega),\ \vec{g}\cdot\nu = 0\},$$

where ν denotes the outward unit normal to $\partial\Omega$ and $\vec{g}\cdot\nu$ is the trace of the normal component of g. Recall from definition 1.3.6 the space $G^A(\Omega)$ and the norm $||\cdot||_{G^A}$ on

$G^A(\Omega)$. It is shown in [10] that $G^A(\Omega) = X = \{v \in L^2(\Omega), \int_\Omega v dx = 0\}$. Thus, $||\cdot||_*$ is a different way to define the norm $||\cdot||_{G^A}$ on the space X, and we have the identity $(G^A(\Omega), ||\cdot||_{G^A}) = (X, ||\cdot||_*)$.

Recall also the space $G^V(\Omega)$ in definition 1.3.5. (This space is isometrically isomorphic to $(W_0^{1,1}(\Omega))^* = W^{-1,\infty}(\Omega)$ under the norm $||\cdot||_{W^{-1,\infty}(\Omega)}$ dual to $|\cdot|_{BV}$). The norm on $G^V(\Omega)$ is defined as $||v||_{G^V} = \inf\{|||\vec{g}|||_\infty, v = \operatorname{div}(\vec{g}), \vec{g} \in L^\infty(\Omega)^2\}$. Now, if $v \in X$, then for any $\vec{g} \in L^\infty(\Omega)^2$ with $v = \operatorname{div}(\vec{g})$, $0 = \int_\Omega v dx = \int_\Omega \operatorname{div}(\vec{g}) = \int_{\partial\Omega} \vec{g} \cdot \nu ds$. This implies $\vec{g} \cdot \nu = 0$. Therefore, based on the above remarks, we conclude that for any $v \in X$, $||v||_* = ||v||_{W^{-1,\infty}(\Omega)}$. Thus, $(X, ||\cdot||_*)$ is a closed subspace of $W^{-1,\infty}(\Omega)$ under the same norm as the ambience space.

Application to image decomposition:

Since at every pixel the light intensity has finite energy, we can assume that an image f belongs to $L^\infty(\Omega) \subset L^2(\Omega)$ (since Ω is bounded). Indeed, it is not too restrictive to assume that $f \in L^2(\Omega)$.

We are interested in decomposing f into $u + v$, with $u \in BV(\Omega)$ and $v := f - u \in X$. The first term u corresponds to a cartoon component, while v corresponds to an additive oscillatory component of zero mean (such as additive noise or texture).

One of the minimization model that we would like to solve is, given $f \in L^2(\Omega)$,

$$\inf\{E(u, v) = |u|_{BV(\Omega)} + \lambda\|v\|_*,\ f = u + v,\ u \in BV(\Omega),\ v \in X\},$$

which is equivalent with

$$\inf_{u \in BV(\Omega)} E(u) = |u|_{BV(\Omega)} + \lambda\|f - u\|_*, \tag{3.9}$$

where

$$\|v\|_* = \sup_{\substack{w\in BV(\Omega)\\ |w|_{BV(\Omega)}\neq 0}} \frac{|\int vwdx|}{|w|_{BV(\Omega)}} = \sup_{\substack{w\in BV(\Omega)\\ |w|_{BV(\Omega)}\neq 0}} \frac{\int vwdx}{|w|_{BV(\Omega)}}.$$

Remark 3.3.1 *We can assume without loss of generality that the data $f \in L^2(\Omega)$ has zero mean, i.e. $\int_\Omega f(x)dx = 0$. Then $f \in X$ and $\|f\|_* < \infty$. Therefore, if we take $u(x) = 0$ for all $x \in \Omega$ in (3.9), then $E(u) = \lambda\|f\|_* < \infty$. This shows that the functional in (3.9) has a finite infimum. Moreover, the functional $\Phi^*(v) = \|v\|_*$ defined in (3.7) for $v \in X$ is convex, lower-semicontinuous and positive homogeneous of degree one (see [8], Chapter 1).*

We do have existence of minimizers for the problem (3.9). Indeed, let u_n be a minimizing sequence such that $E(u_n) \leq E(0) < \infty$. Therefore $u_n \in BV(\Omega) \in L^2(\Omega)$ and $|u_n|_{BV(\Omega)} \leq C$, for all $n \geq 0$. Also, $\|f - u_n\|_* \leq C < \infty$. This implies that $f - u_n \in X$, i.e. $f - u_n$ has zero mean. Applying Poincaré inequality, we deduce that u_n is uniformly bounded in $BV(\Omega)$ and in $L^2(\Omega)$. We then obtain existence of a subsequence, still denoted by u_n, and $u \in BV(\Omega)$, such that u_n converges to u strongly in $L^1(\Omega)$, weekly in $L^2(\Omega)$ and weekly in $BV - w^*(\Omega)$. In conclusion, by the lower semi-continuity of $|\cdot|_{BV(\Omega)}$ and $\|\cdot\|_*$, we deduce that

$$E(u) \leq \liminf_{n\to\infty} E(u_n) = \inf_{w\in BV(\Omega)} E(w),$$

in other words, u is a minimizer. It is posisble to show that there is no uniqueness of minimizers for this model (Y. Meyer [47]).

3.3.1 (BV, BV^*) Minimization Algorithm

We now give an algorithm for solving the minimization problem 3.9. We shall use the notation $|u|_{BV} := \int_\Omega |Du|dx$ for the total variation of u.

- Start with u^0.

• For integers $n \geq 0$, if u^n is known or previously computed, estimate w^n by the maximization process:

$$\|f - u^n\|_* \approx \sup_{\substack{w \in BV(\Omega) \\ |w|_{BV(\Omega)} \neq 0}} \frac{\int (f - u^n) w dx}{\int \sqrt{\epsilon^2 + |\nabla w|^2} dx}.$$

Remark: In the calculation of the supremum, with ϵ-regularization of the total variation, $|w|_{BV}$ will never become zero, because if it does, then w =constant, which makes the quantity $\frac{\int v w dx}{\int \sqrt{\epsilon^2 + |\nabla w|^2} dx} = 0$, hence cannot be the supremum, unless $f - u = 0$.

The associated PDE in $w = w^n$, (to obtain w^n), formally is

$$0 = \frac{f - u^n}{\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx} + \frac{\int_\Omega (f - u^n) w dx}{(\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx)^2} \mathrm{div}\Big(\frac{\nabla w}{|\nabla w|}\Big), \tag{3.10}$$

or

$$0 = (f - u^n) + \frac{\int_\Omega (f - u^n) w dx}{\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx} \mathrm{div}\Big(\frac{\nabla w}{|\nabla w|}\Big), \tag{3.11}$$

with associated natural boundary condition $\frac{\partial w}{\partial \nu} = 0$ on $\partial\Omega$.

• Once w^n is computed, we compute u^{n+1} by minimizing with respect to $u = u^{n+1}$ the energy

$$E(u) = \int_\Omega |\nabla u| dx + \lambda \frac{\int (f - u) w^n dx}{|w^n|_{BV(\Omega)}},$$

which formally yields the associated Euler-Lagrange equation in $u = u^{n+1}$:

$$0 = \frac{\lambda w}{|w|_{BV(\Omega)}} + \mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big), \tag{3.12}$$

or

$$0 = w + \frac{|w|_{BV(\Omega)}}{\lambda} \mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big), \tag{3.13}$$

with $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$.

Remark: Integrating (3.11) in space, we obtain that $\int_\Omega (f - u^n)dx = 0$.

In a time-dependent approach, the main algorithm is summarized as follows: start with $(u^0, w^0) = (u(0,x), w(0,x))$, and solve for $t > 0$

$$\frac{\partial w}{\partial t} = f - u + \frac{\int_\Omega (f-u)w dx}{\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx} \mathrm{div}\Big(\frac{\nabla w}{|\nabla w|}\Big), \tag{3.14}$$

$$\frac{\partial u}{\partial t} = w + \frac{|w|_{BV(\Omega)}}{\lambda} \mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big). \tag{3.15}$$

We note that when $u = u^n$ fixed, in the maximization process in w, if we start with w^0 such that $\int_\Omega (f-u)w^0 dx \geq 0$, then $\int_\Omega (f-u)w dx \geq 0$, because we increase the energy in w with the right choice of $\triangle t$ (in the supremum formula) at any step. If this is a problem, then we can keep the absolute value in the definition of the supremum, to avoid numerical instability.

3.4 Application to Image Denoising and Decomposition

We have applied the (Φ, Φ^*) model introduced above to image denoising and also to cartoon + texture separation. In addition to the convex case where $\Phi(u) = \int_\Omega |Du|$, have also considered the case when Φ is not convex.

3.4.1 (BV, BV^*) denoising and decomposition results

In practice, to solve (3.9) for the (BV, BV^*) model, we use the time dependent equations for gradient ascent in w and gradient descent in u, as given in (3.14)-(3.15).

We will compare the proposed (BV, BV^*) model with the Rudin-Osher-Fatemi (BV, L^2)

model,

$$\inf_{u\in BV(\Omega)} \mathcal{F}(u) = \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega |f-u|^2\, dx. \tag{3.16}$$

Recalling from Section 1.2, a minimizer of (3.16) satisfies the Euler-Lagrange equation,

$$0 = \text{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda(f-u). \tag{3.17}$$

To solve the above stationary equation, we consider the gradient descending method

$$\frac{\partial u}{\partial t} = \frac{1}{\lambda}\text{div}\left(\frac{\nabla u}{|\nabla u|}\right) + (f-u). \tag{3.18}$$

We use explicit central schemes for all the partial differential equations, with the divergence operator in expanded form:

$$\text{div}\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{u_{xx}u_y^2 - 2u_{xy}u_xu_y + u_{yy}u_x^2}{(u_x^2+u_y^2)^{3/2}}. \tag{3.19}$$

Our numerical computations use $\Delta x = \Delta y = 1$, and $\Delta t = 0.001$.

We summarize again the main steps of the algorithm:

1. Start with initial guesses for $u_0 = f$, and w_0, such that $\int_\Omega |\nabla w_0| \neq 0$.

2. Given u_n, and w_n , we then compute w_{n+1} using equation (3.14) for 10 iterations.

3. With u_n, and w_{n+1} known, we then compute u_{n+1} with method (3.15).

4. Repeat step $2-3$ until Max_Iter is reached.

In all of our numerical computations, $Max\_Iter = 1000$. In each figure, $\bar{w}$ denotes the rescaling of w to be in the range $[0, 255]$.

Let $\bar{u}$ be the true image of size $N \times M$, and u be the recovered image. To quantify how good the recovered image is, and to choose the parameter λ, we use the root mean square

error

$$rmse = \frac{\sqrt{\sum_{i,j} |\bar{u}_{i,j} - u_{i,j}|^2}}{NM}.$$

The following experimental results and comparisons show improvement over the (BV, L^2) model.

3.4.2 Denoising and Decomposition with non-convex functional Φ and its dual Φ^*

Similarly, the decomposition model (3.3), with $\phi(t)$ non-convex function, can be minimized by the algorithm given in (3.4)-(3.5). Here we will consider two choices (see [33], [34], [35], [36] for non-convex regularizations):

$$\phi(t) = |t|^p,\ 0 < p < 1, \tag{3.20}$$

and

$$\phi(t) = \frac{|t|^q}{1 + \alpha|t|^q},\ 0 < \alpha < 1,\ q \geq 1. \tag{3.21}$$

Through out our numerical computations, we use $p = 0.75$ in (3.20) and $\alpha = 0.001$ and $q = 2$ in (3.21). The following experimental results with a non-convex potential are again much improved over the ROF model, and we no longer see "geometry" in the v component, both in the denoising and decomposition results.

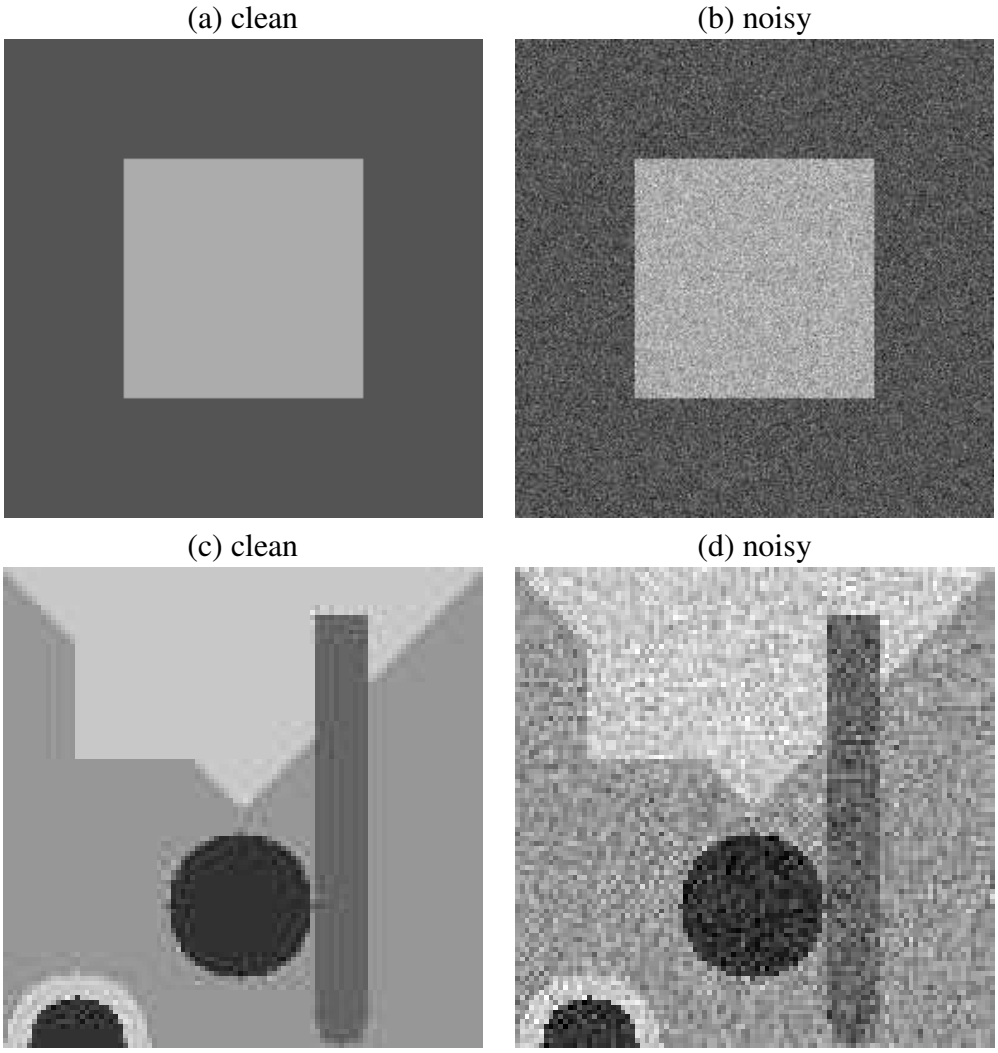


Figure 3.1: Data images.

(a) clean (b) noisy

(c) clean (d) noisy

Figure 3.2: Data images.

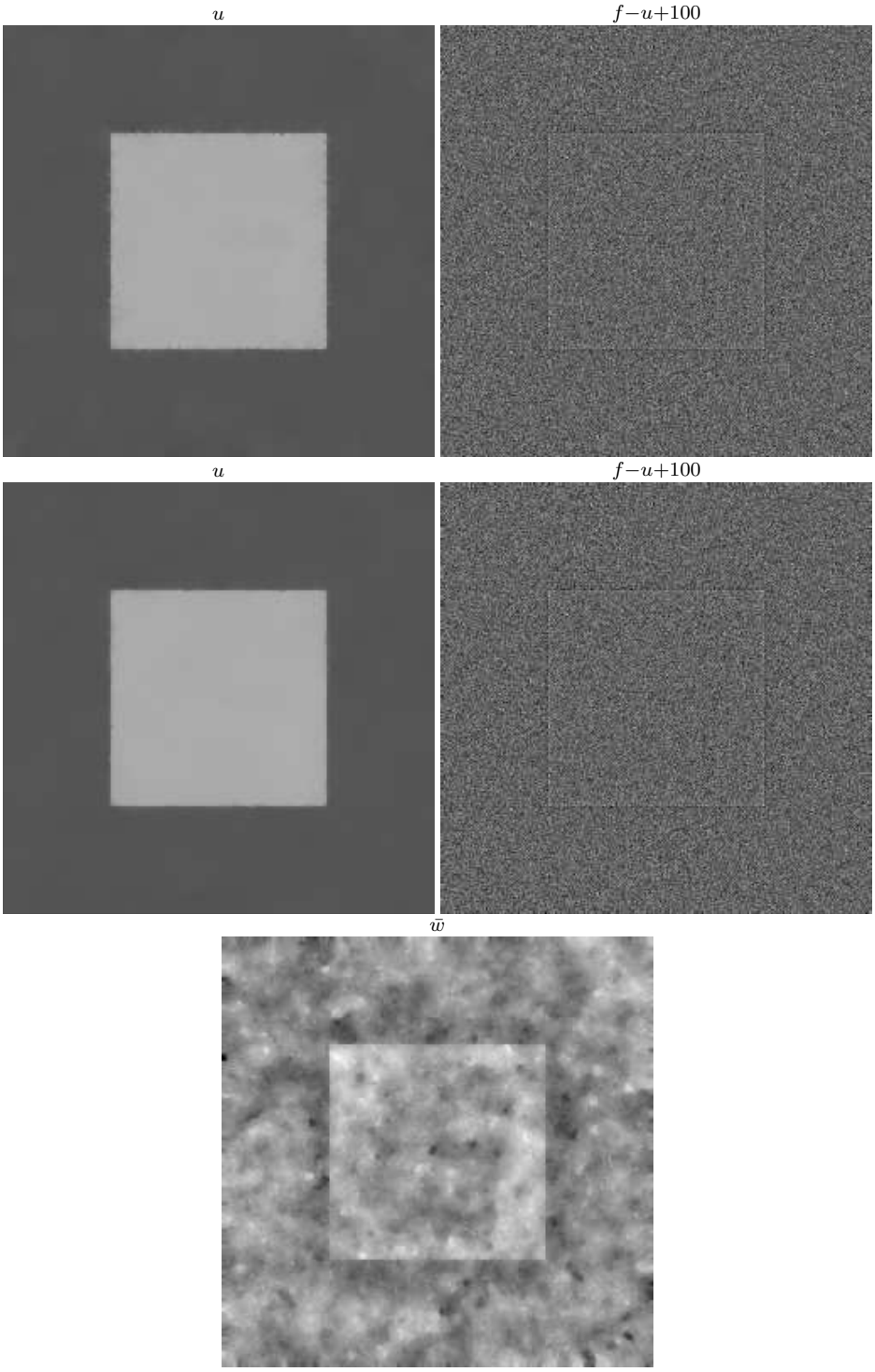


Figure 3.3: Denoising results of our proposed (BV, BV^*) model and ROF model applied to noisy image in Fig. 3.1(b). Top: ROF using (3.18) with $\lambda = 2.04$, $rmse = 0.01020834$. Bottom: (BV, BV^*) using (3.14) and (3.15) with $\lambda = 1900$, $rmse = 0.009918792$.

