

PDE-based image restoration, I: Anti-staircasing and anti-diffusion^{*}

Kisee Joo[†] and Seongjai Kim[‡]

May 16, 2003

Abstract

This article is concerned with simulation issues arising in the PDE-based image restoration such as the total variation minimization (TVM) and its generalizations. In particular, we study the issues of staircasing and excessive dissipation of TVM-like smoothing operators. A strategy of scaling the algebraic system and a non-convex minimization are considered respectively for anti-staircasing and anti-diffusion. Furthermore, we introduce a variable constraint parameter to better preserve image edges. The resulting algorithm has been numerically verified to be efficient and reliable in denoising. Various numerical results are shown to confirm the claim.

Key words. Total variation minimization, anti-staircasing, anti-diffusion, non-convex optimization, constraint parameter.

AMS subject classifications. 35K55, 65M06, 65M12.

^{*}Technical Report #2003-07 Department of Mathematics, University of Kentucky, Lexington, KY 40506

[†]Mokpo National Maritime University, School of Maritime Transportation, 571-2 Chukyo-dong, Mokpo, Chonnam 530-729 South Korea Email: jksjoo@mmu.ac.kr. Visiting: Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506-0027 USA

[‡]Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506-0027 USA Email: skim@ms.uky.edu The work of the author is supported in part by NSF grant DMS-0107210

1. Introduction

Mathematical frameworks employing recent powerful tools of partial differential equations (PDEs) and functional analysis have emerged in image restoration such as denoising and enhancement [1, 5, 7, 12, 13, 16, 18, 20, 22, 27], color image processing [3, 4, 9, 14, 15, 23, 24, 25], and inpainting [8, 10, 17]. These PDE-based techniques have been extensively studied to answer fundamental questions in image processing. They turn out to allow researchers not only to introduce innovative mathematical models but also to analyze and improve traditional algorithms. Good references to work on them are e.g. Aubert-Kornprobst [2] and Osher-Fedkiw [19].

Image restoration is often necessary as a pre-processing for segmentation and compression; good denoising methods have strong demands. However, those PDE-based methods can show some drawbacks unless the governing equations are incorporating appropriate parameters and discretized by suitable numerical schemes. The selected model parameters and numerical schemes should be able to capture characteristics of the image. The development of appropriate *numerical* techniques for the PDE models is another important component of PDE-based approaches. The main goal of this article and the companion one [11] is to develop *efficient* and *reliable* numerical methods for image restoration via the simulation of PDE models such as the *total variation minimization* (TVM) and the *motion by mean curvature* (MMC).

When the standard TVM [21, 22] is applied for image smoothing, the image can be locally flattened due to huge diffusion near critical points where the gradient magnitude of the image is zero, i.e., $|\nabla u| = 0$. Such a phenomenon is called *staircasing* in the literature. Even though the gradient magnitude is regularized as

$$|\nabla^\varepsilon u| = (u_x^2 + u_y^2 + \varepsilon^2)^{1/2}, \quad \text{for some } \varepsilon > 0,$$

the staircasing can happen unless the regularization parameter ε is sufficiently large. But it is not a good idea to set ε large; one can easily observe a larger amount of blur in the restored image as the regularization parameter grows. These conflicting requirements must be dealt with effectively, which is the main goal of the current article. To focus on the main theme, we assume that the noise is Gaussian. In the companion paper [11], we will consider a strategy for the removal of combinations of impulse and Gaussian noises via a MMC-TVM hybridization, efficient numerical schemes for image derivatives, and color image processing in the angle domain with the brightness-chromaticity decomposition.

An outline of the paper is as follows. In Section 2, we review TVM, along with a brief explanation of staircasing. Section 3 introduces a new method for anti-staircasing; the basic idea is to scale the algebraic system for its elements not to exceed to a prescribed number. Such a strategy turns out to preserve a certain degree of convexity and concavity of the image and therefore it is *little* dissipative.

In Section 4, we consider two anti-diffusion strategies, the non-convex minimization ($p = 1 - \omega < 1$) and the variable constraint parameter $\beta(\mathbf{x}, t)$, for the purpose of edge enhancement and edge-preservation. These three anti-staircasing/anti-diffusion techniques (algebraic scaling, non-convex minimization, and variable constraint parameter) are carefully combined to optimize the performance of TVM for the removal of Gaussian noise. In Section 5, we present numerical results for the new strategies we have introduced in this article. It has been numerically verified for various images that the new techniques can remove the noise efficiently, preserving the edges quite satisfactorily. The last section concludes our experiments.

2. Preliminaries

Let u_0 be the observed image of the form

$$u_0 = K * u + e, \quad (2.1)$$

where K is a blurring operator and e denotes the noise. In this article, our main concern is to remove Gaussian noise. Thus we assume that $K = I$, the identity operator, and that e denotes a certain Gaussian noise.

A common denoising technique is to minimize a functional of gradient, as

$$\min_u \mathcal{F}_p(u), \quad \mathcal{F}_p(u) = \int_{\Omega} |\nabla u|^p d\mathbf{x} + \lambda \|u_0 - u\|^2, \quad (2.2)$$

where $\lambda \geq 0$. When $p = 1$, the first term in $\mathcal{F}_1(u)$ is called the *total variation* (TV).

It is often convenient to transform the minimization problem (2.2) into a differential equation, called the *Euler-Lagrange equation*. Recall that for the minimization problem in 2D of the form

$$\min_u \int_{\Omega} f(\mathbf{x}, u, \nabla u) d\mathbf{x},$$

the Euler-Lagrange equation [26, §3.3] reads

$$\left(\frac{d}{dx} \frac{\partial}{\partial u_x} + \frac{d}{dy} \frac{\partial}{\partial u_y} - \frac{\partial}{\partial u} \right) f = 0.$$

When the above is applied for (2.2), we have

$$p \nabla \cdot \left(\frac{\nabla u}{|\nabla u|^{2-p}} \right) - 2\lambda (u - u_0) = 0. \quad (2.3)$$

Parameterizing the descent direction by an artificial time t , the resulting *evolutionary* Euler-Lagrange equation can be formulated as

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|^{2-p}} \right) + \beta(u_0 - u), \quad (2.4)$$

where $\beta (= 2\lambda/p)$ is the *constraint parameter*. Note that the restored image becomes closer to u_0 as β grows.

