UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Adaptive Total Variation Minimizing Image Restoration (Ph.D. Thesis)

David M. Strong

August 1997

CAM Report 97-38

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

UNIVERSITY OF CALIFORNIA

Los Angeles

Adaptive

Total Variation Minimizing

Image Restoration

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

David M. Strong

1997

The dissertation of David M. Strong is approved.

Stan Osher

Bjorn Engquist

Henry Huang

Tony Chan, Committee Chair

University of California, Los Angeles

1997

Table of Contents

List of Figures

VITA

1969	Born in Provo, Utah, USA.
1991-1992	Teaching Assistant, Department of Mathematics, Brigham Young University.
1992	B.S. (Mathematics), Valedictorian of College of Mathematical and Physical Sciences, Brigham Young University.
1992-1996	Teaching Assistant and Research Assistant, Department of Mathematics, University of California, Los Angeles.
1994	M.A. (Mathematics), University of California, Los Angeles.
1996–1997	Staff Scientist, Image Processing Group, Jet Propulsion Laboratory, Pasadena.

PUBLICATIONS AND PRESENTATIONS

Relation of Regularization Parameter and Scale in Total Variation Based Image Denoising, with Tony Chan. UCLA Math Department CAM Report 96-7, February 1996.

Exact Solutions to Total Variation Regularization Problems, with Tony Chan. UCLA Math Department CAM Report 96-41, October 1996.

Spatially and Scale Adaptive Total Variation Based Regularization and Anisotropic Diffusion in Image Processing, with Tony Chan. UCLA Math Department CAM Report 96-46, November 1996.

Feature-driven Adaptive Total Variation Minimizing Image Restoration, with Peter Blomgren and Tony Chan. UCLA Math Department CAM Report 97-32, July 1997.

"Scale Analysis of Total Variation Image Regularization." Talk given at *Geometry-driven Diffusion Conference*, Palo Alto, June 14, 1996.

"Exact Solutions of Total Variation Image Restoration with Applications to Adaptive Regularization." Talk given at *ONR Workshop on Image Processing*, Los Angeles, September 6, 1996.

"Scale-sensitive Adaptive Total Variation Minimizing Image Restoration." Talk given at *SIAM Annual Meetings*, Palo Alto, July 17, 1997.

"Feature-driven Adaptive Total Variation Minimimizing Image Restoration." Talk given at *SPIE Annual Meetings*, San Diego, July 24, 1997.

ABSTRACT OF THE DISSERTATION

Adaptive
Total Variation Minimizing
Image Restoration

by

David M. Strong

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 1997

Professor Tony Chan, Chair

We analyze the exact effects of total variation (TV) minimizing function regularization in R^1, R^2 and R^3. Our more precise understanding of TV regularization enables us to construct more effecitve TV minimizing image restoration schemes, as well as to better understand what types of images (and image degradation) are most effectively improved by TV restoration. We analytically find exact solutions to the nonlinear TV minimizing function regularization problem for simple but important cases, which can be used to better understand the effects of TV regularization for more general cases.

We give formulae that describe qualitatively and quantitatively the effects of TV regularization. Four important results which we prove are: (1) TV regularization of piecewise constant (noise-free or noisy) radially symmetric functions results in piecewise constant functions, with edge location being preserved exactly; (2) function intensity change is inversely proportional to local feature scale,

is independent of original intensity, and is directly proportional to the regularization parameter; (3) for smooth function features, function intensity change is inversely proportional to radial position and directly proportional to the regularization parameter; and (4) TV regularization is local in a certain sense.

We develop two *adaptive* TV image restoration schemes. Both schemes are motivated by and constructed using our theoretical results. In the first scheme we accomplish adaptivity by locally weighting the measure of the total variation of the image. The weighting factor decreases as the relatively likelihood of the presence of an edge in the image increases. The second adaptive image restoration scheme is a multi-step scheme driven by the *scale* of individual image features. Each step involves selectively applying restoration, only where the scale of the image features is smaller than a user-controlled threshhold, $scale_{thresh}$. The end result is an image comprised of features with scale greater than $scale_{thresh}$. The process can be characterized as a scale-sensitive, anisotropic diffusion process. These two adaptive schemes can be viewed as prototypes of an array of adaptive TV image restoration schemes which may be developed in the future as a result of our theoretical results.

CHAPTER 1

Introduction

1.1 The Basic Problem

A canonical problem in image processing is to find $u = u(\vec{x})$, an approximation of the true (e.g. noise-free) image $u_{true} = u_{true}(\vec{x})$, given the measured or observed image $u_0 = u_0(\vec{x})$, where

$$u + \eta = u_0, \tag{1.1}$$

where $\eta = \eta(\vec{x})$ is the noise or other unwanted characteristics or degradation in the image u_0.

There are numerous approaches to finding an estimate u of u_{true}. These include:

- function regularization;
- statistics-based, e.g. Wiener Filter;
- frequency domain manipulation, e.g. Fourier Transforms;
- wavelet decomposition;
- PDE-based methods, e.g. anisotropic diffusion.

There are two common mathematical formulations of the *function regulariza-*

tion approach. The first is the *unconstrained* or *Tikhonov* problem

$$\min_u \quad \{ \tfrac{1}{2}\|u - u_0\|^2 + \alpha\, R(u) \}. \tag{1.2}$$

We note that throughout this Dissertation we use the L_2 norm; that is,

$$\|\cdot\| \equiv \|\cdot\|_{L_2}$$

both in the continuous case as well as the discrete (i.e. vector) case. In (1.2), $\alpha > 0$ is the *regularization parameter* that determines the balance between goodness of fit of u to the measured data u_0 and the amount of regularization applied to the measured data u_0 in finding u. Another common formulation of the function regularization problem is the *noise-constrained* problem

$$\min_u \; R(u) \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2, \tag{1.3}$$

where the error (noise) level σ is assumed to be known.

1.2 Choice of Regularization Functional in Image Restoration

Generally the functional $R(u)$ in (1.2) and (1.3) is taken to have the form $R(u) = \|Qu\|^2$, where Q is a *linear* operator. For example,

$$Q = \begin{cases} I & \text{(the identity operator)} \\ \Delta & \text{(the Laplacian operator)} \\ \nabla & \text{(the gradient operator)} \end{cases} \tag{1.4}$$

are often used, as roughness in a function can be minimized by controlling the function or its derivatives (cf. [15]). Unfortunately, linear operators are limited in their effectiveness, particularly is the function is discontinuous, and *non-linear* operators are often needed. The following (non-linear) total variation regularization functional has recently been proposed in [22] as a choice of $R(u)$ in image restoration:

$$TV(u) \equiv \int |\nabla u(\vec{x})| \, d\vec{x}. \tag{1.5}$$

The functional $TV(u)$ simply measures the total variation of u. We note that his definition of $TV(u)$ is only valid for differentiable functions u. In general, we must interpret $\nabla u(\vec{x})$ as the distribution derivative so that

$$TV(u) \equiv \max_{|w|_\infty \leq 1} \int u(\vec{x}) \, \nabla \cdot w(\vec{x}) \, d\vec{x},$$

in order to be valid definition of the TV functional for non-differentiable functions. However, (1.5) is a more intuitive definition of the intuitive *idea* of total variation, and it will be consequently be the definition of $TV(u)$ to which we will usually refer in this Dissertation.

The main advantage that TV regularization has over other image restoration techniques is that it does not penalize discontinuities in u (we show in this paper that TV regularization preserves *exactly* the location of discontinuous edges), while simultaneously not penalizing *smooth* functions either. TV regularization looks for an approximation u to the original (e.g. noisy) function u_0 which has minimal total variation, but with no particular bias toward a discontinuous or smooth solution.

TV minimizing image restoration can also be viewed as a special, *model* case of PDE-based anisotropic diffusion schemes, which have recently been introduced.

[20] These anisotropic diffusion schemes are becoming quite popular as they are studied and further understood and developed. This dual nature of TV restoration is subsequently described in a bit more detail in this Dissertation. In the context of anistropic diffusion schemes, TV restoration is quite unique because it can be analyzed in its function regularization form, which it turns out allows us to develop some very important and useful theory which describes the effects of TV minimizing image restoration. These results can potentially be extended to other anisotropic diffusion schemes.

1.3 Outline of the Dissertation

In the balance of this Dissertation we prove and discuss exact analytic solutions to the TV regularization problem, and we use this theoretical understanding of TV minimizing regularization to develop two spatially adaptive image restoration schemes. In Chapter 2, we first describe and discuss TV minimizing function regularization. In Chapter 3, we present a detailed analysis of TV regularization, and derive the exact solutions to several simple but very useful TV regularization problems. In Chapter 4, using the theory developed in Chapter 3, we develop a adaptive image restoration scheme in which adaptivity is accomplished using a spatially varying weighting factor in measuring the total variation in the image. In Chapter 5, again using the results of Chapter 3, we develop a second spatially adaptive image restoration scheme which is sensitive to the *scale* of the individual image features. A summary of our results is given in Chapter 6.

CHAPTER 2

Total Variation Minimizing Image Restoration

In this chapter we give a formal and intuitive explanation of TV minimizing image restoration. We also compare it to other methods of image restoration, and discuss TV minimizing image restoration as a special case of anisotropic diffusion. We conclude the chapter by discussing some of the reasons for which a better understanding of TV minimizing image restoration is useful and indeed needed.

2.1 Formal and Intuitive Explanations

There are two common mathematical formulations of the standard total variation minimizing function regularization problem. The first is the *unconstrained* or *Tikhonov* problem

$$\min_u \quad \{ \tfrac{1}{2}\|u - u_0\|^2 + \alpha\, TV(u) \}. \tag{2.1}$$

In (2.1), $\alpha > 0$ is the *regularization parameter* that determines the balance between goodness of fit of u to the measured data and the amount of regularization done to the measured data u_0. Another common formulation of this problem is the *noise-constrained* problem

$$\min_u \; TV(u) \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2, \tag{2.2}$$

where the error (noise) level σ is assumed to be known. We note that solving (2.1) is equivalent to solving (2.2) when $\alpha = \frac{1}{\lambda}$, where λ is the Lagrange multiplier found in solving (2.2), so that the results we develop for (2.1) are equally useful for (2.2).

The main advantage of TV regularization is that it does not penalize discontinuities (i.e. edges) in u. In fact, we show in this Dissertation that TV regularization preserves edges *exactly*. On the other hand, the functional TV has the advantage that it does not penalize *smooth* functions. For a more intuitive explanation of this property of the TV functional, consider the following lemma for R^1, to which Figure 2.1(a) corresponds.

Lemma 2.1 *Given the set of functions $u(x)$ defined on the interval $[0,1]$, the argument of*

$$\min_u TV(u) \tag{2.3}$$

subject to

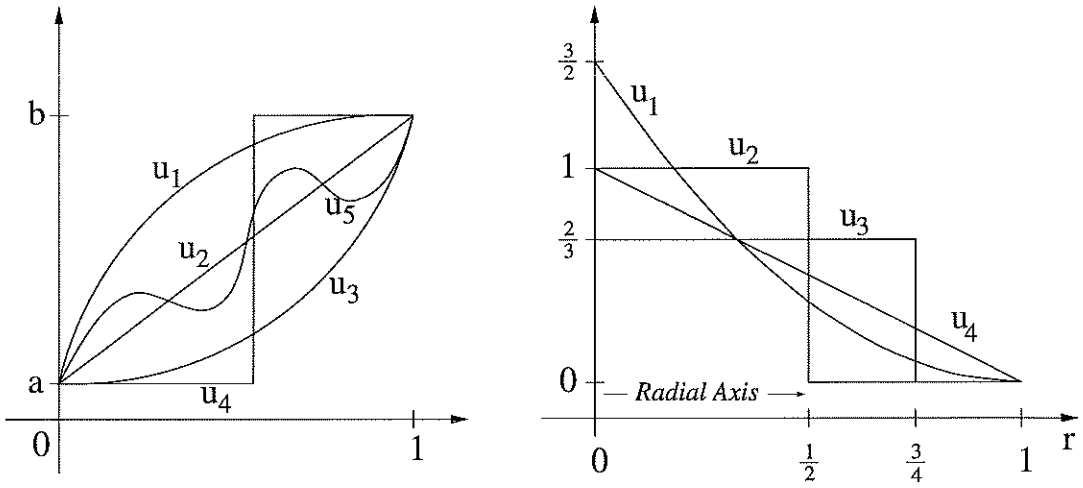
$$u(0) = a \quad \textit{and} \quad u(1) = b \tag{2.4}$$

is any monotone, not necessarily continuous, function $u(x)$ satisfying (2.4).

In Figure 2.1(a) we see that u_1, u_2, u_3 and u_4 would all be acceptable arguments in minimizing (2.3), since

$$TV(u_1) = TV(u_2) = TV(u_3) = TV(u_4) = b - a = \min_u TV(u).$$

Since u_5 is not monotone (in particular, it oscillates), $TV(u_5) > b - a$, and it is consequently not a solution to (2.3).



(a) Lemma 2.1, for R^1

(b) Lemma 2.2, for R^2

Figure 2.1: Functions in R^1 and R^2 with equal total variation, in (a) and (b), respectively.

Lemma 2.1 could be viewed in terms of equivalence classes, where all functions with the same total variation (and which satisfy the necessary conditions, which for Lemma 2.1 were the boundary conditions) form an equivalence class.

For a further intuitive explanation of the TV functional (1.5), we examine this idea for R^2 functions in the following lemma:

Lemma 2.2 *Given the set of radially symmetric, monotonically decreasing functions $u = u(x, y) = u(r)$, defined on the unit circle, such that $u(1)$ is equal to some constant, then all functions with equal integrals $\int_0^1 u(r)\,dr$ have equal total variation.*

PROOF If u is radially symmetric, then with our domain Ω as the unit circle,

$$\begin{aligned} TV(u) &= \int_\Omega |\nabla u(x,y)|\, dx\, dy \\ &= 2\pi \int_0^1 r\, |u_r(r)|\, dr \quad \text{(since } u \text{ is radially symmetric—see Section 3.1.2.1)} \\ &= -2\pi \int_0^1 r\, u_r(r)\, dr \quad \text{(since } u \text{ is monotonically decreasing)} \\ &= 2\pi \,[\int_0^1 u(r)\, dr - u(1)]. \end{aligned}$$

Thus for two functions u_1 and u_2, if $u_1(1) = u_2(1)$ then

$$\int_0^1 u_1(r)\, dr = \int_0^1 u_2(r)\, dr \iff TV(u_1) = TV(u_2).$$

□

In Figure 2.1(b) are the R^1 cross-sections of five radially symmetric R^2 functions, each of which is monotonically non-increasing in radial position r. In this example, for each function we have $u_i(1) = 0$ ($\implies TV(u_i) = 2\pi \int_0^1 u_i(r)\, dr$), and $TV(u_i) = \pi$. For this contrived example, we have used two types of functions that satisfy these conditions $u_i(1) = 0$ and $TV(u_i) = \pi$:

$$u(r) = \begin{cases} a & \text{if } 0 \le r \le b \\ 0 & \text{if } b < r \le 1 \end{cases} \implies ab = \tfrac{1}{2} \qquad \text{(functions } u_2, u_3)$$

$$u(r) = m(1-r)^n \implies \tfrac{m}{n+1} = \tfrac{1}{2} \qquad \text{(functions } u_1, u_4)$$

The TV minimizing approach to image restoration looks for an approximation to the observed image u_0 which has minimal *total variation*. Lemmas 2.1 and 2.2 help to illustrate that in finding an approximation u to u_0, there is no particular bias towards a discontinuous or smooth solution. The measured data u_0, as well

as the regularization parameter α (when solving (2.1)) or the estimated noise level σ^2 (when solving (2.2)), determine the sharpness or smoothness of the restored function u.

2.2 Total Variation Minimizing Anisotropic Diffusion

Total variation minimizing regularization can be viewed as a special case of the more general class of PDE-based anisotropic diffusion schemes, in which time-marching is done using

$$u_t = \nabla \cdot (g(|\nabla u|)\nabla u) \qquad \text{with} \qquad u(\vec{x}, t = 0) = u_0(\vec{x}). \tag{2.5}$$

In solving (2.1), we differentiate (with respect to u) the functional to be minimized, and set th result equal to 0 to get

$$-\alpha \nabla \cdot (\frac{\nabla u}{|\nabla u|}) + (u - u_0) = 0. \tag{2.6}$$

In [22], a time-marching scheme is proposed to solve (2.1) by marching the following PDE to steady state:

$$u_t = \alpha \nabla \cdot (\frac{\nabla u}{|\nabla u|}) - (u - u_0) \qquad \text{with} \qquad u(\vec{x}, t = 0) = u_0(\vec{x}). \tag{2.7}$$

As with other anisotropic diffusion schemes, the first term on the right-hand-side of the time-derivative equation in (2.7) can be viewed as a nonlinear diffusion operator, with diffusion coefficient $\frac{1}{|\nabla u|}$, which diffuses less where the gradient is large (e.g. near edges). One could also diffuse without the fitting constraint,

$u - u_0$, by taking

$$u_t = \nabla \cdot (\frac{\nabla u}{|\nabla u|}) \qquad \text{with} \qquad u(\vec{x}, t = 0) = u_0(\vec{x}). \tag{2.8}$$

In this form we see that TV regularization is a special case of anisotropic diffusion, since (2.8) is simply (2.5) using $g(|\nabla u|) = \frac{1}{|\nabla u|}$).

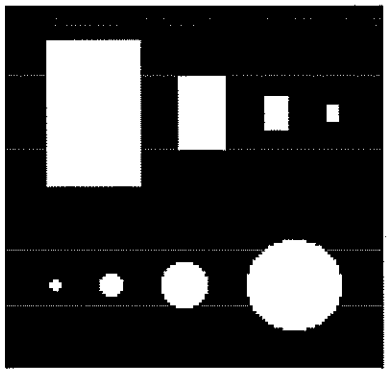
To see the direct connection between TV minimizing function regularization and TV minimizing anisotropic diffusion, notice that we can re-write (2.6) as

$$u = u_0 + \alpha \nabla \cdot (\frac{\nabla u}{|\nabla u|}) \tag{2.9}$$

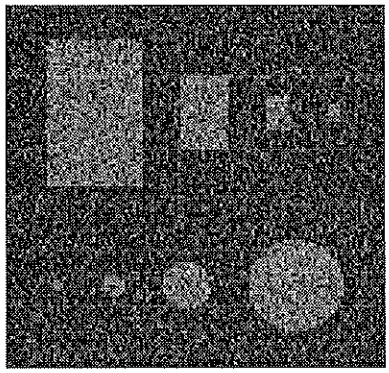
so that solving (2.1) or (2.2) is equivalent to doing a single *implicit* step of time-marching using (2.8) with step size of α. If we time-march *explicitly* using (2.8) for n time steps with time increment $\triangle t$, then the resulting image will be approximately equal to the solution of (2.1) if we have

$$n \triangle t = \alpha. \tag{2.10}$$

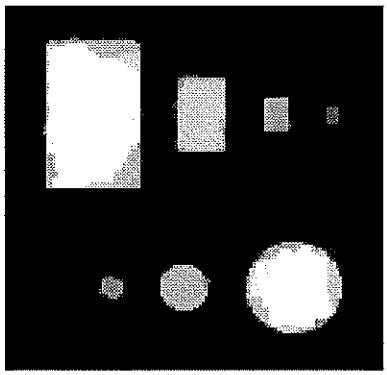
In Figure 2.2 is a simple comparison of TV minimizing regularization, and TV minimizing anisotropic diffusion, using (2.2) and (2.8), respectively. To find the image in (c), we solve the noise-constrained problem, for which we compute a Lagrange multiplier, the inverse of the regularization parameter in (2.1). In finding the image in (d), we use this regularization parameter to choose the number of time-marching steps n (after choosing a time-step size $\triangle t$) in solving (2.8) such that (2.10) holds.



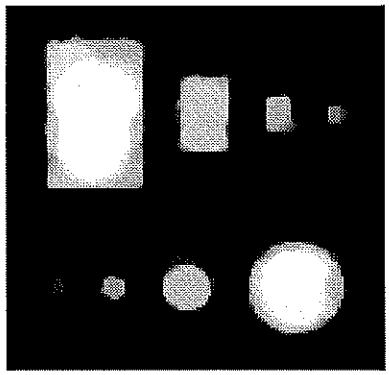
(a) True image



(b) Noisy image, SNR = 0 db



(c) TV minimizing regularization, with noise constraint.



(d) TV minimizing diffusion, with amount of time-marching done to match noise constraint in (c).

Figure 2.2: A comparison of TV regularization and TV diffusion.

2.2.1 Other Ways of Viewing Total Variation Minimzing Function Regularization

In addition to viewing TV minimizing regularization as a special case of anisotropic diffusion, there are other similar ways of interpreting TV regularization, from which it is more obvious that TV regularization is a very natural and intuitive approach to image restoration. We can consider a general function $f(|\nabla u|)$ in the integrand in (1.5). A common choice of $f(|\nabla u|)$ is $f(|\nabla u|) = |\nabla u|^n$. If $n > 1$, then smooth edges are preferred, resulting in loss of sharp discontinuities in the image. On the other hand, if $n < 1$ then sharp edges are preferred, which can lead to unwanted edges, for example staircasing (and moreoever, the minimization problem is no longer even convex). The only unbiased choice would be $f(|\nabla u|) = |\nabla u|$, as is the case for TV regularization.

We can again consider (2.8) back to (2.5), where $g(|\nabla u|)$ is chosen as a decreasing function of $|\nabla u|$. Several choices of $g(|\nabla u|)$ have been used in constructing various anisotropic diffusion schemes. The simplest—and the most unbiased—choice of a non-negative function $g(|\nabla u|)$ decreasing in its non-negative argument is to choose $g(|\nabla u|) = \dfrac{1}{|\nabla u|}$, which is inherent and automatic when doing TV regularization. This also corresponds to choosing $f(|\nabla u|) = |\nabla u|$ in the integrand of (1.5), as just discussed in the previous paragraph.

We can consequently view TV minimizing regularization as a model or canonical case of anisotropic diffusion, of which other anisotropic diffusion schemes are a variation or modification. Because it can be analyzed in its variational form (2.1), it turns out that the effects of TV regularization can be understood more precisely than the effects of other anisotropic schemes. This understanding is in fact subsequently developed in Chapter 3 of this Dissertation by finding exact

solutions to the TV regularization problem by solving the variational formulation of the problem. Additionally, because TV regularization can be viewed as a model anisotropic diffusion scheme, the understanding and schemes we develop for TV regularization can be extended to other anisotropic diffusion schemes.

2.3 Comparison of TV Restoration to Other Techniques

There are several function regularization or noise removal techniques which may be used to denoise, sharpen, smooth, or otherwise enhance data such as a digital image. Many of these are effective when the data is smooth, but tend to perform poorly when the data has steep gradients or, in particular, discontinuities. Conversely, there are also regularization techniques which may perform well in restoring sharp edges, but that perform poorly when restoring functions with gradual or smooth edges by introducing artificial sharpness in the function. Figure 2.3 gives a comparison of TV regularization (solving the constrained formulation (2.2)) to some standard noise removal techniques; these include using the FFT (obtained by computing the FFT of the noisy signal and dropping high frequencies), H1 regularization (solving the constrained formulation) and wavelet decomposition (obtained by using D. Donoho's Wavelab software, using "soft threshholding."). In this example, we have SNR = 0 db. The dashed line represents the true function and the solid line represents the regularized or denoised function. TV restoration is the optimal choice for this example.

TV minimizing regularization techniques have been shown to be very effective in numerical studies (see [22], [26]) as well as in theoretical analysis (see [12]). In [14], it is shown that a unique solution to (2.1) and (2.2) exists under certain conditions.

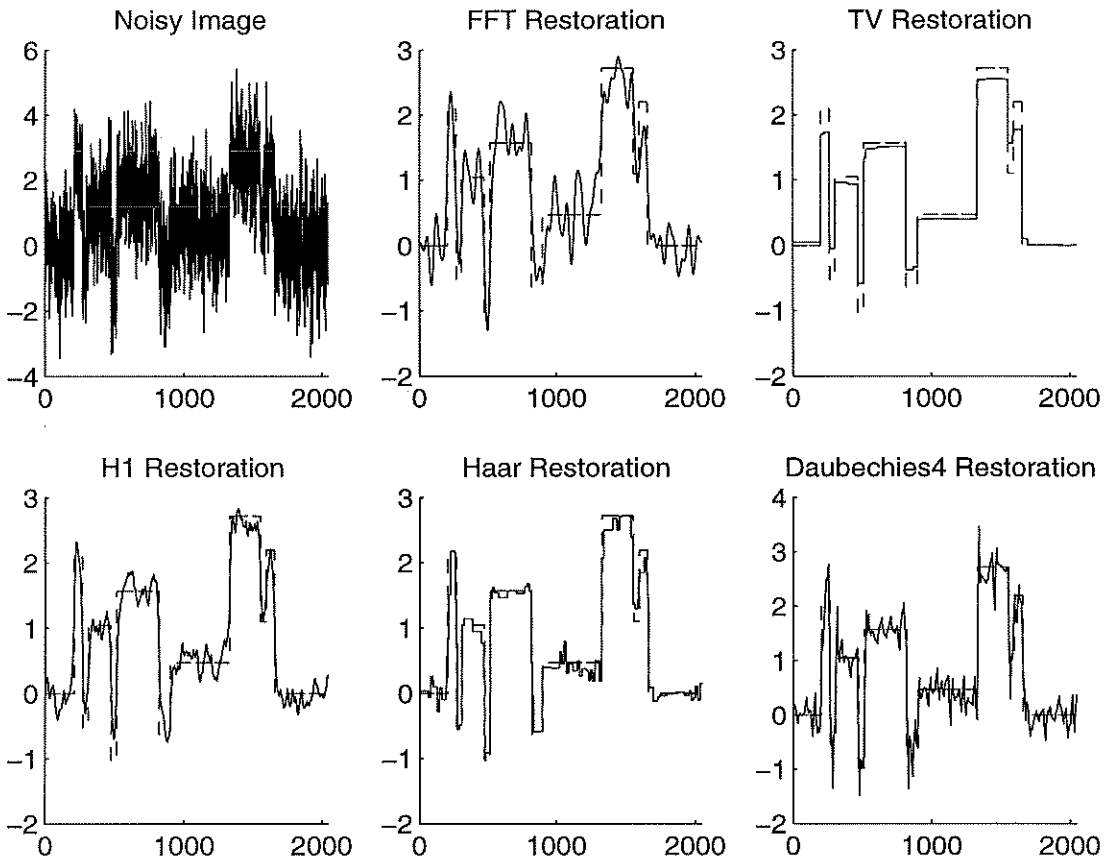


Figure 2.3: A comparison of several standard noise removal techniques in R^1.