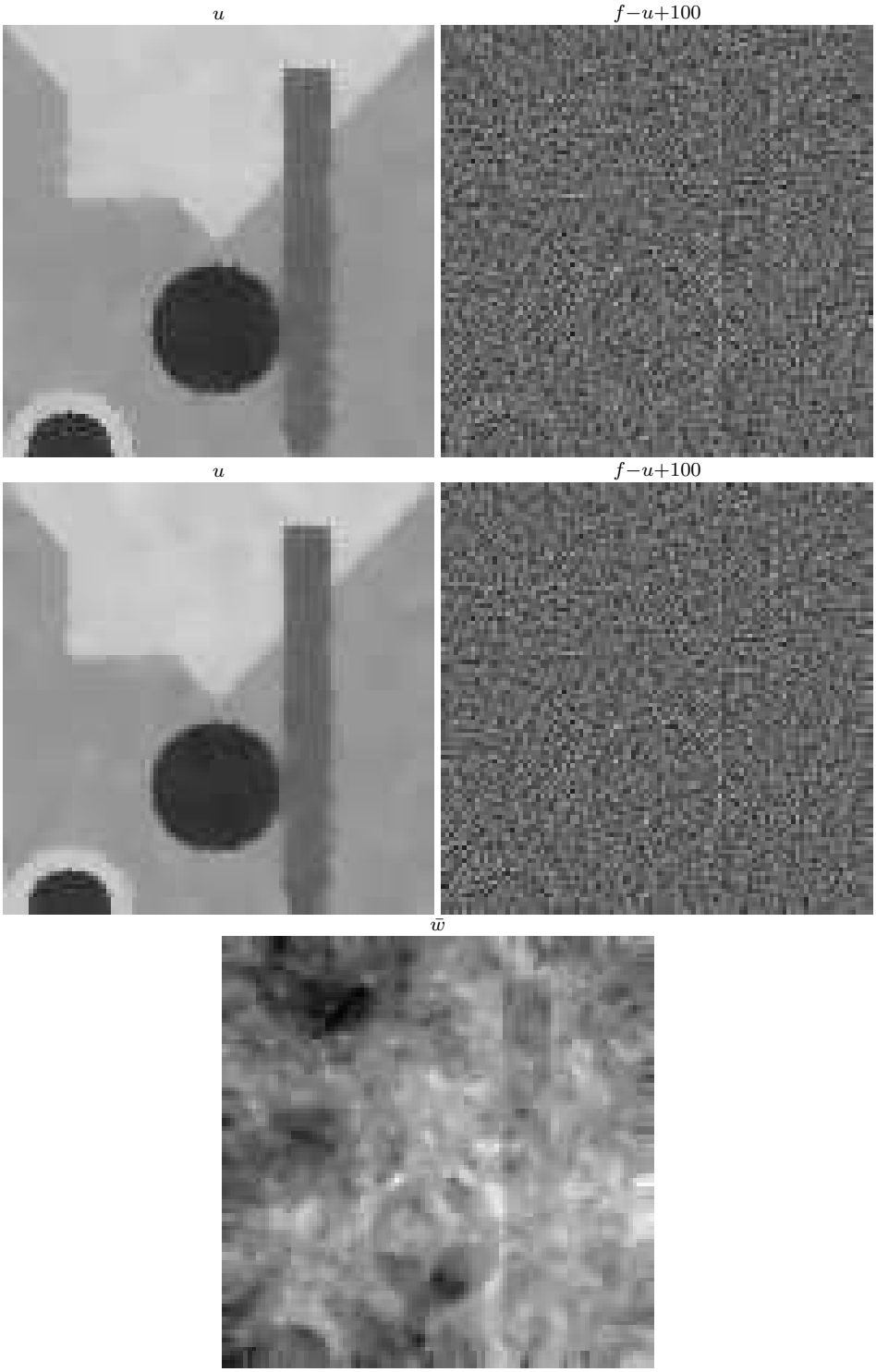


Figure 3.4: Denoising results of our proposed (BV, BV^*) model and ROF model applied to noisy image in Fig. 3.1(d). Top: ROF using (3.18) with $\lambda = 7.92$, $rmse = 0.0645842$. Bottom: (BV, BV^*) using (3.14) and (3.15) with $\lambda = 5000$, $rmse = 0.06406284$.

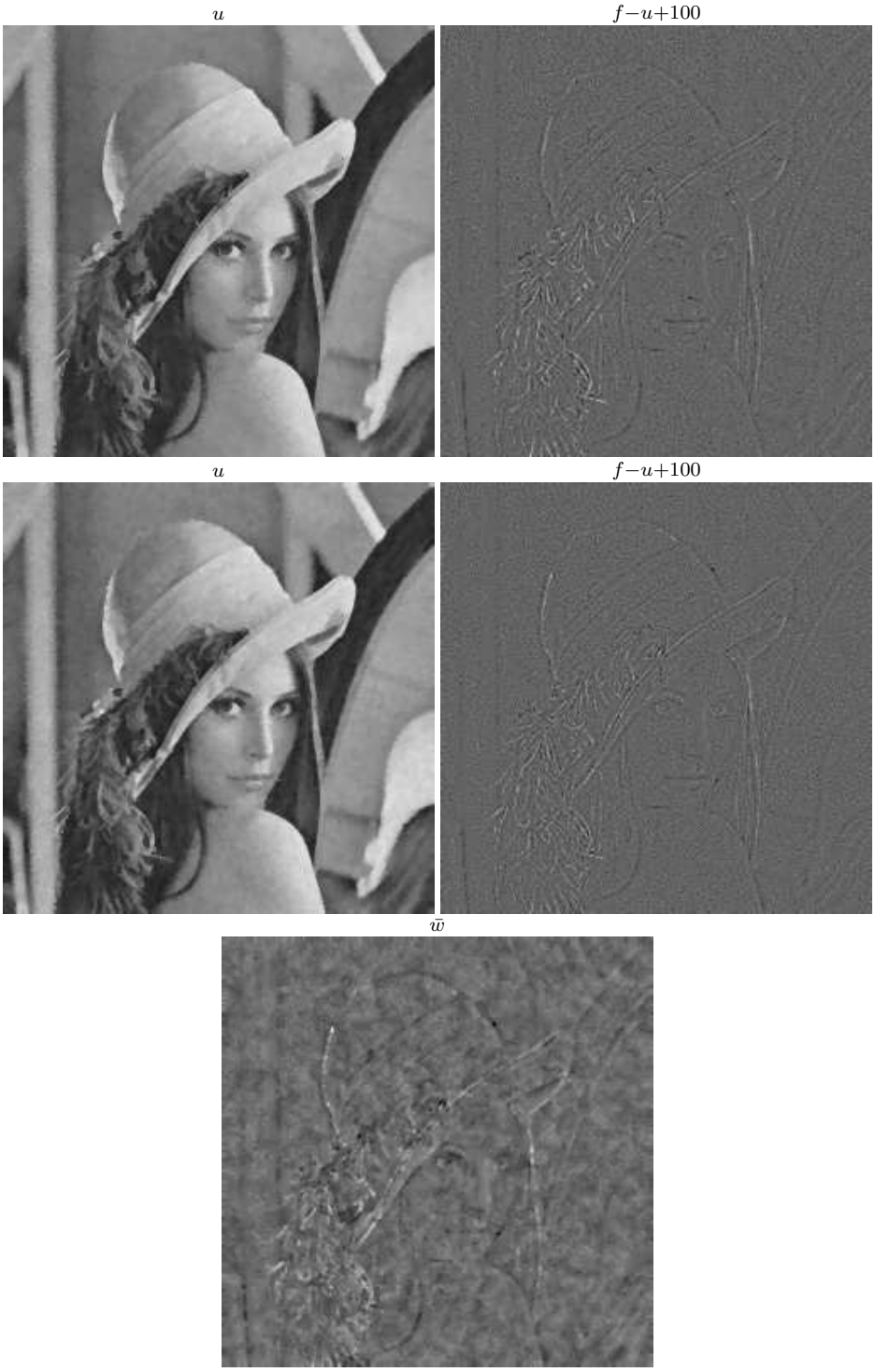


Figure 3.5: Denoising results of our proposed (BV, BV^*) model and ROF model applied to noisy image in Fig. 3.1(b). Top: ROF using (3.18) with $\lambda = 63.75$. Bottom: (BV, BV^*) using (3.14) and (3.15) with $\lambda = 2000$.

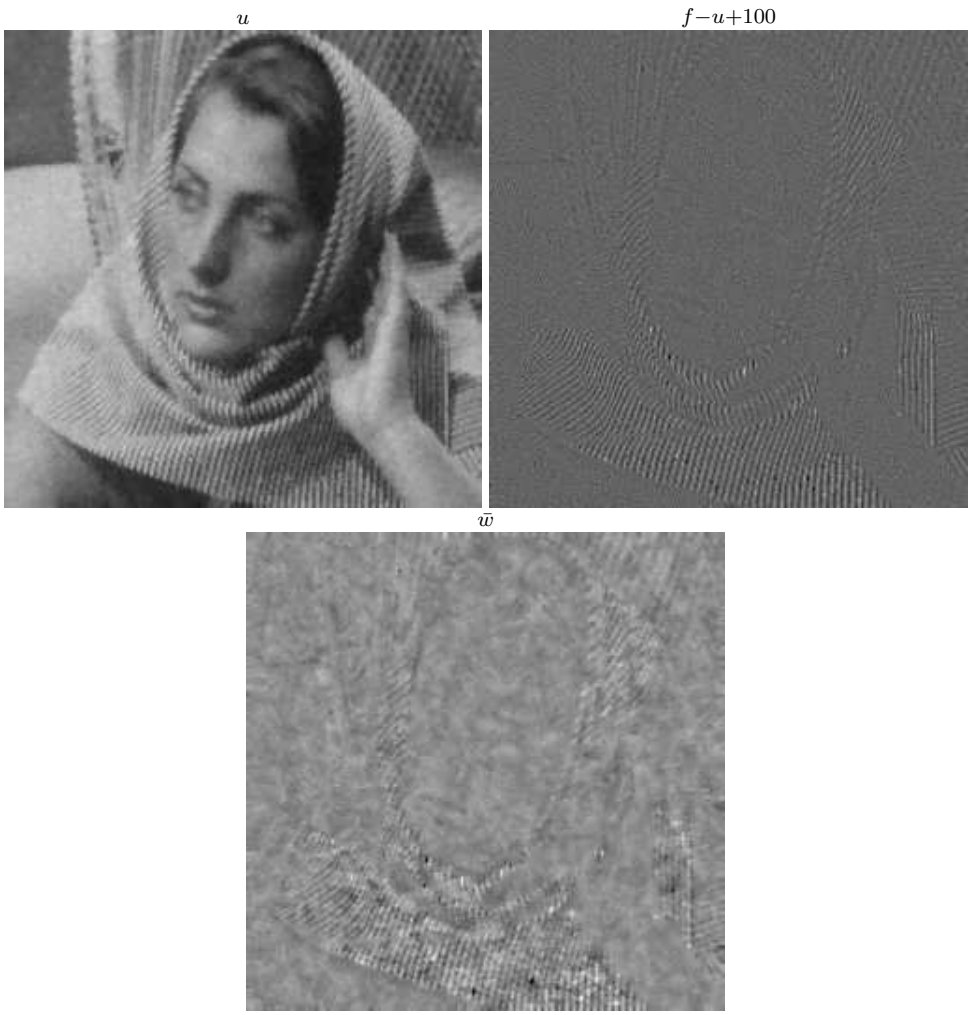


Figure 3.6: Denoising result of (BV, BV^*) model applied to noisy image in Fig. 3.2(d) using (3.14) and (3.15) with $\lambda = 2500$.

Figure 3.7: Decomposition result of our proposed (BV, BV^*) model.

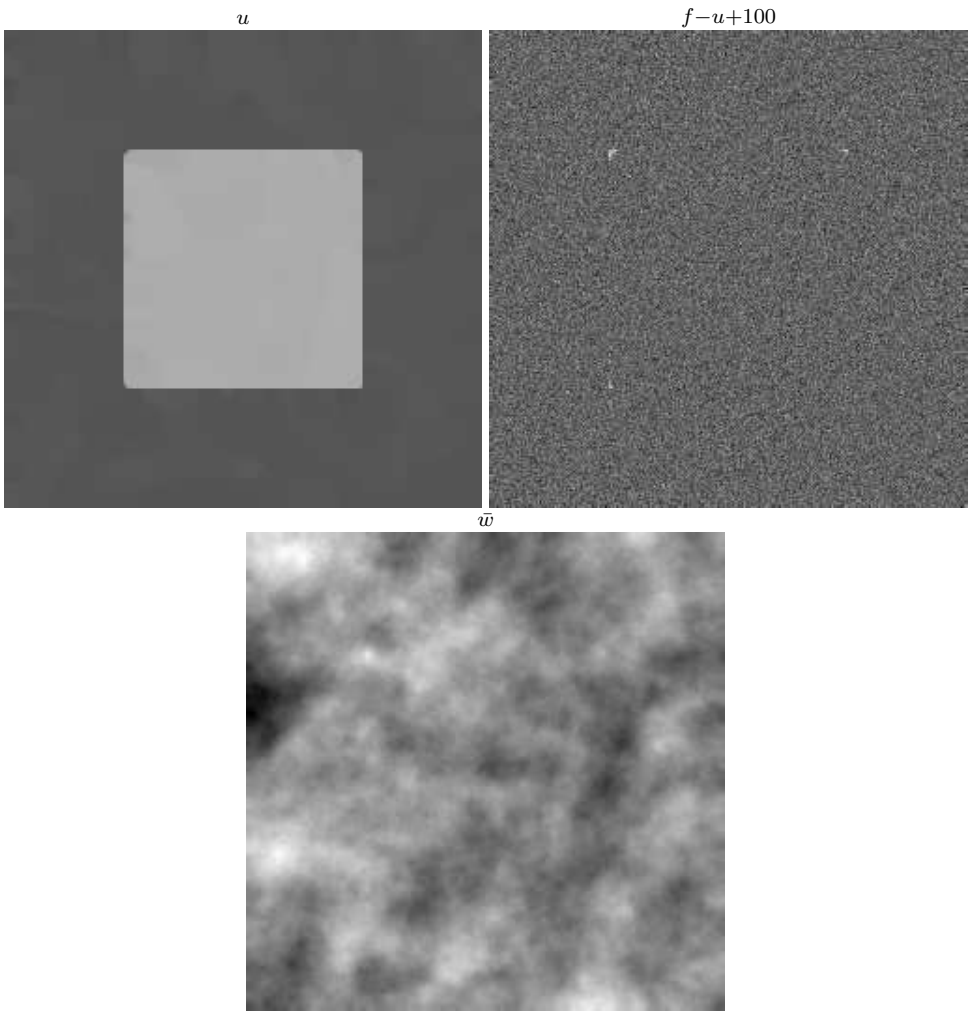


Figure 3.8: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.20) applied to noisy image in Fig. 3.1(b), $rmse = 0.00746253$, $\lambda = 10$.