A very important and interesting case is when $p = 1$. Note that the minimization process with \mathcal{F}_1 in (2.2) does not make the piecewise constant images blurry, because a blurrier version of such images does not reduce the value of the functional. Thus one may expect the same property for the corresponding Euler-Lagrange equation, (2.4) with $p = 1$:

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \beta(u_0 - u). \quad (2.5)$$

However, in practice, the above *total variation minimization* (TVM) introduces “blur” and/or “staircasing” to the images unless the given image u_0 is globally linear.

Since the denominator $|\nabla u|$ cannot be zero in the computation, the term is often regularized as

$$|\nabla u| \approx |\nabla^\varepsilon u| := (u_x^2 + u_y^2 + \varepsilon^2)^{1/2}, \quad (2.6)$$

for some $\varepsilon > 0$ small. In this article, we will call ε the *regularization parameter*. The corresponding minimization problem for the regularization reads

$$\min_u \mathcal{F}_{\varepsilon,p}(u), \quad \mathcal{F}_{\varepsilon,p}(u) = \int_{\Omega} |\nabla^\varepsilon u|^p d\mathbf{x} + \lambda \|u_0 - u\|^2, \quad (2.7)$$

and its Euler-Lagrange equation becomes

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla^\varepsilon u|^{2-p}} \right) + \beta(u_0 - u). \quad (2.8)$$

When (2.8) is discretized, the diffusion contributed by the term $1/|\nabla^\varepsilon u|^{2-p}$ can be large when $p < 2$ and $|\nabla u| \approx 0$. That is,

$$\frac{1}{|\nabla^\varepsilon u|^{2-p}} = \frac{1}{(u_x^2 + u_y^2 + \varepsilon^2)^{(2-p)/2}} \approx \frac{1}{\varepsilon^{2-p}}, \quad \text{for } |\nabla u| \approx 0.$$

This implies that TVM ($p = 1$) can enforce a strong diffusion on relatively flat regions and therefore the image can be transformed into locally constant there; staircasing happens. The removal of such staircasing has been an interesting research topic [3, 4, 6]. A modified approach can be formulated as

$$p = \begin{cases} 2, & \text{if } |\nabla u| \leq g_0, \\ 1, & \text{if } |\nabla u| \geq g_1, \\ \frac{1}{2} \left[\cos \left(\frac{|\nabla u| - g_0}{g_1 - g_0} \pi \right) + 3 \right], & \text{otherwise,} \end{cases} \quad (2.9)$$

where g_0 and g_1 are prescribed constants ($g_0 < g_1$). See §5.1 for a reasonable choice of g_0 and g_1 .

The main goals of the article are

- to understand the staircasing and blurring mechanisms on TVM and
- to suggest effective/efficient techniques for anti-staircasing and anti-diffusion.

We begin with an effective but simple anti-staircasing method, in the following section.

3. Anti-Staircasing and Algebraic Scaling

In the following we will introduce an effective strategy which prevents staircasing. Recall first that the staircasing is based on the large quantities in the algebraic system (corresponding to a large diffusion coefficient). Thus there is a good possibility to achieve anti-staircasing by scaling down the corresponding elements of the algebraic system. We suggest the following: scale the elements of the matrix $A = (a_{ij})$ as

$$\begin{aligned} &\text{Select a value } a_{\max} > 0. \\ &\text{For each row } \mathbf{a}_i \text{ of } A = (a_{ij}), \text{ if } |a_{ii}| > a_{\max}, \mathbf{a}_i \leftarrow \frac{a_{\max}}{|a_{ii}|} \mathbf{a}_i. \end{aligned} \quad (3.1)$$

Or, one can implement a smooth revision, instead of the above abrupt adjustment.

With the above scaling of the algebraic system, there will be no staircasing, when the maximum value a_{\max} is set to be small enough. Also, it turns out to be less dissipative than the variable p as in (2.9).

To verify the above claims, we consider 1D case for simplicity. Let $x_i = ih$ be the grid points (or cell centers). Let the image have a *strictly* convex part at a location, having a local minimum at $i = i_0$. That is, introducing $u_{x,i} = u_x(x_i)$,

$$\cdots < u_{x,i_0-2} < u_{x,i_0-1} < u_{x,i_0} = 0 < u_{x,i_0+1} < u_{x,i_0+2} < \cdots$$

Then, for the grid point x_i , the corresponding row \mathbf{a}_i of A for the curvature term in (2.8) reads

$$\mathbf{a}_i = (0, \cdots, 0, a_{i,i-1}, a_{i,i}, a_{i,i+1}, 0, \cdots, 0), \quad (3.2)$$

where

$$\begin{aligned} a_{i,i-1} &= (u_{x,i-1/2}^2 + \varepsilon^2)^{-1/2}, \\ a_{i,i+1} &= (u_{x,i+1/2}^2 + \varepsilon^2)^{-1/2}, \\ a_{i,i} &= -(a_{i,i-1} + a_{i,i+1}), \end{aligned}$$

where $u_{x,i+1/2}$ is an average of $u_{x,i}$ and $u_{x,i+1}$. When i is near i_0 and ε is small, then $|a_{i,i}|$ must be large. Let us set $a_{\max} = 2$, for example. (Later, we will see it is the most reasonable choice.) Then the scaled \mathbf{a}_i in (3.2) becomes

$$\mathbf{a}'_i := \frac{2}{|a_{i,i}|} \mathbf{a}_i = (0, \cdots, 0, a'_{i,i-1}, -2, a'_{i,i+1}, 0, \cdots, 0), \quad (3.3)$$

where

$$\begin{aligned} a'_{i,i-1} &= \frac{2}{\gamma}(u_{x,i+1/2}^2 + \varepsilon^2)^{1/2}, \\ a'_{i,i+1} &= \frac{2}{\gamma}(u_{x,i-1/2}^2 + \varepsilon^2)^{1/2}, \end{aligned}$$

with γ given as

$$\gamma = (u_{x,i-1/2}^2 + \varepsilon^2)^{1/2} + (u_{x,i+1/2}^2 + \varepsilon^2)^{1/2}.$$

Since $a'_{i,i-1} > 1$ for $i > i_0$ and $a'_{i,i+1} > 1$ for $i < i_0$ (independently on $\varepsilon \geq 0$), the scaled row \mathbf{a}'_i allows a *strictly* convex stationary solution of the image near x_{i_0} . When $i = i_0$, the scaling turns out to be the same as the $(1, -2, 1)$ -difference scheme only if $u_{x,i_0-1} = u_{x,i_0+1}$. For concave parts of the image, one can do a similar analysis. The algebraic scaling keeps a certain degree of convexity and concavity for convex and concave parts, respectively. This feature enforces for the algorithm to be less dissipative than the $(1, -2, 1)$ -difference scheme obtained from the discretization of u_{xx} ($p = 2$).