2.4 Spatially Adaptive TV Minimizing Image Restoration

Image restoration, and more generally function regularization, can be more effective if done in a spatially adaptive way. There is typically a trade-off between noise removal and detail preservation, and functions are typically comprised of multiple features of different spatial scales. In removing noise from a function, the details and/or contrast are often reduced and can be lost completely. A natural approach to partially alleviate this problem is to use spatial adaptivity in restoring the image. In general, less regularization of the image is desired in regions of more detail, while more regularization is appropriate in regions of less detail.

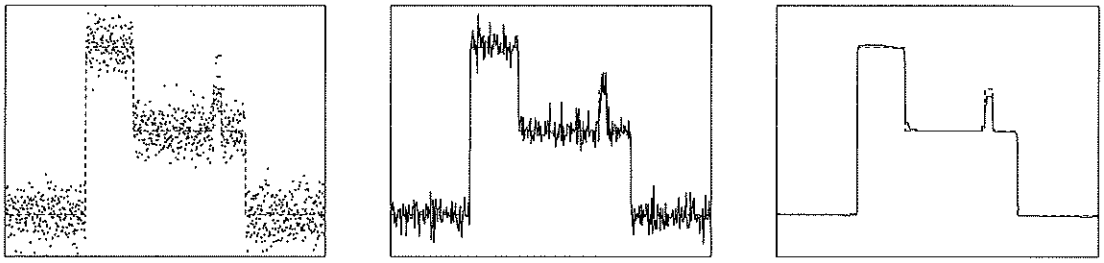
Spatial adaptivity has been studied extensively in image restoration literature (see for example [3] or [4]). At present, the author is aware of a single proposal for an adaptive *TV minimizing* image restoration scheme [21], which to date has not been published. In the context of TV regularization, one natural way of achieving spatial adaptivity would be by modifying equation (2.1) to become

$$\min_u \quad \{ \tfrac{1}{2}\|u - u_0\|^2 + \int \alpha(\vec{x})|\nabla u(\vec{x})|\, d\vec{x} \}. \tag{2.11}$$

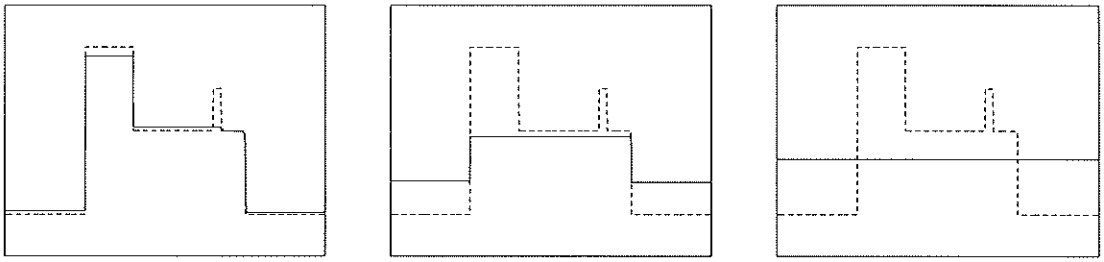
We could also solve a modified form of the constrained problem (2.2),

$$\min_u \int \alpha(\vec{x})|\nabla u(\vec{x})|\, d\vec{x} \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2. \tag{2.12}$$

In both problems, $\alpha(\vec{x})$ can be chosen to vary based on the local features in the function as well as on the information to be extracted from the function. The theory developed in Chapter 3 this Dissertation provides a simple foundation upon which to appropriately base our choice of α in doing adaptive regularization.



(a) True, noisy functions. (b) Denoised, $\alpha = .0001$. (c) Denoised, $\alpha = .001$.



(d) Denoised, $\alpha = .01$. (e) Denoised, $\alpha = .1$. (f) Denoised, $\alpha = 1$.

Figure 2.4: Example 2.1, a demonstration of the effects of choice of regularization parameter α on the restored image. when solving (2.1).

2.5 Need for Better Understanding of TV Regularization

As previously mentioned, numerical results have shown that TV regularization is quite useful in image restoration. However, while a fair amount of effort has been put into developing fast schemes for numerically implementing TV regularization (e.g. [10], [11], [22] or [26]) and into proving the existence of solutions (e.g. [14]), a very limited amount of study has been devoted to understanding more precisely how TV regularization affects an image or other function.

There is a need for a more precise understanding of how TV minimizing regularization affects an image or in general a function. Understanding TV regularization in a more quantitative way is important for two reasons: first, if we understand more precisely how it affects an image (i.e. why and how it works), then we can better exploit the properties of TV regularization which make it a useful approach to image restoration; second, if we *do not* have a reasonable understanding of TV regularization, it can lead to ineffective and even disastrous results.

Example 2.1 Consider the example shown in Figure 2.4, in which we demonstrate the problems that can occur when choosing α in solving (2.1) without an understanding of how its value affects the restored function. In this example the dashed line is the true function, the dotted line is the noisy function u_0 and the solid line is the restored function u. We must find u, the minimizer of (2.1), which depends on u_0 and our choice of α. An unintelligent choice of α can result in too little noise removal or in too much detail loss with increasing values of α. For α too small, not enough smoothing is done to remove the noise, while for α too large, more noise is removed but we lose too much of the detail. As α becomes larger, features with smaller scale are lost due to the condition of trying to lower

the total variation of the function. (In this sense, TV regularization is a type of multi-scale function regularization, [25] with the scale dependent on α.)

The key is to understand how to select an appropriate regularization parameter α, in the cases both when α is constant and when $\alpha(\vec{x})$ is spatially varying, when doing adaptive restoration. A better quantitative understanding of TV regularization could be quite useful, and indeed is somewhat necessary to provide a stronger mathematical justification of using TV minimizing image restoration schemes.

In order to determine the "optimal" value of α, whether constant or spatially varying, it is necessary to establish a relationship between the regularization parameter α and the resulting effects—both qualitative and quantitative—which occurs. In [12], it is suggested that the effectiveness of TV based noise removel depends on the "mass" (essentially the area and intensity of the feature) of the function relative to the total variation of the function. It could be very useful to define a more precise relationship between the amount of smoothing which occurs to an feature of given "mass" or area, and of a given scale (which corresponds to its total variation). By developing a more precise relationship, it becomes possible to make a more intelligent and appropriate choice of the regularization parameter α.

It is due to the nonlinearity of problems (2.1) and (2.2) that TV regularization is especially adept at recovering edges in an image while simultaneously not penalizing smooth edges. It is also because of the nonlinearity that it is impossible for the *general* case—and certainly non-trivial even in simple cases—to develop an analytic theory which describes the effects of TV regularization in a simple and useful way. However, in the following chapter we are able to develop a precise and simple quantitative theory of TV regularization by finding solutions

to the TV minimizing regularization problem in some specific and very important cases. These results are also useful in better understanding and predicting how TV regularization affects images (or other functions) in the general case.

CHAPTER 3

Theoretical Analysis of Total Variation Minimizing Function Regularization

In this chapter we give a more complete, mathematically founded analysis of the effects of TV minimizing function regularization. Our theory shows that

1. TV regularization of a piecewise constant radially symmetric function, whether noise-free or noise-contaminated, results *exactly* in a piecewise constant function, with edge location being preserved *exactly*, under certain conditions;

2. function intensity change is *exactly* inversely proportional to local feature *scale*, (this helps to explain why TV regularization can remove smaller-scaled noise, while leaving larger-scaled features essentially intact) is independent of original intensity, and is directly proportional to the regularization parameter α;

3. for smooth radially symmetric function feaatures, function intensity change is *exactly* inversely proportional to radial position and directly proportional to α;

4. TV regularization is somewhat local in its effects on image features, which not only gives us a better understanding of how TV regularization affects an image, but it is potentially quite useful in developing faster numerical

schemes. This is important because of the relatively high computational costs in solving the TV minimizing function regularization problem, due to the nonlinearity of the problem.

3.1 Exact Solution of the TV Regularization Problem for Piecewise Constant Functions

We first analyze the effects of TV regularization on piecewise constant functions. We do this because image features are often partially or entirely piecewise constant, and because our results can then be extended to more general cases. We consider the unconstrained problem with spatially varying regularization parameter (2.11). This makes the effects of TV regularization more obvious (than if we were to solve (2.12)), since the regularization is directly dependent on the regularization parameter $\alpha(\vec{x})$. Our results, of course, are still equally useful for (2.12), in which a regularization parameter is implicitly present in the form of the Lagrange multiplier.

In this section we prove that TV regularization results in intensity changes that are inversely proportional to the *scale* of individual features and directly proportional to α. In Section 3.1.1 we look at two simple examples to motivate the analytic results which we subsequently prove in Section 3.1.2. In Section 3.1.3 we interpret our theoretical results for for three important practical cases—in R^1, R^2 and R^3, with constant α—and give numerical examples which verify our theory for each of these cases.

3.1.1 Two Motivating Examples

To motivate our subsequent discussion and results, we consider two examples to illustrate the basic effects of TV regularization on piecewise constant functions. We note that a single noise element could be thought of as a piecewise constant function feature with width of a single pixel, so that our results are useful for understanding how TV regularization affects noise, as well as how it affects the image features.

Example 3.1 We first consider a simple, noise-free example in R^1. For this example, the function u_0 is defined as

$$u_0(x) = \begin{cases} 0 & \text{for } x \in \Omega_1 \\ 1 & \text{for } x \in \Omega_2 \\ 0 & \text{for } x \in \Omega_3 \end{cases} \tag{3.1}$$

as shown in Figure 3.1(a). For a piecewise constant function u in R^1, the total variation $TV(u)$ is simply the absolute sum of the jumps in u. Based on intuitition and past experience (and subsequently proven in this chapter), with u_0 as defined as in (3.1), it is reasonable to expect that the regularized function u will be

$$u(x) = \begin{cases} 0 + \delta_1 & \text{for } x \in \Omega_1 \\ 1 + \delta_2 & \text{for } x \in \Omega_2 \\ 0 + \delta_3 & \text{for } x \in \Omega_3 \end{cases} \tag{3.2}$$

where δ_1, $\delta_3 > 0$ and $\delta_2 < 0$, as illustrated in Figure 3.1(a). *Throughout this Dissertation we will generally denote by δ_i the change in function intensity level in region Ω_i due to TV regularization.* For this example we assume that u has this form (3.2) and that α is constant, so that solving for u in (2.1) is equivalent

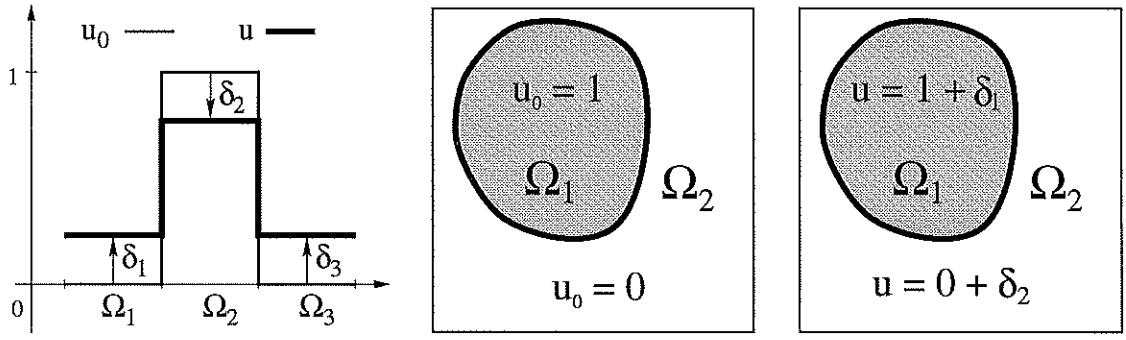
to solving for $\{\delta_i\}_{i=1}^3$ in the problem

$$\min_{\delta_1,\delta_2,\delta_3} \quad \{ \tfrac{1}{2} \sum_{i=1}^{3} |\Omega_i| \delta_i^2 + \alpha (1 + \delta_2 - \delta_1 + 1 + \delta_2 - \delta_3) \}.$$

We differentiate with respect to δ_i and easily find the resulting equations and solutions

$$\begin{aligned} |\Omega_1|\,\delta_1 - \alpha = 0 \quad &\Longrightarrow \quad \delta_1 = \tfrac{1}{|\Omega_1|}\,\alpha, \\ |\Omega_2|\,\delta_2 + 2\alpha = 0 \quad &\Longrightarrow \quad \delta_2 = -\tfrac{2}{|\Omega_2|}\,\alpha, \\ |\Omega_3|\,\delta_3 - \alpha = 0 \quad &\Longrightarrow \quad \delta_3 = \tfrac{1}{|\Omega_3|}\,\alpha. \end{aligned} \tag{3.3}$$

We obtained this result *with the assumption* that u would be as defined in (3.2). Note that in each region Ω_i that δ_i is inversely proportional to the *scale*, or in the R^1 case the width, of the region $|\Omega_i|$, and directly proportional to the regularization parameter α. We will subsequently show that the resulting function u would have been be identical if there had been noise present in u_0, under certain conditions.



(a) Piecewise constant function in R^1, before and after TV regularization.

(b) Piecewise constant function in R^2, *before* TV regularization.

(c) Piecewise constant function in R^2, *after* TV regularization.

Figure 3.1: The basic effects of TV regularization on (noise-free) piecewise constant functions in R^1 and R^2.

Example 3.2 We next consider a simple example in R^2. For this example, the function u_0 is defined as

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

Given two regions Ω_i and Ω_j we define the intersection of their two boundaries as

$$\partial\Omega_{i,j} \equiv \partial\Omega_i \cap \partial\Omega_j.$$

We state a result, a proof for which can be found in [24], which describes the total variation of a piecewise constant function in R^2 or R^3.

Lemma 3.1 *Let Ω be a compact domain in R^2 or R^3. Let u be defined in Ω as*

$$u(\vec{x}) = \begin{cases} U_1 & \textit{for } \vec{x} \in \Omega_1 \\ U_2 & \textit{for } \vec{x} \in \Omega_2 = \Omega - \Omega_1 \end{cases}.$$

Then the total variation of u is

$$TV(u) = |U_1 - U_2|\,|\partial\Omega_{1,2}|.$$

For now we simplify this example by assuming that u assumes the same shape as u_0; that is, we assume that the boundary $\partial\Omega_1$ is not deformed due to TV regularization and that u is piecewise constant, so that

$$u(\vec{x}) = \begin{cases} 1 + \delta_1 & \text{for } \vec{x} \in \Omega_1 \\ 0 + \delta_2 & \text{for } \vec{x} \in \Omega_2 \end{cases}.$$

We also assume for this example that α is constant. The fitting error between u and u_0 is

$$\int_\Omega [u(\vec{x}) - u_0(\vec{x})]^2 \, d\vec{x} = |\Omega_1| \, \delta_1^2 + |\Omega_2| \, \delta_2^2,$$

and according to Lemma 3.1, the total variation of u is $TV(u) = (1 + \delta_1 - \delta_2) \, |\partial\Omega_{1,2}|$. To solve for u, we can solve for $\{\delta_i\}_{i=1}^2$ in

$$\min_{\delta_1, \delta_2} \quad \{ \tfrac{1}{2}(|\Omega_1| \, \delta_1^2 + |\Omega_2| \, \delta_2^2) + \alpha \, (1 + \delta_1 - \delta_2) \, |\partial\Omega_{1,2}| \, \}.$$

We differentiate with respect to each δ_i, and easily find

$$\begin{array}{rcl} |\Omega_1| \, \delta_1 + \alpha \, |\partial\Omega_{1,2}| = 0 & \Longrightarrow & \delta_1 = -\frac{|\partial\Omega_{1,2}|}{|\Omega_1|} \, \alpha \\ |\Omega_2| \, \delta_2 - \alpha \, |\partial\Omega_{1,2}| = 0 & \Longrightarrow & \delta_2 = \frac{|\partial\Omega_{1,2}|}{|\Omega_2|} \, \alpha \end{array} . \tag{3.4}$$

Often Ω_1 is completely contained within Ω, so that $\partial\Omega_{1,2} = \partial\Omega_1$, as in this example. If this is the case, then for an image feature, generically labeled as Ω, which is of constant intensity level, we have (taking δ as a non-negative value)

$$\boxed{\delta = \frac{|\partial\Omega|}{|\Omega|} \, \alpha \, .} \tag{3.5}$$

We define the scale, s, of this image feature as the ratio of the area (volume in R^3) of the feature to its boundary length (surface area in R^3); that is,

$$\boxed{scale = \frac{|\Omega|}{|\partial\Omega|} \, .} \tag{3.6}$$

For example, in R^2 a circle of radius r would have scale

$$scale = \frac{\pi r^2}{2\pi r} = \frac{r}{2},$$

so that larger circles have larger scales which are linearly proportional to r, which makes sense intuitively. With this definition (3.6) of scale, (3.5) can be rewritten as

$$\boxed{\delta = \frac{\alpha}{scale}\,,} \tag{3.7}$$

so that *the change δ in image intensity is inversely proportional to the scale.* The relationship (3.5) was found with assumptions that trivialize the problem. We will subsequently show that (3.5) holds without relying on those assumptions.

This fundamental relationship (3.7) between scale and intensity change helps to explain in a very basic way how and why TV regularization is effective in denoising an image: *TV regularization causes (smaller-scaled) noise to be removed while larger-scaled image features are relatively unaffected.* One example of this is the fact that TV image restoration is especially well suited for denoising images with large-scaled, blocky features, as concluded in [12].

Another useful application of (3.7) is that if we can locally measure the change in intensity level due to TV regularization, then we can find the scale of various image features, by re-writing equation (3.7) as

$$scale = \frac{\alpha}{\delta}\,. \tag{3.8}$$

In Chapter 5, and we use (3.8) to develop automatic scale recognition schemes, which coupled with (3.7) is used to construct adaptive image restoration schemes.

3.1.2 Formulae Describing Effects of TV Regularization on Piecewise Constant Functions: in R^1; in R^2 and R^3 with Radial Symmetry

In this section we develop and prove the mathematical formulae which describe how TV regularization affects a piecewise constant image—noise-free or noisy—

with features (i.e. regions of piecewise constancy) of any given scale, for functions in R^1 and for radially symmetric functions in R^2 and R^3 (we later apply our results to non-radially symmetric functions). In Section 3.1.2.1 we show how the minimization problem (2.11) in $R^d, d = 2, 3,$ can be transformed into a minimization problem in R^1. In Section 3.1.2.2 we prove and discuss the formulae which describe the effects of TV regularization.

3.1.2.1 Radial Symmetry in R^2 and R^3

One special class of functions in R^2 and R^3 are those that are radially symmetric. For the moment, we consider the R^2 case. Because of radially symmetry, we work in polar coordinates:

$$\begin{aligned} x &= r\,\cos\theta, \\ y &= r\,\sin\theta, \\ r &= \sqrt{x^2+y^2}\,, \end{aligned}$$

so that

$$\begin{aligned} r_x &= \frac{x}{r}, \\ r_y &= \frac{y}{r}. \end{aligned}$$

Then since $u_\theta = 0$ (due to radial symmetry), we have

$$u_x = u_r\,r_x + u_\theta\,\theta_x = u_r\,r_x = \frac{x}{r}\,u_r.$$

Similarly,

$$u_y = \frac{y}{r}\,u_r,$$

so that

$$\begin{aligned} |\nabla u(x,y)| &= \sqrt{u_x^2 + u_y^2} \\ &= \sqrt{(\frac{x}{r}\,u_r)^2 + (\frac{y}{r}\,u_r)^2} \\ &= \sqrt{\frac{x^2+y^2}{r^2}\,u_r^2} \\ &= |u_r(r)|. \end{aligned}$$

To simplify the analysis, we take Ω to be the unit circle, and assume that $\alpha(x,y)$ is also radially symmetric, so that

$$\int_\Omega \alpha(x,y)\,|\nabla u(x,y)|\,dx\,dy \;=\; \int_0^1 2\pi\, r\,\alpha(r)\,|u_r(r)|\,dr. \tag{3.9}$$

If u_0 is radially symmetric then the fitting error in the region of the unit circle Ω is given by

$$\int_\Omega [u(x,y) - u_0(x,y)]^2\,dx\,dy \;=\; \int_0^1 2\pi\, r\,[u(r) - u_0(r)]^2\,dr. \tag{3.10}$$

Recall that the problem (2.11) we wish to solve in R^1 (taking the domain to be $[0,1]$) is to find the argument u of

$$\min_u \quad \{ \int_0^1 \tfrac{1}{2}[\,u(x) - u_0(x)\,]^2 + \alpha(x)|u_x(x)|\;dx\;\}. \tag{3.11}$$

In R^2, if u and u_0 are radially symmetric (and assuming that $\alpha(\vec{x}) = \alpha(r)$ is radially symmetric), then by using (3.9) and (3.10) we see that the problem is to find the argument of

$$\min_u \quad \{ \int_0^1 2\pi\, r\,\{\,\tfrac{1}{2}[\,u(r) - u_0(r)\,]^2 + \alpha(r)|u_r(r)|\,\}\;dr\;\}. \tag{3.12}$$

This is easily extended to R^3, in which the problem is find the argument of

$$\min_u \quad \{ \int_0^1 4\pi\, r^2 \, \{ \tfrac{1}{2}[\, u(r) - u_0(r)\,]^2 + \alpha(r)|u_r(r)| \, \}\, dr \, \}. \tag{3.13}$$

The problems in the radially symmetric R^d cases have reduced to problems that are almost exactly the same as the problem in the R^1 case. Problems (3.11) - (3.13) differ only by a multiplicative constant (which has no effect in the minimization problem) and the weighting factor r or r^2 in the integrand.

In general, with radially symmetry, the minimization problem (2.11) in R^d, for $1 \leq d \leq 3$, can be written

$$\boxed{\min_u \quad \{ \int \beta_d\, r^{d-1} \, \{ \tfrac{1}{2}[\, u(r) - u_0(r)\,]^2 + \alpha(r)|u_r(r)| \, \}\, dr \, \}} \tag{3.14}$$

where

$$\beta_d = \begin{cases} 1 & \text{for } d = 1 \\ 2\pi & \text{for } d = 2 \\ 4\pi & \text{for } d = 3 \end{cases} .$$

To make our subsequent notation more obvious (particularly in the subsequent proofs), we point out that given an interval $\Omega_r = [0, r]$ in R^1, or a circle Ω_r in R^2 or sphere Ω in R^3 with radius r, then we have

$$\begin{aligned} |\partial\Omega_r| &= \beta_d\, r^{d-1}, \\ |\Omega_r| &= \frac{\beta_d}{d}\, r^d, \end{aligned}$$

for $1 \leq d \leq 3$. Notice that $|\partial\Omega_r|$ is actually the weighting term in each integrand

in (3.11) - (3.13), so that (3.14) could also be written as

$$\boxed{\min_u \quad \{ \int |\partial\Omega_r| \, \{ \tfrac{1}{2}[\, u(r) - u_0(r)\,]^2 + \alpha(r)|u_r(r)| \, \} \, dr \; \}} \tag{3.15}$$

The notation we use in the subsequent proofs is given in Definition 3.1.

Definition 3.1 (Explanation of notation $|\partial\Omega_{i,i+1}|$ and $|\Omega_i|$) *Let the region $[r_{i-1}, r_i]$ be the R^1 region which corresponds to the radially symmetric region Ω_i in R^d, $d = 2, 3$. Then*

$$|\partial\Omega_{i,i+1}| \;=\; |\partial\Omega_{r_i}| \;=\; \beta_d \, r_i^{d-1},$$

$$|\Omega_i| \;=\; |\Omega_{r_i}| - |\Omega_{r_{i-1}}| \;=\; \frac{\beta_d}{d}\,(r_i^d - r_{i-1}^d).$$

3.1.2.2 Formulae and Corresponding Proofs

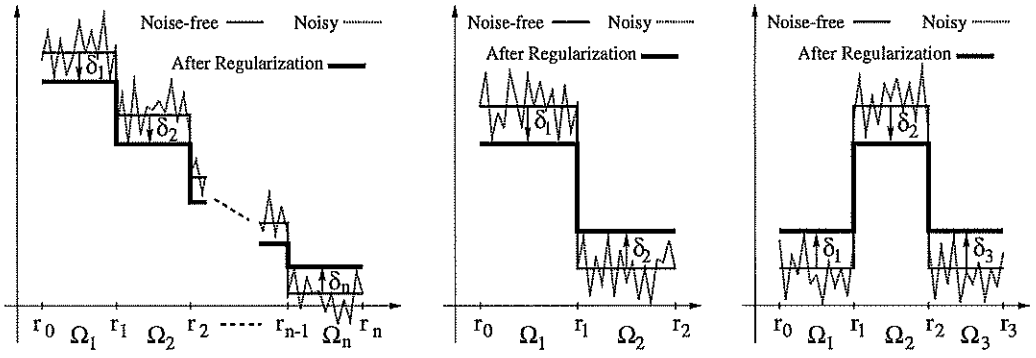
Piecewise constant functions are comprised of three types of features: *extremum*, *steps* and *boundary* regions. We now give formulae which describe the effects of TV regularization on each of these types of features. The basic effects of TV regularization are (1) preserved edge location and (2) change in image intensity which is inversely proportional to local feature scale and directly proportional to the regularization parameter α. Theorem 3.1 describes the effects of TV regularization on a monotone step function, which is illustrated in Figure 3.2(a). Corollary 3.1 describes these results for the special case when there is a single step, illustrated in Figure 3.2(b). Using Corollary 3.1 we prove Theorem 3.2, which describes the effects of TV regularization on a hat function, that is, a function with a single extremum, as illustrated in Figure 3.2(c). Theorem 3.3 summarizes the results of Theorem 3.1, Corollary 3.1 and Theorem 3.2 for general piecewise constant

functions, whether noise-free or noisy.

In Theorem 3.1 we make the following assumptions:

- $u_0(\vec{x}) = u_0(r)$ is radially symmetric;
- $u_0(r)$, without noise, is piecewise constant in each region Ω_i of the domain, so that if noise-free $u_0(r) \equiv U_i$ in Ω_i; however, noise may be present in u_0, in which case $mean(u_0(r)) = U_i$ in Ω_i (see (3.16));
- $\alpha(r)$ is sufficiently large to remove the noise in the image, and is sufficiently small to not remove any edges (see (3.20)).

