Applied Mathematics & Information Sciences
An International Journal

A fast primal-dual method for generalized Total Variation denoising

Francesc Aràndiga, Pep Mulet and Vicent Renau

Departament de Matemàtica Aplicada, Universitat de València, Av. Vicent Andrés Estellés s/n, 46100, Burjassot, Spain.

Received: Jul 8, 2011; Revised Oct. 4, 2011; Accepted Oct. 6, 2011
Published online: 1 Sep. 2012

Abstract: Total Variation denoising, proposed by Rudin, Osher and Fatemi in [22], is an image processing variational technique that has attracted considerable attention in the past fifteen years. It is an advantageous technique for preserving image edges but tends to sharpen excessively smooth transitions. With the purpose of alleviating this *staircase effect* some generalizations of Total Variation denoising have been introduced in [17,18,19]. In this paper we propose a fast and robust algorithm for the solution of the variational problems that generalize Total Variation image denoising [22]. This method extends the primal-dual Newton method, proposed by Chan, Golub and Mulet in [7] for total variation restoration, to these variational problems. We perform some experiments for assessing the efficiency of this scheme with respect to the fixed point method that generalizes the lagged diffusivity fixed point method proposed by Vogel and Oman in [24].

Keywords: image denoising, total variation, staircase effect, primal-dual problem, Newton's method, Huber function, robust estimation.

1. Introduction

The presence of blur and/or noise is inherent to many image acquisition devices. This noise and blurring should be removed in the earlier stages of the image processing, for their presence can be a burden for the later phases of the image processing.

Many algorithms for noise removal and deblurring are based on linear filters (convolutions) and are implemented with Fast Fourier Transforms in the frequency domain. These algorithms do not perform well near discontinuities (edges), giving excessively blurred outputs and Gibbs phenomena. To overcome this difficulty a technique based on the minimization of the Total Variation seminorm subject to some noise constraints is proposed by Rudin, Osher and Fatemi in [22]. The solution of this variational problem is a function of bounded variation admitting discontinuities, in contrast to other variational models that assume some regularity of the solution.

In spite the success of the Total Variation restoration technique, it has some aspects that could be improved, such as the feature that the image output is close to a piecewise constant function (*staircase effect*), thus preventing some natural smooth transitions between image levels.

In [17,18,19] some variational formulations with functionals that generalize the Total Variation functional have been proposed with the goal of taking the best properties of the space of functions of bounded variation and the Sobolev space H^1, which is preserving the strong gradients of the image while, at the same time, smoothing homogeneous regions without discontinuities. Another approach, based on using higher order derivatives in the objective functional is explored in [6].

The Euler-Lagrange equations associated to these generalized Total Variation minimization problems have the same difficulties as the original equations in [22], since they are highly nonlinear elliptic partial differential equations. The simplest scheme for their solution is a time marching scheme, similar to the one proposed in [22]. Since the convergence of this scheme can be slow due to stability constraints in the time step size, an alternative is a fixed point method, analogous to the one proposed in [24].

The convergence of this fixed point scheme is global, much faster than that of the time marching scheme, but it is linear, a fact that can limit its efficiency if full convergence

* Corresponding author: mulet@uv.es

is pursued. A primal-dual Newton method was proposed in [7] for the fast and robust solution of the Euler-Lagrange equations of the minimization of the Total Variation functional. In this paper we propose a similar technique for the generalized Total Variation problems and prove by experiment that, in terms of efficiency, it is highly competitive when compared to the fixed point method.

The organization of this paper is as follows: in section 2 we introduce the problem and the Euler-Lagrange equations associated to it. In section 3 we present some methods for the solution of these nonlinear equations. In section 4 we propose the primal-dual linearization technique for these equations. In section 5 we present some numerical results. Finally, in section 6 we give some concluding remarks.

2. Generalized Total Variation denoising problems

Assume that a gray level image u, regarded as a real function on the unit square Ω, has been degraded by white Gaussian noise n, so that $z = u + N$. The objective of image denoising is the estimation of u from the knowledge of statistical parameters of N, such as its variance σ^2, i.e., $\|N\|^2_{\mathcal{L}^2} = \sigma^2$. From a variational point of view, this problem can be formulated as the minimization of some functional $G(u)$ (that may depend on derivatives of u) subject to the quadratic noise constraint $\|u - z\|^2_{\mathcal{L}^2} = \sigma^2$, where the objective functional F should be chosen according to the *a priori* knowledge of u. This approach is related to Tikhonov regularization [23]. If $G(u) = \int_\Omega |\nabla u|^2 dx$, then the Euler-Lagrange equation is a linear reaction-diffusion equation that can be easily solved by Fourier transforms. On the other hand, the solution is continuous, so the edges of the images would appear blurred.