From an analysis, one can see that the above scaling procedure can be written as an approximation of the following model

$$u_t = \min \left(\frac{a_{\max}}{2}(u_x^2 + \varepsilon^2)^{1/2}, 1 \right) \left(\frac{u_x}{(u_x^2 + \varepsilon^2)^{1/2}} \right)_x + \beta(u_0 - u). \quad (3.4)$$

Thus the parameter a_{\max} must be chosen to be neither too small nor too large; a reasonable choice is $a_{\max} = 2$. In this case, (3.4) can be viewed as a thresholding variant of Marquina and Osher [16]:

$$u_t = |u_x| \left(\frac{u_x}{(u_x^2 + \varepsilon^2)^{1/2}} \right)_x + \beta |u_x| (u_0 - u). \quad (3.5)$$

For high-dimensional problems, one can apply the above algebraic scaling to each of the directions.

4. Anti-Diffusion

4.1. Non-convex minimization

To understand effects of p and the regularization parameter ε in (2.7) and therefore in (2.8), we will first measure the functional $\mathcal{F}_{\varepsilon,p}$, with $\lambda = 0$, for the cartoon images as in Figure 1. Table 1 presents the values of $\mathcal{F}_{\varepsilon,p}$. Consider the first case $\mathcal{F}_{0,2}$; the value is smaller for the blurry image in Figure 1(b), which implies that the minimization with $(\varepsilon, p) = (0, 2)$ invokes blur. Applying the same logic, one can see that $\mathcal{F}_{0,1}$ does not introduce blur during the minimization. But, when we set $(\varepsilon, p) = (0.1, 1)$ as to prevent the denominator become zero in the Euler-Lagrange equation, the minimization algorithm will make the image blurry, because the blurry

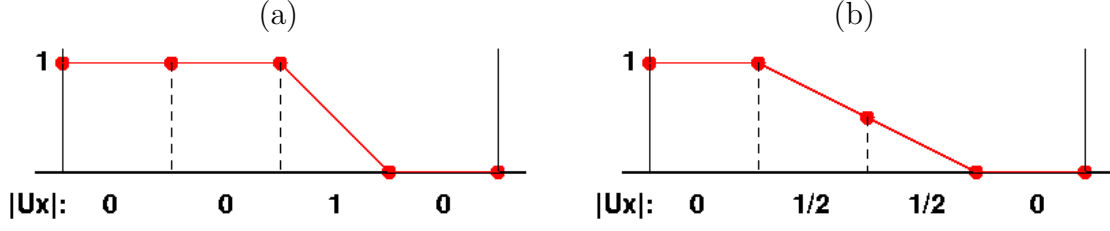


Figure 1: A cartoon image in one-dimensional space (a) and its blurry version (b). Gradient magnitudes are indicated in the bottom of each side.

Table 1: The values of the functional $\mathcal{F}_{\varepsilon,p}$, with $\lambda = 0$, for images in Figure 1.

Figure 1(a)	Figure 1(b)
$\mathcal{F}_{0,2} = 0 + 0 + 1 + 0 = 1$	$\mathcal{F}_{0,2} = 0 + 0.25 + 0.25 + 0 = 0.5$
$\mathcal{F}_{0,1} = 0 + 0 + 1 + 0 = 1$	$\mathcal{F}_{0,1} = 0 + 0.5 + 0.5 + 0 = 1$
$\mathcal{F}_{0.1,1} = 0.1 + 0.1 + \sqrt{1.01} + 0.1 \approx 1.31$	$\mathcal{F}_{0.1,1} = 0.1 + \sqrt{0.26} + \sqrt{0.26} + 0.1 \approx 1.22$
$\mathcal{F}_{0,0.9} = 0 + 0 + 1 + 0 = 1$	$\mathcal{F}_{0,0.9} = 0 + 0.5^{0.9} + 0.5^{0.9} + 0 \approx 1.07$

image reduces the value of the functional. Here an interesting case is when $p = 1 - \omega$, $\omega > 0$. As one can see from the table that when $\omega = 0.1$ ($p = 0.9$), the functional has a smaller value for the sharper image. Thus when $p < 1$, the functional $\mathcal{F}_{0,p}$ can be utilized for edge enhancement, at least for the cartoon image in Figure 1.

We summarize what we have seen with $\mathcal{F}_{\varepsilon,p}$, as follows:

- Heat equation ($p = 2$) makes images blurry.
- TVM itself ($p = 1$ and $\varepsilon = 0$) does not introduce “blur”. But its regularization ($p = 1$ and $\varepsilon > 0$) does.
- When $p = 1 - \omega < 1$, the functional can make the image sharper. Even though there is no mathematical guarantee for existence of the minimizer (because the functional is not convex), the case motivates an *effective* denoising algorithm that can preserve edges satisfactorily.

Yet, one has to choose ε positive to avoid the denominator in the corresponding Euler-Lagrange equation approach to zero, which has a trend to make the image blurry. Consider the following minimization problem

$$\min_u \mathcal{F}_{\varepsilon,1-\omega}(u), \quad (4.1)$$

and its Euler-Lagrange equation

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla \varepsilon u|^{1+\omega}} \right) + \beta(u_0 - u). \quad (4.2)$$

An effective edge-preserving noise removal is expected when the parameters ε and ω are chosen appropriately. As an example, when $(\varepsilon, \omega) = (0.1, 0.2)$ and $\lambda = 0$, one can easily see that

$$\mathcal{F}_{0.1,0.8} \approx 1.479, \quad \mathcal{F}_{0.1,0.8} \approx 1.483,$$

respectively for Figure 1(a) and Figure 1(b). Thus this choice will not blur the sharper image.