Note that no assumptions are made about u, except that it is radially symmetric (which is a natural assumption since u_0 is radially symmetric). We also point out that this theorem is stated and proved for monotonically *decreasing* step functions, but the results (and corresponding proof) are analogous for monotonically *increasing* step functions, *which results we consequently assume without proof.*



(a) Monotone Step Function

(b) Step Function, Single Step.

(c) Hat Function: Single Extremum

Figure 3.2: The effects of TV regularization on noisy piecewise constant functions.

Theorem 3.1 (Monotone Step Function) *Let $u_0(r)$ be the function defined on $[r_0, r_n]$ which corresponds to radially symmetric function $u_0(\vec{x})$ in R^d, for $1 \leq d \leq 3$. Let*

$$U_i \equiv \frac{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, u_0(r) \, dr}{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, dr} \quad \text{for } 1 \leq i \leq n, \tag{3.16}$$

and assume that $U_i \geq U_{i+1}$, $1 \leq i \leq n-1$. Assume that

$$\begin{aligned} \alpha(r_i) &= \alpha_i && \text{for } 1 \leq i \leq n-1, \\ \alpha(r) &\geq \alpha_{max} \equiv \max_{1 \leq i \leq n-1} \{\alpha_i\} && \text{elsewhere in } [r_0, r_n]. \end{aligned} \tag{3.17}$$

Then the solution to (3.14) is given by

$$u(r) = U_i + \delta_i \quad \text{for } r \in [r_{i-1}, r_i], \ 1 \leq i \leq n \tag{3.18}$$

where

$$\delta_i = \begin{cases} \frac{-\alpha_1 |\partial\Omega_{1,2}|}{|\Omega_1|} & \text{for } i = 1 \\ \frac{\alpha_{i-1}|\partial\Omega_{i,i-1}| - \alpha_i |\partial\Omega_{i,i+1}|}{|\Omega_i|} & \text{for } 2 \leq i \leq n-1 \\ \frac{\alpha_{n-1}|\partial\Omega_{n-1,n}|}{|\Omega_n|} & \text{for } i = n \end{cases} \tag{3.19}$$

(see Definition 3.1 for explanation of notation) if the following conditions hold:

$$\begin{aligned} &(i) \quad U_i + \delta_i \geq U_{i+1} + \delta_{i+1} && \text{for } 1 \leq i \leq n-1 \\ &(ii) \quad |\delta_i| \geq \max_{r_{i-1} \leq r \leq r_i} |u_0(r) - U_i| && \text{for } 1 \leq i \leq n \end{aligned} \tag{3.20}$$

To prove Theorem 3.1, we will need the following technical result.

Lemma 3.2 *In R^d, $1 \leq d \leq 3$, let $u(r)$ and $\epsilon(r)$ be functions defined on $[r_0, r_n]$. If $u(r)$ is a monotically decreasing step function with discontinuities at*

$S = \{r_i\}_{i=1}^{n-1}$, *then* [1]

$$\int_{[r_0,r_n]} \beta_d\, r^{d-1}\, \alpha(r)\, |u_r(r) + \epsilon_r(r)|\, dr \;\geq$$
$$\int_{[r_0,r_n]} \beta_d\, r^{d-1}\, \alpha(r)\, |u_r(r)|\, dr + \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)\, |\epsilon_r(r)|\, dr$$
$$- \int_S \beta_d\, r^{d-1}\, \alpha(r)\epsilon_r(r)\, dr.$$

Proof of Lemma 3.2 Let

$$S_1 = \{r_i : \int_{\{r_i\}} |u_r(r)| - \epsilon_r(r)\, dr \geq 0,\ \ 1 \leq i \leq n-1\},$$
$$S_2 = \{r_i : \int_{\{r_i\}} |u_r(r)| - \epsilon_r(r)\, dr < 0,\ \ 1 \leq i \leq n-1\}.$$

Then we have

$$\int_{[r_0,r_n]} \alpha(r)\, |\, u_r(r) + \epsilon_r(r)\, |\, dr$$

$$= \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)\, |\epsilon_r(r)|\, dr \;+\; \int_S \beta_d\, r^{d-1}\, |u_r(r) + \epsilon_r(r)|\, dr$$
(since $u_r(r) = 0$ in $[r_0, r_n] - S$)

$$= \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r) |\epsilon_r(r)|\, dr \;+\; \int_{S_1} \beta_d\, r^{d-1}\, \alpha(r)\, [\, |u_r(r)| - \epsilon_r(r)\,]\, dr$$
$$+ \int_{S_2} \beta_d\, r^{d-1}\, \alpha(r)\, [\, \epsilon_r(r) - |u_r(r)|\,]\, dr$$
(since $\int_{\{r_i\}} u_r(r)\, dr < 0$ for $1 \leq i \leq n-1$)

---

[1] *The integral over a single point* $\int_{r_i} \beta_d\, r^{d-1}\, \alpha(r)\, \epsilon_r(r)\, dr$ *is identical to* $\beta_d\, r_i^{d-1}\, \alpha(r_i)\, jump(\epsilon(r_i))$, *which* $= 0$ *if* $\epsilon(r)$ *is continuous at* r_i. *Integrating over a set of points* S *is analogous.*

$$\geq \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)|\epsilon_r(r)|\, dr \;+\; \int_S \beta_d\, r^{d-1}\, \alpha(r)\,[\,|u_r(r)| - \epsilon_r(r)\,]\, dr$$

$$= \int_{[r_0,r_n]} \beta_d\, r^{d-1}\, \alpha(r)\,|u_r(r)|\, dr + \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)|\epsilon_r(r)|\, dr$$
$$- \int_S \beta_d\, r^{d-1}\, \alpha(r)\,\epsilon_r(r)\, dr.$$

□

Proof of Theorem 3.1 Define

$$f(v) = \int_{r_0}^{r_n} \beta_d\, r^{d-1} \{\tfrac{1}{2}[v(r) - u_0(r)]^2 + \alpha(r)\,|v_r(r)|\}\, dr.$$

We show that $f(u+\epsilon) > f(u)$ for any $\epsilon = \epsilon(r)$ (unless $\epsilon \equiv 0$), where u is as defined in (3.18). Using Lemma 3.2 we have

$$\begin{aligned} f(u+\epsilon) &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \{\, \tfrac{1}{2}[u(r) + \epsilon(r) - u_0(r)]^2 + \alpha(r)|u_r(r) + \epsilon_r(r)|\,\}\, dr \\ &\geq f(u) + g(u,\epsilon) \end{aligned}$$

where

$$\begin{aligned} g(u,\epsilon) &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \epsilon(r)[u(r) - u_0(r)]\, dr + \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)\,|\epsilon_r(r)|\, dr \\ &\quad - \int_S \beta_d\, r^{d-1} \alpha(r)\epsilon_r(r)\, dr + \int_{r_0}^{r_n} \tfrac{1}{2}\, \beta_d\, r^{d-1}\, [\epsilon(r)]^2\, dr \end{aligned}$$

We must show that $g(u,\epsilon) > 0$.

We first define $\tilde{\epsilon}(r)$ to be a modification of ϵ which is continuous at each of

the points in $S = \{r_i\}_{i=1}^{n-1}$. Where $S_i = \{r_j\}_{j=1}^{i}$, define

$$\tilde{\epsilon}(r) \equiv \epsilon(r) - \int_{S_{i-1}} \epsilon_r(r)\, dr \equiv \epsilon(r) - \sum_{j=1}^{i-1} jump(\epsilon(r_j)) \quad \text{for } r \in [r_{i-1}, r_i], \quad 1 \le i \le n$$

(notice that $\tilde{\epsilon}(r) \equiv \epsilon(r)$ for $r \in [r_0, r_1]$) so that

$$\begin{aligned} \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)\, |\epsilon_r(r)|\, dr &= \int_{[r_0,r_n]-S} \beta_d\, r^{d-1}\, \alpha(r)\, |\tilde{\epsilon}_r(r)|\, dr \\ &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \alpha(r)\, |\tilde{\epsilon}_r(r)|\, dr. \end{aligned} \tag{3.21}$$

since $\tilde{\epsilon}(r)$ is continuous at each $r_i \in S$. We also have

$$\begin{aligned} &\int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \epsilon(r)[u(r) - u_0(r)]\, dr \\ &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr \\ &\quad + \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, [\epsilon(r) - \tilde{\epsilon}(r)][u(r) - u_0(r)]\, dr \\ &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr \\ &\quad + \sum_{i=1}^{n} \int_{r_{i-1}}^{r_i} \beta_d\, r^{d-1}\, [\epsilon(r) - \tilde{\epsilon}(r)][u(r) - u_0(r)]\, dr \\ &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr \\ &\quad + \sum_{i=2}^{n} \int_{r_{i-1}}^{r_i} \beta_d\, r^{d-1}\, [\int_{S_{i-1}} \epsilon_r(r)\, dr][u(r) - u_0(r)]\, dr \\ &= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr \end{aligned}$$

$$+ \sum_{i=2}^{n-1} [\,\alpha_{i-1}|\partial\Omega_{i,i-1}| - \alpha_i|\partial\Omega_{i,i+1}|][\int_{S_{i-1}} \epsilon_r(r)\,dr]$$
$$+ \alpha_{n-1}|\partial\Omega_{n-1,n}| \int_{S_{n-1}} \epsilon_r(r)\,dr \qquad \text{(using (3.16), (3.18) and (3.19))}$$

$$= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\,dr + \sum_{i=1}^{n-1} \int_{\{r_i\}} \beta_d\, r^{d-1}\, \alpha(r)\, \epsilon_r(r)\,dr$$

$$= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\,dr + \int_S \beta_d\, r^{d-1}\, \alpha(r)\, \epsilon_r(r)\,dr. \tag{3.22}$$

Next we define

$$\tilde{\tilde{\epsilon}}(r) \equiv \tilde{\epsilon}(r) - \tilde{\epsilon}(r_{n-1})$$

so that

$$\tilde{\tilde{\epsilon}}(r_{n-1}) = 0 \implies \tilde{\tilde{\epsilon}}(r) = \int_{t=r_{n-1}}^{r} \tilde{\tilde{\epsilon}}_t(t)\,dt,$$

and $\tilde{\tilde{\epsilon}}_r(r) = \tilde{\epsilon}_r(r)$ leads to

$$\int_{r_0}^{r_n} \beta_d\, r^{d-1} \alpha(r)|\tilde{\epsilon}_r(r)|\,dr = \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \alpha(r)|\tilde{\tilde{\epsilon}}_r(r)|\,dr. \tag{3.23}$$

Then since

$$\int_{r_0}^{r_n} \beta_d\, r^{d-1}\, [u(r) - u_0(r)]\,dr = 0 \qquad \text{(using (3.16), (3.18) and (3.19))}$$

we have

$$\int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\,dr$$

$$= \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\tilde{\epsilon}}(r)[u(r) - u_0(r)]\,dr$$

$$= \int_{r=r_0}^{r_n} \beta_d \, r^{d-1} \, [\, \int_{t=r_{n-1}}^{r} \tilde{\tilde{\epsilon}}_t(t) \, dt \,] \, [u(r) - u_0(r)] \; dr$$

$$= \int_{t=r_0}^{r_n} \beta_d \, h(t) \, \tilde{\tilde{\epsilon}}_t(t) \, dt$$

where

$$h(t) = \begin{cases} -\int_{r=r_0}^{t} r^{d-1} \, [u(r) - u_0(r)] \, dr & \text{for } r_0 \leq t < r_{n-1} \\ \\ \int_{r=t}^{r_n} r^{d-1} \, [u(r) - u_0(r)] \, dr & \text{for } r_{n-1} \leq t \leq r_n \end{cases} .$$

We next show that

$$|h(t)| \leq \alpha_{max} \, t^{d-1} \quad \text{for } r_0 \leq t \leq r_n$$

so that

$$-|h(t)| \geq -\alpha_{max} \, t^{d-1} \quad \text{for } r_0 \leq t \leq r_n. \tag{3.24}$$

If $r_0 \leq t < r_1$, then

$$|h(t)| = |\int_{r_0}^{t} r^{d-1} \, [u(r) - u_0(r)] \, dr \,|$$

$$= \alpha_1 \frac{r_1^{d-1}(t^d - r_0^d)}{r_1^d - r_0^d}$$

$$\leq \alpha_1 \, t^{d-1}$$

$$\leq \alpha_{max} \, t^{d-1} \qquad \text{(using (3.17)).}$$

If $r_j \le t < r_{j+1},\ 1 \le j \le n-2$, then

$$\begin{aligned}
|h(t)| &\le \{\int_{r_0}^{r_1} + \sum_{k=1}^{j-1}\int_{k}^{k+1} + \int_{r_j}^{t}\}\,\beta_d\, r^{d-1}\,|u(r)-u_0(r)|\,dr \\
&= \alpha_1 r_1^{d-1} + \sum_{k=1}^{j-1}\alpha_k(r_{k+1}^{d-1}-r_k^{d-1}) + \alpha_j(r_{j+1}^{d-1}-r_j^{d-1})\frac{t^d-r_j^d}{r_{j+1}^d-r_j^d} \\
&\qquad \text{(using (3.16), (3.18) and (3.19))} \\
&\le \alpha_{max}\big[r_1^{d-1} + \sum_{k=1}^{j-1}(r_{k+1}^{d-1}-r_k^{d-1}) + (r_{j+1}^{d-1}-r_j^{d-1})\frac{t^d-r_j^d}{r_{j+1}^d-r_j^d}\big] \\
&\qquad \text{(using (3.17))} \\
&\le \alpha_{max}(r_j^{d-1} + t^{d-1} - r_j^{d-1}) \\
&= \alpha_{max}\, t^{d-1}.
\end{aligned}$$

If $r_{n-1} \le t \le r_n$, then

$$\begin{aligned}
|h(t)| &= |\int_{r=t}^{r_n}\beta_d\, r^{d-1}\,[u(r)-u_0(r)]\,dr| \\
&= \alpha_{n-1}\frac{r_{n-1}^{d-1}(r_n^d-t^d)}{r_n^d-r_{n-1}^d} \qquad \text{(using (3.16), (3.18) and (3.19))} \\
&\le \alpha_{n-1}\, t^{d-1} \\
&\le \alpha_{\max}\, t^{d-1} \qquad \text{(using (3.17))}.
\end{aligned}$$

So

$$\int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr$$

$$= \int_{t=r_0}^{r_n} \beta_d\, h(t)\, \tilde{\tilde{\epsilon}}_t(t)\, dt$$

$$\geq -\int_{r_0}^{r_n} \beta_d\, |h(t)|\, |\tilde{\tilde{\epsilon}}_t(t)|\, dt$$

$$\geq -\alpha_{max} \int_{r_0}^{r_n} \beta_d\, t^{d-1}\, |\tilde{\tilde{\epsilon}}_t(t)|\, dt \qquad \text{(using (3.24))}$$

$$\geq -\int_{r_0}^{r_n} \beta_d\, t^{d-1} \alpha(t)\, |\tilde{\tilde{\epsilon}}_t(t)|\, dt \qquad \text{(using (3.17))}$$

$$= -\int_{r_0}^{r_n} \beta_d\, t^{d-1} \alpha(t)\, |\tilde{\epsilon}_t(t)|\, dt \qquad \text{(using (3.23))} \qquad (3.25)$$

Finally we have

$$g(u,\epsilon) = \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \tilde{\epsilon}(r)[u(r) - u_0(r)]\, dr + \int_{r_0}^{r_n} \beta_d\, r^{d-1}\, \alpha(r) |\tilde{\epsilon}_r(r)|\, dr$$
$$+ \int_{r_0}^{r_n} \tfrac{1}{2} \beta_d\, r^{d-1}\, [\epsilon(r)]^2\, dr \qquad \text{(using (3.21) and (3.22))}$$

$$\geq \int_{r_0}^{r_n} \tfrac{1}{2} \beta_d\, r^{d-1}\, [\epsilon(r)]^2\, dr \qquad \text{(using (3.25))}$$

$$> 0 \qquad \text{(unless } \epsilon \equiv 0\text{)}.$$

So we've shown that $f(u+\epsilon) = f(u) + g(u,\epsilon) > f(u)$, unless $\epsilon \equiv 0$. Therefore u is the unique solution.

□

Condition (3.20) in Theorem 3.1 ensures what is subsequently referred to as the *α-condition*, when applying TV regularization to noise-free or noisy piecewise constant functions. The α-condition is defined qualitatively, and must be interpreted quantitatively for each individual problem.

Definition 3.2 (α-condition) *The regularization parameter $\alpha(\vec{x})$ meets the α-condition if:*

(*i*) *$\alpha(\vec{x})$ is sufficiently small that all jumps (not due to noise) in u_0 are present in u;*

(*ii*) *$\alpha(\vec{x})$ is sufficiently large that the noise is completely removed, resulting in a regularized function that is exactly piecewise constant with reduced contrast.*

We discuss a few of the implications of the α-condition. The smaller the jumps are in u_0 (i.e. the noise-free version of u_0), the smaller the upper bound on α must be in order to satisfy condition (i). Similarly, the more noise present, the larger the lower bound on α must be to satisfy (ii). For relatively simple functions, there is often a region for values of α in which (i) and (ii) are both satisfied; however, if there is too much noise relative to the size of the jumps in u_0, then the upper bound for (i) could be smaller than the lower bound for (ii), in which case there is theoretically no region for acceptable α values which give the results of Theorem 3.1 exactly.

For the theorems and corollaries given in this chapter, this region for an acceptable value of α is relatively simple to find, but it can become quite complicated to find this region *exactly* for more complex functions u_0. Consequently,

the α-*condition* is generally an *intuitive* guide rather than a set of strict bounds on $\alpha(x)$. *In practice, these theoretical bounds on* α *generally do not have to be met in order to obtain approximately the predicted results.*

An obvious corollary of Theorem 3.1 is when there is a single discontinuity, as illustrated in Figure 3.2(b).

Corollary 3.1 (Step Function, Single Step) *Let* $u_0(r)$ *be the function defined on* $[r_0, r_2]$ *which corresponds to radially symmetric function* $u_0(\vec{x})$ *in* R^d, $1 \leq d \leq 3$. *Let*

$$U_i \equiv \frac{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, u_0(r) \, dr}{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, dr} \quad \textit{for } i = 1, 2$$

and assume that $U_1 \geq U_2$. *Assume that*

$$\begin{aligned} \alpha(r_1) &= \alpha_1 \\ \alpha(r) &\geq \alpha_1 \quad \textit{elsewhere in } [r_0, r_2]. \end{aligned}$$

Then the solution to (3.14) is given by

$$u(r) = U_i + \delta_i \quad \textit{for } r \in [r_{i-1}, r_i], \quad i = 1, 2$$

where

$$\delta_i = \begin{cases} -\frac{|\partial\Omega_{1,2}|}{|\Omega_1|} \alpha_1 & \textit{for } i = 1 \\ \frac{|\partial\Omega_{1,2}|}{|\Omega_2|} \alpha_1 & \textit{for } i = 2 \end{cases}$$

if the following α-*condition (Definition 3.2) is met:*

$$\min_{r \in [r_0, r_1]} u_0(r) \geq U_1 + \delta_1 \geq U_2 + \delta_2 \geq \max_{r \in [r_1, r_2]} u_0(r).$$

The following Theorem is a natural extension of Corollary 3.1 to the case of a radially symmetric piecewise constant function with a single extremum, as illustrated in Figure 3.2(c).

Theorem 3.2 (Hat Function: Single Extremum) *Let $u_0(r)$ be the function defined on $[r_0, r_3]$ which corresponds to radially symmetric function $u_0(\vec{x})$ in R^d, $1 \le d \le 3$. Let*

$$U_i \equiv \frac{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, u_0(r) \, dr}{\int_{r_{i-1}}^{r_i} \beta_d \, r^{d-1} \, dr}, \quad \text{for } 1 \le i \le 3,$$

and assume that $U_2 \ge U_i$ for $i = 1, 3$. Assume that

$$\begin{aligned} \alpha(r_i) &= \alpha_i && \text{for } i = 1, 2 \\ \alpha(r) &\ge \alpha_{max} \equiv \max_{i=1,2}\{\alpha_i\} && \text{elsewhere in } [r_0, r_3]. \end{aligned}$$

Then the solutions to (3.14) is given by

$$u(r) = U_i + \delta_i \quad \text{for } r \in \Omega_i, \quad i = 1, 2 \tag{3.26}$$

where

$$\delta_i = \begin{cases} \frac{\alpha_1 |\partial\Omega_{1,2}|}{|\Omega_1|} & \text{for } i = 1 \\ -\frac{\alpha_2 |\partial\Omega_{2,3}| + \alpha_1 |\partial\Omega_{1,2}|}{|\Omega_2|} & \text{for } i = 2 \\ \frac{\alpha_2 |\partial\Omega_{2,3}|}{|\Omega_3|} & \text{for } i = 3 \end{cases}$$

if the following α-condition (Definition 3.2) is met:

$$\max_{r \in [r_{i-1}, r_i]} u_0(r) \le U_i + \delta_i \le U_2 + \delta_2 \le \min_{r \in [r_1, r_2]} u_0(r) \quad \text{for } i = 1, 3$$

Proof Choose $\tilde{r} \in (r_1, r_2)$ such that

$$\tilde{r}^d = \frac{\alpha_1 r_1^{d-1} r_2^d + \alpha_2 r_2^{d-1} r_1^d}{\alpha_1 r_1^{d-1} + \alpha_2 r_2^{d-1}},$$

so that

$$\frac{\alpha_1 |\partial\Omega_{1,2}| + \alpha_2 |\partial\Omega_{2,3}|}{|\Omega_2|} = \frac{\alpha_1 |\partial\Omega_{1,2}|}{|\Omega_{2_l}|} = \frac{\alpha_2 |\partial\Omega_{2,3}|}{|\Omega_{2_r}|}. \tag{3.27}$$

where $\Omega_{2_l} \subset R^d$ corresponds to $r \in [r_1, \tilde{r}]$ and where $\Omega_{2_r} \subset R^d$ corresponds to $r \in [\tilde{r}, r_2]$. Define

$$f(v) = \int_{r_0}^{r_3} \beta_d \, r^{d-1} \left\{ \tfrac{1}{2} [v(r) - u_0(r)]^2 + \alpha(r) \, |v_r(r)| \right\} dr. \tag{3.28}$$

Let

$$\begin{aligned} f_1(v) &= \int_{[r_0, \tilde{r})} \beta_d \, r^{d-1} \left\{ \tfrac{1}{2} [v(r) - u_0(r)]^2 + \alpha(r) \, |v_r(r)| \right\} dr, \\ f_2(v) &= \int_{(\tilde{r}, r_3]} \beta_d \, r^{d-1} \left\{ \tfrac{1}{2} [v(r) - u_0(r)]^2 + \alpha(r) \, |v_r(r)| \right\} dr. \end{aligned}$$

Then

$$\begin{aligned} f(v) &= f_1(v) + f_2(v) + \int_{\{\tilde{r}\}} \beta_d \, r^{d-1} \, \alpha(r) |v_r(r)| \, dr \\ \\ \Longrightarrow \min_v f(v) &= \min_v \{ \, f_1(v) + f_2(v) + \int_{\{\tilde{r}\}} \beta_d \, r^{d-1} \, \alpha(r) |v_r(r)| \, dr \, \} \\ \\ &\geq \min_v \{ \, f_1(v) + f_2(v) \, \} \\ \\ &\geq \min_v f_1(v) + \min_v f_2(v). \end{aligned}$$

By Corollary 3.1 and using (3.27) we find that f_1 and f_2 are each minimized by

u in (3.26). Since this u is continuous at $r = \tilde{r}$, then f in (3.28) is minimized by this u.

□

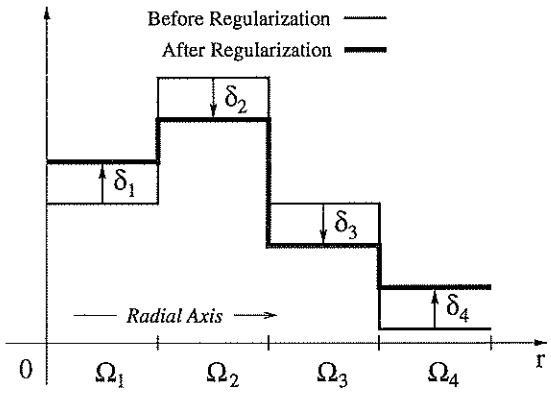
The following theorem is a summary of Theorem 3.1, Corollary 3.1 and 3.2, and can be used to understand and to some extent predict the effects of TV regularization on any piecewise constant function: in R^1, and in R^2 and R^3 with radially symmetry. Figure 3.3 illustrates the effects of TV regularization applied to a general piecewise constant in R^d, with radially symmetry for $d > 1$.

Theorem 3.3 (General Piecewise Constant Function) *Let u_0 be a radially symmetric piecewise constant function, possibly contaminated with noise, in R^d, $1 \leq d \leq 3$ (a noise-free example of which is illustrated in Figure 3.3), and that the α-condition (Definition 3.2) holds. Then the unique argument u to (3.14) is as shown in Figure 3.3, where the change in function intensity (treated as a non-negative value) for each region is given by*

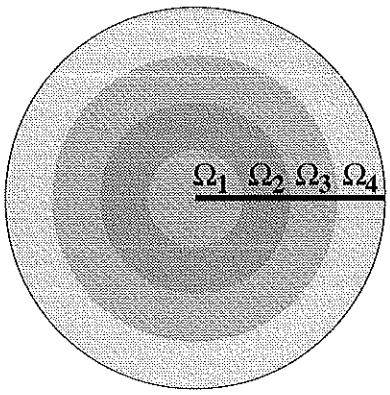
$$\delta_i = \begin{cases} \frac{\alpha_{i-1}|\partial\Omega_{i,i-1}| + \alpha_i|\partial\Omega_{i,i+1}|}{|\Omega_i|} & \textit{Extremum Regions} \quad (\Omega_1 \textit{ and } \Omega_2) \\ \frac{\alpha_{i-1}|\partial\Omega_{i,i-1}| - \alpha_i|\partial\Omega_{i,i+1}|}{|\Omega_i|} & \textit{Step Regions } (\Omega_3) \\ \frac{\alpha_{i-1}|\partial\Omega_{i,i-1}|}{|\Omega_i|} & \textit{Boundary Regions } (\Omega_4) \end{cases} \tag{3.29}$$

The sign of the change in intensity δ_i in each region will be such as to reduce contrast (i.e. total variation) in the function.