To overcome this drawback, the solution of the variational problem

$$\begin{aligned} &\min_u \int_\Omega |\nabla u(x)| dx, \\ &\text{subject to} \frac{1}{2}\int_\Omega (u(x) - z(x))^2 dx = \frac{\sigma^2}{2}, \end{aligned} \tag{1}$$

is proposed in [22] as the denoising output of the image $z \in \mathcal{L}^2(\Omega)$. The Total Variation functional $\int_\Omega |\nabla u(x)| dx$ measures the jumps in the intensity of the image, but do not penalize excessively their existence, so that discontinuous images can be obtained from this formulation. Of course, for these images one needs to use the general definition of the total variation as

$$\begin{aligned} \sup\{ \int_\Omega u \,\mathrm{div}(w) dx \colon w \in (\mathcal{C}^\infty)^2, \\ w|_{\partial\Omega} = 0, |w|_\infty \le 1\}. \end{aligned}$$

or approximate it by $\int_\Omega |\nabla u(x)|_\beta dx$, for u in the Sobolev space $W^{1,1}(\Omega)$, where $|w|_\beta = \sqrt{w \cdot w + \beta}$, $\beta > 0$ is introduced to overcome the non-differentiability of the square root at 0.

The image restoration model based on the TV tends to yield piecewise constant images. Whereas this is certainly useful for many applications, it can be a serious drawback for many others. In particular, when applying this denoising technique to an affine image degraded with noise, the result will invariably be a *staircase*, thus creating oversharpening at smooth transitions.

To prevent this staircase effect, some functionals that penalize small jumps more have been proposed in the literature [17,18,19,6,16,15]. In this paper we consider functionals that generalize total variation, so that the denoising problem reads:

$$\begin{aligned} &\min_u \int_\Omega \phi(|\nabla u(x)|_\beta) dx, \\ &\text{subject to } \frac{1}{2}\int_\Omega (u(x) - z(x))^2 dx = \frac{\sigma^2}{2}, \end{aligned} \tag{2}$$

for non-negative, non-decreasing and convex functions $\phi \in \mathcal{C}^1$ and u belonging to a suitable functional space. The conditions on ϕ to ensure existence and/or uniqueness of solutions of (2) in this continuous setting may be rather involved due to functional analytical details that fall out of the scope of this paper. We refer to [1,2,10] and references therein for some of those details.

We denote by $F(u)$ the Lagrangian of (2) for a fixed $\lambda > 0$

$$\begin{aligned} F(u) = \int_\Omega \big(&\phi(|\nabla u(x)|_\beta) \\ &+ \frac{\lambda}{2}(u(x) - z(x))^2\big) dx - \frac{\lambda}{2}\sigma^2. \end{aligned} \tag{3}$$

The Euler-Lagrange equations for (2) can be obtained by standard Fréchet differentiation of (3):

$$\begin{aligned} F'(u) = -\mathrm{div}\left(\frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}\nabla u\right) \\ + \lambda(u - z) = 0, \quad \frac{\partial u}{\partial \overrightarrow{n}} = 0, \end{aligned} \tag{4}$$

with the Lagrange multiplier λ suitably chosen to enforce the quadratic constraint and a Neumann boundary condition ($\overrightarrow{n}$ denotes the unit normal vector to $\partial\Omega$ pointing in the outward direction). This nonlinear partial differential equation is of elliptic type, since $\phi'(x) > 0$ for $x > 0$.

Assuming $\phi \in \mathcal{C}^2$, one can show, again by using Fréchet differentiation, that the second variation of $F(u)$ in (3) is

$$\begin{aligned} F''(u)(v) = -\mathrm{div}\Big[&\frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}\nabla v \\ &+ G(\nabla u)\frac{\nabla u \cdot \nabla v}{|\nabla u|^2_\beta}\nabla u\Big] + \lambda v, \\ G(\nabla u) = \phi''(|\nabla u|_\beta) &- \frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}. \end{aligned} \tag{5}$$

2.1. Discretized problem

We discretize (2) on the grid points $(ih, jh), i, j = 0, \ldots, n-1$, $h = 1/n$, by using the rectangle rule for integrals and forward differences approximations ∇_x, ∇_y of ∂x and ∂y, respectively, i.e.,

$$\begin{aligned}\nabla_x(u)_{i,j} &= \frac{u_{i+1,j} - u_{i,j}}{h}, i < n-1,\\ \nabla_x(u)_{n-1,j} &= 0,\\ \nabla_y(u)_{i,j} &= \frac{u_{i,j+1} - u_{i,j}}{h}, j < n-1,\\ \nabla_y(u)_{i,n-1} &= 0,\end{aligned}$$

for any function u on the grid. With these choices, the discretization of the ∇ and $-$div operator are

$$\nabla_h(u) = \begin{bmatrix}\nabla_x u\\ \nabla_y u\end{bmatrix},$$
$$-\text{div}_h \begin{bmatrix}w^x\\ w^y\end{bmatrix} = \nabla_x^T w^x + \nabla_y^T w^y,$$

for grid function u, w^x, w^y and the discretization of problem (2) can be written as:

$$\begin{aligned}&\min_u \sum_{i,j=0}^{n-1} \phi(|\nabla_h u_{i,j}|_\beta)dx,\\ &\text{subject to } \frac{1}{2}\sum_{i,j=0}^{n-1}(u_{i,j} - z_{i,j}))^2 dx = \frac{\sigma^2}{2},\end{aligned} \tag{6}$$

where $u_{i,j} \approx u(ih, jh), z_{i,j} = z(ih, jh)$. It is easy to see that the optimality conditions for problem (6) read:

$$-\text{div}_h\left(\frac{\phi'(|\nabla_h u|_\beta)}{|\nabla_h u|_\beta}\nabla_h u\right) + \lambda(u - z) = 0, \tag{7}$$

i.e., the corresponding discretization of the Euler-Lagrange equations (4). Likewise, the Hessian of the Lagrangian $F_h(u)$ of (6) is

$$\begin{aligned}F_h''(u)(v) = -\text{div}_h\Big[&\frac{\phi'(|\nabla_h u|_\beta)}{|\nabla_h u|_\beta}\nabla_h v +\\ &G(\nabla_h u)\frac{\nabla_h u \cdot \nabla_h v}{|\nabla_h u|_\beta^2}\nabla_h u\Big] + \lambda v,\\ G(\nabla_h u) = \phi''(|\nabla_h u|_\beta) &- \frac{\phi'(|\nabla_h u|_\beta)}{|\nabla_h u|_\beta},\end{aligned}$$

that corresponds to the consistent discretization of $F''(u)(v)$. Consequently, we will drop the subscript h from the finite difference operators whenever no confusion can arise.

Assuming that $\phi \in \mathcal{C}^2$, it is easy to check that $F''(u)$ is positive definite and therefore F is strictly convex. Since it is a continuous and coercive function (that is, $\lim_{u\to\infty} |F(u)| = \infty$), there exist a unique minimum of the discretized problem (e.g. see [9,21]).

3. Algorithms for generalized Total Variation denoising

In this section we briefly review some algorithms for the solution of (7). They are direct generalizations of the corresponding algorithms proposed for Total Variation restoration.

3.1. Time Marching

Since (7) is of elliptic type, one can obtain its solution by approximating the steady state of the associated parabolic equation

$$\begin{aligned}u_t &= \text{div}\left(\frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}\nabla u\right) - \lambda(u - z),\\ \frac{\partial u}{\partial \overrightarrow{n}} &= 0, u(x, 0) = z(x),\end{aligned} \tag{8}$$

for the unknown function $u = u(x, t)$. In principle, an explicit Euler scheme could be applied to get the iteration:

$$\begin{aligned}u^{n+1} = u^n + \Delta t\Big(&\text{div}\big(\frac{\phi'(|\nabla u^n|_\beta)}{|\nabla u^n|_\beta}\nabla u^n\big)\\ &- \lambda(u^n - z)\Big), u_{i,j}^n \approx u(ih, jh, n\Delta t).\end{aligned} \tag{9}$$

This simple scheme, proposed and used in the seminal paper [22], has the disadvantage of having severe CFL restrictions on Δt, caused not only by the parabolic nature, but also for the fact that the diffusion coefficient $\frac{\phi'(|\nabla u^n|_\beta)}{|\nabla u^n|_\beta}$ can be very large for small β. Implicit schemes could be applicable, but their difficulty is comparable to the alternatives that appear below, hence it would not be clearly effective when compared to them.

3.2. Fixed point

A natural linearization of (7) consists in the following fixed point iteration:

$$\begin{aligned}&-\text{div}\left(\frac{\phi'(|\nabla u^n|_\beta)}{|\nabla u^n|_\beta}\nabla u^{n+1}\right)\\ &+ \lambda(u^{n+1} - z) = 0, n \geq 0, \quad u^0 = z,\end{aligned} \tag{10}$$

where at each step a linear reaction-diffusion equation has to be solved, with the diffusivity computed from the previous step. This generalized the *lagged diffusivity fixed point* scheme introduced by Vogel and Oman in [24] for Total Variation restoration.

The natural discretization of this scheme converges to the solution of (7) at a linear rate that can deteriorate towards 1 as $\beta \to 0$ (see [8] for a proof).

3.3. Newton primal

In principle, a faster convergence can be obtained with Newton's method, which consists in the iteration:

$$F''(u^n)(\delta u^n) = -F'(u^n),$$
$$u^{n+1} = u^n + \delta u^n.$$

It is a quadratically convergent algorithm (since then we can assure that $F''(u)$ is always non singular), but their convergence is not ensured, unless *line search procedures* are used. But these procedures typically restrict the step-length in a dramatic manner when β is small. *Continuation strategies* on the parameter β can be used to cope with this difficulty (see [5,7]).

4. Newton primal-dual

In [7], the Euler-Lagrange equations of the TV restoration problem, that is, problem (7) with $\phi(x) = x$, are enlarged by introducing the variable (dual) $w = \frac{\nabla u}{|\nabla u|_\beta}$ and the new equation $|\nabla u|_\beta w - \nabla u = 0$. Then, Newton's method is applied to this enlarged system. The resulting algorithm has proven to be very robust and fastly convergent. In the literature one can find other approaches to the TV denoising/restoration problem that use dual variables [3,12,25, 4].