A natural question is: *How can we choose $\varepsilon > 0$ and $\omega > 0$ that minimize “blur” for the minimization problem (4.1) and equivalently for (4.2)?* Unfortunately, we do not know or know of any analytical answer. Here is a guide line we have found from various experiments:

- Set $\varepsilon = 0.01 \sim 0.1$ and $\omega = 0.4 \sim 1$.
- The choice of ε is not much sensitive due to effects of the algebraic scaling, but it is recommended not to set too small. One may select $\varepsilon = 0.05$.
- A good news is that ω seems *little* depending on ε ; the choice $\omega = 0.5 \sim 0.7$ works satisfactorily for most cases.

See the numerical results in §5.2 below.

4.2. Variable constraint parameter $\beta(\mathbf{x}, t)$

The basic mechanism of the TVM denoising is diffusion. Thus the parameter β cannot be too large; it must be small enough to introduce a sufficient amount of diffusion. On the other hand, it should be large enough to keep the details in the image. However, in the literature, it is very often the case that the parameter is chosen constant: $\beta = \beta_0$. With such a naive strategy, the smoothing operator either smears out some important portions more excessively than desired, in particular near edges, or leaves a certain amount of noise in the resulting image. Then, how can we choose the parameter β ? As *an* answer for the question, we suggest the following:

The diffusion iteration begins with a constant parameter $\beta = \beta_0$ and, as the iteration goes on, let it grow wherever diffusion is observed excessively.

We call such a dynamic parameter $\beta (= \beta(\mathbf{x}, t))$ the *operator-driven optimal parameter* (ODOP).

To describe ODOP more systematically, we consider our PDE model formulated as

$$u_t - \nabla \cdot \left(\frac{\nabla u}{|\nabla^\varepsilon u|^{1+\omega}} \right) = \beta(u_0 - u), \quad \beta = \beta(\mathbf{x}, t), \quad (4.3)$$

and its incomplete *Crank-Nicolson alternating direction implicit* (CN-ADI) time-stepping procedure

$$\begin{aligned} \left(1 + \frac{\Delta t}{2} A_1^{n-1}\right) u^* &= \left(1 - \frac{\Delta t}{2} A_1^{n-1} - \Delta t A_2^{n-1}\right) u^{n-1} + \Delta t \beta u_0, \\ \left(1 + \frac{\Delta t}{2} A_2^{n-1}\right) u^n &= u^* + \frac{\Delta t}{2} A_2^{n-1} u^{n-1}, \end{aligned} \quad (4.4)$$

where Δt is the timestep size and

$$A_\ell^{n-1} u^n := -D_{x_\ell} \left(\frac{D_{x_\ell} u^n}{|\nabla_h^\varepsilon u^{n-1}|^{1+\omega}} \right) + \frac{\beta}{2} u^n, \quad \ell = 1, 2.$$

Here $(D_{x_1}, D_{x_2})^T$ is the half-step central difference operator for the gradient ∇ and $|\nabla_h^\varepsilon u^{n-1}|$ denotes the (standard) second-order central scheme approximating $|\nabla^\varepsilon u^{n-1}|$. See the companion paper [11] for details.

Now, we are ready to describe ODOP in terms of the diffusion iteration:

- Begin the CN-ADI iteration (4.4) with $\beta = \beta_0$, a constant.
- For each iteration (or each several iterations) $n \geq 1$,

$$\beta \leftarrow \beta + \gamma S_m(|\nabla_S(u^n - u_0)|), \quad (4.5)$$

where ∇_S is a smoothed approximation of the gradient such as the Sobel operator, S_m denotes a smoother like a few Jacobi iterations, and γ is a scaling factor.

Note that the above ODOP has been designed based on the following intuition. Let us assume that the observed image u_0 involves only Gaussian white noise. Then, far from the edges, the quantity $|\nabla_S(u^n - u_0)|$ must be small, because it is a locally averaged gradient for the difference (random values) between the noisy image and its smooth version. On the other hand, near edges, $|\nabla_S(u^n - u_0)|$ would be relatively large, because u^n can easily involve an excessive blur (structured difference) near edges. Thus, the parameter β grows, as the iteration goes on, wherever diffusion is excessive, in particular, near edges. The outcome of such a dynamic parameter is to minimize the blur near edges.

In practice, one may wish to limit the parameter in a prescribed interval, i.e., $\beta(\mathbf{x}, t) \in [\beta_0, \beta_1]$. Also one may want to update the parameter a few times only during the simulation. Note that the algorithm (4.4) converges fast (less than 10 iterations for most cases). Thus we can update β three times e.g. $n = 2, 4$, and 6 . In the case, the scaling factor γ can be selected such that

$$\gamma \max_{i,j} \{S_m(|\nabla_S(u^n - u_0)|)\} = \lambda(\beta_1 - \beta_0), \quad \lambda = \begin{cases} 0.5, & n = 2, \\ 0.3, & n = 4, \\ 0.2, & n = 6. \end{cases} \quad (4.6)$$

For convenience, let us introduce a new terminology: we will call the TV minimization (4.3)-(4.4) the *enhanced TV minimization* (ETVM) when it incorporates the algebraic scaling (3.1), the non-convex minimization ($\omega > 0$), and the variable constraint parameter (4.5).

Remark. Impulse noise can hardly be removed unless the constraint parameter β is sufficiently small. Furthermore, smoothing such a large perturbation is not a reliable strategy in denoising. Thus, one has to remove impulse noise (if any) from u_0 prior to the application of the TVM-like smoothing. For the removal of impulse noise, one can adopt a numerical procedure for the motion by mean curvature, which can be solved in a similar fashion as for TVM (4.3), applying (4.4). See [11] for details.