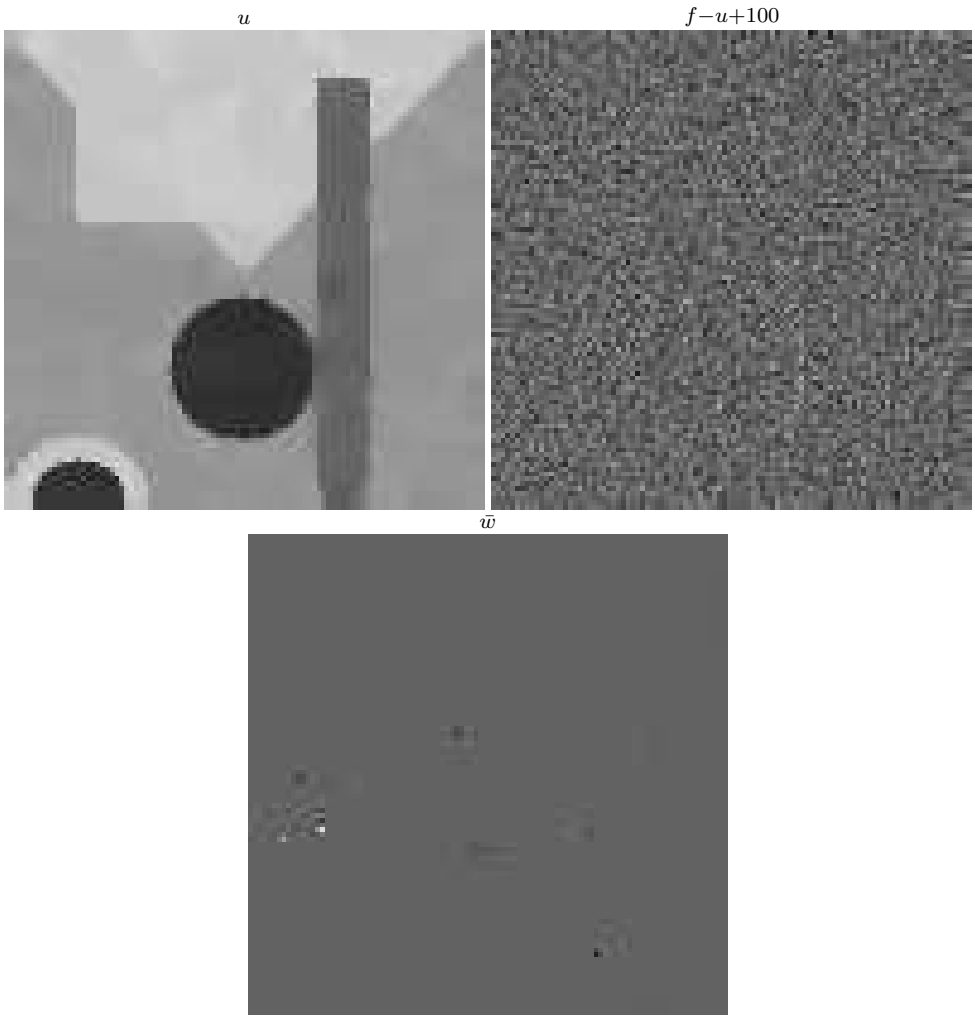


Figure 3.9: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.20) applied to noisy image in Fig. 3.1(d), $rmse = 0.05111486$, $\lambda = 8$.

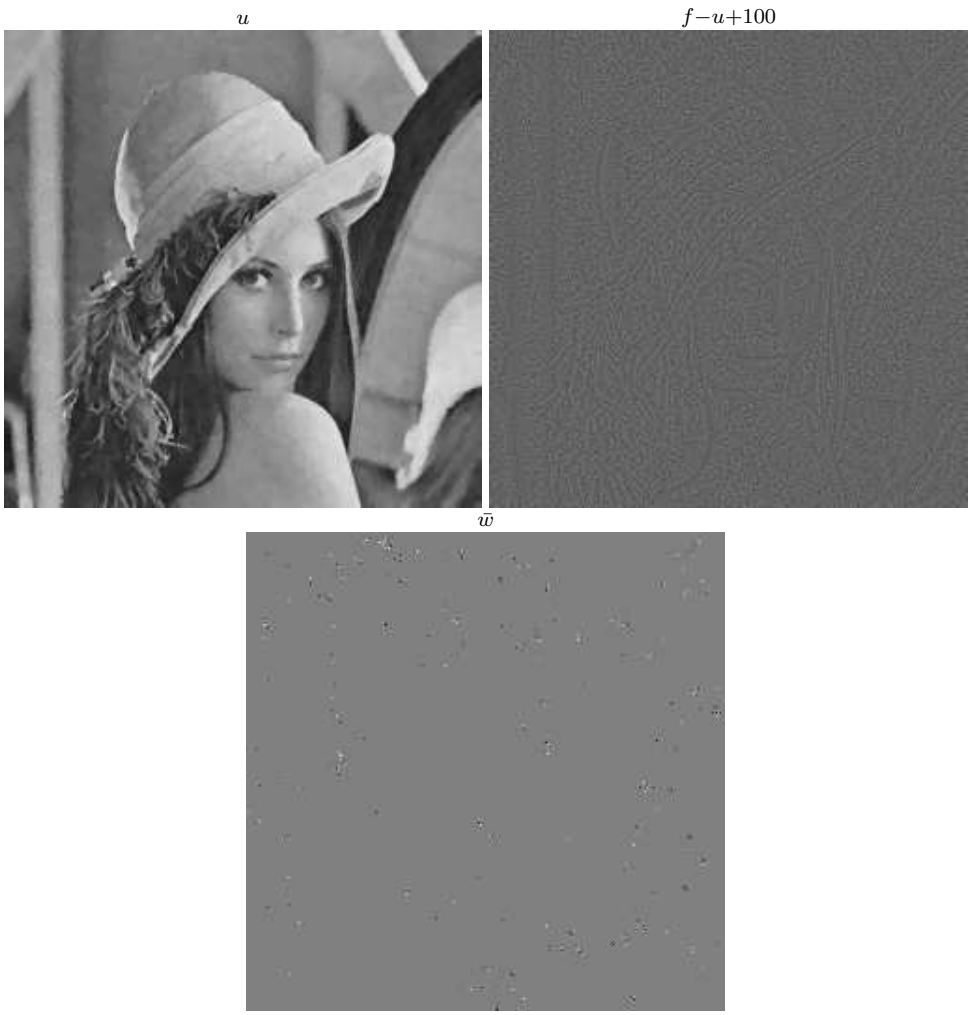


Figure 3.10: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.20) applied to noisy image in Fig. 3.2(b), $rmse = 0.02197276$, $\lambda = 110$.

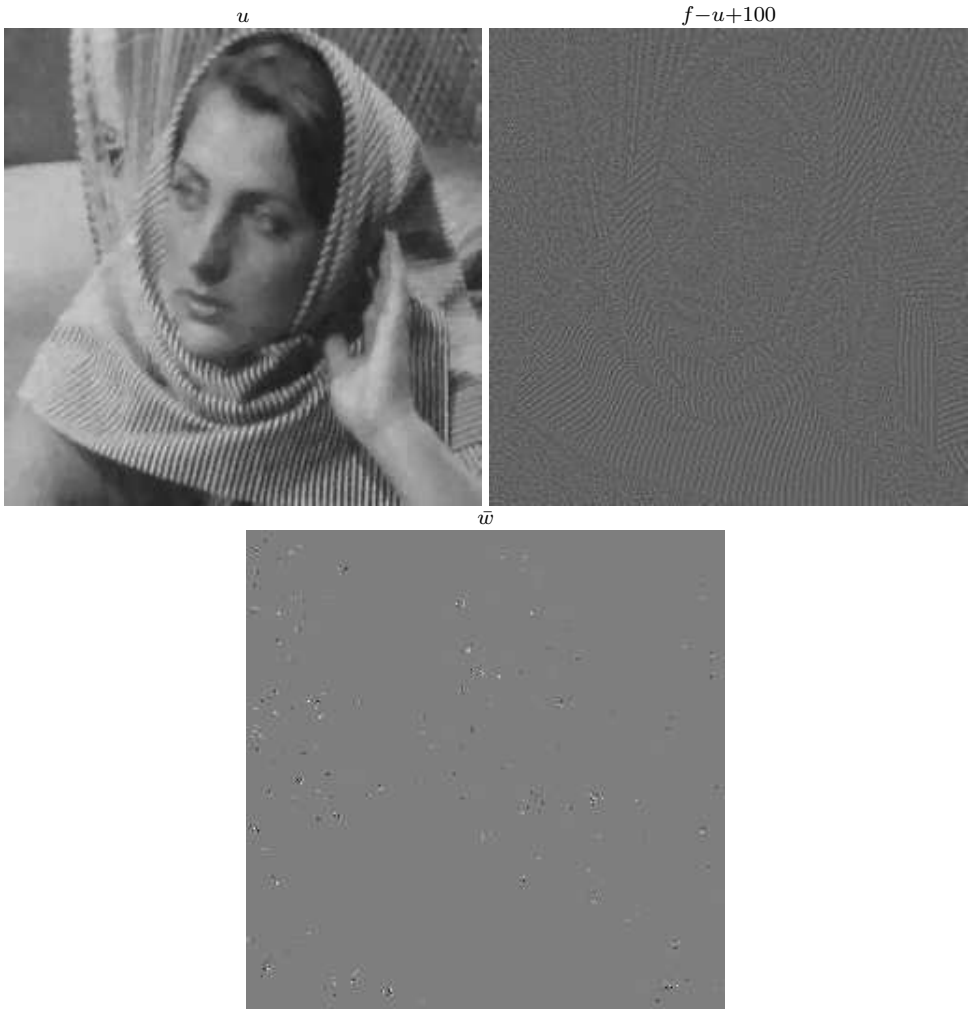


Figure 3.11: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.20) applied to noisy image in Fig. 3.2(d), $rmse = 0.04416518$, $\lambda = 100$.

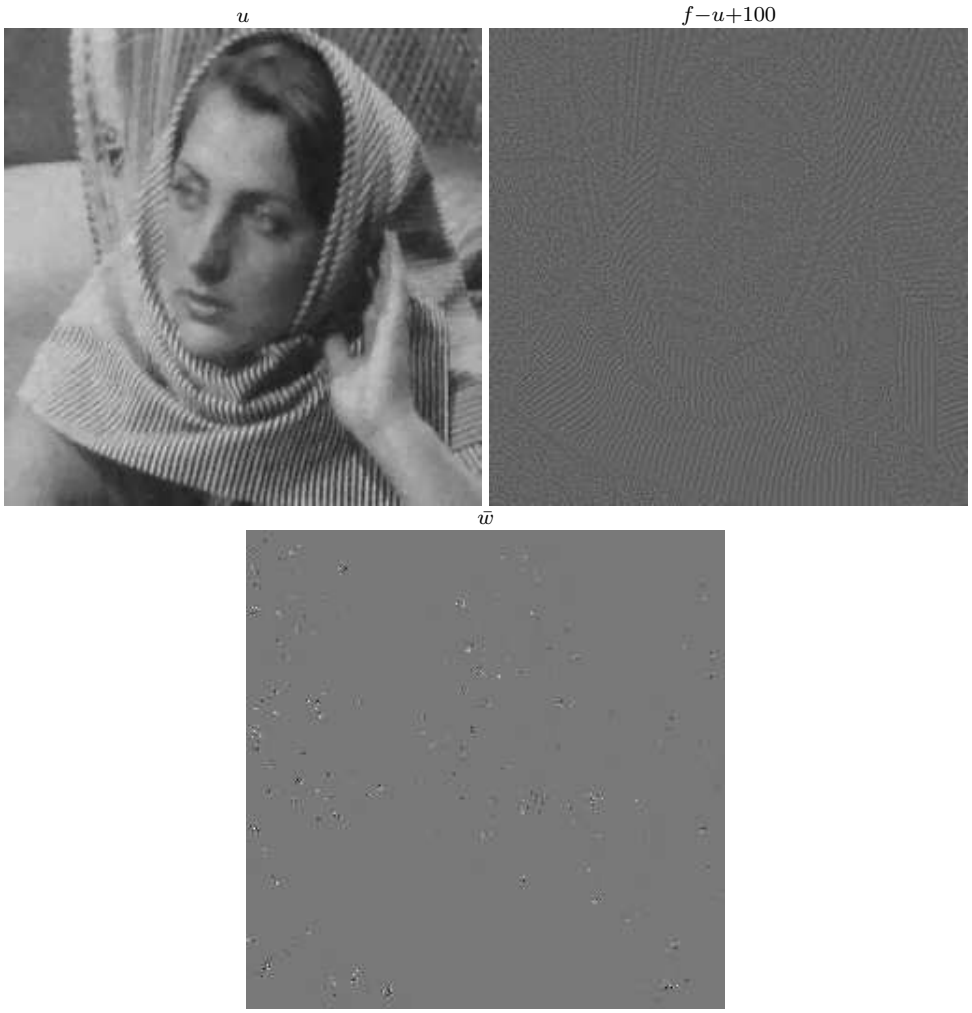


Figure 3.12: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.20) applied to noisy image in Fig. 3.2(d), $rmse = 0.04296311$, $\lambda = 120$.

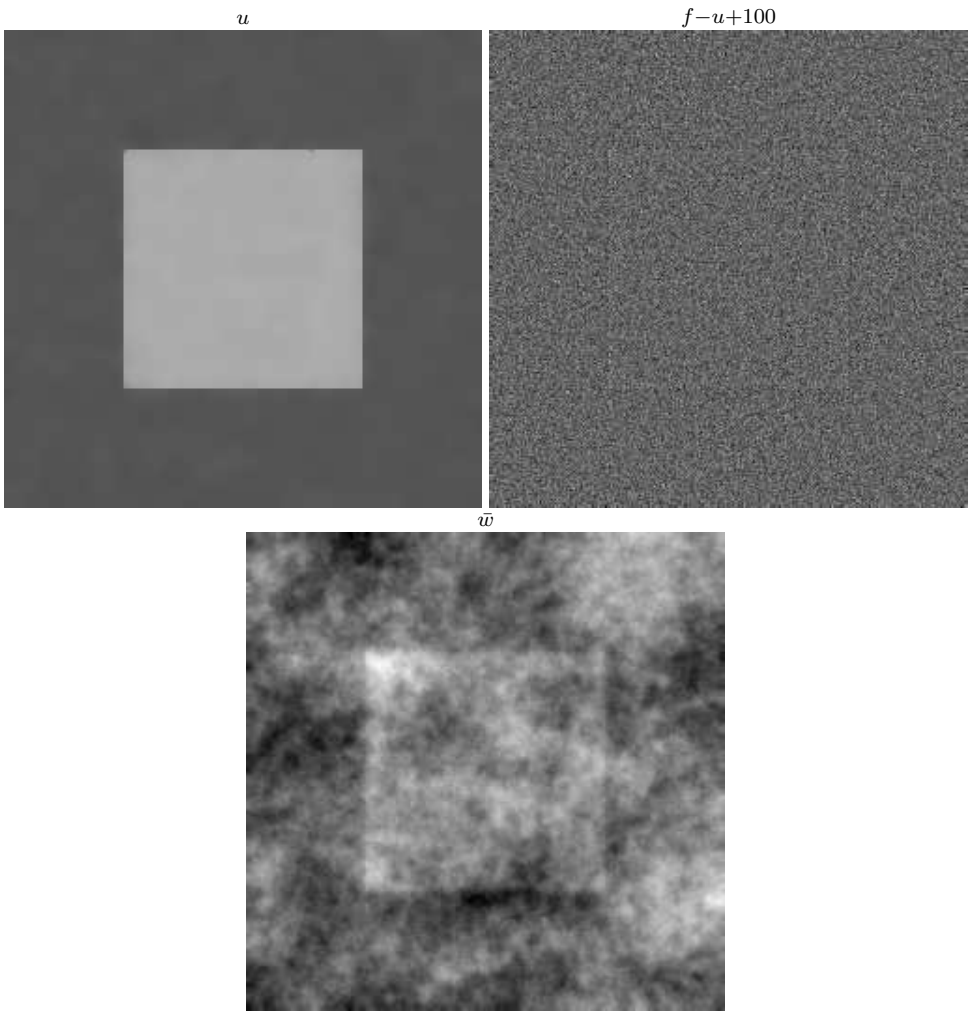


Figure 3.13: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.21) applied to noisy image in Fig. 3.1(b), $rmse = 007092052$, $\lambda = 1$.

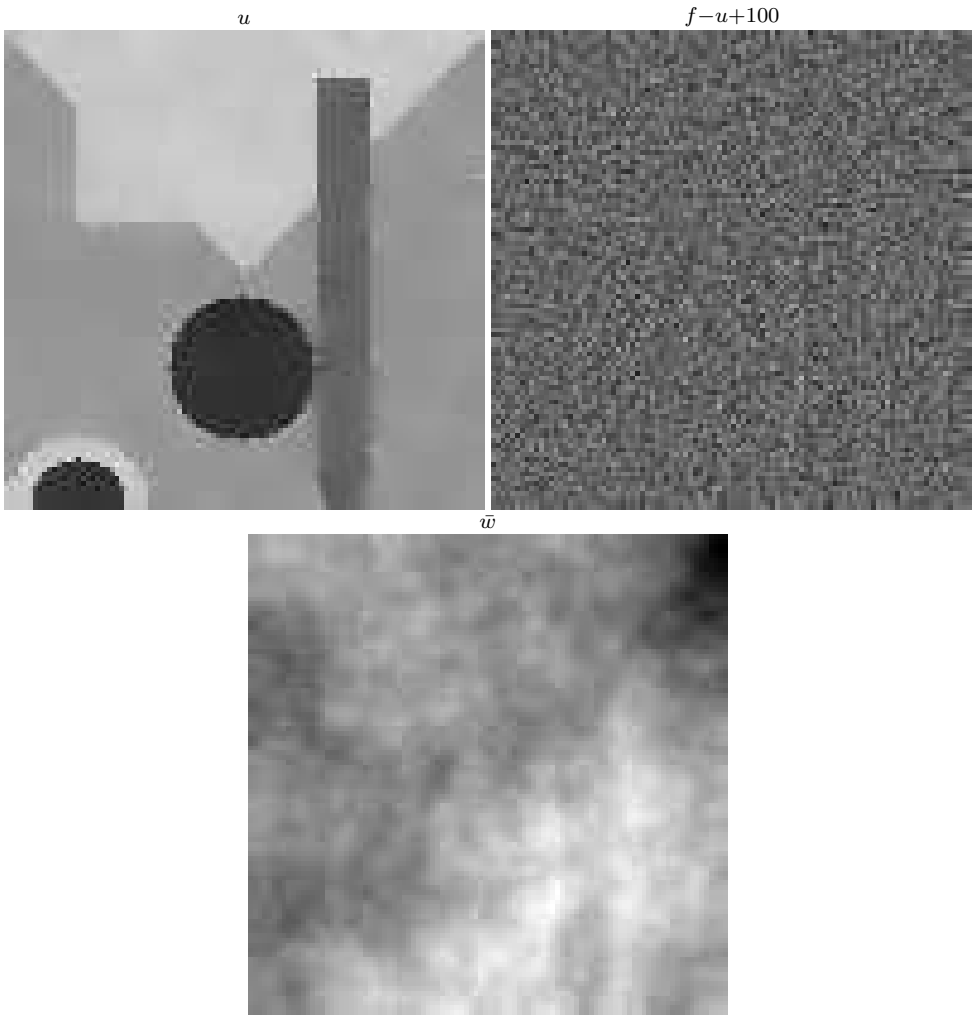


Figure 3.14: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.21) applied to noisy image in Fig. 3.1(d), $rmse = 0.04810694$, $\lambda = 0.4$.