In this paper we propose a primal-dual Newton's method for generalized TV restoration. This method is an extension to the primal-dual method proposed in [7]. Following those ideas, the Euler-Lagrange equation (7) is rewritten by introducing $\frac{\nabla u}{|\nabla u|_\beta}$ as a new variable w:

$$-\text{div}\,(\phi'(|\nabla u|_\beta)w) + \lambda(u - z) = 0, \qquad (11)$$
$$|\nabla u|_\beta w - \nabla u = 0, \qquad (12)$$

and then applying Newton's method to this system:

$$-\text{div}\Big(\phi'(|\nabla u|_\beta)\delta w + \phi''(|\nabla u|_\beta)(\frac{\nabla u}{|\nabla u|_\beta}\cdot\nabla\delta u)w\Big) + \lambda(u-z) = -(-\text{div}\,(\phi'(|\nabla u|_\beta w)) + \lambda(u-z)),$$

$$\frac{\nabla u\cdot\nabla\delta u}{|\nabla u|_\beta}w - \nabla\delta u + |\nabla u|_\beta\delta w = -(|\nabla u|_\beta w - \nabla u).$$

From the latter equation δw can be expressed in terms of δu:

$$\delta w = -(1 + \frac{\nabla u\cdot\nabla\delta u}{|\nabla u|_\beta^2})w + \frac{\nabla u}{|\nabla u|_\beta} + \frac{\nabla\delta u}{|\nabla u|_\beta}, \qquad (13)$$

which, when substituted in the former equation, yields:

$$-\text{div}\,(K(u,w)\nabla\delta u) + \lambda\delta u = \text{div}\left(\frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}\nabla u\right) - \lambda(u-z),$$

$$K(u,w)_{i,j} = \frac{\phi'(|\nabla u_{i,j}|_\beta)}{|\nabla u_{i,j}|_\beta}I_2 + G(\nabla u_{i,j})w_{i,j}\left(\frac{\nabla u_{i,j}}{|\nabla u_{i,j}|_\beta}\right)^T,$$
$$G(\nabla u_{i,j}) = \phi''(|\nabla u_{i,j}|_\beta) - \frac{\phi'(|\nabla u_{i,j}|_\beta)}{|\nabla u_{i,j}|_\beta}. \qquad (14)$$

Due to the presence of the term

$$w_{i,j}\left(\frac{\nabla u_{i,j}}{|\nabla u_{i,j}|_\beta}\right)^T,$$

the 2×2 matrices $K(u,w)_{i,j}$ are symmetric only when $w = \frac{\nabla u}{|\nabla u|_\beta}$ for each (i,j). Of course, the primal Newton's method is recovered with this setting. A key point here is to substitute $K(u,w)$ by its symmetrization (we omit the i,j index)

$$\widetilde{K}(u,w) = \frac{1}{2}(K(u,w) + K(u,w)^T) = \frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}I_2 + \frac{1}{2}G(\nabla u)\left(w\left(\frac{\nabla u}{|\nabla u|_\beta}\right)^T + \frac{\nabla u}{|\nabla u|_\beta}w^T\right),$$
$$G(\nabla u) = \phi''(|\nabla u|_\beta) - \frac{\phi'(|\nabla u|_\beta)}{|\nabla u|_\beta}.$$

The resulting *inexact* Newton's method can then use numerical methods for the solution of the linear systems that appear at each iteration without sacrificing the quadratic convergence of Newton's method (see [14, chap. 5 and 6]), since the matrix

$$-\text{div}\left(\widetilde{K}(u^*,w^*)\nabla\right) + \lambda I \qquad (15)$$

coincides with $F''(u^*)$ at the solution

$$(u^*, w^*),$$

for which $w^* = \frac{\nabla u^*}{|\nabla u^*|_\beta}$. The primal-dual Newton's method that we propose as an extension of the one proposed

Figure 1 Left: original image; right: noisy image.

in [7] reads:

$$\begin{aligned} -\text{div}\left(\widetilde{K}(u^n, w^n)\nabla\delta u^n\right) + \lambda\delta u^n \\ = \text{div}\left(\frac{\phi'(|\nabla u^n|_\beta)}{|\nabla u^n|_\beta}\nabla u^n\right) - \lambda(u^n - z) \\ \delta w^n = -(1 + \frac{\nabla u^n \cdot \nabla\delta u^n}{|\nabla u^n|^2_\beta})w^n \\ + \frac{\nabla u^n}{|\nabla u|^n_\beta} + \frac{\nabla\delta u^n}{|\nabla u^n|_\beta}, \\ u^{n+1} = u^n + \delta u^n, \quad w^{n+1} = w^n + s\delta w^n, \end{aligned} \tag{16}$$

where $u^0 = z$, $w^0 = 0$ and s is computed to enforce $|w^{n+1}|_\infty \leq 1$ in an efficient way by setting

$$\begin{aligned} s &= \min(0.95\min\{s_i, \delta w_i \neq 0\}, 1), \\ s_i &= \frac{-z_i + \sqrt{z_i^2 + (1 - |w_i|^2)|\delta w_i|^2}}{|\delta w_i|^2 + 10^{-16}}, \end{aligned}$$

with $z_i = w_i \cdot \delta w_i$. The advantage of this approach is that there are very efficient numerical methods for sparse symmetric matrices (conjugate gradient or symmlq [20]) that can be therefore applied to the matrix in (16).