5. Numerical Experiments

For numerical experiments, we choose gray-scale images in public domain. In the application of (4.4), we first scale the cell values by multiplying $\sqrt{3}/255$ such that the possible maximum value of u in the program is $\sqrt{3}$, i.e.,

$$u_{ij} \leq u_{\max} := \sqrt{3}. \quad (5.1)$$

(Of course, we scale back the results for picturing.) For the noise, we consider random-valued additive Gaussian noise, which is scaled to be mean-zero and have a certain variance. The iterative algorithm (4.4) is stopped when

$$\max_{i,j} |u_{i,j}^n - u_{i,j}^{n-1}| \leq 0.03. \quad (5.2)$$

5.1. Anti-staircasing

In this subsection, we show numerical verification for the algebraic scaling discussed in §3 as an anti-staircasing strategy. We will test it, incorporated with (4.3)-(4.4), for the removal of Gaussian noise. We set $\omega = 0$ ($p = 1$) and $\beta = \beta_0 = 0.5$, to focus on the effects of the algebraic scaling; choose $\varepsilon = 0.05$. To understand characteristics of the algebraic scaling, we will compare its results with those of the variable- p approach (2.9). Note that the algebraic scaling works with the choice $p = 1$.

In Figure 2, we present the numerical results carried out utilizing the gray-scale Lenna image in 256×256 cells. The original image, Figure 2(a), is perturbed by Gaussian noise of variance 488, as shown in Figure 2(b). The noisy image is denoised by the variable- p approach and the algebraic scaling and the results are depicted in Figure 2(c) and Figure 2(d), respectively. For the variable- p approach (2.9), the constants g_0 and g_1 are selected as

$$g_0 = \frac{1}{16}u_{\max}, \quad g_1 = \frac{1}{4}u_{\max},$$

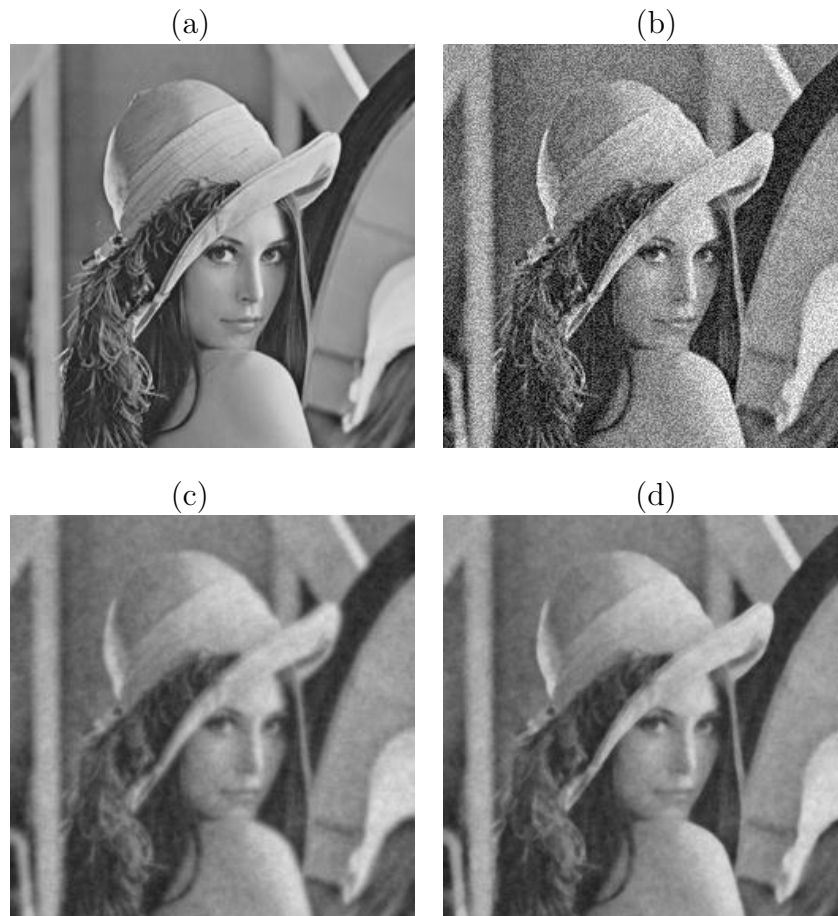


Figure 2: Gray-scale Lenna images in 256×256 cells: (a) the original image, (b) the noisy image perturbed by Gaussian noise of variance 488, (c) the restored image by the variable- p (2.9), and (d) the restored image by the algebraic scaling. Set $\omega = 0$ and $\beta = 0.5$.

where u_{\max} is defined in (5.1). Given the stopping criterion (5.2), the CN-ADI algorithm (4.4) converges in four and three iterations respectively for the variable- p approach and the algebraic scaling. It is clear to see that none of them have invoked staircasing. However, as one can see from the figure, the variable- p approach introduces a lot of dissipation. On the other hand, the algebraic scaling can remove Gaussian noise with much less blur, as we have expected from the analysis in §3.

Such a superior performance of the algebraic scaling (in both anti-staircasing and anti-diffusion) has been observed for all examples we have tested. Here the upshot is that the algebraic scaling has an ability to keep a certain degree of convexity and concavity appeared in the image content. To see robustness of the newly suggested method, we apply it to the House image, with the same algorithm parameters chosen for Figure 2; the results are depicted in Figure 3. The figure gives another exam-

