UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Extensions to Total Variation Denoising

Peter Blomgren
Tony F. Chan
Pep Mulet

September 1997

CAM Report 97-42

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

EXTENSIONS TO TOTAL VARIATION DENOISING [§]

PETER BLOMGREN[*], TONY F. CHAN[†] AND PEP MULET[‡]

Abstract. The Total Variation denoising method, proposed by Rudin, Osher and Fatemi, 92, is a PDE-based algorithm for edge-preserving noise removal. The images resulting from its application are usually piecewise constant, possibly with a *staircase effect* at smooth transitions and may contain significantly less fine details than the original non-degraded image.

In this paper we present some extensions to this technique that aim to improve the above drawbacks, through redefining the Total Variation functional or the noise constraints.

1. Introduction. The degradation of an image is usually unavoidable at the early stages of its processing and it renders difficult and inaccurate the latter phases. The algorithms for image denoising have been mainly based on least squares and, consequently, their results are continuous and do not approximate well images containing edges. To overcome this difficulty a technique based on the minimization of the Total Variation norm subject to some noise constraints is proposed in Ref. [1]. The images resulting of the application of this technique are often piecewise constant, which implies that the finer details in the original image may not be recovered satisfactorily and that *ramps* (affine regions) will give *stairs* (piecewise constant regions). In this paper we present some extensions to the Total Variation image restoration method that aim to improve these drawbacks.

The paper is organized as follows: in section 2 we introduce the problem and the nonlinear equations associated to it. In section 3, the *global* noise constraint of the method is replaced by *local* noise constraints, in order to get a better recovery of the fine details of the image. In section 4 we present changes to the Total Variation functional which aim to solve the *staircase effect*.

2. Total Variation Denoising.

2.1. Degradation model. An image can be interpreted as either a real function defined on Ω, a bounded and open domain of $\mathbb{R}^2$, (for simplicity we will assume Ω to be the unit square henceforth) or as a suitable discretization of this continuous image.

Our interest is to restore an image which is contaminated with noise, in such a way that the process should recover the edges of the image. Let us denote by z the observed image and u the real image. The model of degradation we assume is $u + n = z$, where n is a Gaussian white noise, i.e., the values n_i of n at the pixels i are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2 .

Our objective is to estimate u from statistics of the noise and some *a priori* knowledge of the image (smoothness, existence of edges). This knowledge is incorporated into the formulation by using a functional R that measures the quality of the image

[*]Department of Mathematics, University of California, Los Angeles, CA 90095. E-mail address: blomgren@math.ucla.edu. Supported by grants NSF ASC-92-01266 and ONR-N00014-96-1-0277.

[†]Department of Mathematics, University of California, Los Angeles, CA 90095. E-mail address: chan@math.ucla.edu. Supported by grants NSF ASC-92-01266, ONR-N00014-96-1-0277 and NSF INT9602089

[‡]Department of Mathematics, University of California, Los Angeles, CA 90095 and Departament de Matemàtica Aplicada, Universitat de València, 46100 Spain. E-mail address: mulet@math.ucla.edu, mulet@uv.es. Supported by DGICYT grant PB94-0987 and NSF INT9602089

[§]To appear in Proc. SPIE 97, San Diego

u, in the sense that smaller values of $R(u)$ correspond to better images. The process, in other words, consists in the choice of the best quality image among those matching the constraints imposed by the statistics of the noise.

2.2. Regularization. The usual approach consists in solving the following constrained optimization problem:

$$\begin{aligned} &\min_u R(u) \\ &\text{subject to } ||u-z||^2_{\mathcal{L}^2} = |\Omega|\sigma^2, \end{aligned} \tag{2.1}$$

since $n = z-u$ and $E(\int_\Omega n^2\,dx) = |\Omega|\sigma^2$ ($E(X)$ denotes the expectation of the random variable X) give that $||u-z||^2_{\mathcal{L}^2} = \int_\Omega (u-z)^2\,dx \approx |\Omega|\sigma^2$.