5. Numerical experiments

In this section we compare the convergence speed of the fixed point method and the primal-dual Newton's method for several functions $\phi(x)$. We display in Fig. 1 the original and noisy images that are used for all the experiments. The noise level corresponds to an SNR=2 and $\sigma^2 \approx 571$.

Since we are testing full convergence, the stopping criterion for the (outer) Newton's iteration is a relative decrease of the nonlinear residual (i.e. the right hand side member of (16)) by a factor of 10^{-8}. We use a *truncated* Newton's algorithm, based on the conjugate gradient method with incomplete Cholesky, with a fill-in level of 3, as preconditioner (cf. [11]). The stopping criterion for the n-th inner linear iteration is a relative decrease of the linear residual by a factor of η_n, where we follow the suggestion of [14, Eq. 6.18] and set

$$\eta_n = \begin{cases} 0.01 \text{ if } n = 0, \\ \min(0.01, 0.6||g_n||^2/||g_{n-1}||^2) \end{cases}, \tag{17}$$

where g_n denotes the gradient of the objective function i.e. the right hand side of (16). We point out here that, although we can only ensure the positive definiteness of the matrix in (16) for $n = 0$ (for which $w = 0$ and the matrix is certainly positive definite) or near the minimum u^* (at which it coincides with $F''(u^*)$), the convergence of the conjugate gradient method has been extremely satisfactory in all cases, requiring very few iterations at the early stages of the outer iteration and some more when η_n is set to small numbers when full convergence is near to take place.

For the fixed point iteration, we use the same linear solver with a fixed relative tolerance on the residual set to 0.01.

In the first test we analyze the influence of the parameter β on the convergence speed of the algorithm. We run both algorithms for $\phi(x) = x^{1.1}$, $\lambda = 20.5$, close to the Lagrange multiplier that ensures the quadratic constraint imposed by the noise variance, and $\beta = 1, 10^{-5}, 10^{-10}$ and display their convergence history in terms of the gradient of the objective function versus iterations and CPU time with vertical logarithmic scale in Fig. 2-4. From these pictures, we deduce a quadratic convergence for the primal-dual Newton method and a linear convergence for the fixed point method and that the gradient of the objective function for each iteration of the primal-dual method is smaller or equal than for the fixed point method. Regarding efficiency, the pictures that display CPU time versus nonlinear residual show that the fixed point method is slightly more efficient than the primal-dual method at the early stages, due to the fact that the cost per iteration of the linear solver is about 30% higher for the latter.

	primal-dual		fixed point	
β	iterations	CPU time	iterations	CPU time
1	18	28.5	137	131.9
10^{-5}	23	50.7	150	183.7
1^{-10}	25	57.1	154	391.2

Table 1 Number of iterations and CPU time for the first test with $\phi(x) = x^{1.1}$ and $\beta = 1, 10^{-5}, 10^{-10}$.

As can be seen in Table 1, where we display the final number of iterations and CPU time in each case, the difference between the computational cost of the fixed point method with respect to the primal-dual method grows dramatically as β is reduced.

We perform more tests with another interesting function ϕ, the Huber function, used in robust Estimation (cf. [13]):

$$\phi(x) = \begin{cases} x^2 & x < \xi \\ 2\xi x - \xi^2 & x \geq \xi \end{cases}. \tag{18}$$

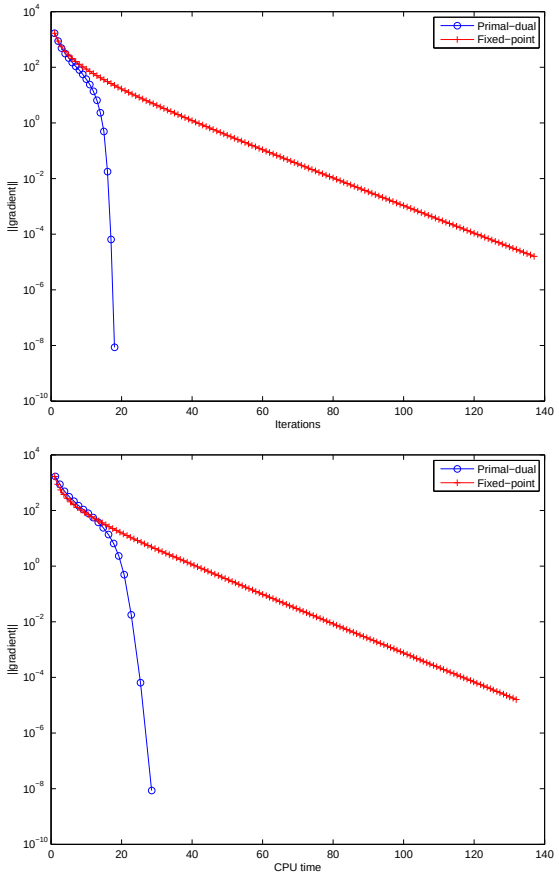

Figure 2 The convergence history in terms of the gradient of the objective function, for $\phi(x) = x^{1.1}$, $\beta = 1$: gradient versus iterations at the top; gradient versus CPU time at the bottom. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