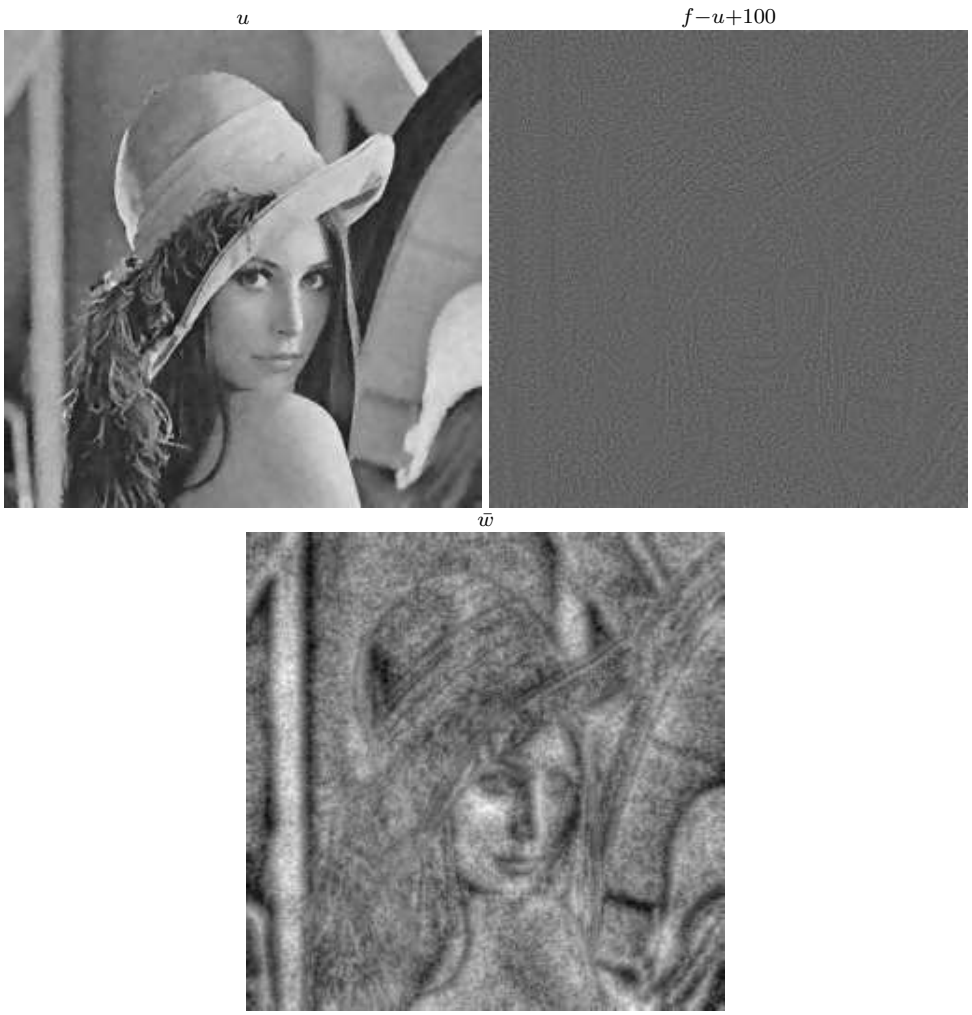


Figure 3.15: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.21) applied to noisy image in Fig. 3.2(b) $rmse = 0.02260758$, $\lambda = 5$.

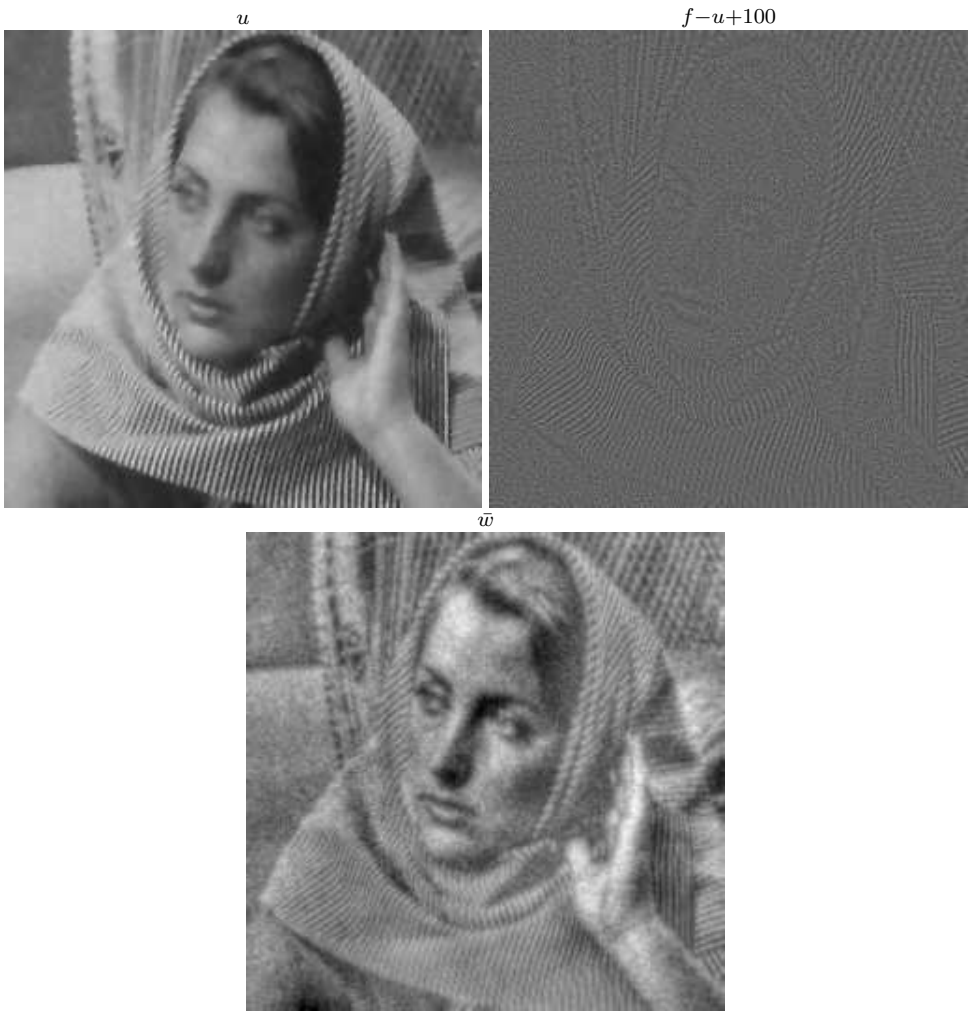


Figure 3.16: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.21) applied to noisy image in Fig. 3.2(d) $rmse = 0.04596249$, $\lambda = 7$.

Figure 3.17: Denoising result of (Φ, Φ^*) model with non-convex Φ in (3.21) applied to noisy image in Fig. 3.2(d) $rmse = 0.04297547$, $\lambda = 9$.

3.5 Application to Image Deblurring

The above model (3.9) can also be applied to image deblurring. We give a description of the algorithm in the section below. Following the description, we present numerical results from our experiments.

3.5.1 Description of the algorithm

Let $K : L^2(\Omega) \to L^2(\Omega)$ be a linear and continuous operator (sometimes, we have to assume in addition that K does not anihilate constants). If we assume the degradation model $f = Ku + n$, where K denotes a blurring operator and n denotes additive noise of zero mean, then we propose the following restoration model:

$$\inf_{u \in BV(\Omega)} E(u) = |u|_{BV(\Omega)} + \lambda \|f - Ku\|_*, \tag{3.22}$$

or

$$\inf_{u \in BV(\Omega)} E(u) = \Phi(u) + \lambda \Phi^*(f - Ku), \tag{3.23}$$

with Φ defined in (3.6). We apply the same steps in the minimization algorithm as before, and the associated time-dependent Euler-Lagrange equations formally are:

$$\frac{\partial w}{\partial t} = \frac{f - Ku}{\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx} + \frac{\int_\Omega (f - Ku) w dx}{(\int_\Omega \sqrt{\epsilon^2 + |\nabla w|^2} dx)^2} \mathrm{div}\Big(\frac{\nabla w}{|\nabla w|}\Big), \tag{3.24}$$

$$\frac{\partial u}{\partial t} = K^* w + \frac{|w|_{BV(\Omega)}}{\lambda} \mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big), \tag{3.25}$$

where K^* denotes the adjoint operator of K. Note that here, we no longer have to invert K^*Ku as it is obtained when the fidelity term is $\|f - Ku\|_2^2$.

3.5.2 (BV, BV^*) Image deblurring results

We apply the deblurring model (3.22), and solve the coupled system (3.24)-(3.25). The blur operator K is given by a convolution with a 5x5 (symmetric) blurring mask (or kernel k) of the form:

$$\frac{1}{273} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 7 & 4 & 1 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 7 & 26 & 41 & 26 & 7 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 1 & 4 & 7 & 4 & 1 \\ \hline \end{array}$$

The equation in u from (3.25) is discretized using a semi-implicit scheme, as in [12, 64], while the equation in w from (3.24) is discretized by a fully explicit scheme. We run 30 iterations in w for every iteration in u. Also, $\triangle x = \triangle y = 1$. The root mean square error ($RMSE = \frac{\sqrt{\sum_{i,j}(u_0(i,j)-u(i,j))^2}}{MN}$ for an image of size MxN) is being used to measure the quality of the restoration, and to find the optimal parameter λ.

Below, for the purpose of validation, we give experimental results obtained with our proposed (BV, BV^*) deblurring method and those obtained with the Rudin-Osher model [56].

I. For results in Figure 3.18 (Deblurring the synthetic image) :

(i) Original RMSE = 0.0812219

(ii) ROF : 6000 iterations, RMSE = 0.03262219, λ = 13.5

(iii) (BV,BV*) : 1000 iterations, RMSE = 0.03408236, λ = 24,000

II. For results in Figure 3.19 (Denoising and deblurring) :

(i) Original RMSE = 0.105052

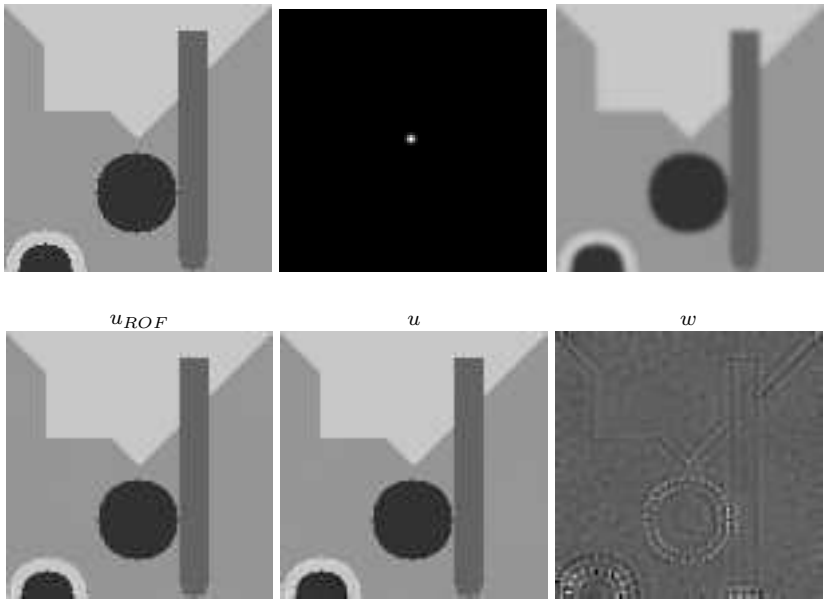


Figure 3.18: A synthetic image, the blurring kernel represented as an image, and the blurry version of the synthetic image (top). Deblurring with ROF and deblurring with (BV,BV*)-model (bottom).

(ii) ROF : 6000 iterations, RMSE = 0.0573248, λ = 0.4

(iii) (BV,BV*) : 6000 iterations, RMSE = 0.06060835, λ = 10,000

III. For results in Figure 3.20 (Deblurring the office image)

(i) Original RMSE = 0.111223

(ii) ROF : 3000 iterations, RMSE = 0.04248943, λ = 60

(iii) (BV,BV*) : 6000 iterations, RMSE = 0.04450807, λ = 250,000

3.6 Conclusions

We have presented numerical algorithms for minimizing functional energies consisting of the sum of a functional Φ and its dual Φ^*. We have shown in the particular case when the

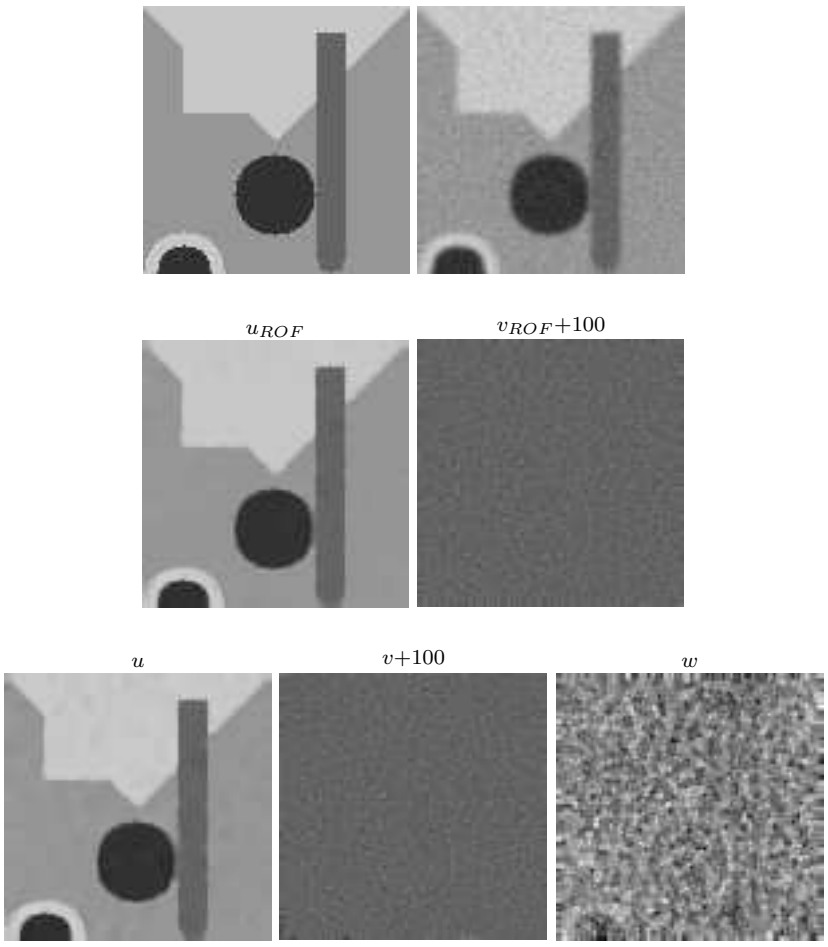


Figure 3.19: A synthetic image and its noisy-blurred version (top). Reconstruction using ROF (middle). Reconstruction using (BV,BV*)-model (bottom).