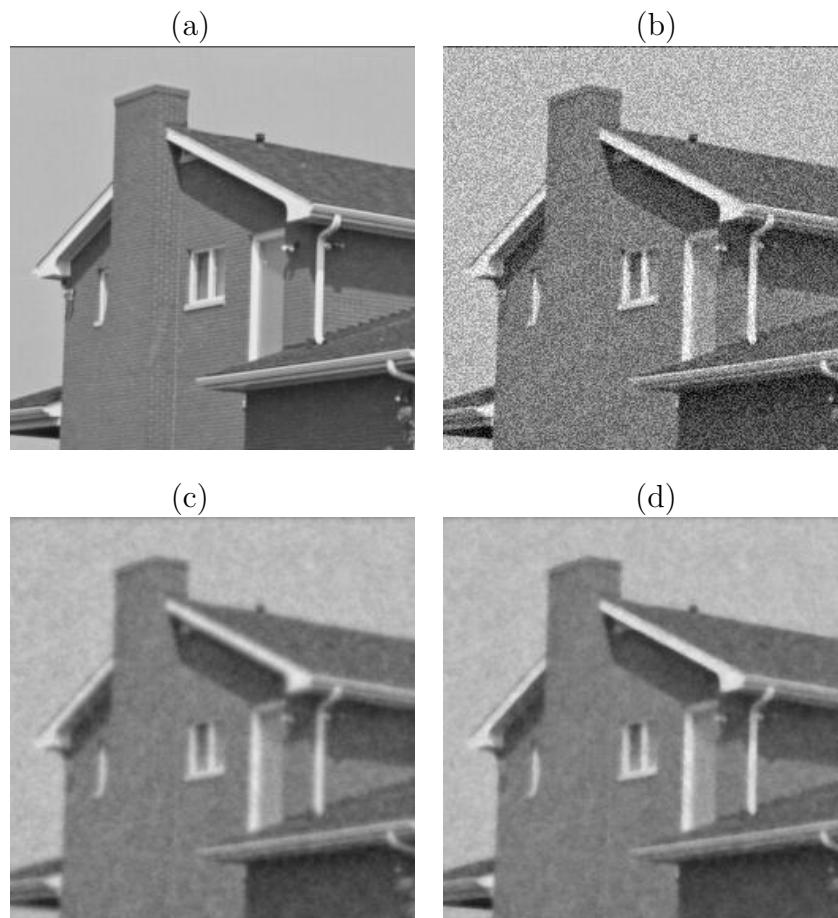


Figure 3: House images in 256×256 cells: (a) the original image, (b) the noisy image perturbed by Gaussian noise of variance 488, (c) the restored image by the variable p (2.9), and (d) the restored image by the algebraic scaling. Both take 4 iterations each for the convergence. Set $\omega = 0$ and $\beta = 0.5$.

ple (among many others) with which one can convince the superior behavior of the algebraic scaling.

5.2. Anti-diffusion

In this subsection, we verify the anti-diffusion schemes introduced in §4 to minimize nonnecessary diffusion of the denoising algorithm. For anti-staircasing, we utilize the algebraic scaling for the rest of the article. We begin with an example in which ω varies, while the constraint parameter β is kept constant; the experiment is designed to get insights for an appropriate choice of ω .

Figure 4 contains restored images from a noisy Lenna image contaminated by a Gaussian noise of variance 488 (as in Figure 2(b)), with the choices: (top) $\omega = 0.5$, (mid) $\omega = 1$, and (bottom) $\omega = 2$. For this example, we set $\beta = 0.5$, a constant.



Figure 4: Gray-scale Lenna images in 256×256 cells: The images are restored from Figure 2(b) with (top) $\omega = 0.5$, (mid) $\omega = 1$, and (bottom) $\omega = 2$. Set $\beta = 0.5$ for all cases.

The algorithm (4.4) takes 4, 9, and 14 iterations respectively for $\omega = 0.5$, 1, and 2. One can see from the figure that when $\omega > 0$, the restored image involves less dissipation than the one for the case $\omega = 0$ (see Figure 2(d)). However, when we set $\omega = 2$, an overshooting is observed; the restored image has been transformed to show locally linear portions. This result proves another kind of staircasing which is observable when e.g. the Perona-Malik model [20] is applied for image enhancement. For the choices $\omega = 0.5$ and $\omega = 1$, the restored images keep clear details with *little* introduction of diffusion and no staircasing. We have found from various experiments that the restored image shows the best quality when

$$\omega = 0.5 \sim 1.2,$$

for various constant constraint parameters β .

In Figure 5, we present numerical results for variable β suggested in (4.5)-(4.6), with the choice of

$$\beta_0 = 0.3, \quad \beta_1 = 4.0.$$

We utilize the same noisy image in Figure 2(b), which involves Gaussian noise of variance 488. For a comparison purpose, we first collect images from previous experiments:

- Figure 5(a): Figure 2(d) ($p = 1$, $\omega = 0$, and $\beta = 0.5$)
- Figure 5(b): Figure 4 (mid) ($p = 1$, $\omega = 1$, and $\beta = 0.5$)

Figure 5(c) is the restored image with $p = 2$ ($\omega = -1$) and the variable β . Note that the diffusion operator for this experiment is the heat equation. However, it is clear to see that the result is either better than or comparable with the ones recovered by the standard non-staircasing TVM ($\omega = 0$) and the edge-enhanced TVM ($\omega = 1$) that incorporate constant constraint parameters β . Corresponding to Figure 5(c), the parameter β in the final is depicted in Figure 5(d), where larger values indicate the regions that are actively smeared out by the heat diffusion when β has not grown enough.

The above example shows the main idea of ODOP: set β bigger wherever an excessive dissipation occurs. Also one should notice that the ODOP can be applied for relatively simpler diffusion operators such as the heat flow and yet it can possibly result in a better recovery than TVM not incorporating ODOP. In the next example, we verify the performance of the ETVM which combines non-convex minimization ($\omega > 0$) and ODOP, along with the algebraic scaling (3.1), for an improved edge-preserving noise removal.

In Figure 6, we present numerical results for the ETVM. For experiments, we select the gray-scale Elaine image in 256×256 cells: the original image in Figure 6(a)