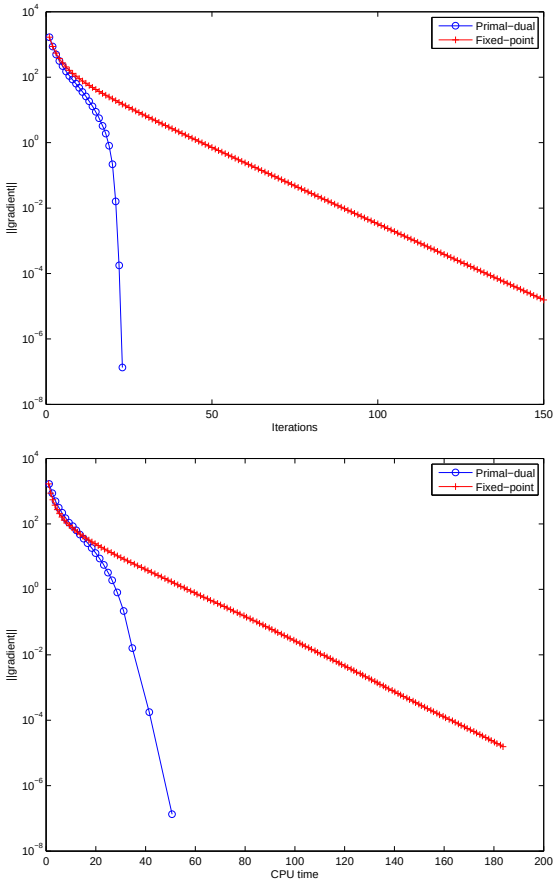

Figure 3 The convergence history in terms of the gradient of the objective function, for $\phi(x) = x^{1.1}$, $\beta = 10^{-5}$: gradient versus iterations at the top; gradient versus CPU time at the bottom. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

This function is convex and continuously differentiable, but the second derivative is discontinuous at ξ. We set

$$\phi''(\xi) = \phi''(\xi^+) = 0.$$

This function is designed to penalize small jumps more than the total variation functional, as prescribed in [19] to address the staircase effect.

The picture at the left side of figure 5 displays a piece of the original picture. This piece contains smooth parts (although the original picture contains some noise) and edges. We plan to run the method that we propose and the fixed point method for the Huber function ϕ with several threshold parameters ξ and compare their (common) output with that obtained with Total Variation denoising, shown at the right side of figure 5. We also compare the convergence history of both methods in terms of the nonlinear residual against iterations and CPU time. The parameter λ has been set to ensure, in every case, the constraint:

$$\int_\Omega (u(x) - z(x))^2 \, dx = \sigma^2 \approx 571.$$

The threshold parameters we have chosen for the Huber function are $\xi = 1000$ (results in Figures 6 and 7), $\xi = 500$ (results in Figures 8 and 9) and $\xi = 100$ (results in Figures 10 and 11). These results show edges that are as well defined as those obtained with the Total Variation functional. As expected, the higher the threshold parameter ξ is set, the less staircase effect in the denoised images. The conclusions, in terms of the convergence velocity, that can be drawn from these tests are similar to the ones obtained previously: the primal-dual method converges much faster than the fixed point method, although the latter is slightly more efficient than the former in the early stages.

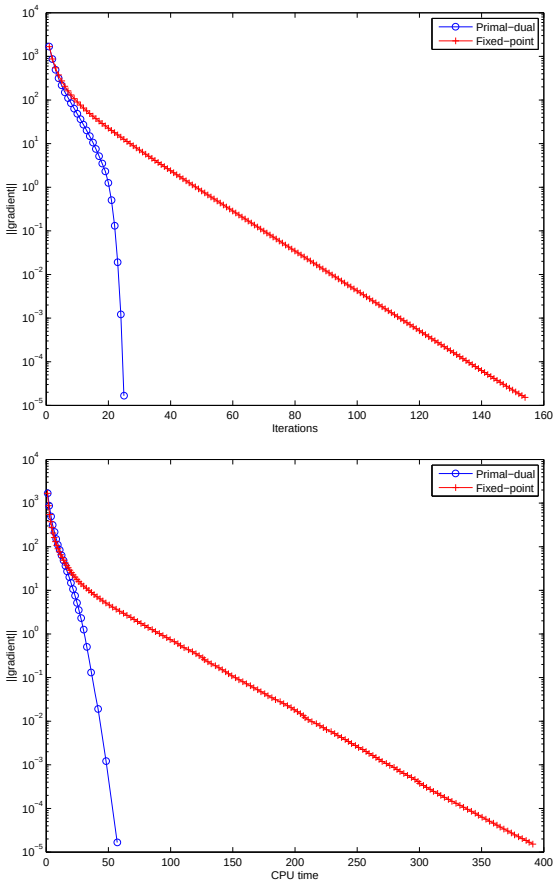

Figure 4 The convergence history in terms of the gradient of the objective function, for $\phi(x) = x^{1.1}$, $\beta = 10^{-10}$: gradient versus iterations at the top; gradient versus CPU time at the bottom. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

Figure 5 Enlarged views of a part of the original image (left) and denoised image with the function $\phi(x) = x$ (Total Variation denoising) with $\beta = 1$

Figure 6 Denoised image with the Huber function ϕ with $\xi = 1000$ (left) and enlarged view of a part of the denoised image.