Examples of regularization functionals that can be found in the literature are, $R(u) = \|u\|_{\mathcal{L}^2}, \|\Delta u\|_{\mathcal{L}^2}, \|\nabla u \cdot \nabla u\|_{\mathcal{L}^2}$, where ∇ is the gradient and Δ is the Laplacian, see Refs. [2, 3]. The drawback of using these functionals is that they do not allow discontinuities in the solution.

2.3. Total Variation regularization. In Ref. [1], it is proposed to use as regularization functional the so-called *Total Variation norm* or *TV-norm*:

$$TV(u) = \int_\Omega |\nabla u|\,dx = \int_\Omega \sqrt{u_x^2 + u_y^2}\,dx. \tag{2.2}$$

The TV norm does not penalize discontinuities in u, and thus allows us to recover the edges of the original image. The restoration problem can be thus written as

$$\begin{aligned} &\min_u \int_\Omega |\nabla u|\,dx, \\ &\text{s.t. } \frac{1}{2}\Big(\int_\Omega (u-z)^2\,dx - |\Omega|\sigma^2\Big) = 0. \end{aligned} \tag{2.3}$$

We refer to this problem as the *globally constrained* problem. Its Lagrangian is

$$\int_\Omega |\nabla u|\,dx + \frac{\lambda}{2}\Big(\int_\Omega (u-z)^2\,dx - |\Omega|\sigma^2\Big)$$

and its Euler-Lagrange equations, assuming homogeneous Neumann boundary conditions, are:

$$\begin{aligned} 0 &= -\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda(u-z) \\ 0 &= \frac{1}{2}\Big(\int_\Omega (u-z)^2\,dx - |\Omega|\sigma^2\Big). \end{aligned} \tag{2.4}$$

Since the first equation is not well defined at points where $\nabla u = 0$, due to the presence of the term $1/|\nabla u|$, it is common to slightly perturb the Total Variation norm functional to become:

$$\int_\Omega \sqrt{|\nabla u|^2 + \beta}\,dx\,dy, \tag{2.5}$$

where β is a small positive parameter. In Ref. [4] it is shown that the solutions of the perturbed problems converge to the solution of (2.3) when $\beta \to 0$. The main difficulty that equation (2.4) poses is the linearization of the highly nonlinear term $\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right)$.

A number of methods have been proposed to solve (2.4). For the simplicity of the exposition, let us assume that λ is fixed. L. Rudin, S. Osher and E. Fatemi [1] used a time marching scheme to reach the steady state of the parabolic equation:

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda(u - z), \quad u(x, 0) = z(x). \tag{2.6}$$

This method can be slowly convergent due to stability constraints. C. Vogel and M. Oman [5] proposed the following fixed point iteration to solve the Euler-Lagrange equation:

$$-\nabla \cdot \left(\frac{\nabla u^{k+1}}{|\nabla u^k|} \right) + \lambda(u^{k+1} - z) = 0. \tag{2.7}$$

Starting with $u_0 = z$, at each step u^{k+1} is obtained as the solution of the linear differential equation (2.7), whose coefficients are computed from u^k. This method is robust but only linearly convergent.

Due to the presence of the highly nonlinear term $\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right)$, Newton's method does not work satisfactorily, in the sense that its domain of convergence is very small. This is especially true if the regularizing parameter β is small. On the other hand, if β is relatively large then this term is well behaved. So it is natural to use a continuation procedure starting with a large value of β and gradually reducing it to the desired value. T. Chan, R. Chan and H. Zhou proposed in Ref. [6] such an approach. Although this method is locally quadratically convergent, the choice of the sequence of subproblems to solve is crucial for its efficiency. The authors have not succeeded in finding a fully satisfactory selection procedure, although some heuristics can be used.

In Ref. [7] the above Euler-Lagrange equations are restated in such a way that it results in a more *globally linear* system of equations, to which Newton's method can be applied satisfactorily, giving a very robust and locally quadratically convergent algorithm. We give more details of this approach in the next section.