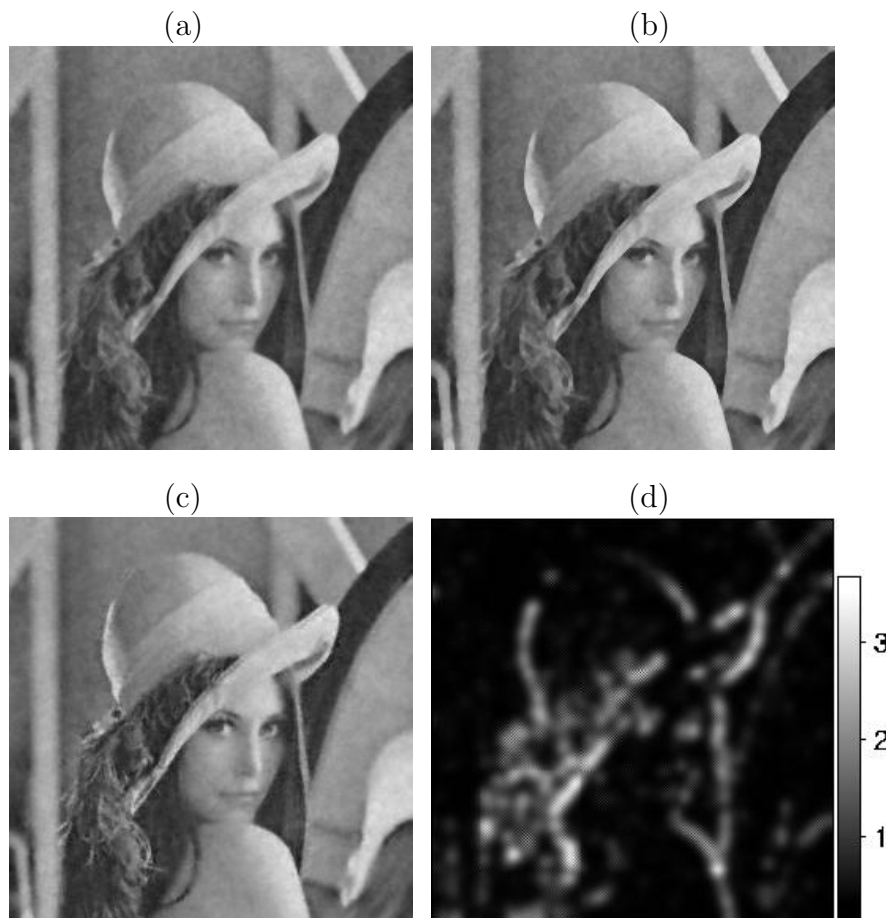


Figure 5: Gray-scale Elaine images in 256×256 cells: (a) the image in Figure 2(d) ($p = 1$, $\omega = 0$, and $\beta = 0.5$), (b) the image in Figure 4 (mid) ($p = 1$, $\omega = 1$, and $\beta = 0.5$), (c) the restored image from the noisy Lenna (Gaussian noise of variance 488) in Figure 2(b), with $p = 2$ ($\omega = -1$) and $\beta = \beta(\mathbf{x}, t)$ as in (4.5)-(4.6), (d) β in the final. Set $\beta_0 = 0.3$ and $\beta_1 = 4.0$.

is perturbed by Gaussian noise of variance 338, as depicted in Figure 6(b). Set $\omega = 0.5$ and $\beta_0 = 0.3$ and $\beta_1 = 3.0$. The CN-ADI iteration (4.4) converges in six iterations resulting in the recovered image in Figure 6(c). The residual (at $n = 6$)

$$\max_{i,j} |u_{\text{orig}} - u_{i,j}^n|,$$

the absolute difference between the original image and the restored image, is depicted in Figure 6(d). As one can see from the figure, details are clearly preserved in the restored image and the residual does not show any significant structural contents, as we have wished and expected. For various examples, we have seen similarly satisfactory performances of the ETVM.

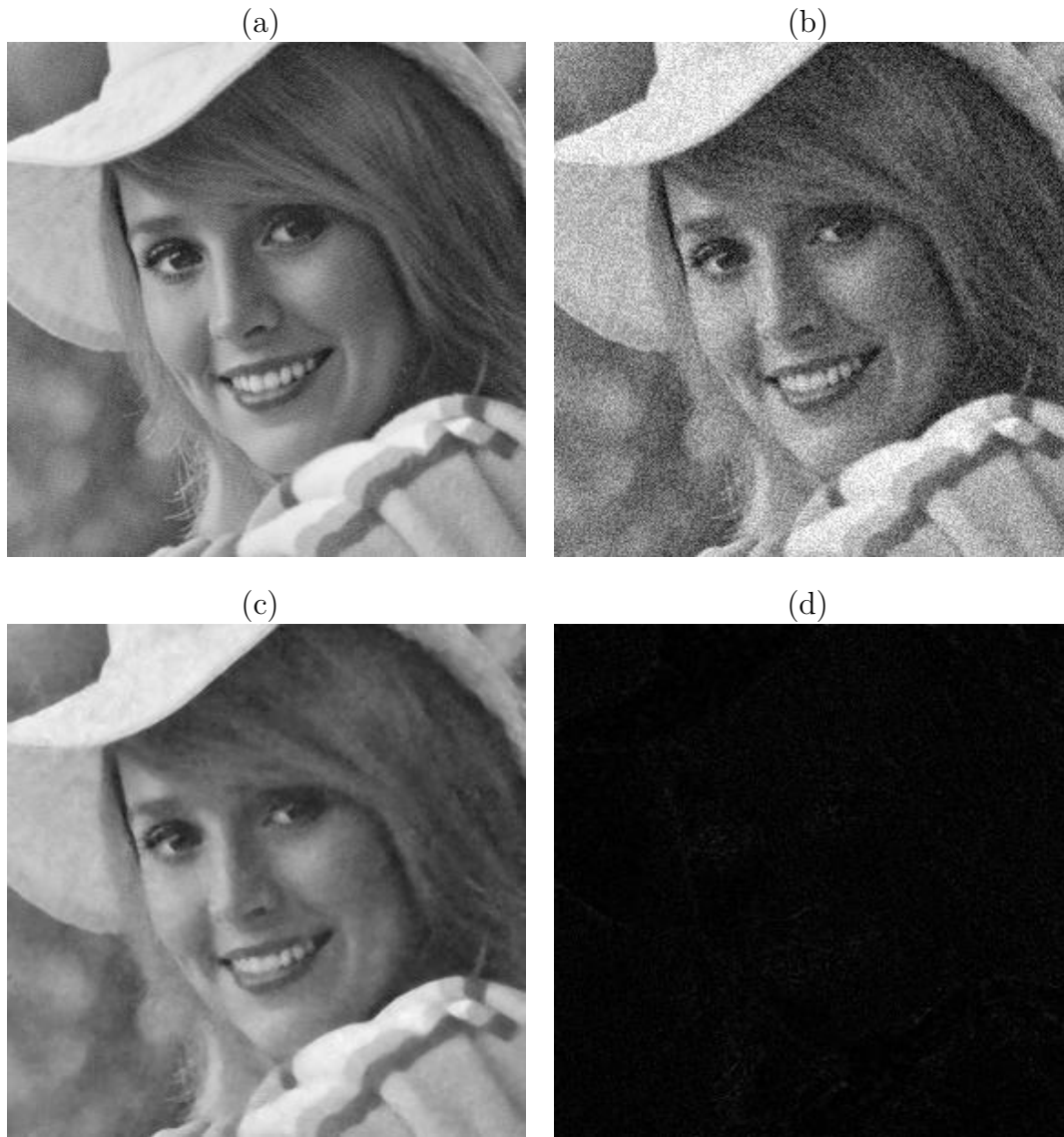


Figure 6: Gray-scale Elaine images in 256×256 cells: (a) the original image, (b) the noisy image perturbed by Gaussian noise of variance 338, (c) the restored image by the ETVM ($\omega = 0.5$, $\beta_0 = 0.3$, and $\beta_1 = 3.0$), and (d) the residual, the absolute difference between the original image and the restored image.