Proof This theorem can easily be proved for any piecewise constant function by dividing the function into its three types of components (extremum, step and



(a) An R^1 function, or the R^1 cross-section of (b), before and after TV regularization.



(b) An R^2 image, or the R^2 cross-section of an R^3 image, corresponding to (a).

Figure 3.3: The effects of TV regularization on piecewise constant, radially symmetric functions in R^2 and R^3.

boundary regions), as was done in proving Theorem 3.2.

□

We give a brief intuitive explanation of these results for the important case where $\alpha(\vec{x}) \equiv \alpha$ is constant, in order to make it easier to understand the effects of TV regularization. For the boundary regions (in Figure 3.3, Ω_4),

$$\delta_i = \frac{|\partial\Omega_{i,i-1}|}{|\Omega_i|}\alpha.$$

The $|\partial\Omega_{i,i-1}|$ in the numerator corresponds to the boundary size of the jump between Ω_i and its single neighbor Ω_{i-1} that affects the total variation of u.

Similarly, in extremum regions (Ω_1 and Ω_2), we have

$$\delta_i = \frac{|\partial\Omega_{i,i-1}| + |\partial\Omega_{i,i+1}|}{|\Omega_i|}\,\alpha = \frac{|\partial\Omega_i|}{|\Omega_i|}\alpha,$$

due to the boundary size of the jumps between Ω_i's two neighboring regions Ω_{i-1} and Ω_{i+1} which affect the total variation of u. Finally, in the step region, as δ_3 increases, the variation between Ω_2 and Ω_3 *increases* while the variation between Ω_4 and Ω_3 *decreases* (or conversely), with the net change in total varation of u being proportional to the *difference* of the boundary lengths, $|\partial\Omega_{3,4}| - |\partial\Omega_{2,3}|$. The precise formula then is

$$\delta_3 = \frac{|\partial\Omega_{3,4}| - |\partial\Omega_{2,3}|}{|\Omega_3|}\,\alpha.$$

(Notice that in the extremum regions, the total variation is proportional to the *sum* of the two boundary sizes.) In each region, the greater the change in intensity δ_i the more the total variation is decreased, but at the cost of increasing the fitting error. Hence a balance is found between decreasing the total variation and increasing the fitting error, depending on the value of α.

Our results show that the regularized image is the same when restoring a noise-free image and when restoring the noisy version of that image, if the mean of the noise is 0, for sufficiently large values of α. This is subsequently illustrated in Figures 3.4 - 3.6. Theorem 3.3 can be applied to any noise-contaminated piecewise constant function in $R^d, 1 \leq d \leq 3$.

3.1.3 Three Special Cases of Theorem 3.3

We conclude Section 3.1 by giving three practical corollaries of Theorem 3.3 which give the formulae (3.29) for the case when $\alpha(\vec{x}) \equiv \alpha$ is constant in R^1, R^2 and R^3.

We note that in R^1 the "size" of a boundary between two regions is $|\partial\Omega_{i,i+1}| = 1$.

3.1.3.1 Theorem 3.3 in R^1, with Constant α

Corollary 3.2 (R^1 **Piecewise Constant, Constant** α) *In* R^1 *let the conditions of Theorem 3.3 be satisfied. Let* $\Omega_i = [x_{i-1}, x_i]$ *and assume that* $\alpha(x) \equiv \alpha$ *is constant. Then (3.29) is given by*

$$\delta_i = \begin{cases} \dfrac{2}{x_i - x_{i-1}}\alpha & \textit{Extremum Regions} \\ \dfrac{0}{x_i - x_{i-1}}\alpha \equiv 0 & \textit{Step Regions} \\ \dfrac{1}{x_i - x_{i-1}}\alpha & \textit{Boundary Regions} \end{cases} \tag{3.30}$$

Corollary 3.2 can be applied to any (noise-free or noisy) piecewise constant function in R^1. We give a single example of this.

Example 3.3 For this example we define the original image u_0 as

$$u_0(x) = \begin{cases} 0.5 & \text{for } 0.00 \le x < 0.25 \quad (\Omega_1) \\ 1.0 & \text{for } 0.25 \le x < 0.50 \quad (\Omega_2) \\ 0.5 & \text{for } 0.50 \le x < 0.75 \quad (\Omega_3) \\ 0.0 & \text{for } 0.75 \le x \le 1.00 \quad (\Omega_4) \end{cases} . \tag{3.31}$$

The results of solving (2.1) using $\alpha = 0.01$ are given in Table 3.1 and FigureFigureNumerical1D.

In Figure 3.4(a) we give the numerical results in applying TV regularization to the function (3.31), solving (2.1) using $\alpha = 0.01$. In Figure 3.4(a), we also apply

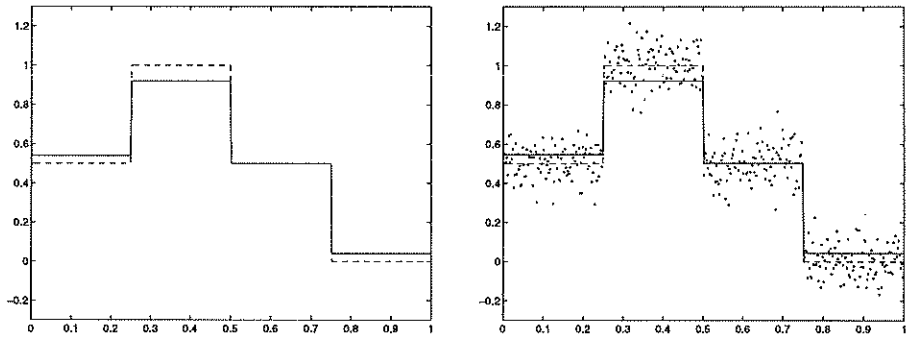
Region	Type	u_0	u	$u - u_0$	Predicted δ using (3.30)
Ω_1	Boundary	0.50	0.54	0.04	$\frac{1}{0.25-0.00}(0.01) = 0.04$
Ω_2	Extremum	1.00	0.92	-0.08	$\frac{-2}{0.50-0.25}(0.01) = -0.08$
Ω_3	Step	0.50	0.50	0.00	$\frac{0}{0.75-0.50}(0.01) = 0.00$
Ω_4	Boundary	0.00	0.04	0.04	$\frac{1}{1.00-0.75}(0.01) = 0.04$

Table 3.1: Example 3.3: TV regularization of an R^1 piecewise constant function.

TV regularization to a noisy version of this image and obtain identical results (in the step region, Ω_3, we obtain only approximate agreement in the results of the noise-free and the noisy cases). In the two figures, the dashed line is the noise-free function, the solid line is the regularized or denoised function, and in (b) the dotted line is the noisy function. Example 3.3 shows that the regularized image is the same when restoring a noise-free image and when restoring a noisy version of that same image, for sufficiently large values of α (except in step regions, where in general we would have *approximate* agreement between the regularized function u in the noise-free and noisy cases). In this example the *α-condition* is not actually met, which demonstrates that in practice the bounds on α required by the *α-condition* need not be strictly met to approximately (or even exactly) obtain the results predicted by (3.30).

3.1.3.2 Theorem 3.3 in R^2, with Constant α

Corollary 3.3 (R^2 Piecewise Constant, Constant α) *In R^2 let the conditions of Theorem 3.3 be satisfied. Let $\Omega_i \subset R^2$ correspond to $r \in [r_{i-1}, r_i]$, and assume that $\alpha(r) \equiv \alpha$ is constant. Then (3.29) is given by*



(a) R^1, noise-free case. (b) R^1, noisy case.

Figure 3.4: Example 3.3: TV regularization of an R^1 piecewise constant function.

$$\delta_i = \begin{cases} \dfrac{2}{r_i - r_{i-1}}\alpha & \textit{Extremum Regions} \\ \dfrac{2}{r_i + r_{i-1}}\alpha & \textit{Step Regions} \\ \dfrac{2r_{i-1}}{r_i^2 - r_{i-1}^2}\alpha & \textit{Boundary Regions} \end{cases} \tag{3.32}$$

As for the R^1 case, for the R^2 case Corollary 3.3 can be applied to any (noise-free or noisy) piecewise constant function in R^2. We give a single example of this.

Example 3.4 For this example we define the original image $u_0(r) = u_0(x, y)$ as

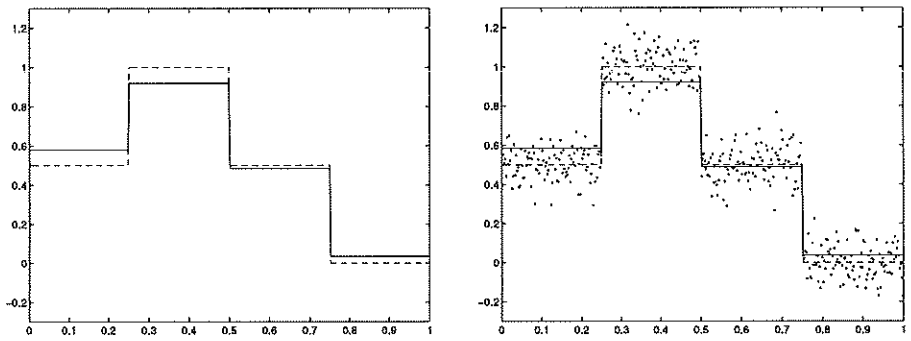
$$u_0(r) = \begin{cases} 0.5 & \text{for } 0.00 \leq x < 0.25 \quad (\Omega_1) \\ 1.0 & \text{for } 0.25 \leq x < 0.50 \quad (\Omega_2) \\ 0.5 & \text{for } 0.50 \leq x < 0.75 \quad (\Omega_3) \\ 0.0 & \text{for } 0.75 \leq x \leq 1.00 \quad (\Omega_4) \end{cases} . \tag{3.33}$$

The results of solving (2.1) using $\alpha = 0.01$ are given in Table 3.2 and Figure 3.5.

Region	Type	u_0	u	$u - u_0$	Predicted δ using (3.32)
Ω_1	Extremum	0.500	0.580	0.080	$\frac{2}{0.25-0.00}(0.01) = 0.080$
Ω_2	Extremum	1.000	0.920	-0.080	$\frac{2}{0.50-0.25}(0.01) = -0.080$
Ω_3	Step	0.500	0.484	-0.016	$\frac{2}{0.75+0.50}(0.01) = -0.016$
Ω_4	Boundary	0.000	0.034	0.034	$\frac{2(0.75)}{1.00^2-0.75^2}(0.01) \approx 0.034$

Table 3.2: Example 3.4: TV Regularization of an R^2 radially symmetric piecewise constant function.

In Figure 3.5(a) we give the numerical results in applying TV regularization to the function 3.33, solving 2.1) using $\alpha = 0.01$. In Figure 3.5(a), we also apply TV regularization to a noisy version of this image and obtain identical results. In the two figures, the dashed line is the noise-free function, the solid line is the regularized or denoised function, and in (b) the dotted line is the noisy function. Our results show that the regularized image is the same when restoring a noise-free image and the noisy version of that same image, for sufficiently large values of α In this example the *α-condition* is not actually met, which demonstrates that in practice the bounds on α required by the *α-condition* need not be strictly met to approximately (or even exactly) obtain the results predicted by (3.32).



(a) R^2, noise-free case. (b) R^2, noisy case.

Figure 3.5: Example 3.4: TV Regularization of an R^2 radially symmetric piecewise constant function.

3.1.3.3 Theorem 3.3 in R^3, with Constant α

Corollary 3.4 (R^3 Piecewise Constant, Constant α) *In R^2 let the conditions of Theorem 3.3 be satisfied. Let $\Omega_i \subset R^2$ correspond to $r \in [r_{i-1}, r_i]$, and assume that $\alpha(r) \equiv \alpha$ is constant. Then (3.29) is given by*

$$\delta_i = \begin{cases} \dfrac{3(r_i^2 + r_{i-1}^2)}{r_i^3 - r_{i-1}^3}\,\alpha & \textit{Extremum Regions} \\ \dfrac{3(r_i^2 - r_{i-1}^2)}{r_i^3 - r_{i-1}^3}\,\alpha & \textit{Step Regions} \\ \dfrac{3r_{i-1}^2}{r_i^3 - r_{i-1}^3}\,\alpha & \textit{Boundary Regions} \end{cases} \tag{3.34}$$

As for the R^1 and R^2 cases, for the R^3 case Corollary 3.4 can be applied to any (noise-free or noisy) piecewise constant function in R^3. We give a single example of this.

Example 3.5 For this example we define the original image $u_0(r) = u_0(x, y, z)$ as

$$u_0(r) = \begin{cases} 0.5 & \text{for } 0.00 \le x < 0.25 \quad (\Omega_1) \\ 1.0 & \text{for } 0.25 \le x < 0.50 \quad (\Omega_2) \\ 0.5 & \text{for } 0.50 \le x < 0.75 \quad (\Omega_3) \\ 0.0 & \text{for } 0.75 \le x \le 1.00 \quad (\Omega_4) \end{cases} . \tag{3.35}$$

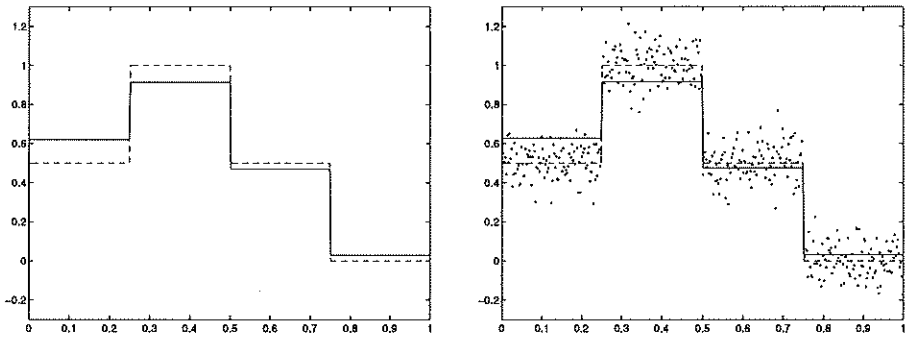
The results of solving (2.1) using $\alpha = 0.01$ are given in Table 3.3 and Figure 3.6.

Region	Type	u_0	u	$\lvert u - u_0 \rvert$	Predicted δ using (3.34)
Ω_1	Extremum	0.500	0.620	0.120	$\frac{3(0.25^2+0.00^2)}{0.25^3-0.00^3}(0.01) = 0.120$
Ω_2	Extremum	1.000	0.014	-0.086	$\frac{3(0.50^2+0.25^2)}{0.50^3-0.25^3}(0.01) \approx -0.086$
Ω_3	Step	0.500	0.468	-0.032	$\frac{3(0.75^2-0.50^2)}{0.75^3-0.50^3}(0.01) \approx -0.032$
Ω_4	Boundary	0.000	0.029	0.029	$\frac{3(0.75^2)}{1.00^3-0.75^3}(0.01) \approx 0.029$

Table 3.3: Example 3.5: TV Regularization of an R^3 radially symmetric piecewise constant function.

In Figure 3.6(a) we give the numerical results in applying TV regularization to the function 3.35, solving 2.1) using $\alpha = 0.01$. In Figure 3.6(a), we also apply TV regularization to a noisy version of this image and obtain identical results. In the two figures, the dashed line is the noise-free function, the solid line is the regularized or denoised function, and in (b) the dotted line is the noisy function. Our results show that the regularized image is the same when restoring a noise-free image and the noisy version of that same image, for sufficiently large values of α (except in step regions, where in general we would have *approximate* agreement between the regularized function u in the noise-free and noisy cases). In this example the *α-condition* is not actually met, which demonstrates that in

practice the bounds on α required by the *α-condition* need not be strictly met to approximately (or even exactly) obtain the results predicted by (3.34).



(a) R^3, noise-free case. (b) R^3, noisy case.

Figure 3.6: Example 3.5: TV Regularization of an R^3 radially symmetric piecewise constant function.

3.2 Exact Solution of the TV Regularization Problem for Smooth Radially Symmetric Functions

In this section we extend the theory developed in Section 3.1 to *non*-piecewise constant features. Our results are again developed for radially symmetric functions in R^d, $1 \leq d \leq 3$. To simplify the problem we consider the case where $\alpha(\vec{x}) \equiv \alpha$ is constant.

3.2.1 Smooth Functions as the Limit of Piecewise Constant Functions

We first consider the portion of a radially symmetric function with constant, decreasing slope, as illustrated in Figure 3.7(b). We can take this function as the limit of the step function shown in Figure 3.7(a). In Section 3.1 we found that δ_i

in 3.7(a) is given by

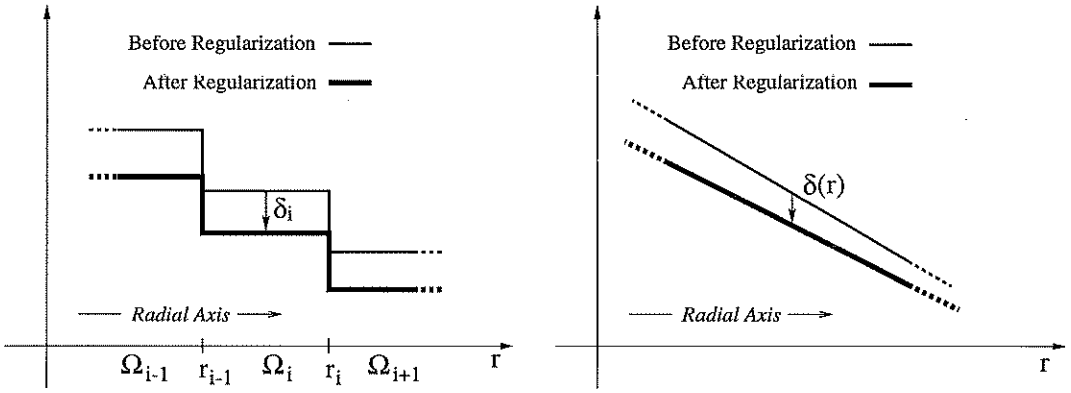
$$\delta(r) = \frac{|\partial\Omega_{i,i+1}| - |\partial\Omega_{i,i-1}|}{|\Omega_i|}\alpha = \begin{cases} \dfrac{1-1}{r_i - r_{i-1}}\alpha & \text{in } R^1 \\ \dfrac{2\pi r_i - 2\pi r_{i-1}}{\pi r_i^2 - \pi r_{i-1}^2}\alpha & \text{in } R^2 \\ \dfrac{4\pi r_i^2 - 4\pi r_{i-1}^2}{\frac{4}{3}\pi r_i^3 - \frac{4}{3}\pi r_{i-1}^3}\alpha & \text{in } R^3 \end{cases}$$

As we take the limit to the smooth case, we let $r_{i-1} \to r_i$, and find that for $r \in [r_{i-1}, r_i]$, where we treat $\delta(r)$ as a non-negative value,

$$\boxed{\delta(r) = \begin{cases} 0 & \text{in } R^1 \\ \dfrac{1}{r}\alpha & \text{in } R^2 \\ \dfrac{2}{r}\alpha & \text{in } R^3 \end{cases}} \tag{3.36}$$

3.2.2 Smooth Functions in R^1

We next consider the effects of TV on the "smooth" function shown in Figure as illustrated in Figure 3.8. We refer to it as "smooth" in the sense that it is continuous, with no sharp discontinuous edges. (Although this function is not actually smooth is the usual sense, the results we develop using this example are applicable to functions which *are* smooth in the usual sense. In addition, the notion of smoothness is actually somewhat moot in the *discrete* numerical implementation of the problem anyway.) By viewing this function as the limit of a step function, the results developed in Section 3.1 and in (3.36) would predict



(a) Piecewise constant case.

(b) Smooth case.

Figure 3.7: The effects of TV regularization in the smooth case taken as the limit of the piecewise constant case.

the regularized function shown in Figure 3.8. The function $f(x)$ which describes the function $u_0(x)$ for $x_1 \leq x \leq x_2$ is

$$f(x) = \frac{x - x_2}{x_1 - x_2} = m(x - x_2), \quad \text{where } m = \frac{1}{x_1 - x_2},$$

so that $f(x_1) = 1$ and $f(x_2) = 0$. For convenience our domain is $[0, 1]$.

For the R^1 case we can find explicit equations for δ_i and $\tilde{x}_i$. Note that in Figure 3.8

$$\tilde{x}_i = x_i - \frac{\delta_i}{m}, \quad \text{for } i = 1, 2, \tag{3.37}$$

where we will have $\delta_1 < 0 < \delta_2$. Assuming that α is not large enough to flatten out the function, that is $1 + \delta_1 > 0 + \delta_2$, then the fitting error would be

$$\begin{aligned} \int_0^1 [u(x) - u_0(x)]^2 \, dx &= \int_0^{x_1} \delta_1^2 \, dx + \int_{x_1}^{\tilde{x}_1} [f(x) - (1 + \delta_1)]^2 \, dx \\ &\quad + \int_{\tilde{x}_2}^{x_2} [\delta_2 - f(x)]^2 \, dx + \int_{x_2}^1 \delta_2^2 \, dx \end{aligned}$$

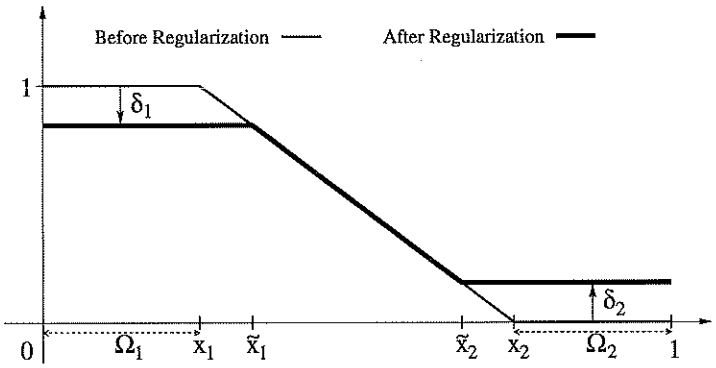


Figure 3.8: A "smooth" function, before and after TV regularization.

$$= |\Omega_1|\delta_1^2 - \frac{\delta_1^3}{3m} + \frac{\delta_2^3}{3m} + |\Omega_2|\delta_2^2.$$

The total variation of u would be $TV(u) = 1 + \delta_1 - \delta_2$. So the problem to solve is

$$\min_{\delta_1,\delta_2} \{ \tfrac{1}{2} [|\Omega_1|\delta_1^2 - \frac{\delta_1^3}{3m} + \frac{\delta_2^3}{3m} + |\Omega_2|\delta_2^2] + \alpha(1 + \delta_1 - \delta_2) \},$$

which by differentiating with respect to each δ_i gives the equations

$$\begin{aligned} |\Omega_1|\delta_1 - \frac{\delta_1^2}{2m} + \alpha &= 0 \\ |\Omega_2|\delta_2 + \frac{\delta_2^2}{2m} - \alpha &= 0. \end{aligned}$$

By taking the appropriate value of δ_i in solving each quadratic equation, we find that

$$\begin{aligned} \delta_1 &= +m|\Omega_1| + \sqrt{(m|\Omega_1|)^2 + 2m\alpha}\,, \\ \delta_2 &= -m|\Omega_2| - \sqrt{(m|\Omega_2|)^2 + 2m\alpha}\,. \end{aligned} \tag{3.38}$$

To see the similarity to the piecewise constant case, notice that as $m \to \infty$, we have

$$\delta_1 \to -\frac{\alpha}{|\Omega_1|} \quad \text{and} \quad \delta_2 \to \frac{\alpha}{|\Omega_1|},$$

as found for the piecewise constant case in Section 3.1.

We consider an example in R^1.

Example 3.6 Using the parameters

$x_1 = 0.25$	$m = -2$
$x_2 = 0.75$	$\alpha = 0.03$

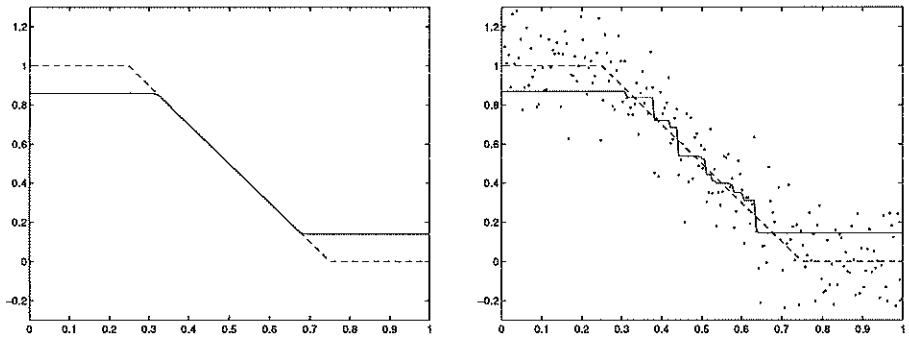
(3.39)

we find that the values predicted by (3.37) and (3.38),

$\delta_1 = -0.139$	$\tilde{x}_1 = 0.320$
$\delta_2 = +0.139$	$\tilde{x}_2 = 0.680$

(3.40)

agree with the regularized function found by solving (2.1), as shown in Figure 3.9(a), where the dashed line is the original function and the solid line is the regularized function. We do not give a formal proof to validate these results, as they are a natural extension of results given in Section 3.1) for the R^1 piecewise constant case. For comparison, we apply TV regularization to a noisy version of the function, and observe similar results in Figure 3.9(b), except for the unwanted effect of "staircasing," which is caused by the presence of noise. (This staircasing effect is well-known and is currently being studied [6].) In (b) the dashed line is the true function, the dotted line is the noisy function and the solid line is the regularization function.



(a) R^1, noise-free case. (b) R^1, noisy case.

Figure 3.9: Example 3.6: TV regularization of a "smooth" function in R^1.

3.2.3 Smooth Radially Symmetric Functions in R^2 and R^3

We next consider a radially symmetric function R^2 or R^3 as illustrated in Figure 3.10, analogous to the R^1 function just considered in Section 3.2.2. The function $f(r)$ which describes the function $u_0(r)$ for $r_1 \leq r \leq r_2$ is

$$f(r) = \frac{r - r_2}{r_1 - r_2} = m(r - r_2), \quad \text{where } m = \frac{1}{r_1 - r_2},$$

so that $f(r_1) = 1$ and $f(r_2) = 0$. For convenience we again take the domain to be $r \in [0, 1]$.