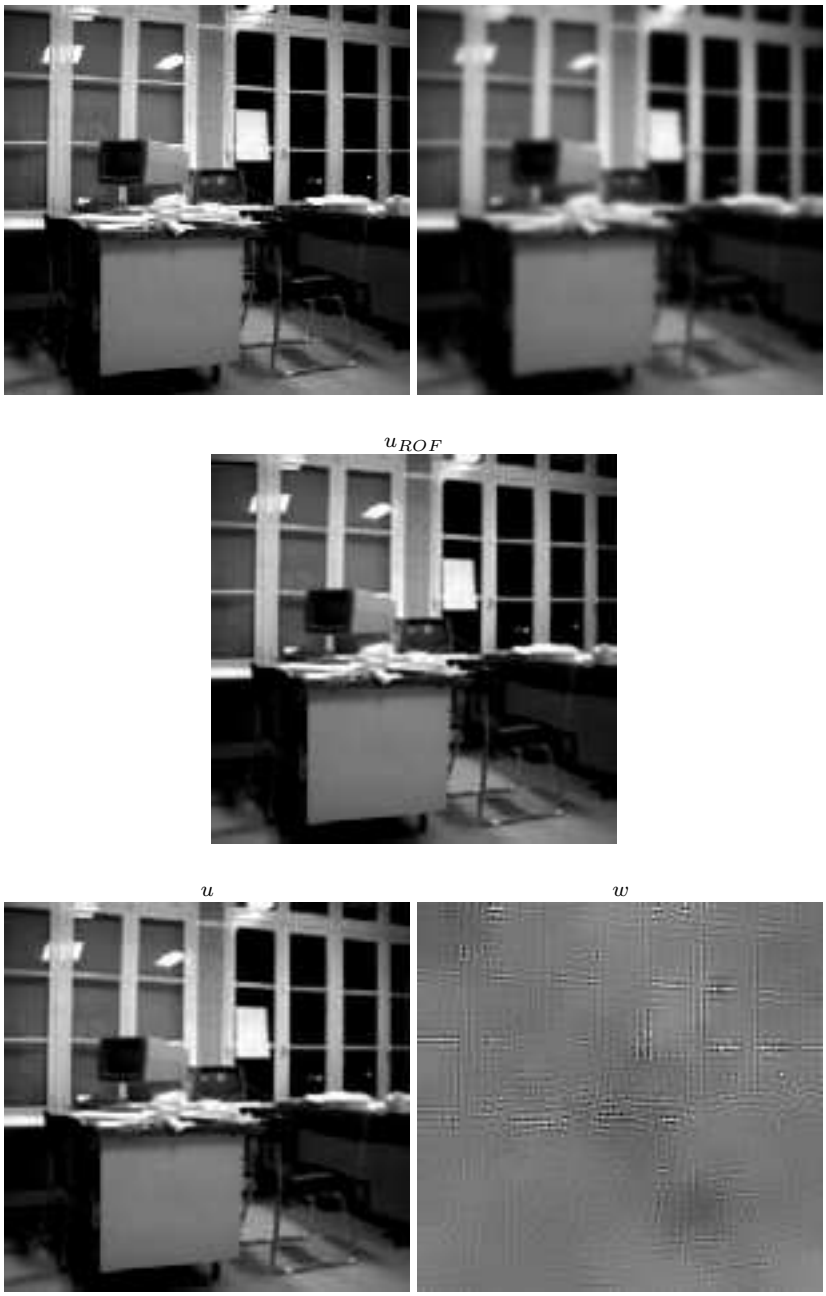


Figure 3.20: Original image of an office and a blurry version (top). Deblurring with ROF (middle). Deblurring with (BV,BV*)-model (bottom).

functional is the total variation of functions in the quotient space $\mathcal{BV}(\Omega) = BV(\Omega)/P_0(\Omega)$, then its dual coincides with the norm $||\cdot||_{W^{-1,\infty}(\Omega)}$ on a closed subspace of $W^{-1,\infty}(\Omega)$. We have validated our algorithms with experimental results of image denoising, image decomposition, and image deblurring. For image denoising and decomposition, we have considered the cases when Φ is convex and nonconvex. For image deblurring with or without noise, we have considered the case when Φ is the total variation. Our results show an improvement over the (BV, L^2) model in cartoon + texture separation, especially for the case of non-convex Φ.

PART II

IMAGE SEGMENTATION

This part of the dissertation is devoted to image segmentation. We are presented with an edge detection problem arising from the analysis of image data collected by streak cameras during scientific experiments at Los Alamos National Lab (LANL). We consider the two well known models in image segmentation, Mumford-Shah [50] and Chan-Vese [18], for our edge detection problem.

In the next chapter, we give a description of our research problem. Then we review in details the Mumford-Shah and Chan-Vese models. We show how the Chan-Vese model performs when applied to our problem. And finally, we present a modification of the Chan-Vese model that better fits our purpose and improves overall performance.

This research is part of a summer project sponsored by the Los Alamos National Lab which was supervised by Professor Andrea Bertozzi of UCLA and Larry Hill of LANL.

Chapter 4

Edge Detection in Streak-Camera Images

4.1 Description of the Research Problem

During a detonation of explosives experiment, data is collected with a rotating lense, which deflects a ray of lights across the surface of a film. The result is a two dimensional image on film. We then scan the films to obtain digital streak-camera images, such as those in Figure 4.1. The vertical direction in the image is the time dimension, and the horizontal is the spatial dimension. Each line across the spatial dimension corresponds to one rotation in the lense. We want to warn the readers that this description of streak-camera image acquisition is only the general idea. It is not at all a thorough description.

The task at hand to determine the edge in the image (which in many cases describes the detonation waves). We want to detect the edge as accurately as possible, so that we can further determine various information concerning the waves. For example, we would like to be able to determine the rate of change and the curvature of the edge.

Since these images consist of two homogeneous regions of approximately the same

gray scale intensity, we consider the Chan-Vese "Active Contour Without Edges" model for our problem. In this chapter, we will describe in details the Chan-Vese model. But, to better understand the Chan-Vese model, we will first study the Mumford-Shah model [50] in Section 4.2. In Section 4.3.4, we will give some numerical results to illustrate the drawback in application of Chan-Vese model to our problem. In Section 4.4, we present a modified version of the Chan-Vese model adapted to our problem, together with the improved numerical results.

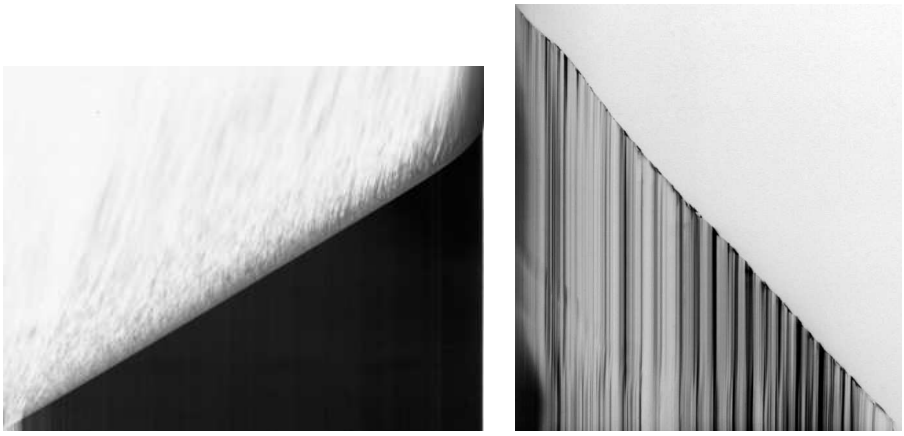


Figure 4.1: Examples of streak-camera images.

4.2 Mumford-Shah Segmentation Model

Throughout this chapter, we will denote by $I : \Omega \to \mathbb{R}^2$ the two-dimensional gray-scale image in consideration. Solid objects in an image generally appear as regions where the values of I vary smoothly, and the boundary of these regions appear in the image as a result of marked changes in I. In other words, edges appear because of discontinuities in the intensity function. Thus, the problem of edge detection is actually the problem of finding a discontinuity set that corresponds to the edges. Unfortunately, oscillatory patterns such as texture and noise also belong to the discontinuity set of I. The question then is how to

formulate an automated process that detects edges (boundary of objects) and ignore small oscillatory details (texture and noise).

D. Mumford and J. Shah suggest to solve the problem by finding an optimal approximation u of I such that: (i) u is piecewise smooth in Ω, and (ii) the set of discontinuities of u, denoted by K, is the union of piecewise smooth boundaries of disjoint connected open subsets of Ω. More precisely, $\Omega \backslash K = \cup R_j$, where $R_j \subset \Omega$ are open connected sets with piecewise smooth boundary, and u is continuously differentiable on each R_j. In [50] the authors propose the following variational approach to image segmentation:

$$\text{Minimize } E(u,K) = \int_\Omega (u-I)^2 dx + \alpha \int_{\Omega\backslash K} |\nabla u|^2 dx + \beta \mathcal{H}^1(K), \tag{4.1}$$

where I is assumed to belong to $L^\infty(\Omega)$, $\alpha, \beta > 0$ are fixed parameters, K varies in the class of one dimensional closed subsets in Ω, and u varies in $C^1(\Omega\backslash K)$.

The first term in (4.1) demands that u approximates I, the second term that u is smooth in $\Omega\backslash K$, and the third term that length of K be minimized.

As observed in [11], the difficulty in studying problem (4.1) is that it involves two unknowns of different nature: u is a function defined on two-dimensional space, while K is a one-dimensional set. The map $R \mapsto \mathcal{H}^1(\partial R)$ is not lower semicontinuous with respect to any compact topology, therefore it is not possible to prove existence of solutions for (4.1) by direct method of the calculus of variations. A variant formulation of (4.1) is necessary. The idea is to set $K = S_u$, the set of jumps discontinuities of u. Then problem (4.1) has the following weak formulation:

$$\text{Minimize } F(u) = \int_\Omega (u-I)^2 dx + \alpha \int_{\Omega\backslash S_u} |\nabla u|^2 dx + \beta \mathcal{H}^1(S_u). \tag{4.2}$$

Existence of solutions for (4.2) in the space $SBV(\Omega) \cap L^\infty(\Omega)$ is a direct consequence

of Theorem 1.1.11 (see also [19]). The minimizers u_F of F are piecewise C^1 functions, and (see [4]) if we set $K = \Omega \cap \overline{S_u}$, then (u_F, K) solves (4.1), (see also [24]).

For the purpose of practical computation of a minimizer, L. Ambrosio and V.M. Tortorelli [5, 6] have proposed two approximations for the Mumford-Shah functional. The simpler of the two is the following elliptic approximation [6]:

$$G_\varepsilon(u,v) = \int_\Omega \Big[\varepsilon|\nabla v|^2 + \alpha(v^2 + o_\varepsilon)|\nabla u|^2 + \frac{(v-1)^2}{4\varepsilon} + \beta|u - I|^2\Big]dxdy, \qquad (4.3)$$

where $\varepsilon \downarrow 0$, and o_ε is any non negative infinitesimal quantity approaching 0 faster than ε.

The authors have shown in [6] that if $w_\varepsilon = (u_\varepsilon, v_\varepsilon)$ minimizes G_ε, then (passing to subsequences) u_ε is an approximation of a minimizer u of the Mumford-Shah problem in (4.1), and $v_\varepsilon \to 1$ as $\varepsilon \to 0^+$ in the $L^2(\Omega)$-topology.

Here, $v_\varepsilon \leq 1$. Indeed, it is equal to less than 1 only in a small neighborhood of S_u, which shrinks as $\varepsilon \to 0^+$.

Minimizing (4.3) with respect to u and v, we obtained the associated Euler-Lagrange equations

$$\begin{cases} u &= \frac{\alpha}{\beta}\text{div}\Big[(v^2 + o_\varepsilon)\nabla u\Big] + I & \text{in } \Omega, \\ v\Big(\frac{1}{4\varepsilon} + \alpha|\nabla u|^2\Big) &= \varepsilon\triangle v + \frac{1}{4\varepsilon} & \text{in } \Omega, \\ (v^2 + o_\varepsilon)\frac{\partial u}{\partial \vec{n}} &= 0 & \text{on } \partial\Omega, \\ \frac{\partial v}{\partial \vec{n}} &= 0 & \text{on } \partial\Omega. \end{cases}$$

In [50], Mumford and Shah have also considered the particular case of (4.1), the case of minimal partition problem in which u is constrained to be piecewise constant. In this case, the segmentation problem reduces to:

$$\begin{aligned} &\text{Minimize } E_0(u,K) = \textstyle\int_\Omega (u-I)^2 dx + \beta \mathcal{H}^1(K), \\ &\text{subject to } \ \nabla u = 0 \ \text{ in } \ \Omega \backslash K. \end{aligned} \tag{4.4}$$

Proof of existence of solutions for (4.4) can be found in [48].

Ambrosio and Tortorelli have also proposed an approximation to (4.4) [6]:

$$E_\varepsilon(u,v) = \int_\Omega \Big[\varepsilon |\nabla v|^2 + (M_\varepsilon v^2 + o_\varepsilon)|\nabla u|^2 + \frac{(v-1)^2}{4\varepsilon} + \beta |u-I|^2\Big] dxdy, \tag{4.5}$$

where $M_\varepsilon \to +\infty$ as $\varepsilon \to 0^+$, and o_ε is as before.

The associated E-L equations in this case will be:

$$\begin{cases} u &= \frac{1}{\beta}\text{div}\Big[(M_\varepsilon v^2 + o_\varepsilon)\nabla u\Big] + I, \\ v\Big(\frac{1}{4\varepsilon} + M_\varepsilon |\nabla u|^2\Big) &= \varepsilon \triangle v + \frac{1}{4\varepsilon}, \end{cases}$$

A variant formulation of (4.4), in the level-set approach, has been done by Chan and Vese [18]. We will study this in the next section.

4.3 Chan-Vese Active Contours Without Edges

The Chan-Vese model is derived from the Mumford-Shah model by restricting the Mumford-Shah energy to the space of binary piecewise constant functions. In the level-set formulation, the model is written in an elegant way as a minimization problem of a functional energy. We now give a full description of the Chan-Vese model and implement the model in application to our problem of edge detection.

4.3.1 Description of the Chan-Vese Model

The basic idea in the Chan-Vese Active Contours Without Edges model [18] is to partition the pixels in an image into two sets: one contains pixels belonging to the objects in the image and the other contains pixels belonging to the background. Given an image $I : \Omega \rightarrow [0, L]$, the Chan-Vese model searches for a binary piecewise-constant function u such that

(a) u best approximates I,

(b) the boundary of the two level sets of u is the desired segmentation contours C, and

(c) $u = \begin{cases} c_1 = \text{ average of } I \text{ on the 'inside' of } C \\ c_2 = \text{ average of } I \text{ on the 'outside' of } C \end{cases}$

The Chan-Vese model is the following variational problem:

$$\begin{aligned} \inf_{c_1,c_2,C} E(c_1, c_2, C) \quad := \quad & \mu \,\text{Length}(C) + \nu \,\text{Area}(inside(C)) \\ & + \lambda_1 \int_{inside(C)} |I(x,y) - c_1|^2 dxdy \\ & + \lambda_2 \int_{outside(C)} |I(x,y) - c_2|^2 dxdy, \end{aligned} \tag{4.6}$$

where $\mu > 0$, $\nu > 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, are parameters which can be adjusted to weigh the importance of each of the terms in the energy.

The parameter μ in the Length term controls the smoothness and total length of the contours C. When μ is small, it allows the contours C to have larger length, hence it can be oscillatory. When μ is large, C has to have small length, hence no oscillations. So when we want to detect many small objects, choose μ to be small because the contours C need to wrap around many regions in the image resulting in larger length. Similarly, choose μ large when we want to detect few large objects.

In practice, the parameter ν is set to 0, so the energy can be written without the Area term. This is because Area$(inside(C)) \leq const \cdot$(Length$(C))^{n/(n-1)}$ when C lies in $\Omega \subset R^n$ $(n > 1)$. Hereafter, we will remove the Area term completely from our discussion of the model. The last two terms in the energy are simply least-square fitting of two constants c_1, c_2 to the values of I in the regions $inside(C)$ and $outside(C)$, respectively.

4.3.2 Level-Set Formulation

In this section, we review the level set formulation of the Chan-Vese model. The level set representations offer unambiguous reference to the 'inside' and the 'outside' of the contour C. It is as follows (see [51] for details on level set methods):

Let $\phi : \Omega \rightarrow R$ be a Lipschitz function such that

- $C = \{(x, y) \in \Omega \ : \phi(x, y) = 0\}$,
- $inside(C) = \{(x, y) \in \Omega \ : \phi(x, y) > 0\}$, and
- $outside(C) = \{(x, y) \in \Omega \ : \phi(x, y) < 0\}$.

Example: Assume the contour C is the unit circle $C = (x, y) \in \mathbb{R}^2, x^2 + y^2 = 1$. Then $inside(C)$ is the unit disk (see Fig. 4.3.2) and the level set function $\phi(x, y)$ can be defined as the signed distance function

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

Using the Heaviside function H and the one-dimensional Dirac measure δ_0, namely

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0, \end{cases} \qquad \delta_0(z) = \frac{d}{dz} H(z),$$

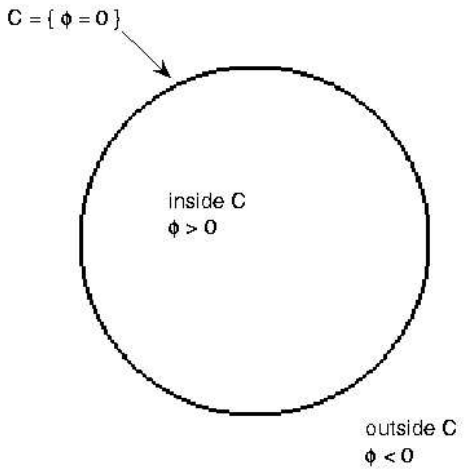


Figure 4.2: Illustration of different regions defined by the unit circle C.

we can express the length term as

$$\text{Length}(\{\phi = 0\}) = \int_\Omega |\nabla H(\phi(x,y))| dxdy = \int_\Omega \delta_0(\phi(x,y))|\nabla\phi(x,y)|dxdy.$$

Then we have the following mathematical expression for the Chan-Vese model:

$$\begin{aligned} \inf_\phi \inf_{c_1,c_2} E(c_1,c_2,\phi) \;\; := \;\; & \mu \int_\Omega \delta_0(\phi(x,y))|\nabla\phi(x,y)|dxdy \\ & +\lambda_1 \int_\Omega |I(x,y)-c_1|^2 H(\phi(x,y))dxdy \\ & +\lambda_2 \int_\Omega |I(x,y)-c_2|^2 (1-H(\phi(x,y)))dxdy. \end{aligned} \tag{4.7}$$

4.3.3 Algorithm and Numerical Approximation

We apply the following iterative algorithm to find a minimizer for (4.7):

- Starting with any initial level set function ϕ^0.
- For $n = 0, 1, 2, \ldots,$ do

1. Solve $\inf_{c_1,c_2} E(c_1, c_2, \phi^n)$.

   Obtain $c_1(\phi^n), c_2(\phi^n) = argminE(c_1, c_2, \phi^n)$.