6. Conclusions

The TV minimization (TVM) often involves staircasing and/or an excessive dissipation during the noise removal. For anti-staircasing, we have introduced the algebraic scaling; it turns out not only to prevent staircasing effectively but also to preserve a certain degree of convexity and concavity of the image contents. The algebraic scaling has been numerically verified to outperform to the variable- p approach. To prevent a nonnecessary diffusion of TVM, we have considered a non-convex minimization and the operator-driven optimal parameter (ODOP). Although the non-convex minimization does not guarantee the existence of the minimizer, it has performed for anti-diffusion with a great satisfaction. The basic idea of ODOP is to let the constraint parameter β grow wherever an excessive dissipation occurs. These three newly-introduced methods (the algebraic scaling, a non-convex minimization, and ODOP) have been combined to enhance the TVM; we call the resulting algorithm the *enhanced TV minimization* (ETVM). The ETVM has been tested along with various images to prove its effectiveness and robustness in the removal of Gaussian noise. It has performed *very satisfactorily* for the edge-preserving noise removal.

Acknowledgment

The authors would like to express sincere thanks to Prof. Sung Ha Kang, University of Kentucky, for her constructive comments.

References

- [1] L. ALVAREZ, P. LIONS, AND M. MOREL, *Image selective smoothing and edge detection by nonlinear diffusion. II*, SIAM J. Numer. Anal., 29 (1992), pp. 845–866.
- [2] G. AUBERT AND P. KORNPORST, *Mathematical Problems in Image Processing*, no. 147 in Applied Mathematics Sciences, Springer-Verlag, New York, 2002.
- [3] P. V. BLOMGREN, *Total Variation Methods for Restoration of Vector Valued Images*, (*Ph.D. thesis*), UCLA Dept. of Math. CAM 98-30, (1998).
- [4] P. V. BLOMGREN AND T. F. CHAN, *Color TV: Total Variation Methods for Restoration of Vector Valued Images*, IEEE Trans. Image Processing, 7 (1998), pp. 304–309.
- [5] F. CATTE, P. LIONS, M. MOREL, AND T. COLL, *Image selective smoothing and edge detection by nonlinear diffusion.*, SIAM J. Numer. Anal., 29 (1992), pp. 182–193.

- [6] A. CHAMBOLLE AND P. L. LIONS, *Image recovery via Total Variational minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
- [7] T. CHAN, S. OSHER, AND J. SHEN, *The digital TV filter and nonlinear denoising*, Technical Report #99-34, Department of Mathematics, University of California, Los Angeles, CA 90095-1555, October 1999.
- [8] T. F. CHAN, S. H. KANG, AND J. SHEN, *Euler's Elastica and Curvature Based Inpaintings*, SIAM Journal on Applied Mathematics, 63 (2002), pp. 564–592.
- [9] T. F. CHAN AND J. SHEN, *Variational restoration of non-flat image features: Models and algorithms*, SIAM Journal of Applied Mathematics, 61 (2000), pp. 1338–1361.
- [10] —, *Mathematical Models for Local Deterministic Inpaintings*, SIAM Journal of Applied Mathematics, 62 (2001), pp. 1019–1043.
- [11] K. JOO AND S. KIM, *PDE-based image restoration, II: Numerical schemes and color image denoising*. (preprint).
- [12] S. KIM, *Edge-preserving noise removal, Part I: Second-order anisotropic diffusion*. (Submitted to SIAM J. Sci. Comput.).
- [13] S. KIM, I. IACOB, AND M. TYNAN, *MinBAD: The minimum-biased anisotropic diffusion for noise removal*, Technical Report #2002-06, Department of Mathematics, University of Kentucky, Lexington, KY 40506, 2002.
- [14] S. KIM AND S. H. KANG, *Implicit procedures for PDE-based color image denoising via brightness-chromaticity decomposition*, Technical Report #2002-07, Department of Mathematics, University of Kentucky, Lexington, KY 40506, 2002.
- [15] R. KIMMEL AND N. SOCHEN, *Orientation Diffusion or How to Comb a Porcupine ?*, Special issue on PDEs in Image Processing, Computer Vision, and Computer Graphics, J. Visual Comm. Image Representation, 13 (2002), pp. 238–248.
- [16] A. MARQUINA AND S. OSHER, *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal*, SIAM J. Sci. Comput., 22 (2000), pp. 387–405.
- [17] S. MASNOU AND J. M. MOREL, *Level Lines based Disocclusion*, Proceedings of 5th IEEE Int'l Conf. on Image Process., Chicago, 3 (1998), pp. 259–263.
- [18] M. NITZBERG AND T. SHIOTA, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. on Pattern Anal. Mach. Intell., 14 (1992), pp. 826–833.

- [19] S. OSHER AND R. FEDKIW, *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag, New York, 2003.
- [20] P. PERONA AND J. MALIK, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. on Pattern Anal. Mach. Intell., 12 (1990), pp. 629–639.
- [21] L. RUDIN AND S. OSHER, *Total variation based image restoration with free local constraints*, Proc. 1st IEEE ICIP, 1 (1994), pp. 31–35.
- [22] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
- [23] G. SAPIRO AND D. RINGACH, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE Trans. Image Processing, 5 (1996), pp. 1582–1586.
- [24] N. SOCHEN, R. KIMMEL, AND R. MALLADI, *A geometrical framework for low level vision*, IEEE Trans. Image Processing, 7 (1998), pp. 310–318.
- [25] B. TANG, G. SAPIRO, AND V. CASELLES, *Color image enhancement via Chromaticity Diffusion*, Technical report, ECE University of Minnesota, (1999).
- [26] R. WEINSTOCK, *Calculus of Variations*, Dover Publications, Inc., New York, 1974.
- [27] Y. YOU, W. XU, A. TANNENBAUM, AND M. KAVEH, *Behavioral analysis of anisotropic diffusion in image processing*, IEEE Trans. Image Process., 5 (1996), pp. 1539–1553.