In Figure 3.10, we can find $\tilde{r}_1$ as a function of r_1, m, δ_1 and α by using (3.36), as shown in Figure 3.10. For R^2 we have

$$\begin{aligned} \frac{\alpha}{\tilde{r}_1} &= f(\tilde{r}_1) - (1 + \delta_1) = m(\tilde{r}_1 - r_1) - \delta_1 \\ \Longrightarrow \alpha &= m\,\tilde{r}_1^2 - m\,r_1\,\tilde{r}_1 + \delta_1\,\tilde{r}_1. \end{aligned}$$

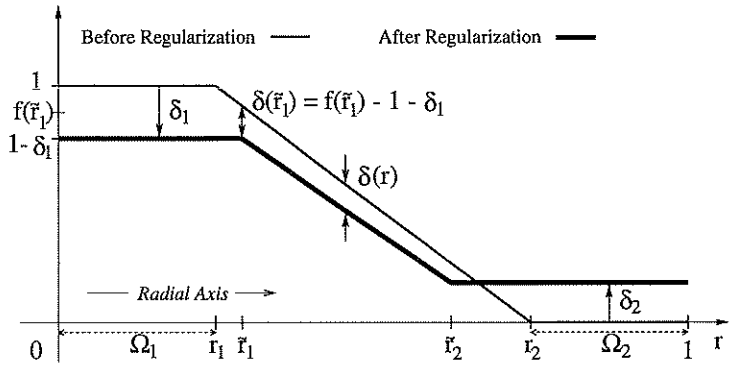


Figure 3.10: The R^1 cross-section of an R^2 or R^3 radially symmetric function, before and after TV regularization.

and solving $m\,\tilde{r}_1^2 - m\,r_1\,\tilde{r}_1 + \delta_1\,\tilde{r}_1 - \alpha = 0$ for $\tilde{r}_1$ (and similarly for $\tilde{r}_2$) gives

$$\tilde{r}_i = \frac{m\,r_i + \delta_i - \sqrt{(\delta_i + m\,r_i)^2 + 4m\alpha}}{2m} \quad \text{for } i = 1, 2. \tag{3.41}$$

In R^3, we similarly find that

$$\tilde{r}_i = \frac{m\,r_i + \delta_i - \sqrt{(\delta_i + m\,r_i)^2 + 8m\alpha}}{2m} \quad \text{for } i = 1, 2. \tag{3.42}$$

For this example in R^2 or R^3, we can write equation (2.1) as a function of δ_1 and δ_2. We can find equations for the fitting error e_i as functions of δ_1 and δ_2 in each of the five subregions in the domain (see Figure 3.11). In R^2, for example, we would have

$$e_i(\delta_1, \delta_2) = \int_{r_{i-1}}^{r_i} 2\pi\, r\, [u_0(r) - u(r)]^2\, dr,$$

so that that in the region $[0, r_1]$ we could find that

$$e_1(\delta_1, \delta_2) \;=\; \int_0^{r_1} 2\pi\, r\, [u_0(r) - u(r)]^2\, dr$$

$$= \int_0^{r_1} 2\pi\, r\, \delta_1^2\, dr$$

$$= \pi r_1^2 \delta_1^2.$$

Also, the total variation of u can be written as a function of δ_1 and δ_2. For example, in R^2 we have

$$\begin{aligned} TV(u) &= \int_0^1 2\pi\, r\, |\, u_r\, |\, dr \\ &= \int_{\tilde{r}_1}^{\tilde{r}_2 1} 2\pi\, r\, (-\frac{\alpha}{r^2} - m)\, dr \\ &= 2\pi\, [\, \frac{m}{2}(\tilde{r}_1^2 - \tilde{r}_2^2) + \alpha \ln \frac{\tilde{r}_1}{\tilde{r}_2}\,]. \end{aligned}$$

Since $\{\tilde{r}_i\}$ are functions of $\{\delta_i\}$ (and of other known parameters), the problem (2.1) then can be written as a function of δ_1 and δ_2:

$$\min_{\delta_1, \delta_2} \quad \{\, \tfrac{1}{2} \sum_{i=1}^{5} e_i(\delta_1, \delta_2) + \alpha\, TV(u(\delta_1, \delta_2))\, \}. \tag{3.43}$$

The R^2 and R^3 cases are more complicated than the R^1 case, and we cannot find an explicit analytic equation for the optimal values of δ_i. We can, however, numerically verify that equations (3.41) - (3.43) accurately describe the effects of TV regularization on the function shown in Figure 3.10 by computing the values of δ_1 and δ_2 which minimize (3.43), where $\{\tilde{r}_i\}$ are functions of $\{\delta_i\}$ using (3.41) and (3.42). It is reasonable to assume that we can numerically find unique

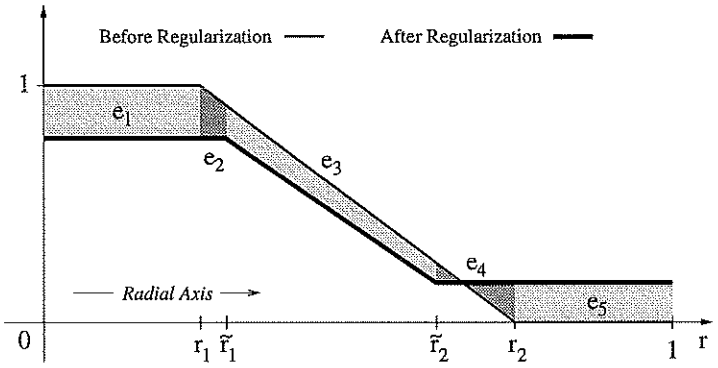


Figure 3.11: The fitting error e_i in each of the five regions created by r_1, r_2, $\tilde{r}_1$ and $\tilde{r}_2$.

optimal values for δ_1 and δ_2 since in [14] it is shown that a unique minimizer to (2.1) exists under certain typical conditions. If our assumptions about the form of u, as illustrated in Figure 3.10, are correct, then the resulting functions u in solving (3.43) should be identical to the functions found when solving (2.1).

Example 3.7 In R^2 we use the parameters

$r_1 = 0.25$	$m = -2$
$r_2 = 0.75$	$\alpha = 0.04$

(3.44)

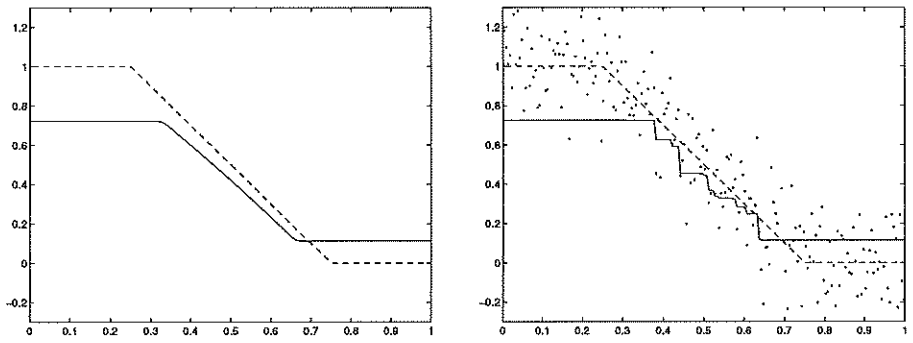
and find that the values found by solving (3.43), using (3.41),

$\delta_1 = -0.278$	$\tilde{r}_1 = 0.328$
$\delta_2 = +0.113$	$\tilde{r}_2 = 0.663$

(3.45)

agree with the regularized function found by solving (2.1), as shown in Figure 3.12(a), in which we see the R^1 cross-sections of R^2 and R^3 radially symmetric functions. The dashed line is the noise-free function and the solid line is the

regularized function. In Figure 3.12(b) are the results of applying TV regularization to the function from (a) which has been contaminated with Gaussian noise. Results for the noisy R^2 case are shown in Figure 3.12(b), in which the dashed line is the noise-free function, the dotted line is the noisy function and the solid line is the regularized function. We note that the staircasing observed in Figure 3.12(b) is exaggerated, as we took the noise to be symmetric (in order to solve this R^2 problem strictly as an R^1 problem). In the actual R^2 problem, the noise of course is generally not radially symmetric, and the staircasing although still present, is not nearly as pronounced.



(a) R^2, noise-free case. (b) R^2, noisy case.

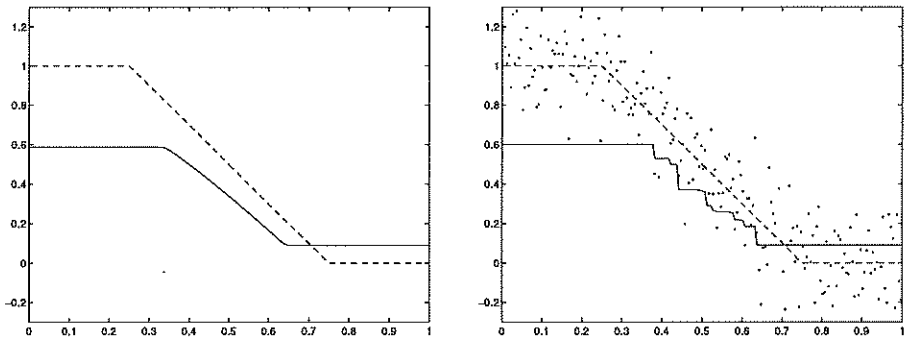
Figure 3.12: Example 3.7: TV regularization of a "smooth" radially symmetric function in R^2.

Example 3.8 In R^3, again using the parameters (3.44), we find δ_1 and δ_2 by miniminizing (3.43), and use these values to compute $\tilde{r}_1$ and $\tilde{r}_2$. Solving (3.43) using (3.42) gives the values

δ_1	$=$	-0.413	$\tilde{r}_1$	$=$	0.338
δ_2	$=$	$+0.089$	$\tilde{r}_2$	$=$	0.643

(3.46)

which agrees with the function found by solving (2.1), as shown in Figure 3.13(a). Results for the noisy R^3 case are shown in Figure 3.13(b). The dashed line is the noise-free function, the solid line is the regularized function, and in (b) the dotted line is the noisy function. Once again, the staircasing is exaggerated, due to the radially symmetric noise.



(a) R^3, noise-free case. (b) R^3, noisy case.

Figure 3.13: Example 3.8: TV regularization of a "smooth" radially symmetric function in R^3.

3.3 Agreement Between Theory and Numerical Solutions in Non-radially Symmetric Case

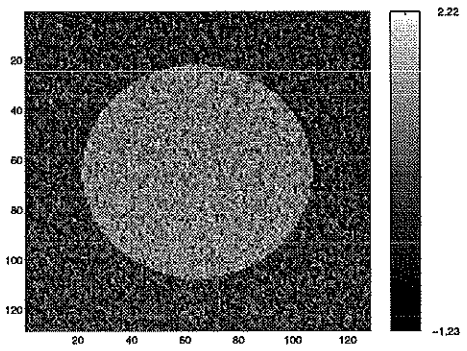
So far, we have considered some special cases of the *continuous* TV regularization problem, namely piecewise constant and simple "smooth" functions, assuming radial symmetry. We have been able to find analytic solutions for these various cases. In the general case, however, we are not able to explicitly find analytic solutions. In addition, in practice we must of course *numerically* solve the *discrete* version of the continuous problem, for example, when denoising a digital image.

In this section we briefly examine the agreement between the theory and the numerical solutions to the discrete problem for more general images. In examining the general case, we limit ourselves to the three images shown in Figure 3.14, in order to keep the discussion relatively simple. Our primary purpose has been to develop a simple and precise theory for the special cases which we have treated, which may be used to understand how TV regularization affects the various types of images dealt with in the general case. It is not our purpose in this Dissertation to attempt to exhaustively examine the effects of TV regularization in the general case.

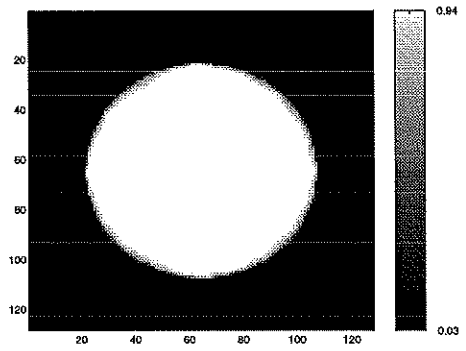
Example 3.9 We wish to examine the agreement between the change in intensity in the feature of interest and the change *predicted* by (3.5). The feature of interest Ω in the first image, shown in Figure 3.14(a), is a noise-contaminated circle of radius $\frac{1}{3}$ in a unit square domain. For this example we use $\alpha = 0.01$, so that in the rectangular region the change in intensity $\delta_{predicted}$ as predicted by (3.7) and (3.29) should be

$$\delta_{predicted} = \frac{|\partial\Omega|}{|\Omega|}\alpha = \frac{2\pi(\frac{1}{3})}{\pi(\frac{1}{3})^2}(0.01) = 0.06, \tag{3.47}$$

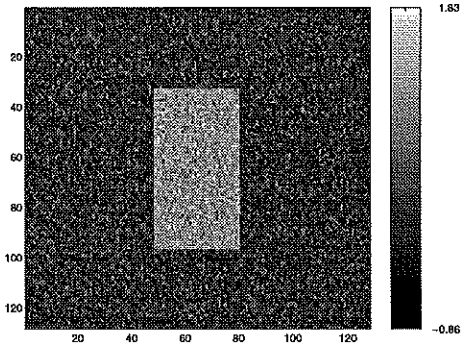
so that in the circular region the intensity level should be 0.94 after regularization. This is nearly exactly the case, as shown in Figure 3.14(b) (notice the grayscale bar to the right of each image). Notice that the boundary of the circle has been slightly smoothed. This is due to the slightly discrepency between the continuous and discretized problems. If Ω is the circular region in the image, then where $\delta_{avg} = \int_\Omega u_0(\vec{x}) - u(\vec{x})\,d\vec{x}$ is the *average* amount of intensity change (over the entire circular feature), it turns out that $\delta_{avg} \approx \delta_{predicted} \approx \delta(\vec{x})$. If desired, this boundary smoothing in the discrete case can be completely overcome by employing an edge-preserving numercial approximation scheme, such as the



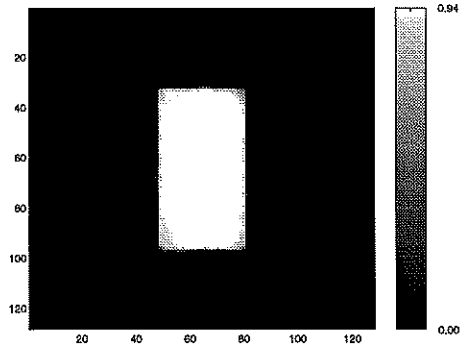
(a) Noisy image, before regularization.



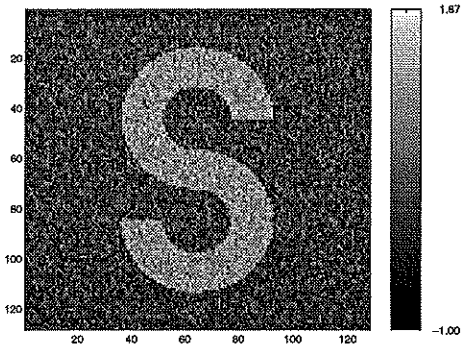
(b) After regularization.



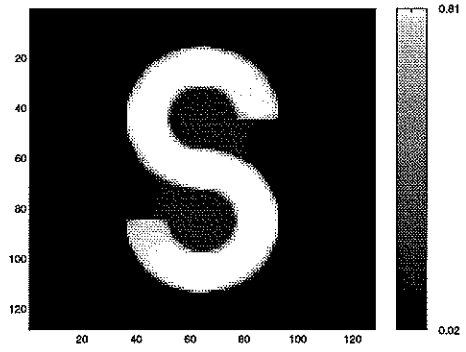
(c) Noisy image, before regularization.



(d) After regularization.



(e) Noisy image, before regularization.



(f) After regularization.

Figure 3.14: The results of applying TV Regularization to general piecewise constant functions. Refer to Table 3.4.

minmod scheme used in [22].

Shape	α	Predicted $\delta = \frac{\lvert\partial\Omega\rvert}{\lvert\Omega\rvert}\alpha$	Computed δ_{avg}
Circle	0.010	0.060	0.060
Rectangle	0.005	0.060	0.056
"S"	0.010	0.180	0.193

Table 3.4: Refer to Figure 3.14 for the images to which this table corresponds. A comparison of predicted changes in intensity to the results found numerically in applying TV regularization to functions with smooth and sharp boundaries. The numerical results nearly exactly match the predicted results, with the slight discrepency being due to discretizing the continuous function and/or the deformation of the boundaries.

We give two more examples of images to which we apply TV regularization: a rectangle of dimensions $\frac{1}{4}$ x $\frac{1}{2}$ and an "S"-shaped object. These two images are both non-radially symmetric. We point out here that in the non-radially symmetric case, TV minimizing techniques tend to smooth out rough or oscillatory boundaries in a function (see [12]), and can consequently result in the deformation of boundaries in a function. This boundary deformation occurs because the total variation of a feature is directly proportional to its boundary size, e.g. for a piecewise constant feature (see Lemma 3.1), so that one way of minimizing the total variation of the feature is to reduce its boundary size. (As we have shown, in the continous radially symmetric case this boundary deformation *does not* occur.)

For each function, we find the theoretically predicted value for the change in intensity, $\delta = \dfrac{|\partial\Omega|}{|\Omega|}\alpha$, as well as the change of intensity when solving the problem numerically. The results are given in Figure 3.14 and Table 3.4.

For these three examples, the slight discrepencies between predicted and computed changes in intensity level are due to the boundary deformation and/or the discretization of the continuous image. Overall, the agreement between the theory and the numerical results is almost exact. The theory which helps us understand the continuous problem thus is of almost equal use for understanding the numerical, discrete problem.

3.4 Summary of Theoretical Results

In this chapter, we have given formulae, proven analytically and verified numerically, which describe the effects of total variation minimizing function regularization for several special cases in R^d, $1 \leq d \leq 3$. These results are quite useful not only for describing TV regularization in these specific cases, but also in better understanding and predicting how TV regularization will affect images (or other functions) in the general case.

We have shown that TV regularization causes piecewise constant radially symmetric features to remain *exactly* piecewise constant, with edge location being preserved *exactly*, and that change in function intensity is inversely proportional to the local feature scale and directly proportional to the regularization parameter (see (3.7) and Theorem 3.3). In addition to the formulae given for the general case, we have given these results for three specific practical cases: in R^1, R^2 and R^3 with constant α. Our results demonstrate that TV image restoration is especially effective for restoring images with piecewise constant features. For smooth (i.e. non-piecewise constant) functions we have shown that intensity level change is inversely proportional to the radial distance from the center of the feature and directly proportional to the regularization parameter (see (3.36)). We have shown that for general (i.e. smooth) radially symmetric features, the basic shape

of the feature is retained, and noise, if present, is essentially removed. Our formulae help explain how and why TV image restoration can remove (smaller-scaled) noise while leaving relatively intact larger-scaled image features. Our formulae also help to demonstrate the localness of TV regularization, which is useful in understanding how TV regularization affects an image, and is potentially quite useful in developing faster numerical schemes for solving the TV regularization problem.

CHAPTER 4

Feature-driven Adaptive Total Variation Minimizing Image Restoration

4.1 Introduction

In this chapter, we propose a spatially adaptive TV minimizing image restoration scheme where the adaptivity is realized by using a weighted TV functional

$$TV_\alpha(u) \equiv \int \alpha(\vec{x})|\nabla u(\vec{x})|\, d\vec{x}. \tag{4.1}$$

The corresponding noise-constrained minimization problem would then be

$$\min_u TV_\alpha(u) \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2. \tag{4.2}$$

The detailed analysis of TV minimizing function regularization in Chapter 3 provides a theoretical justification for this approach.

We choose $\alpha(\vec{x})$ based on the likelihood of the presence of an edge between any two neighboring discrete image locations. The weighting factor is chosen to be inversely proportional to the likelihood of the presence of an edge. This allows for less regularization where edges are present and more regularization where there are no edges, which results in better overall noise removal and detail preservation. The results are generally good, particularly for images with piecewise constant

image features.

In the balance of the chapter, we present our basic ideas in R^1 and give results for restoring noisy R^1 images, after which we extend our scheme to R^2 and give results for restoring noisy R^2 images.

4.2 Adaptive Image Restoration in R^1

We first discuss the weighted TV norm (4.1) in R^1. We next discuss how to choose the weighting factor, first by using *a priori* information and then automatically. Results of restoring noisy R^1 images are then given.

4.2.1 Weighted TV Norm in R^1

To accomplish adaptive TV miniminizing image restoration, we replace (1.5) with the weighted TV norm (4.1). In R^1, the discrete version of (1.5) is

$$TV(u) = \sum_{i=1}^{n-1} |u_x(i+\tfrac{1}{2})| = \sum_{i=1}^{n-1} |u_{i+1} - u_i|, \tag{4.3}$$

where $u_x(i+\frac{1}{2})$ represents $\frac{du(x)}{dx}|_{x=x_{i+\frac{1}{2}}}$. The weighted TV functional (4.1) is analogously

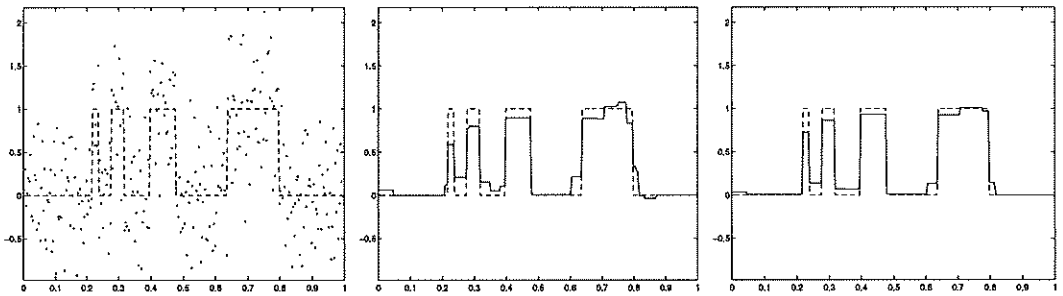
$$TV_\alpha(u) = \sum_{i=1}^{n-1} \alpha_{i+\frac{1}{2}} |u_{i+1} - u_i|. \tag{4.4}$$

The theoretical results of Chapter 3 leads to idea of choosing $\alpha_{i+\frac{1}{2}}$ to be *smaller* where there is an edge between u_i and u_{i+1}, and conversely choosing $\alpha_{i+\frac{1}{2}}$ to be *larger* where there is no edge. The motivation is to better preserve edges by allowing the variation in the image which is due to the edges, and to better remove noise where edges are not present by penalizing the variation in the image that is due to the noise. In regions of relatively moderate intensity change (i.e. smooth,

non-piecewise constant image features), the choice of $\alpha_{\cdot,i+\frac{1}{2}}$ would be somewhere in between. This relationship can be written

$$\alpha_{i+\frac{1}{2}} \sim \frac{1}{\text{Likelihood of an edge between positions } i \text{ and } i+1 \text{ in } u_{true}}.$$

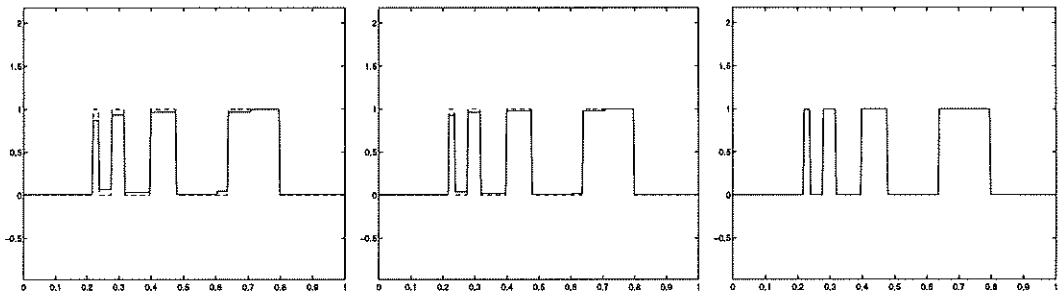
4.2.2 Adaptive Restoration Based on Likelihood of Presence of Edges



(a) Noisy image.

(b) Restored image, $\alpha_{flat} : \alpha_{edge} = 1 : 1$.

(c) Adaptively restored, $\alpha_{flat} : \alpha_{edge} = 2 : 1$.



(d) Adaptively restored, $\alpha_{flat} : \alpha_{edge} = 5 : 1$.

(e) Adaptively restored, $\alpha_{flat} : \alpha_{edge} = 10 : 1$.

(f) Adaptively restored, $\alpha_{flat} : \alpha_{edge} = 100 : 1$.

Figure 4.1: Example 4.1: adaptive TV minimizing image restoration solving (4.2), using *a priori* information about edge location to determine $\alpha_{flat} : \alpha_{edge}$.

Figure 4.1	$\alpha_{flat} : \alpha_{edge}$	ISNR
(a)	(Noisy Image)	0.00 dB
(b)	1:1	11.78 dB
(c)	2:1	16.09 dB
(d)	5:1	22.72 dB
(e)	10:1	28.21 dB
(f)	100:1	47.62 dB

Table 4.1: The improved SNR for each of the adaptively restored images (shown in Figure 4.1) found by solving (4.2), where the ratio of the weighting of (4.1) in the flat regions to the weighting at the edges is given by $\alpha_{flat} : \alpha_{edge}$. We note that the case in which $\alpha_{flat} : \alpha_{edge} = 1 : 1$ is simply the standard (non-adaptive) case.

Example 4.1 In Figure 4.1 we give an example of adaptive restoration for a piecewise constant R^1 function, contaminated with Gaussian noise, SNR = 0 dB, where the edge locations are known *a priori*, to demonstrate the effectiveness of the our idea for choosing $\alpha_{i+\frac{1}{2}}$. In finding each of the restored images shown in Figures 4.1(b) - (f), we have chosen

$$\alpha_{i+\frac{1}{2}} = \begin{cases} \alpha_{edge} & \text{if there } \textit{is} \text{ an edge between positions } i \text{ and } i+1 \text{ in } u_{true} \\ \alpha_{flat} & \text{if there } \textit{is not} \text{ an edge between positions } i \text{ and } i+1 \text{ in } u_{true} \end{cases}$$

where $\alpha_{edge} \leq \alpha_{flat}$. We note that in solving the *constrained* problem (4.2), is is not the actual values of $\{\alpha_{i+\frac{1}{2}}\}$ that are important, but rather their *relative* values, that is, relative to each other. We solve (4.2) using different ratios for $\alpha_{flat} : \alpha_{edge}$, where $TV_\alpha(u)$ is as in (4.4). We find, as expected, that the adaptive restoration scheme is more effective for larger ratios of $\alpha_{flat} : \alpha_{edge}$. Because of the exact information which we had about edge location in this contrived example,

we have

$$u \longrightarrow u_{true} \quad \text{as} \quad \frac{\alpha_{flat}}{\alpha_{edge}} \longrightarrow \infty, \tag{4.5}$$

if the mean of the noise is 0 in each of the piecewise constant regions (which we artificially enforced for this example). Figure 4.1 illustrates (4.5). Note that (b) is simply the standardly (non-adatively) restored image, solving 2.2). The results are better for larger ratios of $\alpha_{flat} : \alpha_{edge}$. Table 4.1 gives the ISNR for each of the restored images.