2. Solve $\inf_{\phi} E(c_1(\phi^n), c_2(\phi^n), \phi)$ via Steepest Gradient Descent.

   Obtain $\phi^{n+1} = argminE(c_1(\phi^n), c_2(\phi^n), \phi)$.

3. Solve $\inf_{c_1,c_2} E(c_1, c_2, \phi^{n+1})$.

   Obtain $c_1(\phi^{n+1}), c_2(\phi^{n+1}) = argminE(c_1, c_2, \phi^{n+1})$.

4. Stop if solution $(c_1(\phi^{n+1}), c_2(\phi^{n+1}), \phi^{n+1})$ is stationary.

   If not, continue.

This process is convergent to a stationary point, since the energy is strictly decreasing at each step.

Steps 1 and 3 can be solved directly from the Euler-Lagrange equations associated with minimizing E in (4.7) with respect to c_1 and c_2.

The Euler-Lagrange equation associated with minimizing E with respect to c_1 is

$$\int_{\Omega} I(x,y)H(\phi(x,y))dxdy - c_1 \int_{\Omega} H(\phi(x,y))dxdy = 0.$$

If $\int_{\Omega} H(\phi(x,y))dxdy = 0$, we have no constraints on c_1. Otherwise,

$$c_1 = \frac{\int_{\Omega} I(x,y)H(\phi(x,y))dxdy}{\int_{\Omega} H(\phi(x,y))dxdy}. \tag{4.8}$$

Similarly,

$$c_2 = \frac{\int_{\Omega} I(x,y)(1 - H(\phi(x,y)))dxdy}{\int_{\Omega}(1 - H(\phi(x,y)))dxdy}, \tag{4.9}$$

provided $\int_{\Omega}(1 - H(\phi(x,y)))dxdy > 0$. Otherwise, c_2 can be any value.

Remark: The optimality conditions for c_1 and c_2 dictate that c_1 and c_2 be the average of I 'inside' and 'outside' of C, respectively. This is expected because we know that the solution to a least-square fitting has to be the average.

To solve Step 2, we need to compute the first variation of E with respect to ϕ. Since the Heaviside function H is not differentiable at 0, we approximate it with a smoother H_ϵ. In addition, we add an artificial time variable $t > 0$ to ϕ to parameterize the descent direction. Let $\vec{n}$ and $\partial\phi/\partial\vec{n}$ denote the outward unit normal to the boundary $\partial\Omega$ and the outward normal derivative of ϕ on $\partial\Omega$, respectively. Then, the evolution PDE for ϕ is:

$$\begin{cases} \frac{\partial\phi}{\partial t} = \delta_\epsilon(\phi)\Big[\mu div\Big(\frac{\nabla\phi}{|\nabla\phi|}\Big) - \lambda_1(I-c_1)^2 + \lambda_2(I-c_2)^2\Big] & \text{in } \Omega\times(0,\infty), \\ \phi(0,\cdot) = \phi^0(\cdot) & \text{in } \Omega, \\ \frac{\delta_\epsilon(\phi)}{|\nabla\phi|}\frac{\partial\phi}{\partial\vec{n}} = 0 & \text{on } \partial\Omega\times(0,\infty), \end{cases}$$

where the Neumann boundary condition is implicit (see [64]).

Finally, the ϕ^{n+1} in Step 2 is computed via the following semi-implicit scheme:

$$\frac{\phi^{n+1}-\phi^n}{\triangle t} = \delta_\epsilon(\phi^n)\Big[\mu div\Big(\frac{\nabla\phi^{n+1}}{|\nabla\phi^n|}\Big) - \lambda_1(I-c_1(\phi^n))^2 + \lambda_2(I-c_2(\phi^n))^2\Big] \quad (4.10)$$

Numerical Approximation:

We give here the discretization of (4.10) as given in the original Chan-Vese paper [18]. Basically, only the second order term in (4.10) needs to be addressed. Recall the following standard notations:

- $\triangle x,\ \triangle y$ are spatial step size
- $\phi_{i,j} = \phi(i\triangle x, j\triangle y), 1 \le i \le M,\ 1 \le j \le N.$
- $D_x^-\phi_{i,j} = \phi_{i,j} - \phi_{i-1,j},\ \ D_x^+\phi_{i,j} = \phi_{i+1,j} - \phi_{i,j},\ \ D_x^0\phi_{i,j} = \phi_{i+1,j} - \phi_{i-1,j},$

$$D_y^- \phi_{i,j} = \phi_{i,j} - \phi_{i,j-1}, \;\; D_y^+ \phi_{i,j} = \phi_{i,j+1} - \phi_{i,j}, \;\; D_y^0 \phi_{i,j} = \phi_{i,j+1} - \phi_{i,j-1}.$$

Following the ideas in [57], we apply the following discretization for the second order term in (4.10):

$$\begin{aligned} div\Big(\frac{\nabla \phi^{n+1}}{|\nabla \phi^n|}\Big) &= \frac{1}{\triangle x} D_x^- \left(\frac{(D_x^+ \phi^{n+1})/\triangle x}{\sqrt{\delta^2 + \left(\frac{D_x^+ \phi^n}{\triangle x}\right)^2 + \left(\frac{D_y^0 \phi^n}{2\triangle y}\right)^2}} \right) \\ &\quad + \frac{1}{\triangle y} D_y^- \left(\frac{(D_y^+ \phi^{n+1})/\triangle y}{\sqrt{\delta^2 + \left(\frac{D_x^0 \phi^n}{2\triangle x}\right)^2 + \left(\frac{D_y^+ \phi^n}{\triangle y}\right)^2}} \right), \end{aligned}$$

where we add a small δ to avoid dividing by zero.

To obtain the Neumann boundary condition, we extend the boundary points of ϕ^n. This makes ϕ_{ext}^n become $(M+2) \times (N+2)$. Then we compute ϕ^{n+1} at all grid points using ϕ_{ext}^n. This semi-implicit scheme is unconditionally stable (see [12] for details).

We initialize ϕ^0 by a periodic function whose zero level set is many small circles (see FIG. 4.3). Distributing the initial curves everywhere in the image has the advantage of faster detection of the segmentation edge.

For a smooth approximation of the Heaviside function H, we use

$$H_\epsilon(z) = \frac{1}{2}\Big(1 + \frac{1}{\pi} \arctan(\frac{z}{\epsilon})\Big). \tag{4.11}$$

Because the energy F is nonconvex, there may be many local minima, hence the solution may depend on the initial ϕ^0. With this choice of H_ϵ, the Euler-Lagrange equation in ϕ can act on all level curves of ϕ. This allows the attainment of the global minimizer independent of the initial curve.

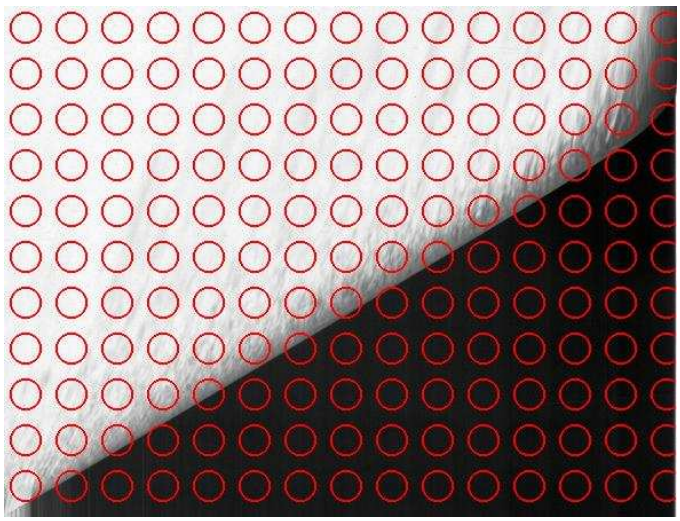


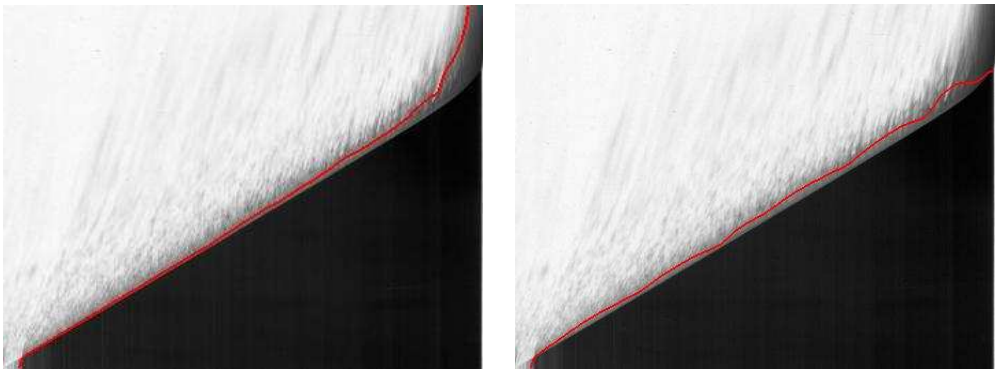
Figure 4.3: Initial Contours.

Although in theory the numerical scheme is unconditionally stable, in practice we set $\triangle t = 0.1$ to control the speed down the gradient flow. This helps when the image is highly blurred and degraded. We set $\epsilon = \triangle x = \triangle y = 1$ pixel and solve the linear system at each time step by Gauss-Seidel iterative method.

4.3.4 Numerical Results

We first apply the Chan-Vese model to the streak-camera images in Fig. 4.1. The results are shown in Fig. 4.4 and Fig. 4.5.

In Fig. 4.4 the segmentation completely misses the edge toward the right of the image. Recall that the Chan-Vese model is simultaneously minimizing the length of the contours and obtaining the best-fit of the two (dark and bright) regions in the image to their averages. Since the region in the upper right corner of the image is so much closer to the (bottom) darker region than the bright (upper left) region, the optimal segmentation obtained by the Chan-Vese model failed to recognize the edge as the human eyes can. As an attempt to solve this problem, we decrease the value of the parameter λ_1 to relax the constraint on

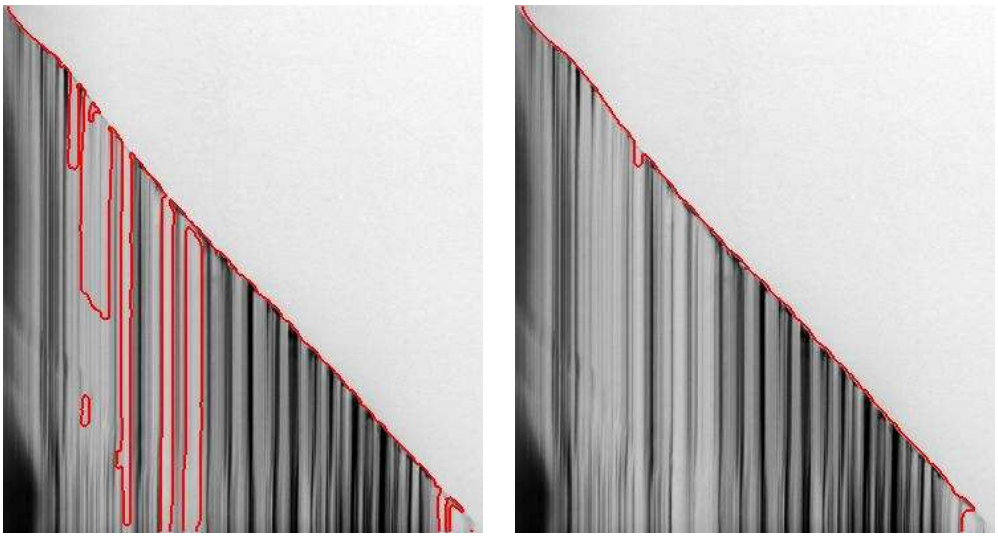


(a) (b)

Figure 4.4: Segmentation result of Chan-Vese model. The parameters are: for (a) $\mu = 0.8 \cdot max^2$, $\lambda_1 = 0.5$, $\lambda_2 = 1$, for (b) $\mu = 2.2 \cdot max^2$, $\lambda_1 = 0.05$, $\lambda_2 = 1$ (where max=maximum intensity value of image I).

the best-fit of the regions to their averages and increase the value of μ to impose that the length of the segmentation contour be shorter. An example is shown in Figure 4.4(b). We can see that the strong regularization on the length of the contour makes it almost a straight line. And in trying to best-fit the regions to their averages, the contour misses the true edge completely.

Similar behaviors can be seen in the results in Fig. 4.5. When the regularization parameter μ is small, the contour length is large, resulting in erroneous segmentation. We increased the value in μ and got better result. However, we still get the unsatisfactory segmentation. The contour caves into regions called *'dropped-out'* regions. Further increase in μ will make the contour disappear, i.e. we get the trivial solution.



(a) (b)

Figure 4.5: Segmentation result of Chan-Vese model. The parameters are : (a) $\mu = 0.05 \cdot max^2$, $\lambda_1 = 1$, $\lambda_2 = 1$ and (b) $\mu = 0.15 \cdot max^2$, $\lambda_1 = 1$, $\lambda_2 = 1$ (where max=maximum intensity value of image I).

4.4 1-D Modified Chan-Vese Model

4.4.1 Formulation of the Model

Since our interest lies solely in detecting the edge that characterizes wave front, we consider defining our level set function ϕ via the graph of a function $f : [0, a] \to [0, b]$. That is, we restrict the Chan-Vese energy to the space of functions defined as follows: assuming that $\Omega = (0, a) \times (0, b)$, let $\phi(x, y) = y - f(x)$ for all $(x, y) \in \Omega$. The zero level set of ϕ is exactly the curve $y = f(x)$. In this setting, we produce the following simplified energy

whose unknown is in one dimension:

$$\begin{aligned}\tilde{E}(c_1, c_2, f) &= \frac{\mu}{2}\int_0^a |f'(x)|^2 dx \\ &\quad +\lambda_1 \int_\Omega \Big(I(x,y) - c_1\Big)^2 H(y - f(x)) dxdy \\ &\quad +\lambda_2 \int_\Omega \Big(I(x,y) - c_2\Big)^2 \Big(1 - H(y - f(x))\Big) dxdy,\end{aligned} \tag{4.12}$$

Anologous to the two dimensional case, c_1 and c_2 are the averages of I in the region above and below the graph of $f(x)$, respectively. However, instead of imposing a penalty on the length of the contour, which is $length := \int_0^a \sqrt{1 + |f'(s)|^2} ds$, we impose a stronger regularizing condition on f, requiring it to be in the Sobolev space $H^1((0, a))$. (We shall clarify our preference of H^1 regularization at the end of Section 4.4.2). Another difference in this modified model from the original Chan-Vese model is that we are able to compute exactly the length of the contour instead of an approximation via the use of H_ϵ.

The optimality conditions for c_1 and c_2 are similar to above case in two dimension:

$$c_1 = \frac{\int_\Omega I(x,y) H(y - f(x)) dxdy}{\int_\Omega H(y - f(x)) dxdy}, \quad c_2 = \frac{\int_\Omega I(x,y)(1 - H(y - f(x))) dxdy}{\int_\Omega (1 - H(y - f(x))) dxdy}, \tag{4.13}$$

provided the denominators are nonzero, i.e. when the intersection between the graph of f and the open rectangle $(0, a) \times (0, b)$ is nonempty. In the case that $f(x) \leq 0$ for all $x \in [0, a]$, then the optimal c_1 is the average of I in Ω, and c_2 can be anything. Similar, if $f(x) \geq b$ for all x, then the optimal c_2 is the average of I in Ω, and c_1 can be anything.

Since the given image I is defined on the rectangle $\Omega = (0, a) \times (0, b)$, how the graph of f behave outside of Ω does not matter to our minimization problem in (4.12). Therefore, it makes sense to search for a solution in the space $Y = H^1((0, a)) \cap \{f(x) |\ 0 \leq f(x) \leq b\ \forall x \in (0, a)\}$. To prove existence of a minimizer for problem (4.12), we recall a definition and a theorem:

Definition 4.4.1 (Hölder Continuous with Exponent γ) *Assume $U \subset \mathbb{R}^n$ is open and $0 < \gamma \leq 1$. A function $f : U \to \mathbb{R}$ is said to be Hölder continuous with exponent γ if there is a constant C such that*

$$|f(x) - f(y)| \leq C|x - y|^\gamma \text{ for all } x, y \in U.$$

The space of all such Hölder continuous functions with exponent γ, denoted by $C^{0,\gamma}(U)$, is a Banach space under the γ^{th}-Hölder norm

$$||f||_{C^{0,\gamma}(\bar{U})} \stackrel{\text{def}}{=} ||f||_{C(\bar{U})} + \sup_{\substack{x,y \in U \\ x \neq y}} \left\{ \frac{|f(x) - f(y)}{|x - y|^\gamma} \right\}.$$

Theorem 4.4.1 (General Sobolev Inequality [29]) *Let U be a bounded open subset of $\mathbb{R}^n$, with a C^1 boundary. Assume $u \in W^{k,p}(U)$. If $k > \frac{n}{p}$, then $u \in C^{0,\gamma}(\bar{U})$, where*

$$\gamma = \begin{cases} \lfloor \frac{n}{p} \rfloor + 1 - \frac{n}{p}, & \text{if } \frac{n}{p} \text{ is not an integer,} \\ \text{any positive number} < 1, & \text{if } \frac{n}{p} \text{ is an integer,} \end{cases}$$

where $\lfloor \cdot \rfloor$ is rounding down to the nearest integer. In addition, we have the estimate

$$||u||_{C^{0,\gamma}(\bar{U})} \leq C||u||_{W^{k,p}(U)},$$

where the constant C depends only on k, p, n, γ, and U.