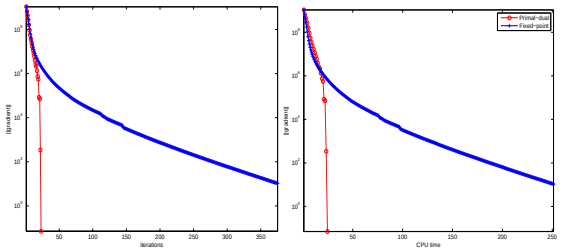

Figure 7 Convergence history for the Huber function ϕ with $\xi = 1000$ with $\beta = 1$ and $\lambda = 30000$: top, iterations versus nonlinear residual; bottom, CPU time versus nonlinear residual. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

Figure 8 Denoised image with the Huber function ϕ with $\xi = 500$ (top) and enlarged view of a part of the denoised image.

6. Conclusions

Image restoration based on the minimization of the Total Variation functional is a reliable technique for recovering sharp edges of the restored images, but typically tends to over-sharpen smooth transitions between gray levels, turning them into piecewise constant functions (staircase effect). In this paper we have proposed a numerical method for solving generalized Total Variation problems, that aim to alleviate this staircase effect, in a fast and robust manner. We have shown that it is highly competitive compared to the lagged diffusivity fixed point method, as our experiments illustrates. The extension of the ideas herein to other

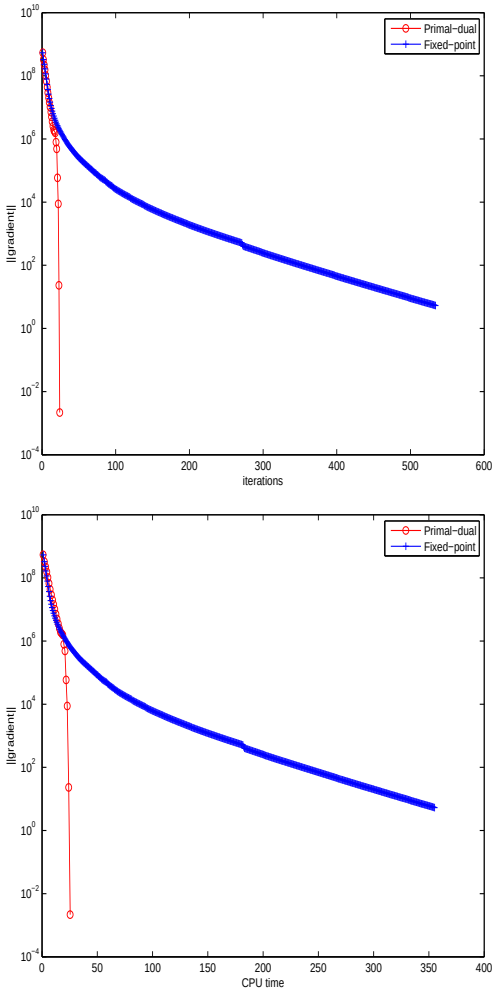

Figure 9 Convergence history for the Huber function ϕ with $\xi = 500$ with $\beta = 1$ and $\lambda = 15800$: left, iterations versus nonlinear residual; right, CPU time versus nonlinear residual. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

Figure 10 Denoised image with the Huber function ϕ with $\xi = 100$ (left) and enlarged view of a part of the denoised image.

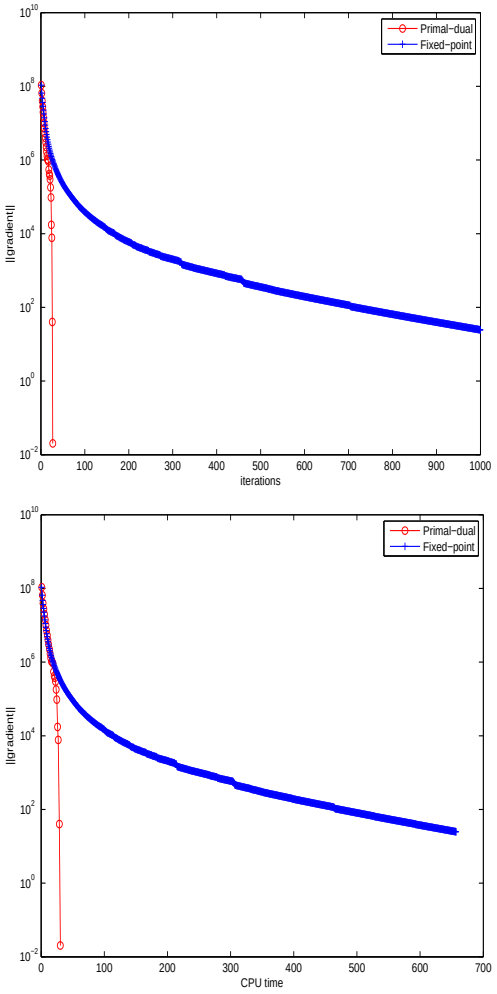

Figure 11 Convergence history for the Huber function ϕ with $\xi = 100$ with $\beta = 1$ and $\lambda = 3300$: left, iterations versus nonlinear residual; bottom, CPU time versus nonlinear residual. The results of primal-dual method are displayed with $\circ$ and those for the fixed point method with $+$.

functionals and to image deblurring is a matter of future research.