4.2.3 Automatically Defining $\{\alpha_{i+\frac{1}{2}}\}$ in R^1

In defining $\{\alpha_{i+\frac{1}{2}}\}$ in the preceeding example we knew *a priori* the location of the edges. In general we do not have to this information *a priori* (otherwise the problem is often already solved). Here our approach in determining the likelihood of an edge at any given location is based on examining the size of the jump between neighboring pixels of a partially or fully (non-adaptively) restored image. The reason for this is that a partially restored image can give us valuable information which can be used to determine the weighting factor α, which we can then use to restore the noisy image with a spatially varying balance between noise removal and fit to the original data. In Chapter 3 we found that the effects of TV minimizing image restoration are inversely proportional to scale, so that smaller-scaled features are sometimes compromised or completely lost in our attempt to remove noise from the image. Thus partial noise removal (i.e. a partially restored image) can potentially give us better information about the more detailed or smaller-scaled features than could a completely restored image, in which the levels of both noise and detail are decreased. The algorithm for implementing this approach is given in Table 4.2.

The first step is to solve the standard minimization problem (2.2) using a

The FATV Scheme in R^1:
Feature-driven Adaptive Total Variation Minimizing Image Restoration in R^1

1. Find $\tilde{u}$ by solving
$$\min_{\tilde{u}} TV(\tilde{u}) \quad \text{subject to} \quad \|\tilde{u} - u_0\|^2 = \tilde{\sigma}^2,$$
where $0 < \tilde{\sigma}^2 \leq \sigma^2$.

2. For $1 \leq i \leq n-1$, define
$$\alpha_{i+\frac{1}{2}} = \frac{1}{|\tilde{u}_{i+1} - \tilde{u}_i| + \epsilon},$$
where $\epsilon > 0$.

3. Find the adaptively restored image u by solving
$$\min_{u} TV_\alpha(u) \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2.$$

Table 4.2: Feature-driven Adaptive TV minimizing (FATV) image restoration scheme in R^1.

(possibly) smaller estimated noise level $\tilde{\sigma}^2 \leq \sigma^2$. Thus by solving (2.2) using $\tilde{\sigma}^2$ we expect to remove much of the noise while preserving the image features, particularly if the image is detailed. This information can then be used to define $\{\alpha_{i+\frac{1}{2}}\}$ in Step 2. In Step 2, the parameter ϵ is used both to stabilize the numerical problem and to give us more control over how adaptive the scheme will be. The larger ϵ is, the less variation there is in the weighting factor $\alpha(\vec{x})$ in (4.1). For ϵ very large, the adaptive scheme essentially becomes the standard scheme (2.2). On the other hand, if ϵ is chosen to be extremely small, the weighting factor $\alpha(\vec{x})$ may be too adaptive, resulting in an image where discontinuous edges are artificially introduced, due to the corruption of the image from the noise. The

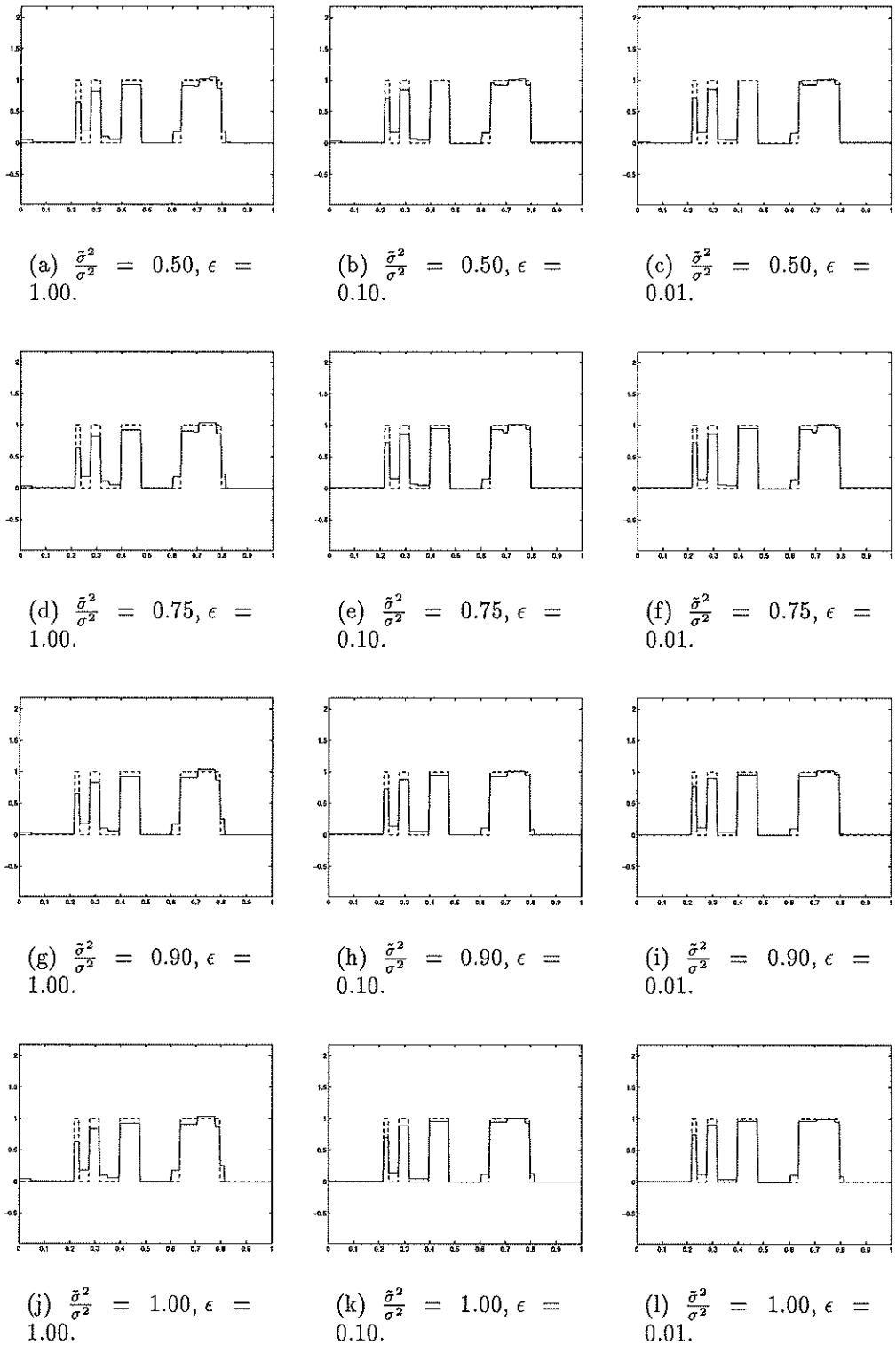
choice of ϵ will be directly related to the range of the grayscale values in the image. Appropriate choices of ϵ will be briefly addressed in our subsequent analysis of our results in R^2. With $\{\alpha_{i+\frac{1}{2}}\}$ defined, in Step 3 we again solve the constrained minimization problem, this time using the weighted TV functional.

We briefly comment about the extra computational cost of the adaptive scheme. In the standard TV minimizing restoration scheme, we solve the constrained minimization problem a single time. In the adaptive scheme we solve a constrained minimization problem twice. Thus the adaptive scheme could be about twice as expensive to apply as the standard scheme. However, we note that the partially restored image $\tilde{u}$ found in Step 1 can be used as a good initial guess for the iterative scheme used to find u in Step 3.[1] Because of this good initial guess for solving the minimization problem in Step 3, the computational work of the adaptive scheme as compared to the standard scheme is actually increased by *less* than a factor of 2.

4.2.4 Numerical Results in R^1

Example 4.2 To illustrate the effects of this scheme, we again consider the noisy image from the previous example. We use the true and noisy images from the previous example, as shown in Figure 4.1(a), and we can compare the results of our scheme to the results of standard (i.e. non-adaptive) TV restoration, found in the previous example and shown in Figure 4.1(b). In Figure 4.2 are the adaptively restored images found by using the values $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.50, 0.75, 0.90, 1.00$, and $\epsilon = 1.00, 0.10, 0.01$. Compare these results to the non-adaptively restored image given in Figure 4.1(b). Table 4.3 gives the ISNR for each of the restored

[1]The numerical problem that arises from solving the TV minimizing restoration problem (2.2) or (4.2) is nonlinear, and hence must be solved with an iterative solver, for which an initial estimate or guess is needed.



(a) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.50, \epsilon = 1.00$. (b) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.50, \epsilon = 0.10$. (c) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.50, \epsilon = 0.01$.

(d) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.75, \epsilon = 1.00$. (e) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.75, \epsilon = 0.10$. (f) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.75, \epsilon = 0.01$.

(g) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.90, \epsilon = 1.00$. (h) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.90, \epsilon = 0.10$. (i) $\frac{\tilde{\sigma}^2}{\sigma^2} = 0.90, \epsilon = 0.01$.

(j) $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.00, \epsilon = 1.00$. (k) $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.00, \epsilon = 0.10$. (l) $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.00, \epsilon = 0.01$.

Figure 4.2: Example 4.2: Feature-driven Adaptive TV minimizing (FATV) restoration of a noisy R^1 image.

images. For this example the results are better for larger values of $\frac{\tilde{\sigma}^2}{\sigma^2}$ and smaller values of ϵ.

$\frac{\tilde{\sigma}^2}{\sigma^2}$ \ ϵ	1.00	0.10	0. 01
0.50	13.63 dB	15.58 dB	15.93 dB
0.75	13.55 dB	15.96 dB	16.45 dB
0.90	13.66 dB	16.86 dB	18.20 dB
1.00	13.56 dB	16.72 dB	18.23 dB

Table 4.3: Example 4.2: the improved SNR for the adaptively restored images, using various ratios of $\frac{\tilde{\sigma}^2}{\sigma^2}$ an various values of ϵ. The ISNR for the standard (non-adaptive) restored image is 11.78 dB, as shown in Table 4.1, row (b). The restored images, shown in Figure 4.2, were found using the scheme given in Table 4.2.

4.3 Adaptive Image Restoration in R^2

4.3.1 Weighted TV Norm in R^2

We now extend our discussion to images in R^2. For a discrete R^2 image $\{u_{i,j}\}$, $1 \leq i, j \leq n$, the most natural way of discretizing (1.5) is

$$\begin{aligned} TV(u) &= \tfrac{1}{2}\{ \sum_{i=1}^{n-1}\sum_{j=1}^{n} \sqrt{[u_x(i+\tfrac{1}{2},j)]^2 + [u_y(i+\tfrac{1}{2},j)]^2} \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{n-1} \sqrt{[u_x(i,j+\tfrac{1}{2})]^2 + [u_y(i,j+\tfrac{1}{2})]^2} \ \} \end{aligned}$$

where

$$\begin{aligned}
u_x(i+\tfrac{1}{2},j) &\equiv \tfrac{\partial u(x,y)}{\partial x}|_{(x,y)=(x_{i+\frac{1}{2}},y_j)},\\
u_y(i+\tfrac{1}{2},j) &\equiv \tfrac{\partial u(x,y)}{\partial y}|_{(x,y)=(x_{i+\frac{1}{2}},y_j)},\\
u_x(i,j+\tfrac{1}{2}) &\equiv \tfrac{\partial u(x,y)}{\partial x}|_{(x,y)=(x_i,y_{j+\frac{1}{2}})},\\
u_y(i,j+\tfrac{1}{2}) &\equiv \tfrac{\partial u(x,y)}{\partial y}|_{(x,y)=(x_i,y_{j+\frac{1}{2}})}.
\end{aligned} \tag{4.6}$$

It is easy to see that the first and fourth terms in (4.6) should be discretized as

$$\begin{aligned}
u_x(i+\tfrac{1}{2},j) &= u_{i+1,j}-u_{i,j},\\
u_y(i,j+\tfrac{1}{2}) &= u_{i,j+1}-u_{i,j}.
\end{aligned}$$

One natural approach to discretizing the second and third terms in (4.6) would be

$$\begin{aligned}
u_y(i+\tfrac{1}{2},j) &= \tfrac{1}{4}(u_{i,j+1}+u_{i+1,j+1}-u_{i,j-1}-u_{i+1,j-1}),\\
u_x(i,j+\tfrac{1}{2}) &= \tfrac{1}{4}(u_{i+1,j}+u_{i+1,j+1}-u_{i-1,j}-u_{i-1,j+1}).
\end{aligned}$$

In this chapter we use the *minmod* scheme [22] in order to better preserve boundaries of image features:

$$\begin{aligned}
&u_y(i+\tfrac{1}{2},j) =\\
&\quad minmod[\,\tfrac{1}{2}(u_{i,j+1}+u_{i+1,j+1}-u_{i,j}-u_{i+1,j}),\\
&\qquad \tfrac{1}{2}(u_{i,j}+u_{i+1,j}-u_{i,j-1}-u_{i+1,j-1})\,]\\
&u_x(i,j+\tfrac{1}{2}) =\\
&\quad minmod[\,\tfrac{1}{2}(u_{i+1,j}+u_{i+1,j+1}-u_{i,j}-u_{i,j+1}),\\
&\qquad \tfrac{1}{2}(u_{i,j}+u_{i,j+1}-u_{i-1,j}-u_{i-1,j+1})\,]
\end{aligned}$$

where

$$minmod(a,b) = \frac{sign(a) + sign(b)}{2} \min(|a|, |b|).$$

The discrete weighted TV functional in R^2 is analogous to the discrete weighted TV functional (4.4) in R^1:

$$\begin{aligned} TV_\alpha(u) &= \tfrac{1}{2}\{ \sum_{i=1}^{n-1}\sum_{j=1}^{n} \alpha_{i+\frac{1}{2},j}\sqrt{[u_x(i+\tfrac{1}{2},j)]^2 + [u_y(i+\tfrac{1}{2},j)]^2} \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{n-1} \alpha_{i,j+\frac{1}{2}}\sqrt{[u_x(i,j+\tfrac{1}{2})]^2 + [u_y(i,j+\tfrac{1}{2})]^2}\ \} \end{aligned}$$

Our adaptive restoration scheme for R^2 is given in Table 4.4. It is analogous to the adaptive restoration scheme for R^1 given in Table 4.2.

We point out that as in the R^1 case, the partially restored image found in Step 1 can be used as a first guess in the iterative numerical scheme used to find the adaptively restored image in Step 3, so that the extra computational cost of the adaptive scheme is generally not significant.

4.3.2 Numerical Results in R^2

Example 4.3 We apply our adaptive TV minimizing restoration scheme, as well as the standard scheme for comparison, to five noisy test images: a cross, a triangle, a circle, a square and a hemisphere. The resulting images are found in Figures 4.3 - 4.5. Errors between the true image and each of the original (noisy) u_0, standard restored $u_{standard}$, and adaptively restored $u_{adaptive}$ images are given in Table 4.5.

As demonstrated in Figures 4.3 and 4.4 and Table 4.5, our scheme is superior to standard TV minimizing image restoration for denoising piecewise constant images. At the same time, our adaptive scheme has similar effects as standard

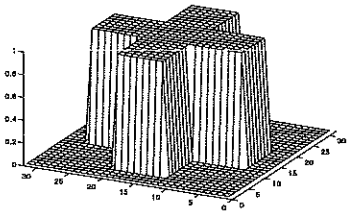
The FATV Scheme in R^2:
Feature-driven Adaptive Total Variation Minimizing Image Restoration in R^2

1. Find $\tilde{u}$ by solving
$$\min_{\tilde{u}} TV(\tilde{u}) \quad \text{subject to} \quad \|\tilde{u} - u_0\|^2 = \tilde{\sigma}^2,$$
where $0 < \frac{\tilde{\sigma}^2}{\sigma^2} \leq 1$.

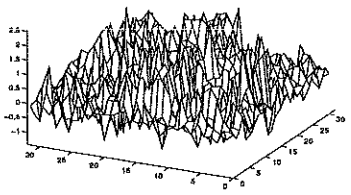
2. For $1 \leq i \leq n-1$, $1 \leq j \leq n$, define
$$\alpha_{i+\frac{1}{2},j} = \frac{1}{|\tilde{u}_{i+1,j} - \tilde{u}_{i,j}| + \epsilon},$$
and for $1 \leq i \leq n$, $1 \leq j \leq n-1$, define
$$\alpha_{i,j+\frac{1}{2}} = \frac{1}{|\tilde{u}_{i,j+1} - \tilde{u}_{i,j}| + \epsilon},$$
where $\epsilon > 0$.

3. Find the adaptively restored image u by solving
$$\min_{u} TV_\alpha(u) \quad \text{subject to} \quad \|u - u_0\|^2 = \sigma^2.$$

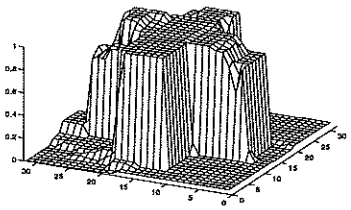
Table 4.4: Feature-driven Adaptive TV minimizing (FATV) image restoration scheme in R^2.



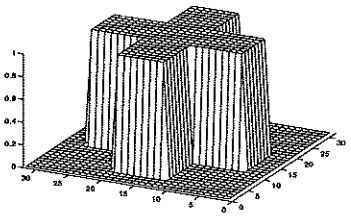
(a) True Image.



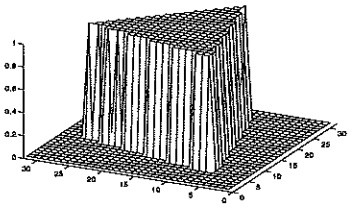
(b) Noisy image, SNR = 0 dB.



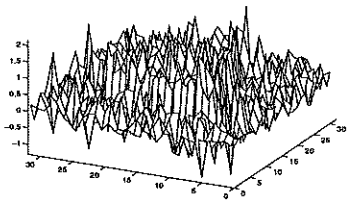
(c) Non-adaptively restored.



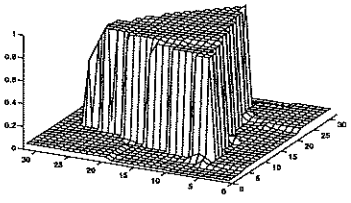
(d) Adaptively restored using $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.0$, $\epsilon = 0.1$.



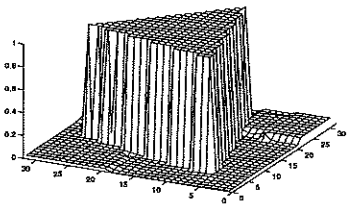
(e) True Image.



(f) Noisy image, SNR = 0 dB.

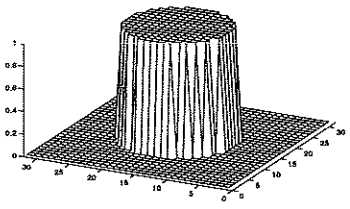


(g) Non-adaptively restored.

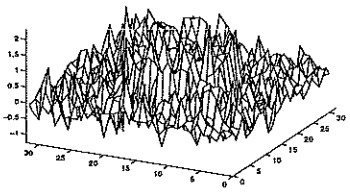


(h) Adaptively restored using $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.0$, $\epsilon = 1.0$.

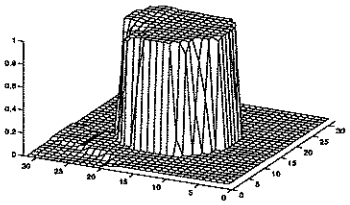
Figure 4.3: Example 4.3: image restoration using the FATV Scheme.



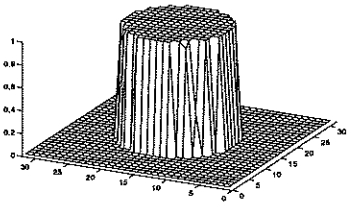
(a) True Image.



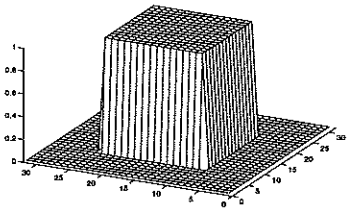
(b) Noisy image, SNR = 0 dB.



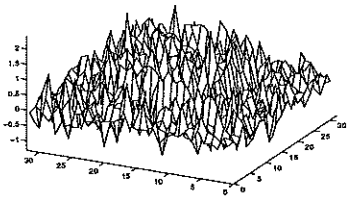
(c) Non-adaptively restored.



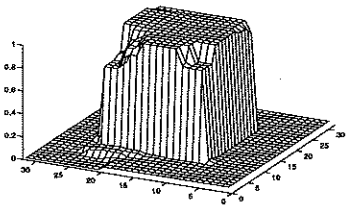
(d) Adaptively restored using $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.0$, $\epsilon = 1.0$.



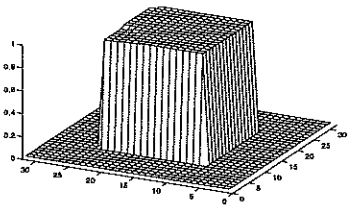
(e) True Image.



(f) Noisy image, SNR = 0 dB.



(g) Non-adaptively restored.

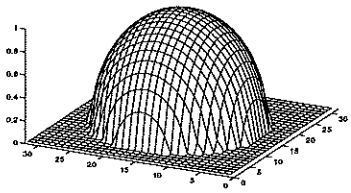


(h) Adaptively restored using $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.0$, $\epsilon = 1.0$.

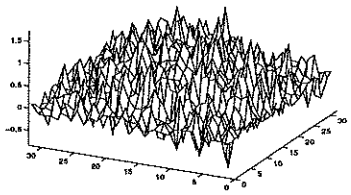
Figure 4.4: Example 4.3: image restoration using the FATV Scheme.

Image	ISNR, $u_{standard}$	ISNR, $u_{adaptive}$	ϵ
Cross	8.00 dB	25.80 dB	0.1
Triangle	12.75 dB	19.17 dB	1.0
Circle	13.75 dB	17.35 dB	1.0
Square	9.73 dB	17.12 dB	1.0
Hemisphere	0.39 dB	0.34 dB	1.0

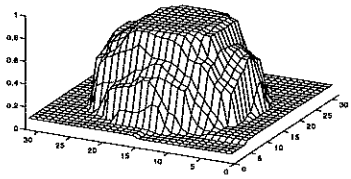
Table 4.5: Example 4.3: the improved SNR for both the standard (non-adaptive) $u_{standard}$ and adaptively restored $u_{adaptive}$ images, which are shown in Figures 4.3, 4.4 and 4.5.



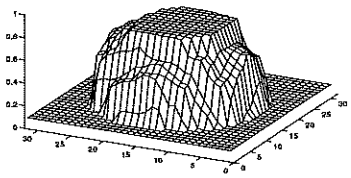
(a) True Image.



(b) Noisy image, SNR = 3 dB.



(c) Non-adaptively restored.



(d) Adaptively restored using $\frac{\tilde{\sigma}^2}{\sigma^2} = 1.0$, $\epsilon = 1.0$.

Figure 4.5: Example 4.3: image restoration using the FATV Scheme.

TV restoration for denoising smooth images, as demonstrated in Figure 4.5.

We found that relatively conservative values of ϵ were appropriate in restoring these R^2 test images. For most of the examples we used a value of $\epsilon = 1.0$, except for the case of denoising the cross, in which for this example $\epsilon = 0.1$ gave better results. We conclude that for an image with grayscale values ranging between 0 and 1, a good range would be $0.1 \leq \epsilon \leq 1.0$, with ϵ being closer to 1 to be more conservative. Of course, if the grayscale range of the image is greater than 0 to 1, ϵ could be chosen larger, in a linearly dependent way. We note that larger values of ϵ also result in a more stable numerical problem.

4.4 Summary

In this chapter, we have given a feature-driven spatially adaptive total variation minimizing image restoration scheme, where the adaptivity is realized by weighting the measure of total variation of the image. A spatially varying weighting factor is chosen to be inversely proportional to the likelihood of there being an edge (i.e. discontinuity) between two neighboring pixels. The approach given in this chapter is to determine the weighting factor by examining the size of the jumps between neighboring pixel values in a partially restored image. To control the effect of the weighting factor, as well as to improve the stability of the resulting numerical problem, a parameter is chosen (which herein is labelled ϵ). We found that the appropriate value of ϵ should be approximately equal or slightly less than the grayscale range of the image. Our adaptive TV minimizing image restoration scheme proved to be quite effective and superior to standard (non-adaptive) TV minimizing restoration in restoring piecewise constant image features. The adaptive and standard schemes were quite similar in restoring smooth image features. The extra numerical cost of solving the adaptive restora-

tion problem is not great relative to the cost of solving the standard restoring problem.

CHAPTER 5

Scale-driven Adaptive Total Variation Minimizing Image Restoration

In this chapter we use the theory developed in Chapter 3 to develop a spatially adaptive TV minimizing image restoration scheme which is driven by the local *scale* in the image. The end result is that the image is selectively restored to a precisely-controlled level of detail; that is, the processed image will be comprised of features with scale no smaller than a controlled threshhold. This scale threshhold is selected by the user, and may be constant or spatially varying, depending on the type of image to be restored and the purpose of the restoration.

When using our scheme, the user has explicit control over the minimum scale to be present in the restored image. The image manipulation is done selectively; that is, the restoration is only done where needed, in order to preserve as completely as possible the original features, even though they have been noise-degraded. Moreoever, inherent to TV image restoration, there is no blurring or shifting of important edge information, as edge location is preserved exactly. Our scheme can also be characterized as a spatially-selective anisotropic diffusion scheme, as our results demonstrate.