Theorem 4.4.2 (Existence of Solution) *Suppose $\Omega = (0, a) \times (0, b)$, and define the set $Y = H^1((0, a)) \cap \{0 \leq f(x) \leq b, \ \forall x \in (0, a)\}$. If $I(x, y) \in L^2(\Omega)$, then the minimization problem (4.12) has a solution in Y.*

Proof: For any $f \in Y$, denote by $\Gamma(f) = \{(x, f(x)), 0 < x < a\}$ (the graph of f in Ω), $\chi_1 = H(y - f(x))$ and $\chi_2 = 1 - H(y - f(x))$ the characteristic functions of two disjoints

sets E_1, E_2 in Ω above and below the curve $\Gamma(f)$, (with $E_1 \cup E_2 \cup \Gamma(f) = \Omega$). Then we can write problem (4.12) as

$$\inf_{f\in Y} \tilde{E}(f) = \frac{\mu}{2}\int_0^a |f'(x)|^2 dx + \sum_{i=1}^{2} \lambda_i \int_\Omega \Big(I(x,y) - c_i(\chi_i)\Big)^2 \chi_i \, dxdy, \tag{4.14}$$

where $c_i(\chi_i) := (\int_\Omega I(x,y)\chi_i \, dxdy)/(\int_\Omega \chi_i \, dxdy)$.

Since $\tilde{E}(0) = \lambda_1 \int_\Omega (I - I_\Omega)^2 dxdy < \infty$, ($I_\Omega$ is the mean of I in Ω), the infimum is finite and we can construct a minimizing sequence $f_k \in Y$, (i.e. $\lim_{k\to\infty} \tilde{E}(f_k) = \inf_{f\in Y} \tilde{E}(f)$), such that $\tilde{E}(f_k) \le B$ for all k.

By General Sobolev Inequality, the space $H^1((0,a))$ consists of $\frac{1}{2}^{\text{th}}$-Hölder continuous functions. (In fact, since we are in one dimension, f_k are absolutely continuous on $(0,a)$, see [29], pp.246). Since $\tilde{E}(f_k) \le B$, $||f_k||_{C^{0,\frac{1}{2}}([0,a])} \le C$, for some $C > 0$. Thus (see [31], pp.138), there is a subsequence, still denoted as f_k, and a $\frac{1}{2}^{\text{th}}$-Hölder continuous function f_0 such that $f_k \to f_0$ uniformly on $[0,a]$, and

$$\sup_{x\in(0,a)} |f_k(x) - f_0(x)| + \sup_{\substack{x,y\in(0,a)\\ x\neq y}} \frac{|(f_k(x) - f_0(x)) - (f_k(y) - f_0(y))|}{|x-y|^{1/2}} \to 0, \tag{4.15}$$

as $k \to \infty$.

Since $f_k \in H^1((0,a))$, the usual derivative of f_k exists a.e. Then, for a.e. $x \in (0,a)$:

$$\begin{aligned} \lim_{k\to\infty} f_k'(x) &= \lim_{k\to\infty} \left(\lim_{h\to 0} \tfrac{f_k(x+h) - f_k(x)}{h}\right) \\ &= \lim_{h\to 0} \left(\lim_{k\to\infty} \tfrac{f_k(x+h) - f_k(x)}{h}\right) \\ &= f_0'(x), \end{aligned}$$

where we are able to interchange the order of the limits because the limit exists with respect to each h and k, and in addition, f_k are absolutely continuous and satisfy (4.15).

Clearly $f_0' \in L^2((0,a))$. Therefore we have shown that there exists $f_0 \in Y$ such that

$f_k \to f_0$ uniformly on $[0,a]$ and $f_k' \to f_0'$ a.e. on $(0,a)$. This clearly implies $\int_0^a |f'|^2 \leq \liminf_{k\to\infty} \int_0^a |f_k'|^2$.

Denote $\chi_{1,k} = H(y - f_k(x))$, $\chi_{1,0} = H(y - f_0(x))$, $\chi_{2,k} = 1 - H(y - f_k(x))$, and $\chi_{2,0} = 1 - H(y - f_0(x))$. Then, $f_k \to f_0$ uniformly on $[0,a]$ implies $\chi_{i,k} \to \chi_{i,0}$ pointwise in Ω, for $i = 1, 2$. Clearly $c_i(\chi_{i,k}) \to c_i(\chi_{i,0})$ as $k \to \infty$, (just apply Dominated Convergence theorem). Therefore, we have

$$\tilde{E}(f_0) \leq \liminf_{k\to\infty} \tilde{E}(f_k),$$

i.e. f_0 is a minimizer to (4.12). □

Formally solving problem (4.12) yields the following evolution equation for f in the Gradient Descent direction of $\tilde{E}$:

$$\begin{cases} \frac{\partial f}{\partial t} = \mu f''(x) + \lambda_1 \int_0^b (I(x,y) - c_1)^2 \delta_\epsilon(y - f(x))dy \\ \qquad\qquad -\lambda_2 \int_0^b (I(x,y) - c_2)^2 \delta_\epsilon(y - f(x))dy, \quad x \in [0,a], \\ f'(0) = f'(a) = 0. \end{cases} \tag{4.16}$$

where $\delta_\epsilon := \frac{d}{dz} H_\epsilon$, the derivative of the approximated Heaviside function.

We solve equation (4.16) using the fixed-point method given in the last section, where the discretization in time is:

$$\begin{cases} \frac{f^{n+1} - f^n}{\triangle t} = \mu (f^n(x))'' + \lambda_1 \int_0^b (I(x,y) - c_1)^2 \delta_\epsilon(y - f^n(x))dy \\ \qquad\qquad -\lambda_2 \int_0^b (I(x,y) - c_2)^2 \delta_\epsilon(y - f^n(x))dy, \quad x \in [0,a], \\ (f^{n+1})'(0) = (f^{n+1})'(a) = 0. \end{cases} \tag{4.17}$$

Notice that the two fitting terms that are functions of the x-variable are integrals in the y-direction. In our numerical computation, for each x, we compute the integral using the approximated δ_ϵ. We initialize f^0 by a constant, and as in the two-dimensional case, we

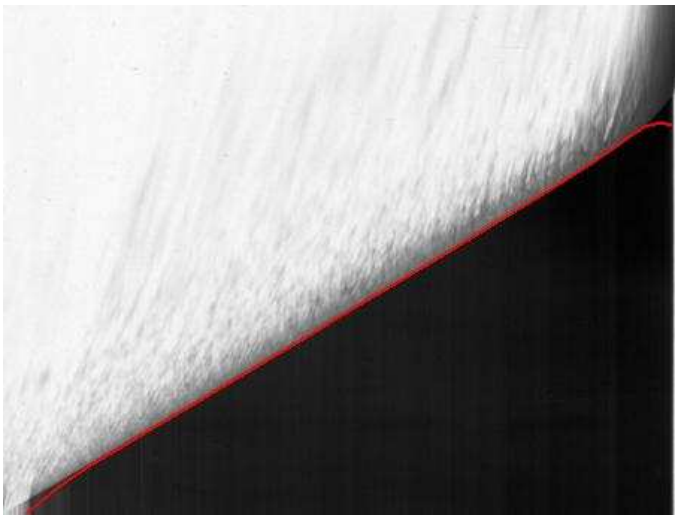


Figure 4.6: Segmentation result of 1D modified Chan-Vese model. The parameters are $\mu = 1.5 \cdot max^2$, $\lambda_1 = 1.0$, $\lambda_2 = 0.1$ (where max=maximum intensity value of image I).

take $\triangle t = 0.1$ and $\epsilon = \triangle x = \triangle y = 1$ pixel.

4.4.2 Numerical Results

We test the 1D modified model with the same image and are able to detect more accurately the edge, see Figure 4.6. However, on the right side of the image where the true edge is barely visible, the computed segmentation still misses the true edge. This is one known short coming in computer vision when compared to the human eyes.

Another advantage of the 1D modified Chan-Vese model is the ability to detect the wave-front edge in images that have texture patterns. In Figure 4.7 we show an example of segmentation by this 1D model. For these types of streak-camera images, the original Chan-Vese model would fail to detect the desired edge.

Remark on regularization by minimal length

We have also considered for our 1D Modified Chan-Vese model regularization by min-

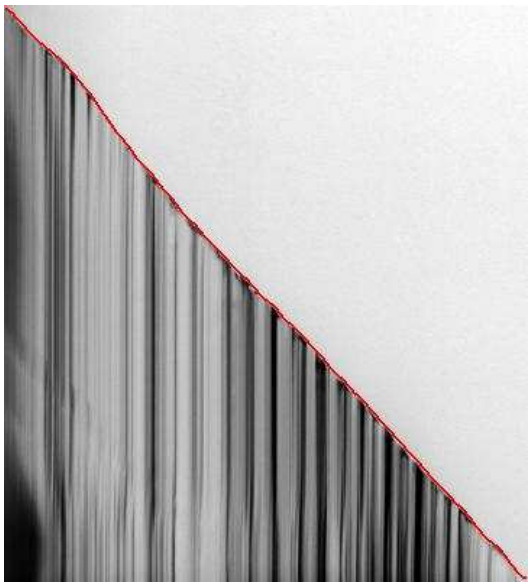


Figure 4.7: Segmentation result of 1D modified Chan-Vese model. The parameters are $\mu = 0.2 \cdot max^2$, $\lambda_1 = 1.0$, $\lambda_2 = 2.0$ (where max=maximum intensity value of image I).

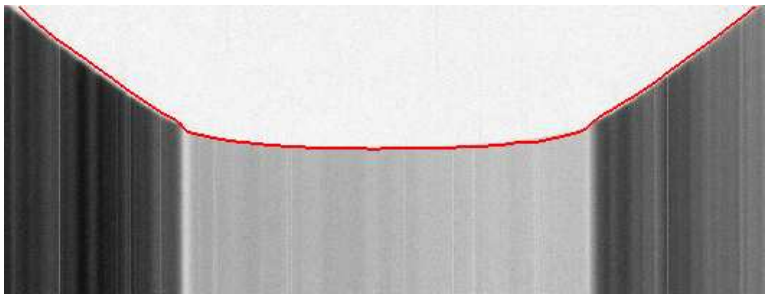


Figure 4.8: Segmentation result of 1D modified Chan-Vese model. The parameters are $\mu = 0.5 \cdot max^2$, $\lambda_1 = 0.8$, $\lambda_2 = 2.75$ (where max=maximum intensity value of image I).

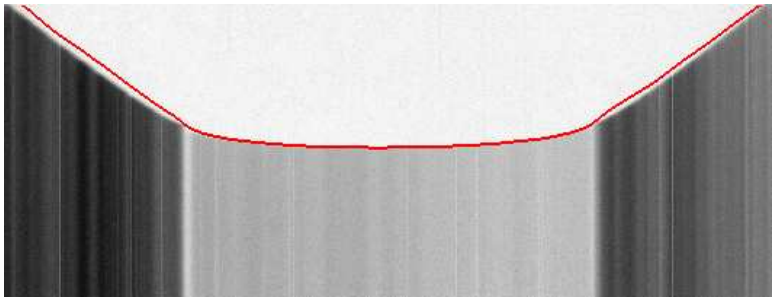


Figure 4.9: Segmentation result of 1D modified Chan-Vese model with minimal length regularization. The parameters are $\mu = 0.4 \cdot max^2$, $\lambda_1 = 0.1$, $\lambda_2 = 0.8$ (where max=maximum intensity value of image I).

imal length, in which case the regularizing term is

$$\int_0^1 \sqrt{1 + |f'|^2} dx.$$

Then the associated PDE has the term

$$\frac{f''}{(1 + |f'|^2)^{3/2}}$$

in place of f''. A segmentation result is shown in Figure 4.9. Comparing to that in Figure 4.8, the contour tends to be straight line segments. This is expected since the penalty is on the length of the graph of f. As a result, the contour misses much of the edge to be detected. Therefore, we conclude that the H^1 regularization as chosen is better for our application than minimal length regularization.

4.5 Conclusions

We have examined and applied the Chan-Vese "Active Contour Without Edges" segmentation model to our problem of detecting the detonation wave edge in streak-camera im-

ages. We have studied the drawbacks of the Chan-Vese model in our application and have made modifications to the model by reducing it from a two dimensional problem to a one dimensional problem. The modified 1D version performs much faster and yields more satisfactory results, especially when texture patterns are present in the image. Moreover, the strong regularization on f prevents the contour from breaking, hence we are able to detect the edge in images with three different regions of different gray scale such as that in Figure 4.8. However, at regions in the image where the sharpness of the edge is not well-resolved (but can still be seen by the human eyes) the 1D version also fails to detect the edge.

Finally, we would like to mention that the idea of using of the graph of a function to detect an edge in two dimensional image can be generalized to detect a surface in three dimension. In this case, the function $f(x, y)$ is defined on a two dimensional domain, and its graph $\{(x, y, f(x, y))\}$ is a hypersurface in three dimension.

Appendix

Lemma For every $u \in L^{n/(n-1)}(\Omega)$ and for every $\vec{p} \in L^n(\Omega)^n$ with $\mathrm{div}\vec{p} \in L^n(\Omega)$, we can define a distribution on Ω, denoted $\vec{p} \cdot Du$, via the formula

$$\langle \vec{p} \cdot Du, \varphi \rangle = -\int_\Omega (u \mathrm{div}\vec{p})\varphi dx - \int_\Omega u(\vec{p} \cdot \nabla\varphi)dx, \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{18}$$

Furthermore, the mapping $(u, \vec{p}) \to \vec{p} \cdot Du$ is bilinear weakly continuous from $L^{n/(n-1)}(\Omega) \times X$ into $\mathcal{D}'(\Omega)$, the space of distributions on Ω, where

$$X = \{\vec{p} \in L^n(\Omega)^n, \mathrm{div}\vec{p} \in L^n(\Omega)\}, \tag{19}$$

which is a Banach Space with the natural norm

$$||\vec{p}||_X = ||\vec{p}||_n + ||\mathrm{div}\vec{p}||_n. \tag{20}$$

Proof: The right-hand side of (18) clearly makes sense for $u, \vec{p}, \varphi$ in the hypotheses. Moreover, for fixed u and $\vec{p}$, this expression depends linearly and continously on φ under the norm $|| \cdot ||_{C^1(\bar{\Omega})}$. Hence $\vec{p} \cdot Du$ is defined as a distribution (and coincides with the function $\vec{p}(x) \cdot \nabla u(x)$ if $\vec{p}$ and u are smooth enough).

For φ fixed, if $u_m \to u$ in $L^{n/(n-1)}(\Omega)$ and $\vec{p}_m \to \vec{p}$ in X, then by Hölder's Inequality,

$$-\int_\Omega (u_m \mathrm{div}\vec{p}_m)\varphi dx - \int_\Omega u_m(\vec{p}_m \cdot \nabla\varphi)dx \longrightarrow -\int_\Omega (u\mathrm{div}\vec{p})\varphi dx - \int_\Omega u(\vec{p} \cdot \nabla\varphi)dx,$$

This establishes the continuity of the mapping $(u, \vec{p}) \to \vec{p} \cdot Du$. $\square$

Lemma Under the assumptions of Lemma 1 and if moreover $u \in BV(\Omega)$ and $\vec{p} \in L^\infty(\Omega)^n$, then $\vec{p} \cdot Du$ is a bounded signed measure with

$$\int_\Omega |\vec{p} \cdot Du| \leq ||\vec{p}||_\infty \cdot \int_\Omega |Du|, \tag{21}$$

and the following generalized Green's formula holds

$$\int_\Omega \vec{p} \cdot Du = \int_\Gamma u(\vec{p} \cdot \nu) d\mathcal{H}^{n-1} - \int_\Omega u \mathrm{div}(\vec{p}) dx, \tag{22}$$

Proof: First recall that $BV(\Omega) \subset L^{n/(n-1)}(\Omega)$ (by Poincaré's inequality), and $\vec{p} \cdot Du$ makes sense by Lemma 1. For (21), let $\vec{p}$ and u be fixed. Let $\vec{p}_{ext}$ denote the function obtained by extending $\vec{p}$ by 0 outside Ω. Let $\vec{p}_m \in C(\bar{\Omega})^n$ be an approximating sequence for $\vec{p}$ obtained by mollification of the function $\vec{p}_{ext}$ and restriction (of the mollified functions) to Ω. Then

$$||\vec{p}_m||_{C(\bar{\Omega})^n} \leq ||\vec{p}||_{L^\infty(\Omega)^n},$$
$$\vec{p}_m(x) \to \vec{p}(x) \quad \text{a.e. in } \Omega,$$
$$\mathrm{div}\vec{p}_m \to \mathrm{div}\vec{p} \quad \text{in } L^n_{\mathrm{loc}}(\Omega).$$

If $\varphi \in \mathcal{D}(\Omega)$, then by (18) and these convergence results, we have

$$\int_\Omega (\vec{p}_m \cdot Du)\varphi dx \to \langle \vec{p} \cdot Du, \varphi \rangle, \text{ as } m \to \infty,$$

and,

$$\begin{aligned} \left| \int_\Omega (\vec{p}_m \cdot Du)\varphi dx \right| &\leq ||\vec{p}_m||_{C(\bar{\Omega})^n} ||\varphi||_{C(\bar{\Omega})} \int_\Omega |Du| \\ &\leq ||\vec{p}||_{L^\infty(\Omega)^n} ||\varphi||_{C(\bar{\Omega})} \int_\Omega |Du| \end{aligned}$$

so that

$$|\langle \vec{p} \cdot Du, \varphi \rangle| \leq ||\vec{p}||_{L^\infty(\Omega)^n} ||\varphi||_{C(\bar{\Omega})} \int_\Omega |Du|,$$

for all $\varphi \in \mathcal{D}(\Omega)$, and (21) is proved.