Acknowledgments:

This research was partially supported by Spanish MCINN MTM2008-00974 and MINECO MTM2011-22741.

References

[1] R. ACAR AND C. R. VOGEL, *Analysis of total variation penalty methods for ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217–1229.

[2] F. ANDREU, C. BALLESTER, V. CASELLES, AND J. M. MAZÓN, *Minimizing total variation flow*, Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, 331 (2000), pp. 867 – 872.
[3] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97. Special issue on mathematics and image analysis.
[4] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.
[5] R. H. CHAN, T. F. CHAN, AND H. M. ZHOU, *Continuation method for total variation denoising problems*, Tech. Rep. 95-18, University of California, Los Angeles, 1995.
[6] T. CHAN, A. MARQUINA, AND P. MULET, *High-order total variation-based image restoration*, SIAM J. Sci. Comput., 22 (2000), pp. 503–516.
[7] T. F. CHAN, G. H. GOLUB, AND P. MULET, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20 (1999), pp. 1964–1977 (electronic).
[8] T. F. CHAN AND P. MULET, *On the convergence of the lagged diffusivity fixed point method in total variation image restoration*, SIAM J. Numer. Anal., 36 (1999), pp. 354–367 (electronic).
[9] I. EKELAND AND R. TEMAM, *Analyse convexe et problèmes variationnels*, Dunod, Paris, 1974.
[10] E. GIUSTI, *Minimal Surfaces and Functions of Bounded Variations*, Birkhäuser, 1984.
[11] G. GOLUB AND C. VAN LOAN, *Matrix computations*, 2^{nd} *ed.*, The John Hopkins University Press, 1989.
[12] M. HINTERMÜLLER AND G. STADLER, *An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration*, SIAM J. Sci. Comput., 28 (2006), pp. 1–23 (electronic).
[13] P. J. HUBER, *Robust estimation of a location parameter*, Ann. Math. Statist., 35 (1964), pp. 73–101.
[14] C. T. KELLEY, *Iterative Methods for Linear and Nonlinear Equations*, vol. 16 of Frontiers in Applied Mathematics, SIAM, 1995.
[15] F. LI, C. SHEN, J. FAN, AND C. SHEN, *Image restoration combining a total variational filter and a fourth-order filter*, Journal Of Visual Communication And Image Representation, 18 (2007), pp. 322–330.
[16] M. LYSAKER AND X. TAI, *Iterative image restoration combining total variation minimization and a second-order functional*, International Journal Of Computer Vision, 66 (2006), pp. 5–18.
[17] M. NIKOLOVA, *Local strong homogeneity of a regularized estimator*, Siam J. Appl. Math., 61 (2000), pp. 633–658.
[18] ———, *Weakly constrained minimization. application to the estimation of images and signals involving constant regions*, Technical Report CMLA, (2001).
[19] P. M. P. BLOMGREN, T. F. CHAN AND C. K. WONG, *Total variation image restoration: Numerical methods and extensions*, Los Angeles, (1997).
[20] C. C. PAIGE AND M. A. SAUNDERS, *Solutions of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
[21] T. ROCKAFELLAR, *Convex analysis*, SIAM, 1970.
[22] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
[23] A. N. TIKHONOV AND V. Y. ARSENIN, *Solutions of Ill-Posed Problems*, John Wiley, New York, 1977.
[24] C. R. VOGEL AND M. E. OMAN, *Iterative methods for total variation denoising*, SIAM J. Sci. Statist. Comput., 17 (1996), pp. 227–238.
[25] M. ZHU, S. J. WRIGHT, AND T. F. CHAN, *Duality-based algorithms for total-variation-regularized image restoration*, Comput. Optim. Appl., 47 (2010), pp. 377–400.

Francesc Arandiga was born in Canals, Spain, in 1963. He received the Ph.D. degree in mathematics from the University of Valencia, Valencia, Spain, in 1992. From 1994 to 2010 he has been associate professor and since 2010 full professor of applied mathematics at the University of Valencia. He has been head and secretary of the Department of Applied Mathematics at the University of Valencia. He has also been coordinator of the graduate program in Computer Science and Computational Mathematics. His research interest include image processing, multiresolution analysis, and wavelets. He has published more than 30 research articles in reputed international journals of mathematical and engineering sciences.

Pep Mulet obtained his M.Sc. in 1989 and his Ph.D. in 1992, both in mathematics from the University of Valencia, Spain. He is currently associate professor at the applied mathematics department of the University of Valencia. He is now head of this department. His research interest include image processing, adaptive scientific computing and numerical analysis of partial differential equations. He has published more than 30 research articles in reputed international journals of mathematical and engineering sciences.

Vicent Renau obtained his M.Sc. in 2004 and his Ph.D. in 2011, both in mathematics from the University of Valencia, Spain. He is currently mathematics professor at the High School Serra d'Irta d'Alcal de Xivert, Castellón, Spain His research interest include image processing and multiresolution analysis.