The balance of this chapter is as follows. In Section 5.1 we discuss the feature of automatic scale recognition which is inherent in TV regularization and which will be a important idea used in constructing our scheme. In Section 5.2.1 we

discuss the basic ideas and theory used to construct our scheme. The basic algorithm is given and discussed in Section 5.2.2, while a modified, more robust version of the algorithm is given and discussed in Section 5.3. In Section 5.4 we give results of applying our scheme to noisy images in R^1 and R^2. In Section 5.5 we give a summary of our results and discuss future work motivated by the results of this chapter.

5.1 Automatic Scale Recognition

Our scheme is constructed using the theory developed in Chapter 3, in which we give exact analytic solutions to the TV image restoration problem for specific cases. We found that the basic effect of TV image restoration, when solving

$$\min_u \quad \{ \tfrac{1}{2}\|u - u_0\|^2 + \alpha TV(u) \}, \tag{5.1}$$

is to change the intensity level of each feature, depending on the *scale* of the feature, in such a way as to reduce the contrast in the image. The change in intensity level δ is inversely proportional to *scale* and directly proportional to the regularization parameter α, as given in Chapter 3:

$$\boxed{\delta = \frac{\alpha}{scale}} \tag{5.2}$$

In general, equation (5.2) is important because of the insight it gives us into how and why TV regularization restores an image: noise (which in general can be thought of as some combination of relatively small-scaled features) is removed, while larger-scaled features are left relatively intact. The balance between noise removal and detail preservation in solving (5.1) is determined by the choice of α.

The scale of an image feature is defined in Chapter 3 as the ratio of the size a feature's area to the size of its boundary. That is, where Ω represents an image feature,

$$\boxed{scale = \frac{|\Omega|}{|\partial\Omega|}.} \tag{5.3}$$

For example, in a two-dimensional image, a circle with radius r would have scale

$$scale = \frac{\pi r^2}{2\pi r} = \frac{r}{2},$$

so that the scale is linearly increasing with the radius r. We note that this definition of scale is given for piecewise constant image features. Although this definition (5.3) does not extend exactly to non-piecewise constant image features, this has no detrimental effect on the results of our scheme.

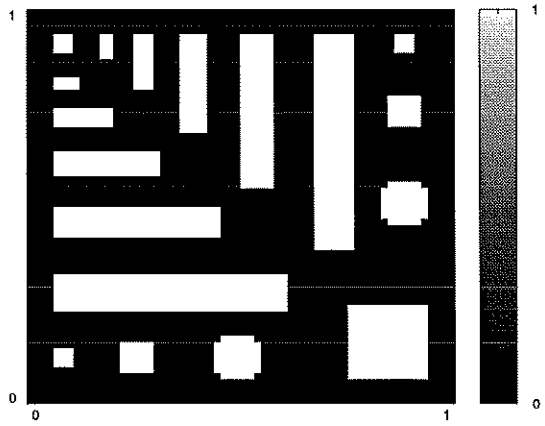
For a noise-free image with distinct piecewise constant features, we can find the scale of each image feature by solving the TV minimizing image restoration problem (5.1), using (5.2), rewritten as

$$\boxed{scale = \frac{\alpha}{\delta}.} \tag{5.4}$$

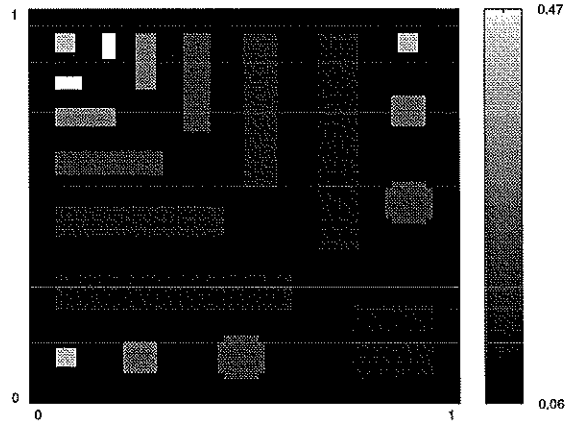
We give an example to illustrate how this is done.

Example 5.1 In Figure 5.1(a) is an image $u_0(\vec{x})$ comprised of simple piecewise constant features. To find the scale throughout the image, we solve the standard TV minimizing image restoration problem (5.1) to find $u(\vec{x})$. We then compute

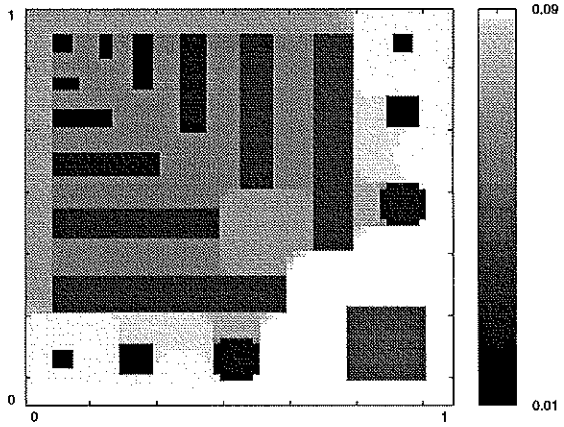
$$\delta(\vec{x}) = |u(\vec{x}) - u_0(\vec{x})|,$$



(a) Original image, u_0.



(b) The change in intensity level, $\delta(\vec{x}) = |u_0(\vec{x}) - u(\vec{x})|$.



(c) The scale of the image features, $scale(\vec{x}) = \frac{\alpha}{\delta(\vec{x})}$.

Figure 5.1: Example 5.1: automatic scale recognition by solving (5.1).

which is plotted in Figure 5.1(b). Finally, we use (5.4) to find

$$\boxed{scale(\vec{x}) = \frac{\alpha}{\delta(\vec{x})},}$$

which is plotted in Figure 5.1(c).

This feature of automatic scale recognition is used to construct our adaptive image restoration scheme, as discussed in the next section.

5.2 The SATV Scheme: Scale-driven Adaptive Total Variation Minimizing Image Restoration

5.2.1 Underlying Ideas of the Scheme

The purpose of our scheme is to resolve an image to a precise, user-controlled level of detail. For example, this level of detail might correspond to the scale of the image features of interest to the user. In this case, our scheme preserves the features of interest (i.e. those with scale larger than the selected threshhold), while removing the noise and/or other extraneous detail in the image. The scale throughout the image is determined while the restoration is being done, as discussed in the previous section.

In standard TV minimizing image restoration, we solve (5.1) in a *single* step, with an appropriate value of α (which is done automatically if solving the noise-constrained problem (2.2)). In our scheme, the approach is to regularize the image gradually and selectively in a process of multiple steps. This is accomplished by

repeatedly solving

$$\min_{u_{new}} \quad \{ \tfrac{1}{2}\|u_{new} - u_{old}\|^2 + \alpha\, TV(u_{new}) \}, \tag{5.5}$$

repeatedly, updating u_{old} at each step, as subsequently described. Initially we set $u_{old} = u_0$. Because this is a multi-step process, we use a smaller value of α in solving (5.5). For example, if in solving (5.1) we used the value of α, then using $\frac{\alpha}{10}$ in our multi-step scheme would require approximately 10 steps.

In each step, after finding u_{new} by solving (5.5), in theory we determine $scale(\vec{x})$, the scale of the features in the current image, as described in the previous section. We then allow the regularization to occur where the scale is *smaller* than the desired threshhold, $scale_{thresh}$; however, the regularization is not applied where the scale of an image feature is *larger* than the scale threshhold. In other words, where $\tilde{u}_{new}$ is the adaptively or selectively restored image, after solving (5.5) we determine $scale(\vec{x})$ and set

$$\tilde{u}_{new}(\vec{x}) = \begin{cases} u_{new}(\vec{x}) & \text{if } scale(\vec{x}) < scale_{thresh} \\ \\ u_{old}(\vec{x}) & \text{if } scale(\vec{x}) \geq scale_{thresh} \end{cases}.$$

So in a single step of this process we start with our current image u_{old}, and eventually find $\tilde{u}_{new}$, a selectively restored image where the regularization is applied only in certain areas (where $scale(\vec{x}) < scale_{thresh}$). At the conclusion of each selective restoration step, we redefine $u_{old} = \tilde{u}_{new}$, and repeat the step, solving (5.5) in each step, until convergence of the scheme. Convergence occurs when

$$\max_{\vec{x}} scale(\vec{x}) \geq scale_{thresh},$$

that is, when all features which are of scale smaller than the desired threshhold (such as noise) have been regularized from the image.

In theory, in each step of this process we check at each location in the image to see if

$$scale(\vec{x}) < scale_{thresh}.$$

In practice, we do not actually compute $scale(\vec{x})$ throughout the image. At each step it is easy to compute

$$\delta(\vec{x}) = |u_{new}(\vec{x}) - u_{old}(\vec{x})|.$$

This actually gives us the information we need without explicitly computing the scale, since (5.2) leads to the relationship

$$\begin{aligned} scale(\vec{x}) < scale_{thresh} \quad &\Longleftrightarrow \quad \delta(\vec{x}) > \delta_{thresh}, \\ \text{where} \qquad \delta_{thresh} &= \frac{\alpha}{scale_{thresh}}. \end{aligned} \tag{5.6}$$

So at each step, rather than computing $scale(\vec{x})$, we need only compute $\delta(\vec{x}) = |u(\vec{x}) - u_0(\vec{x})|$, in order to extract the desired information about scale of the various image features.

5.2.2 The Algorithm

In Table 5.1, we give the algorithm of the SATV Scheme for the R^2 case. It is obvious how this scheme would be constructed in the R^1 and R^3 cases.

5.2.3 A Single Iteration of Step 3

The SATV Scheme:
Scale-driven Adaptive Total Variation Minimizing Image Restoration in R^2

1. Choose $scale_{thresh}$, the desired scale threshhold, and compute δ_{thresh} using equation (5.6).

2. Choose α, to use for each iteration of Step 3.

3a. Given u_{old} and α, find u_{new}, the solution to
$$\min_{u_{new}}\{ \tfrac{1}{2}\|u_{new} - u_{old}\|^2 + \alpha\, TV(u_{new}) \}$$
(use $u_{old} = u_0$ for the first iteration of Step 3).

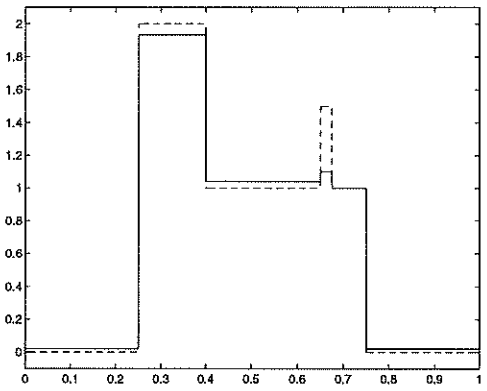
3b. For $1 \le i \le m, 1 \le j \le n$,
$$\delta(i,j) = |u_{new}(i,j) - u_{old}(i,j)|.$$

3c. If $\max_{i,j} \delta(i,j) \le \delta_{thresh}$,
then end the algorithm (else proceed to Step 3d).

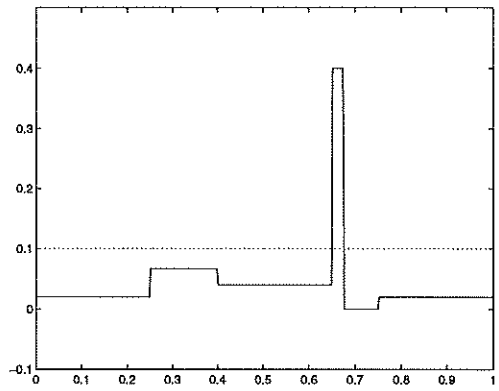
3d. For $1 \le i \le m, 1 \le j \le n$,
if $\delta(i,j) > \delta_{thresh}$,
then $\tilde{u}_{new}(i,j) = u_{new}(i,j)$,
else $\tilde{u}_{new}(i,j) = u_{old}(i,j)$.

3e. Update $u_{old} = \tilde{u}_{new}$, and repeat Step 3.

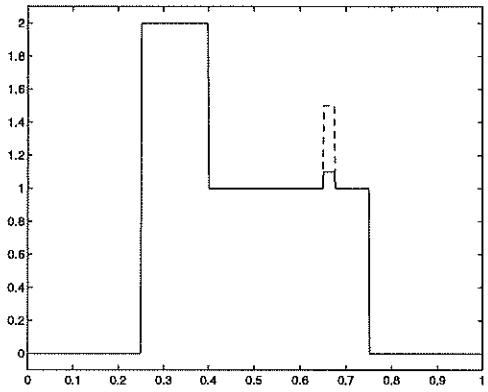
Table 5.1: Scale-driven Adaptive TV minimizing (SATV) image restoration scheme in R^2.



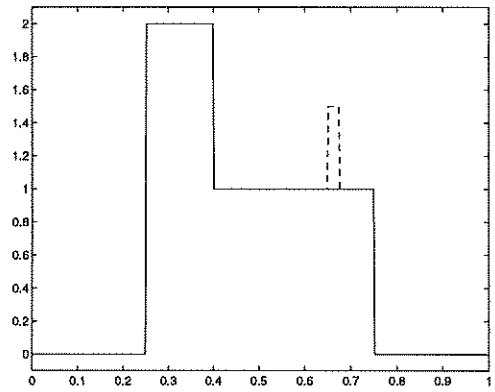
(a) Images u_{old} (solid line) and u_{new} (dotted line), which is found in Step 3a.



(b) Absolute change in intensity $\delta(x) = |u(x) - u_0(x)|$, found in Step 3b (solid line), and $\delta_{thresh} = 0.1$ (dotted line).



(c) $\tilde{u}_{new}$, the modified version of u_{new}, as found in Step 3d.



(d) The final image, after convergence of the scheme.

Figure 5.2: Example 5.2: an illustration of a *single* iteration of Step 3 in the SATV Scheme.

Example 5.2 To illustrate the effects of the Scale-driven Adaptive TV Minimizing Scheme (SATV Scheme), we examine a single iteration of the scheme when applied to a simple, noise-free function in R^1. In Figure 5.2(a) are the current image u_{old} and the regularized image u, found in Step 3a. In Figure 5.2(b), we plot $\delta(x) = |u(x) - u_0(x)|$, to determine where to apply the regularization (i.e., where to allow the changes from u_0 to u in Step 3a, to remain), depending on the scale in the image. We use the relationship (5.6) to do this. Finally, in Figure 5.2(c) we see the modified regularized image, $\tilde{u}_{new}$, which is equal to u_{new} in the regions where $scale(x) < scale_{thresh}$ (this is only in the region of the narrowest column, located between 0.6 and 0.7 in the domain of the image), and equal to u_{old} elsewhere. In 5.2(d) is how the image would appear upon convergence of the scheme.

5.2.4 Choosing $scale_{thresh}$ and α

The choices of parameters made in Steps 1 and 2 must be guided by understanding of the effects of Step 3, as analyzed above. In choosing $scale_{thresh}$ (Step 1), the following should be kept in mind: a larger value of $scale_{thresh}$ will result in less detail being present in the restored image (which includes better noise removal), and conversely, a smaller value of $scale_{thresh}$ will result in more detail (possibly including more noise) being present in the restored image. The choice should be made based mostly on the purpose of the restored image, but with some consideration of the type and level of noise in the image.

In choosing α (Step 2), the following should be kept in mind: a larger value of α will result in faster convergence to the restored image (i.e. fewer iterations of Step 3 to reach the final image), but with less sensitivity to detail; conversely, a smaller value of α will result in better sensitivity to detail preservation, but will

require more steps of the algorithm to reach the final image. For choosing both $scale_{thresh}$ and α, a bit of experience in applying the SATV Scheme will provide useful insight into how to choose appropriate values for a particular image.

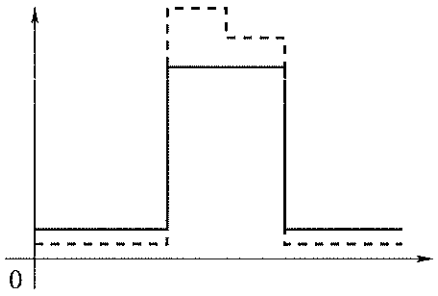
5.2.5 A Numerical Implementation Consideration in R^2 and R^3

Step 3a involves finding the solution to (5.5), the standard TV restoration problem. As shown in Chapter 3, standard TV image restoration naturally preserves the edges exactly for *radially symmetric* image features, but boundary deformation can occur for non-radially symmetric features, such as those with sharp corners. To overcome this effect, if desired, in applying TV restoration to two- or three-dimensional images, one might employ a numerical approximation scheme which is better at preserving edges for non-radially symmetric features, such as the *minmod* scheme used in [22]. In general this is recommended, although for some tasks it may be that this effect of boundary smoothing is desired. Again, experience will help in determining which approach to use.

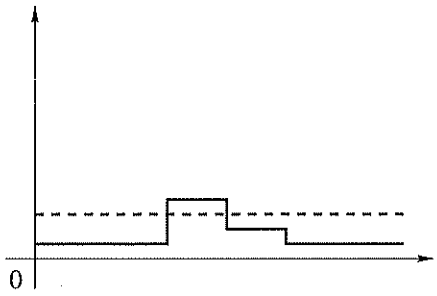
5.3 The Look-ahead Scale-driven Adaptive Total Variation Minimizing (LSATV) Image Restoration Scheme

In its current form, the SATV Scheme given in Table 5.1 is somewhat simplistic. Specifically, there is the possibility that, depending on the noise in the image, some image features of scale greater than the desired threshhold might be removed. An example of this breakdown is given below by examining single iteration of Step 3 of the SATV Scheme.

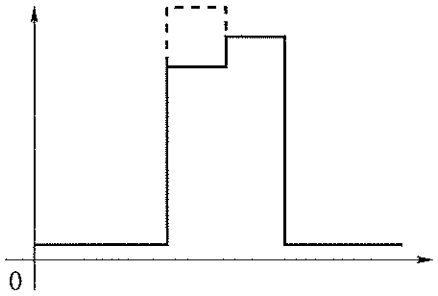
Example 5.3 In Figure 5.3(a) are the current image u_{old} and the regularized image u_{new}, found by solving (5.5). In this example we are interested in preserving



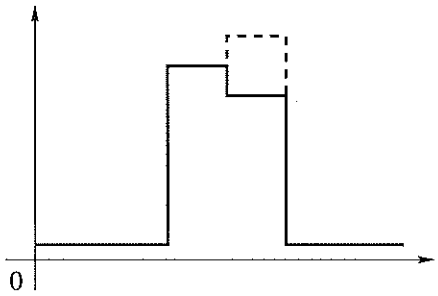
(a) Images u_{old} (solid line) and u_{new} (dotted line), which found in Step 3a of SATV Scheme.



(b) Absolute change in intensity $\delta(x) = |u(x) - u_0(x)|$, found in Step 3b (solid line) and δ_{thresh} (dotted line).

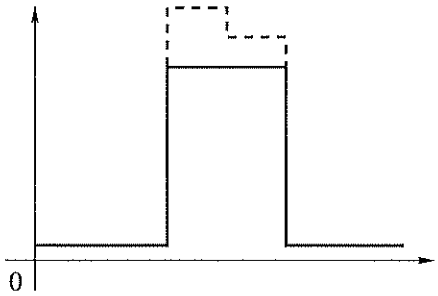


(c) $\tilde{u}_{new}$ (solid line), the modified version of u_{new}, as found in Step 3d.



(d) The results of another iteration of Step 3.

(e) The final image, after convergence of the SATV Scheme.



(f) The final image, after convergence of the LSATV Scheme, subsequently given in Table 5.2.

Figure 5.3: Example 5.3: an illustration of the breakdown in the SATV Scheme.

the single column as a whole, but not the division of the column in to two parts. In Figure 5.3(b), we plot $\delta(x) = |u(x) - u_0(x)|$, to determine where to apply the regularization (i.e., where to allow the changes, due to regularization in Step 3a, to remain), depending on the scale in the image. We use the relationship (5.6) to do this. The δ_{thresh} shown in Figure 5.3(b) corresponds to wanting to preserve image features with scale at least as great as the width of the entire column. Finally, in Figure 5.3(c) we see the modified regularized image, $\tilde{u}_{new}(\vec{x})$, which is equal to $u_{new}(\vec{x})$ in the regions where $scale(x) < scale_{thresh}$ and equal to $u_{old}(\vec{x})$ elsewhere. In Figure 5.3(d) gives the results after another iteration of Step 3 in the SATV Scheme. In (d) the old image (dashed line) is the selectively restored image found in the previous step, shown in Figure 5.3(c). In 5.3(e) is how the image would appear upon convergence of the scheme. The feature, which has a scale greater than $scale_{thresh}$ (which means we want to keep this feature, has been lost, due to the oscillatory-like behavior of the SATV Scheme. Figure 5.3(f) shows the final image which would be found if the LSATV Scheme, a more robust version of the SATV scheme, were used. The LSATV Scheme is given in Table 5.2. The difference between the LSATV and SATV schemes is that in the LSATV Scheme, a look-ahead step is included in the algorithm which checks to see if a certain image feature is actually part of a larger feature. The look-ahead idea is to take the selectively restored image $\tilde{u}_{new}$, apply standard TV regularization to it, and use the results to modify $\tilde{u}_{new}$ a second time before proceeding with another full iteration of the scheme. In this look-ahead step, we use the information to more accurately determine how the scale throughout the image relates to $scale_{thresh}$. A careful analysis of the LSATV Scheme will clarify the purpose of the scheme. The LSATV Scheme is the SATV Scheme, with Steps 3e and 3f added, and Step 3g being a modified version of Step 3e from the SATV Scheme, in order to accomplish this look-ahead approach. This

essentially doubles the work, but will remedy this type of breakdown which might otherwise occur.

5.4 Numerical Examples in R^1 and R^2

Example 5.4 To demonstrate in some detail how the Look-ahead Scale-driven ATV Scheme will affect a noisy image, we consider the R^1 example illustrated in Figures 5.4 - 5.6. The true and noisy images are shown in Figure 5.4(a). The subsequent images show the adaptively restored image at each step of the process. Figure 5.6(c) is the final image produced by our algorithm, along with the true image for comparison. Also for comparison we give Figure 5.6(d), which is the image found using standard TV restoration. In applying the LSATV Scheme, we used a scale threshhold that would remove all features of scale smaller than the column of width 0.05, the approximate width of the smallest column in the true (noise-free) image.

Example 5.5 In this example we apply the LSATV Scheme to a noisy R^2 image. We choose $scale_{thresh}$ to be equal to the smallest-scaled feature in the image, so our scheme will remove the noise in the image while preserving the features, particularly the smaller ones, as well as possible. In Figures 5.7(a), (b), (c) and (d) are the true, noisy, standardly restored and adaptively restored images, respectively.

As demonstrated in Figure 5.7, the adaptive scheme does a better job than standard TV restoration in preserving the smaller-scaled features, but it also tends to leaves the image looking somewhat "patchy." This "patchiness" is due to the noise (more specifically, the fact that even if the mean of the noise throughout

The LSATV Scheme:
Look-ahead Scale-driven Adaptive Total Variation Minimizing Image Restoration in R^2

1. Choose $scale_{thresh}$, the desired scale threshhold, and compute δ_{thresh} using equation (5.6).

2. Choose α, to use for each iteration of Step 3.

3a. Given u_{old} and α, find u_{new}, the solution to
$$\min_{u_{new}}\{ \tfrac{1}{2}\|u_{new} - u_{old}\|^2 + \alpha TV(u_{new}) \}$$
(use $u_{old} = u_0$ for the first iteration of Step 3).

3b. For $1 \le i \le m, 1 \le j \le n$,
$$\delta_1(i,j) = |u_{new}(i,j) - u_{old}(i,j)|.$$

3c. If $\max_{i,j} |\delta(i,j)| \le \delta_{thresh}$,
then end the algorithm (else proceed to Step 3d).

3d. For $1 \le i \le m, 1 \le j \le n$,
if $\delta_1(i,j) > \delta_{thresh}$,
then $\tilde{u}_{new}(i,j) = u_{new}(i,j)$,
else $\tilde{u}_{new}(i,j) = u_{old}(i,j)$.

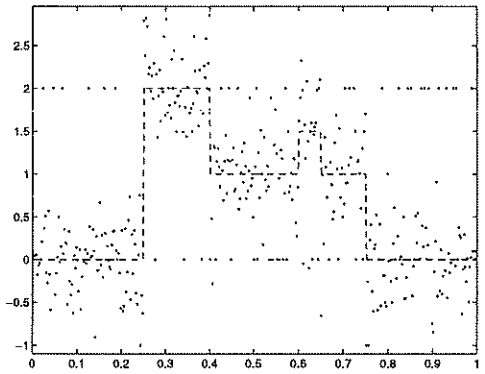
3e. Given $\tilde{u}_{new}$ and α, find $u_{lookahead}$, the solution to
$$\min_{u_{lookahead}} \{ \tfrac{1}{2}\|u_{lookahead} - \tilde{u}_{new}\|^2 + \alpha TV(u_{lookahead}) \}$$

3f. For $1 \le i \le m, 1 \le j \le n$,
$$\delta_2(i,j) = |u_{lookahead}(i,j) - \tilde{u}_{new}(i,j)|.$$

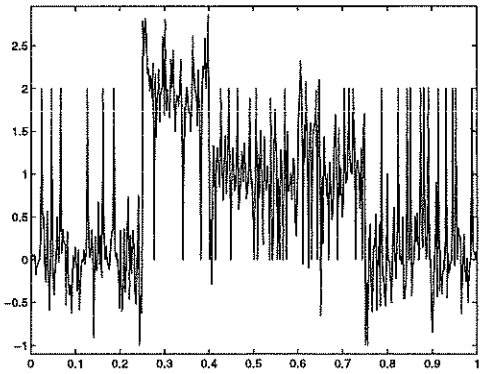
3g. For $1 \le i \le m, 1 \le j \le n$,
if $\max\{\delta_1(i,j), \delta_2(i,j)\} > \delta_{thresh}$,
then $\tilde{u}_{new}(i,j) = u_{new}(i,j)$,
else $\tilde{u}_{new}(i,j) = u_{old}(i,j)$.

3h. Update $u_{old} = \tilde{u}_{new}$, and repeat Step 3.

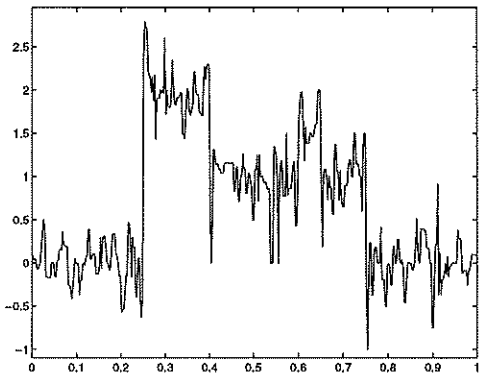
Table 5.2: Look-ahead Scale-driven Adaptive TV minimizing (LSATV) image restoration scheme in R^2.