A detailed proof for (22) can be found in ([8], pp. 315-317). □

Theorem The space $G^V(\Omega)$ is isometrically isomorphic to $W^{-1,\infty}(\Omega)$ under the dual norm $||\cdot||_{W^{-1,\infty}}$ to the norm $||\cdot||_{W_0^{1,1}} = \int_\Omega |\nabla u| dx$.

Proof: (This proof is provided by the second author of [40])

We first recall a definition. Two normed spaces $\mathcal{X}, \mathcal{Y}$ are said to be isometrically isomorphic if there is a bijective linear operator $L : \mathcal{X} \to \mathcal{Y}$ such that $||L(x)||_{\mathcal{Y}} = ||x||_{\mathcal{X}}$ for all $x \in \mathcal{X}$.

Let $P : W_0^{1,1} \to (L^1(\Omega))^2$ be a linear operator defined by $Pu = \nabla u$ for all $u \in W_0^{1,1}$. Denote by W the range of P. Then W is a closed subspace of $(L^1(\Omega))^2$, since $W_0^{1,1}$ is a Banach space. Since $||Pu||_{(L^1(\Omega))^2} = \int_\Omega |\nabla u| dx := ||u||_{W_0^{1,1}}$, P is an isometric isomorphism from $W_0^{1,1}(\Omega)$ to W.

Suppose $L \in W^{-1,\infty}$. Define $L^* : W \to \mathbb{R}$ by $L^*(Pu) = L(u)$ (this is well defined since P is an isometric isomorphism). Then $L^* \in W'$ and $||L^*||_{W'} = ||L||_{W^{-1,\infty}}$. By the Hahn-Banach Extension theorem ([2], pp. 6), there is a norm preserving extension of L^*, (still denoted by L^*), to all of $(L^1(\Omega))^2$.

By the Riesz Representation theorem ([2], pp. 47), the dual of $(L^1(\Omega))^2$ is identified with $(L^\infty(\Omega))^2$. Hence, there exists $\vec{g} \in (L^\infty(\Omega))^2$ such that $L^*(\vec{u}) = \int_\Omega (\vec{u} \cdot \vec{g}) dx$, for all $\vec{u} \in (L^1(\Omega))^2$. Therefore, for all $u \in W_0^{1,1}(\Omega)$, we have $L(u) = L^*(Pu) = \int_\Omega (\nabla u \cdot \vec{g}) dx = -\int_\Omega u \mathrm{div}(\vec{g}) dx$. Moreover,

$$||L||_{W^{-1,\infty}} = ||L^*||_{W'} = ||L^*||_{(L^1(\Omega)^2)'} = ||\vec{g}||_{L^\infty(\Omega)^2} = ||\sqrt{(g_1)^2 + (g_2)^2}||_\infty. \qquad (23)$$

On the other hand, for any $\vec{g}$ such that $L(u) = \int_\Omega (\nabla u \cdot \vec{g}) dx$ for all $u \in W_0^{1,1}$, clearly $|L(u)| \leq ||\vec{g}||_{L^\infty(\Omega)^2} \int_\Omega |\nabla u| dx$. Therefore, $||L||_{W^{-1,\infty}} := \inf\{||\vec{g}||_{L^\infty(\Omega)^2}\} = \min\{||\vec{g}||_{L^\infty(\Omega)^2}\}$, over all such $\vec{g}$. The infimum is attained at $\vec{g}$ in (23).

Now, suppose $T \in \mathcal{D}'(\Omega)$ is defined by $T = -\mathrm{div}(\vec{g})$, for some $\vec{g} \in L^\infty(\Omega)^2$, (i.e. $T(\phi) = \int_\Omega (\vec{g} \cdot \nabla \phi) dx$ for any $\phi \in \mathcal{D}(\Omega)$). Then $L \in W^{-1,\infty}(\Omega)$ characterized by $L(u) = \int_\Omega (\nabla u \cdot \vec{g}) dx$ is an extension of T to $W_0^{1,1}(\Omega)$. Moreover, this extension is unique: let $u \in W_0^{1,1}(\Omega)$ and $\{\phi_k\}_k \in \mathcal{D}(\Omega)$ such that $||\phi_k - u||_{W_0^{1,1}} \to 0$ as $k \to \infty$. Observe that $|T(\phi_k) - T(\phi_j)| \leq |\int_\Omega (\vec{g} \cdot \phi_k) - (\vec{g} \cdot \phi_j) dx| \leq \int |\vec{g} \cdot \nabla(\phi_k - \phi_j)| dx \leq ||\vec{g}||_{L^\infty(\Omega)^2} ||\phi_k - \phi_j||_{W_0^{1,1}} \to 0$ as $k, j \to \infty$. Therefore $\{T(\phi_k)\}_k$ is a Cauchy sequence in $\mathbb{R}$, and so converges to a limit which we denote by $L(u)$, (since it is clear that if $\{\varphi_k\}_k \in \mathcal{D}(\Omega)$ with $||\varphi_k - u||_{W_0^{1,1}} \to 0$ then $T(\phi_k) - T(\varphi_k) \to 0$). Therefore, the functional L thus defined is uniquely identified with T. And we have just shown that $G^V(\Omega)$ is isometrically isomorphic to $W^{-1,\infty}$ in the dual norm $||\cdot||_{-1,\infty}$ as defined above. $\square$

Bibliography

[1] R. ACAR AND C.R. VOGEL, *Analysis of bounded variation penalty methods of ill-posed problems*, Inverse Problems, 10: pp. 1217-1229, 1994.

[2] R.A. ADAMS, Sobolev Spaces, Second Ed., Academic Press, 2003.

[3] L. AMBROSIO, *A compactness theorem for a new class of functions of bounded variation*, Bolletino U.M.I. (4) VII: pp. 857-881, 1989.

[4] L. AMBROSIO, *Existence theory for a new class of variational problems*, Archive for Rational Mechanics and Analysis, 111: pp. 291-322, 1990.

[5] L. AMBROSIO AND V.M. TORTORELLI, *Approximation of functionals depending on jumps by elliptic functionals via* Γ*–Convergence*, Comm. on Pure and Appl. Math., Vol. XLIII: pp. 999-1036, 1990.

[6] L. AMBROSIO AND V.M. TORTORELLI, *On the approximation of free discontinuity problems*, Bollettino U.M.I. (7) 6-D: pp. 105-123, 1992.

[7] F. ANDREU-VAILLO, C. BALLESTER, V. CASELLES, AND J.M. MAZON, *Minimizing total variation flow*, C.R. Acad. Sci. Paris Sér. I Math., 331: pp. 867-872, 2000.

[8] F. ANDREU-VAILLO, V. CASELLES, AND J.M. MAZON, Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, Progress in Mathematics Vol. 223, Birkhauser, 2004.

[9] L. ALVAREZ, P.-L. LIONS, AND J.-M. MOREL, *Image selective smoothing and edge detection by non-linear diffusion 2*, SINUM 29 (3): pp. 845-866, 1992.

[10] G. AUBERT AND J.-F. AUJOL, *Modeling very oscillating signals. Application to image processing*, AMO 51 (2): pp. 163-182, 2005.

[11] G. AUBERT AND P. KORNPROBST, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, Applied Mathematical Sciences 147, Springer-Verlag, 2000.

[12] G. AUBERT AND L. VESE, *A variational method in image recovery*, SIAM Journal on Numerical Analysis, 34 (5): pp. 1948-1979, 1997.

[13] J.-F. AUJOL, G. AUBERT, L. BLANC-FERAUD, A. CHAMBOLLE, *Image decomposition application to SAR images*, LNCS 2695: pp. 297-312, 2003.

[14] J.-F. AUJOL, G. AUBERT, L. BLANC-FERAUD, A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, Journal of Mathematical Imaging and Vision 22 (1): pp. 71-88, 2005.

[15] J.-F. AUJOL, A. CHAMBOLLE, *Dual norms and image decomposition models*, IJCV 63(1): pp. 85-104, 2005.

[16] F. Catte, P.-L. Lions, J.-M. Morel, and T. Coll, *Image selective smoothing and edge-detection by nonlinear diffusion*, SINUM 29 (1): pp. 182-193, 1992.

[17] A. CHAMBOLLE AND P.L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76: pp. 167-188, 1997.

[18] T.F. CHAN AND L.A. VESE, *Active contours without edges*, IEEE-IP, 10 (2) : pp. 266–277, 2001.

[19] G. DAL MASO, J.M. MOREL, S. SOLIMINI, *A variational method in image segmentation. Existence and approximation results*,

[20] I. DAUBECHIES, G. TESCHKE, *Wavelet-Based Image Decompositions by Variational Functionals*, Proc. SPIE Vol. 5266, p. 94-105, Wavelet Applications in Industrial Processing; Frederic Truchetet; Ed., Feb. 2004.

[21] I. DAUBECHIES AND G. TESCHKE, *Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising*, Appl. and Comp. Harm. Anal., 19 (1): pp. 1-16, 2005.

[22] R. DAUTRAY AND J.L. LIONS, Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 2, Springer-Verlag, 1988.

[23] R. DAUTRAY AND J.L. LIONS, Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 3, Springer-Verlag, 1988.

[24] E. DE GIORGI, M. CARRIERO, A. LEACI, *Existence theorem for a maximum problem with free discontinuity set*, Archive for Rational Mechanics and Analysis, 108: pp. 195-218, 1989.

[25] T. DE PAUW, *On SBV dual*, Indiana Univ. Math. J. 47(1): pp. 99-121, 1998.

[26] F. DEMENGEL AND R. TEMAM, *Convex functions of a measure and applications*, Indiana University Mathematics Journal, Vol. 33 No. 5: pp. 673-709, 1984.

[27] I. EKELAND AND R. TEMAM, *Convex Analysis and Variational Problems*, American Elsevier Publishing Company, Inc. New York, 1976.

[28] S. ESEDOGLU AND S.J. OSHER, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Communications on Pure and Applied Mathematics, 57 (12): pp. 1609-1626, 2004.

[29] L.C. Evans, *Partial Differential Equations*, Americal Mathematical Society, Providence, RI, 2000.

[30] L.C. EVANS AND R.F. GARIEPY, Measure Theory and Fine Properties of Functions, CRC Press, 1992.

[31] G.B. FOLLAND, Real Analysis : Modern Techniques and Their Applications, Second Ed., Wiley & Sons, Inc., 1999.

[32] J. GARNETT, T. LE, L. VESE, *Image decomposition using bounded variation and generalized homogeneous Besov spaces*, UCLA CAM Report 05-57, October 2005.

[33] D. GEMAN AND S. GEMAN, *Bayesian image analysis* in Disordered Systems and Biological Organization, E. Bienenstock, F. Fogelman and G. Weisbuch Eds., Berlin, Springer Verlag, 1986, Vol. F20.

[34] D. GEMAN, G. REYNOLDS, *Constrained Restoration and the Recovery of Discontinuities*, IEEE Tr. on PAMI, Vol. 14, No. 3: pp. 367-383, 1992.

[35] D. GEMAN, C. YANG, *Nonlinear Image Recovery with Half-Quadratic Regularization*, IEEE TIP, Vol. 4, No. 7, 1995.

[36] S. GEMAN AND D.E. MCCLURE, *Statistical methods for tomographic image reconstruction*, in Proc. 46th Sess. Int. Stat. Inst. Bulletin ISI, vol. 52, 1987.

[37] G. GILBOA, N. SOCHEN AND Y.Y. ZEEVI, *Estimation of the optimal variational parameter via SNR analysis*, LNCS 3459: pp. 230-241, 2005.

[38] S. KINDERMANN, S. OSHER, AND J. XU, *Denoising by BV-Duality*, UCLA CAM Report 05-21, April 2005.

[39] R. KOHN AND R. TEMAM, *Dual spaces of stresses and strains, with applications to Hencky Plasticity*, Appl. Math. Optim. 10: pp. 1-35, 1983.

[40] T. LE AND L. VESE, *Image decomposition using total variation and div(BMO)*, Multiscale Model. Simul., Vol. 4 No. 2 : pp. 390-423, 2005.

[41] S.E. LEVINE, *An Adaptive Variational Model for Image Decomposition*, in Energy Minimization Methods in Computer Vision and Pattern Recognition: 5th International Workshop, EMMCVPR 2005, St. Augustine, FL, USA, November 9-11, 2005. Proceedings. Editors: Anand Rangarajan, Baba Vemuri, Alan L. Yuille, LNCS 3757, pp. 382 - 397, 2005.

[42] L. LIEU AND L. VESE, *Image Restoration and Decompostion via Bounded Total Variation and Negative Hilbert-Sobolev Spaces*, UCLA CAM Report 05-33, May 2005.

[43] J.L. LIONS AND E. MAGENES, *Nonhomogeneous Boundary Problems and Applications*, Springer-Verlag, Heidelberg, NY, 1972.

[44] A. MARQUINA AND S. OSHER, *A new time dependent model based on level set motion for nonlinear deblurring and noise removal*, LNCS 1682: pp. 429-434, 1999.

[45] A. MARQUINA AND S. OSHER, *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal*, SIAM Journal on Scientific Computing 22 (2): pp. 387-405, 2000.

[46] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, Univ. Lecture Ser. 22, AMS, Providence, RI, 2002.

[47] Y. MEYER, Public Lecture, CMLA ENS-Cachan, Workshop "Contenu informatif des images numériques", 2004.

[48] J.-M. MOREL AND S. SOLIMINI, *Variational Methods in Image Segmentation*, Progress in Nonlinear Differential Equations and Their Applications, Vol. 14, Birkhäuser Boston, 1995.

[49] D. MUMFORD AND B. GIDAS, *Stochastic models for generic images*, Quart. Appl. Math., 59 : pp. 85-111, 2001.

[50] D. MUMFORD AND J. SHAH, *Optimal approximations by piecewise-smooth functions and associated variational problems*, Communications on Pure and Applied Mathematics 42 (5): pp. 577-685, 1989.

[51] S. OSHER AND R. FEDKIW, Level Set Methods and Dynamic Implicit Surfaces, AMS Vol. 153, Springer-Verlag, 2003.

[52] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multi. Model. Simul., Vol. 1 No. 3 : pp. 349-370, 2003.

[53] P. PERONA AND J. MALIK, *Scale-space and edge detection using anisotropic diffusion*, IEEE Transactions on PAMI, 12 (7): pp. 629-639, 1990.

[54] P.T. ROCKAFELLAR, Convex Analysis, Princeton Univ Press, Princeton, NJ, 1970.

[55] S. ROUDENKO, *Noise and Texture Detection in Image Processing*, Preprint, 2004.

[56] L.I. RUDIN AND S. OSHER, *Total variation based image restoration with free local constraints*, Proceedings of IEEE ICIP-94, Vol. 1: pp. 31-35, 1994.

[57] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60: pp. 259-268, 1992.

[58] G. RUSSO AND P. SMEREKA, *A remark on computing distance functions*, J. Comp. Phys., 163: pp. 51-67, 2000.

[59] D.M. STRONG AND T.F. CHAN, *Exact solutions to total variation regularization problems*, UCLA CAM Report 96-41, October 1996.

[60] R. TEMAM, *Dual variational principles in mechanics and physics*, in "Semi-Infinite Programming and Applications," (A.V. Fiacco & K.O. Kortanec, Eds.), Lecture Notes in Economics and Mathematical Systems 215, Springer-Verlag, Berlin, 1983.

[61] R. TEMAM AND G. STRANG, *Functions of Bounded Deformation*, Archive for Rational Mechanics and Analysis, 75 (1): pp. 7-21, 1980.

[62] E. TADMOR, S. NEZZAR, AND L. VESE, *A multiscale image representation using hierarchical (BV, L2) decompositions*, Multiscale Modeling & Simulation 2(4) : pp. 554-579, 2004.

[63] J.-L. STARCK, M. ELAD, AND D.L. DONOHO, *Image Decomposition: Separation of Texture from Piece-Wise Smooth Content*, SPIE annual meeting, 3-8 August 2003, San Diego, California, USA.

[64] L. VESE, *A study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optm., 44 : pp. 131-161, 2001.

[65] L. VESE, S. OSHER, *Modeling Textures with Total variation Minimization and Oscillating patterns in Image Processing*, Journal of Scientific Computing, 19(1-3) : pp. 553-572, 2003.

[66] L. VESE AND S. OSHER, *Image denoising and decomposition with total variation minimization and oscillatory functions*, J. Math. Imaging Vision, 20: pp. 7-18, 2004.

[67] W.P. ZIEMER, Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation, Springer-Verlag, 1989.