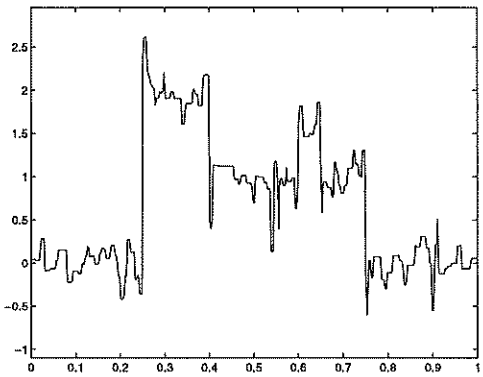
(a) Original image (dashed line) and noisy image (dotted line)



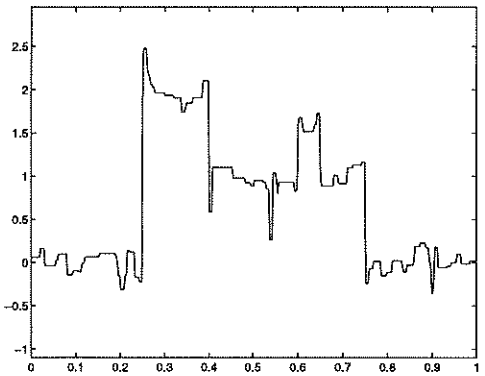
(b) Noisy image.



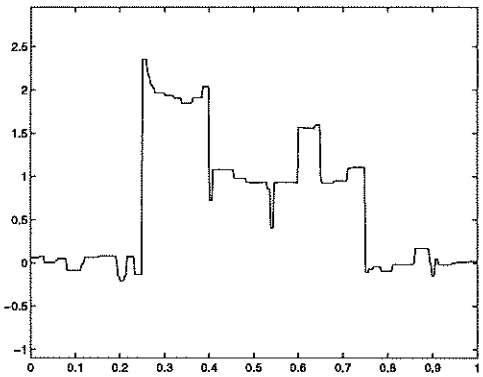
(c) After 1 iteration of LSATV Scheme.



(d) After 2 iterations of LSATV Scheme.

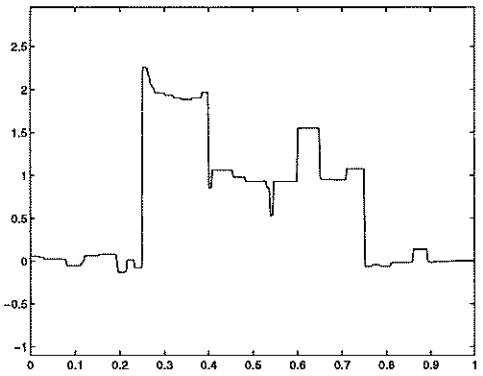


(e) After 3 iterations of LSATV Scheme.

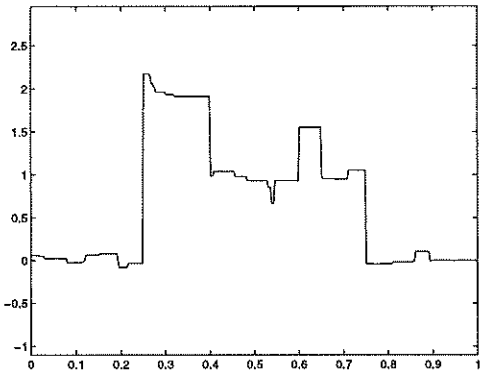


(f) After 4 iterations of LSATV Scheme.

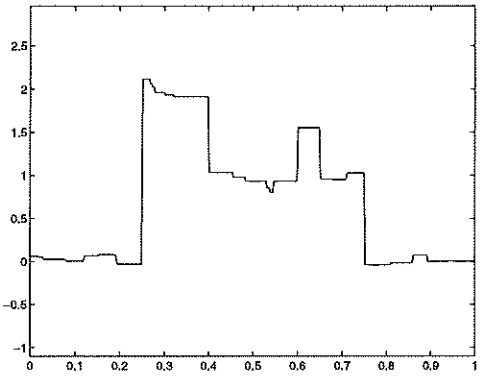
Figure 5.4: Example 5.4: the original image, and the first four steps of restoration using LSATV Scheme.



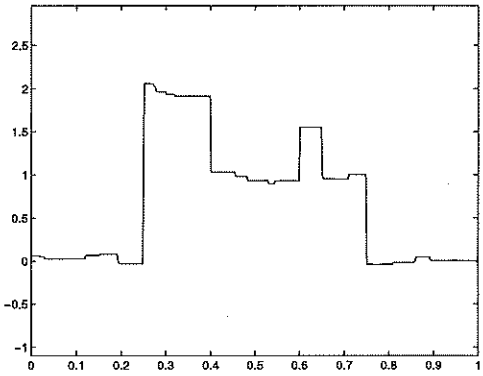
(a) After 5 iterations of LSATV Scheme.



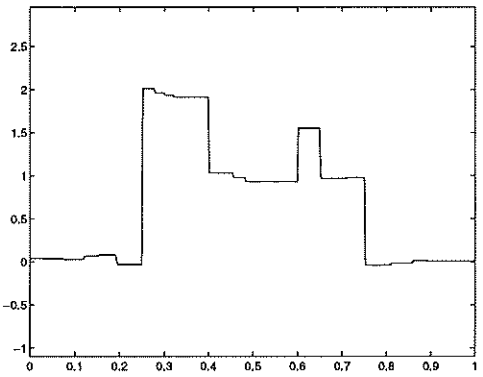
(b) After 6 iterations of LSATV Scheme.



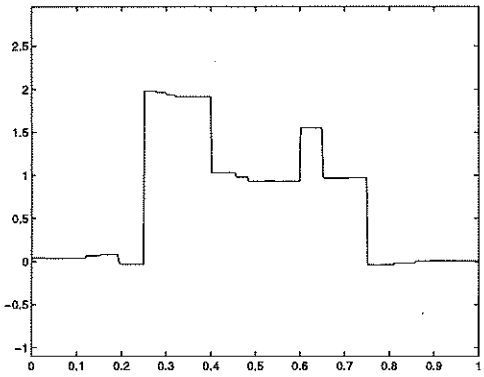
(c) After 7 iterations of LSATV Scheme.



(d) After 8 iterations of LSATV Scheme.

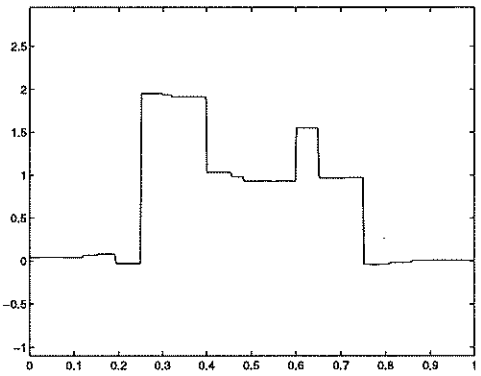


(e) After 9 iterations of LSATV Scheme.

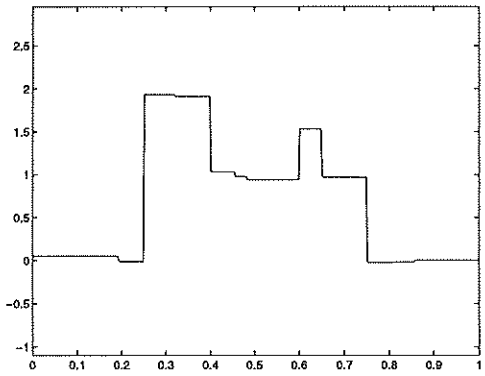


(f) After 10 iterations of LSATV Scheme.

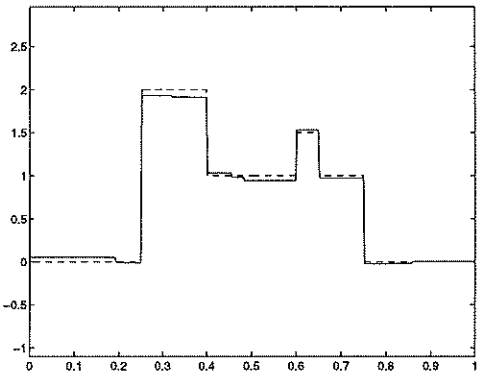
Figure 5.5: Example 5.4: the middle six steps of restoration using LSATV Scheme.



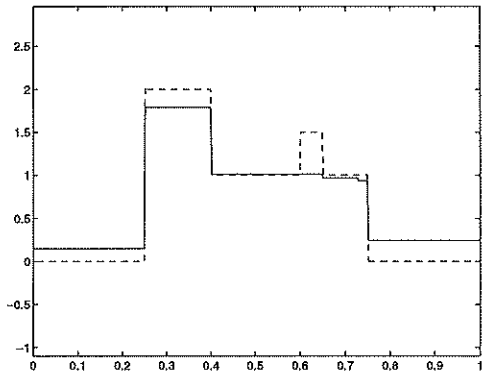
(a) After 11 iterations of LSATV Scheme.



(b) After 12 iterations of LSATV Scheme.

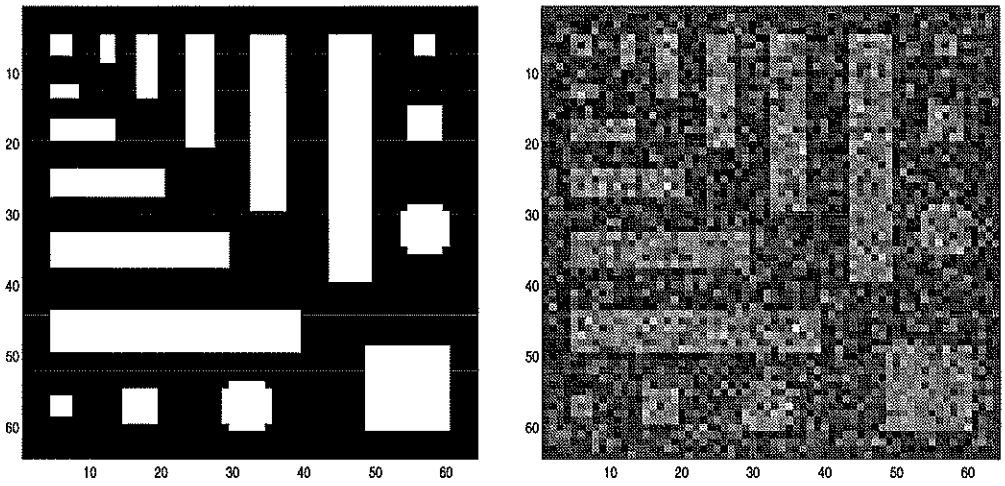


(c) True and adaptively restored images.



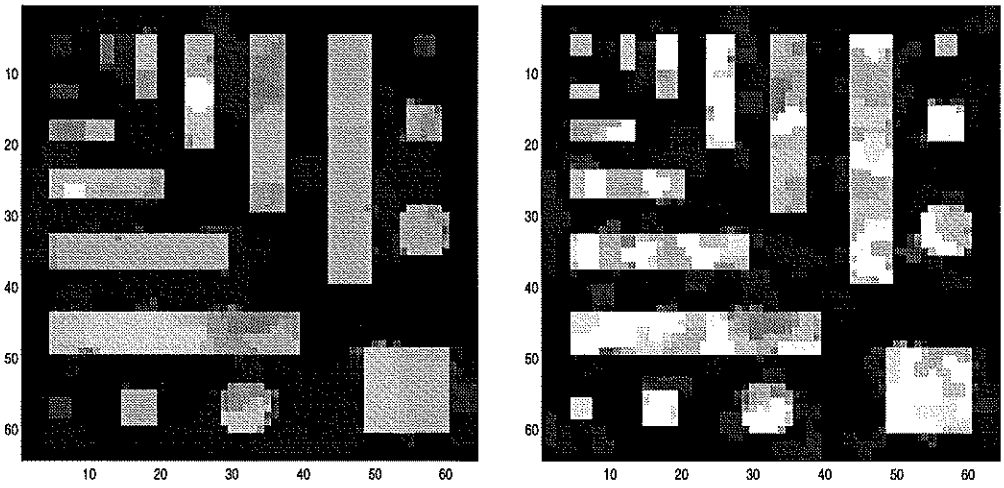
(d) True and standardly restored images.

Figure 5.6: Example 5.4: the final two steps of restoration using LSATV Scheme, and the final restored image, which is compared to the standardly restored image.



(a) Original Image

(b) Noisy Version of (a).



(c) Standard restoration of (b).

(d) Adaptive restoration of (b) using $scale_{thresh} \approx 0.01$.

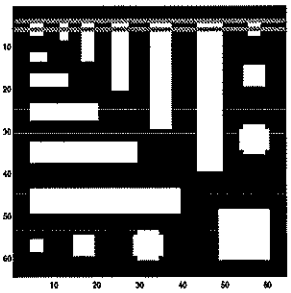
Figure 5.7: Example 5.5: LSATV Scheme applied to a noisy R^2 image.

the image is 0, subsets of the noise will generally have nonzero means), as well as the fact that regularization halts in any location of the image when a feature with scale larger than the scale threshhold has been found in that location of image. This effect and methods for overcoming it are currently under study.

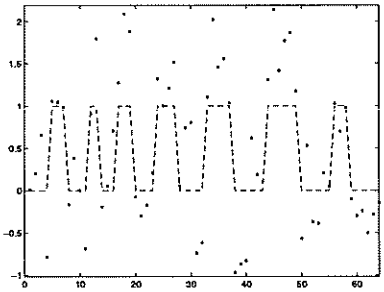
In order to make the results more obvious, we give a cross-section of the images in Figure 5.7. Figure 5.8(a) shows the location of the cross-sections, and Figure 5.8(b) shows the cross-sections of the true and noisy images at that location. In Figure 5.8(c) we see the cross-sections of the true and *adaptively* restored images and in 5.8(d) are the cross-sections of the true and *standardly* restored images. These R^1 cross-sections give us a clearer and more favorable comparison of the LSATV Scheme and the standard TV Scheme.

5.5 Summary

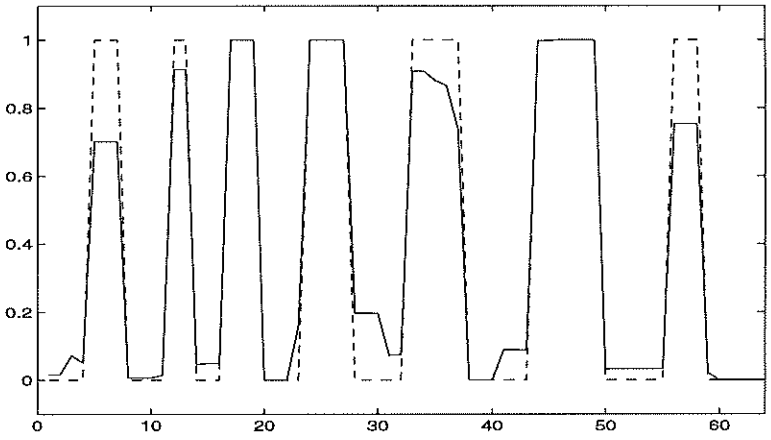
We have used the theory developed in Chapter 3 to develop a spatially adaptive TV minimizing image restoration scheme which is driven by the local scale in the image. This is a multi-step process which involves selective restoration of the image at each step. At each step, the current image is restored, and using the automatic scale recognition feature of TV minimizing function regularization, we determine where to allow the restoration to occur based on the scale in the current image. This process is repeated until convergence, updating the current image with the selective-restored image at each step. Convergence is reached when the processed image is comprised only of features with scale no smaller than a controlled threshhold. This scale threshhold is selected by the user, and may be constant or spatially varying, depending on the type of image to be restored and the purpose of the restoration. The end result is an image comprised of features



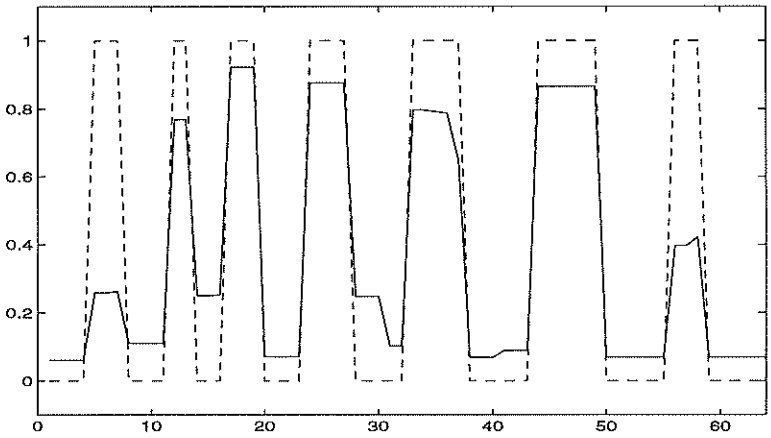
(a) Location of cross-section.



(b) Cross-section of true and noisy images.



(c) Adaptively restored image (solid line).



(d) Standardly restored image (solid line).

Figure 5.8: Example 5.5: a cross-section of the results shown in Figure 5.7.

with scale greater or equal to the given scale threshhold. Our scheme can also be characterized as a spatially-selective, TV minimizing anisotropic diffusion scheme.

CHAPTER 6

Summary

We have analyzed the exact effects of total variation (TV) minimizing function regularization in R^1, R^2 and R^3. This more precise understanding of TV regularization has enabled us to more intelligently construct TV minimizing image restoration schemes, as well as to better understand what types of images (and what types of image degradation) are most effectively improved by TV restoration.

We have analytically found exact solutions to the nonlinear TV regularization problem for simple but important cases, which can be used to better understand the effects of TV regularization for more general cases. Our results were found by solving the unconstrained (Tikhonov) formulation of the TV minimizing function regularization problem, in which a regularization parameter (which may be spatially varying) is chosen to determine the balance between goodness of fit to the original (e.g. observed, noisy) function and the amount of regularization to be done to the function. Our results are equally useful for the noise-constrained formulation of the problem.

We gave exact formulae that describe the effect of TV regularization when applied to radially symmetric image features. These formulae are also applied to more general functions. Four important results which we proved are

- TV regularization of piecewise constant (noise-free or noisy) radially sym-

metric functions results in a piecewise constant function, with edge location being preserved exactly;

- function intensity change due to TV regularization is inversely proportional to local feature scale, is independent of original intensity, and is directly proportional to the regularization parameter;
- for smooth radially symmetric function features, function intensity change is inversely proportional to radial position and directly proportional to the regularization parameter;
- TV regularization is local in a certain sense.

We have used our theoretical results to develop two spatially-*adaptive* TV minimizing image restoration schemes. Each of our schemes is quite effective in image restoration in the examples given in this Dissertation. Both schemes are built around the theory developed in our analysis of TV restoration.

In the first scheme we accomplished adaptivity by locally weighting the measure or computation of the total variation of the image. We do this based on the relative likelihood of the presence of an edge in each discrete location throughout the image. This likelihood is a function of the difference of adjacent neighbors in both the x- and y-directions (in R^2 images). The weighting factor is chosen to be decreasing as the relatively likelihood of the presence of an edge increases. This allows for less regularization where edges are present and more regularlization where there are no edges, which results in better detail preservation and better overall noise removal.

The second adaptive TV minimizing image restoration scheme is driven by the scale of individual features in the image. This is a multi-step approach which involves selectively applying restoration in each step, only where the scale of the

image features is smaller than a user-controlled threshhold. This causes noise to be removed while preserving image features as well as possible. The end result is an image comprised of features with scale greater or equal to the given scale threshhold. The process can be characterized as a scale-sensitive, TV minimizing anisotropic diffusion process, where the diffusion occurs only to features of scale smaller than a user-specified threshhold (e.g. noise).

The theory developed in this Dissertation is useful both in understanding (and thus justifying) TV minimizing image restoration, as well as using this understanding in developing more effective TV image restoration schemes. The two adaptive schemes which we have developed can be viewed as prototypes of an array of adaptive image restoration schemes which will be developed in the future as a result of our theoretical results.

Future work motivated by the results in this Dissertation includes:

- In regards to our theoretical analysis:
  - Extend theoretical results to TV minimzing anisotropic diffusion;
  - Develop theory to understand/predict results when numerical implentation involves using the *minmod* scheme;
  - Develop numerical schemes which take advantage of the localness of TV regularization;
  - Develop theory which describes how the choice of β^2 affects the restored image.
- For feature-driven adaptive restoration schemes (FATV):
  - Explore other ways of choosing $\alpha(\vec{x})$ in $TV_\alpha(u) = \int \alpha(\vec{x}) \left|\nabla u(\vec{x})\right| d\vec{x}$.

- For scale-driven adaptive image regstoration schemes (SATV/LSATV):
  - Allowing $scale(\vec{x})$ to be spatially varying, for example depending on the statistical properties of the image;
  - Developing Scale-driven Adaptive Total Variation Minimizing Anisotropic Diffusion, with (2.8), using the automatic scale recognition feature of TV function regularization to determine where to apply the diffusion process;
  - Interactive restoration;
  - Numerical concerns, e.g. computational efficiency;
  - Overcoming patchiness in the image caused the the scheme (e.g. this might be dealt with by using a spatially varying $scale_{thresh}(\vec{x})$).

References

[1] R. Acar and C. Vogel. "Analysis of Boundary Variation." *Inverse Problems*, Volume 10 (1994), pp. 1217-1229.

[2] L. Alvarez and J. Morel. "Formalization and Computational Aspects of Image Analysis." *Acta Numerica*, pp. 1-59, 1994.

[3] M. R. Banham and A. K. Katsaggelos. "Spatially adaptive wavelet-based multiscale restoration." Technical report, Dept. of E.E. and C.S., Northwestern University, 1995. IEEE Transactions on Image Processing, April 1996.

[4] J. Biedmond and R. Lagendijk. *Iterative Identification and Restoration of Images.* Kluwer Academic Publishers, Dordrecht, 1991.

[5] P. Blomgren and T. Chan. "Total Variation and Beyond." Talk at *SIAM Annual Meeting*, Stanford, 1997.

[6] P. Blomgren and T. Chan. "Color TV: Total Variation Methods for Restoration of Vector Valued Images." UCLA Math Department CAM Report 96-5, February 1996.

[7] P. Blomgren, T. Chan, P. Mulet and C.K. Wong. "Total Variation Image Restoration: Numerical Methods and Extensions," To appear in the *Proceedings of ICIP '97.*

[8] P. Blomgren, T. Chan and P. Mulet. "Extensions to Total Variation Denoising." To appear in the *Proceedings of SPIE*, Volume 3162 (1997).

[9] T. Chan and C.K. Wong. "Total variation blind deconvolution." UCLA Math Department CAM Report 96-45, November 1996.

[10] T. Chan, G. Golub and P. Mulet. *A Nonlinear Primal-Dual Method for Total Variation-based Image Restoration*, in ICAOS '96, 12th Int'l Conf. on Analysis and Optimization of systems: Images, wavelets and PDE's, Paris, June 26-28, 1996, M. Berger, R. Deriche, I. Herlin, J. Jaffre and J. Morel, eds., no. 219 in Lecture Notes in Control and Information Sciences, 1996, pp. 241-252.

[11] Tony F. Chan, Pep Mulet. "Iterative Methods for Total Variation Image Restoration." UCLA Math Department CAM Report 96-38, October 1996. To appear as a chapter in *Iterative Methods in Scientific Computing*, proceedings of the *Winter School on Iterative Methods in Scientific Computing*, (R. Chan, T. Chan, G. Golub, editors), Springer Verlag, 1997.

[12] D. Dobson and F. Santosa. "Recovery of Blocky Images from Noisy and Blurred Data." *SIAM Journal of Applied Mathematics*, Volume 56 (1996), pp. 1181-1198.

[13] D. Dobson and O. Scherzer. *Analysis of Regularized Total Variation Penalty Methods for Denoising*, Inverse Problems, Volume 12 (1996), pp. 601-617.

[14] D. Dobson and C.R. Vogel. "Convergence of An Iterative Method for Total Variation Denoising." *SIAM Journal of Numerical Analysis*, to appear.

[15] R.C. Gonzalez and R.E. Woods. *Digital Image Processing.* Addison Wesley, 1993.

[16] M. Hanke and P.C. Hansen. "Regularization Methods for Large Scale Problems." *Survey of Mathematics in Industry*, Volume 3 (1993), pp. 253-315.

[17] P. Hansen. "Analysis of discrete ill-posed problems by means of the L-curve." *SIAM Review*, Volume 34 (1992), pp. 561-580.

[18] P.L. Lions, S. Osher, and L. Rudin. "Denoising and Deblurring Algorithms with Constrained Nonlinear PDE's." *SIAM Journal of Numerical Analysis*, submitted.

[19] J.M. Morel and S. Solimini. *Variational Methods for Image Segmentation.* Birkhauser, Boston, 1995.

[20] P. Perona and J. Malik. "Scale Space and Edge Detection using Anisotropic Diffusion." *IEEE Transactions on Pattern Analysis Macine Intelligence.* Volume 12 (1990), pp. 629-639.

[21] L. Rudin. "MTV-multiscale total variation principle for a PDE-based solution to non-smooth ill-posed problem." Technical Report, Cognitech, Inc. Talk presented at the Workshop on Mathemaatical Methods in Computer Vision, University of Minnesota, September 11-15, 1995.

[22] L. Rudin, S. Osher, and E. Fatemi. "Nonlinear Total Variation Based Noise Removal Algorithms." *Physica D*, Volume 60 (1992), pp. 259-268.

[23] Guillermo Sapiro and Dario L. Ringach. *IEEE Transactions on Image Processing*, to appear.

[24] G. Strang. "L^1 and L^∞ Approximation of Vector Fields in the Plane." *Lecture Notes in Num. Appl. Anal.*, Volume 5 (1982), pp. 273-288.

[25] B.M. ter Haar Romeny, Editor. *Geometry-driven Diffusion in Computer Vision.* Kluwer Academic Publishers, Dordrecht, 1994.

[26] C.R. Vogel and M.E. Oman. "Iterative Methods for Total Variation Denoising." *SIAM Journal on Scientific Computing*, Volume 17 (1996), pp. 227-238.