3. Local Constraints.

3.1. Motivation. A usual effect of the Total Variation denoising method is a reduction of contrast in regions of the same pixel intensity, this reduction depending on the size of the region: the bigger the region the smaller the intensity change. As a consequence, small details may disappear in the solution if the noise level is sufficiently high, whereas large features remain. Of course, the sizes of those details are relative, in the sense that a small feature can be regarded as a large one of the result of clipping the rest of the image and only retaining a small neighborhood of the feature. This suggests that a good strategy for recovering fine details of an image could be the division of Ω into subregions $\Omega_1, \ldots, \Omega_m$ and the solution of the following problem in each subregion:

$$\begin{aligned} &\min_u \int_{\Omega_i} |\nabla u| \, dx, \\ &\text{s.t. } \frac{1}{2}\Big(\int_{\Omega_i} (u - z)^2 \, dx - |\Omega_i|\sigma^2 \Big) = 0, \quad i = 1, \ldots, m \end{aligned} \tag{3.1}$$

This approach has a serious drawback: the solutions corresponding to two adjacent regions would not match at their intersection and this would certainly create artifacts

at those pixels. Of course, this effect is due to the fact that this approach consists in the solution of independent *local* problems.

3.2. Local constraints. We propose an intermediate approach consisting in the solution of a minimization problem with the same objective function as (2.3) and constrained by the set of constraints of the local problems. To state it more precisely, the proposed problem is:

$$\begin{aligned} &\min_u \int_\Omega |\nabla u|\, dx, \\ &\text{s.t. } \frac{1}{2}\Big(\int_{\Omega_i} (u-z)^2\, dx - |\Omega_i|\sigma^2\Big) = 0,\ i = 1, \ldots, m. \end{aligned} \tag{3.2}$$

We refer to it as the *locally constrained problem.* Its Lagrangian is:

$$\int_\Omega |\nabla u|\, dx + \sum_{i=1}^m \frac{\lambda_i}{2}\Big(\int_{\Omega_i} (u-z)^2\, dx - |\Omega_i|\sigma^2\Big),$$

and its Euler-Lagrange equations, assuming homogeneous Neumann boundary conditions, are:

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) + \sum_{i=1}^m \lambda_i P_i(u-z) \tag{3.3}$$

$$0 = \frac{1}{2}\Big(\int_{\Omega_i} (u-z)^2\, dx - |\Omega_i|\sigma^2\Big),\ i = 1, \ldots, m \tag{3.4}$$

where P_i denotes the restriction operator given by $P_i(v) = v|_{\Omega_i}$.

3.3. Linearization of the Euler-Lagrange equations. Similarly to the globally constrained case, the highly non-linear term $\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$ causes numerical difficulties when linearizing the equations by Newton's method. This difficulty can be alleviated by the introduction of an auxiliary variable w, which is constrained to verify the equation

$$w = \frac{\nabla u}{|\nabla u|}.$$

Equation (3.3) is then replaced by the two equations:

$$0 = -\nabla \cdot w + \sum_{i=1}^m \lambda_i P_i(u-z) \tag{3.5}$$

$$0 = |\nabla u| w - \nabla u. \tag{3.6}$$

Equation (3.5) is now mildly non-linear and (3.6) is more linear than its equivalent equation

$$0 = w - \frac{\nabla u}{|\nabla u|}.$$

In this fashion, we have eliminated part of the non-linearity of (3.3) *before* applying Newton's method.

The Newton's method equations in block matrix form are:

$$\begin{bmatrix} -\nabla\cdot & \sum_{i=1}^m \lambda_i P_i & P(u-z) \\ |\nabla u| & (\frac{\nabla u w^T}{|\nabla u|} - 1)\nabla & 0 \\ (P(u-z))^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta w \\ \delta u \\ \delta\lambda \end{bmatrix} \\ = - \begin{bmatrix} -\nabla\cdot w + \sum_{i=1}^m \lambda_i P_i(u-z) \\ |\nabla u| w - \nabla u \\ [\frac{1}{2}(\int_{\Omega_i}(u-z)^2\,dx - |\Omega_i|\sigma^2)] \end{bmatrix}, \tag{3.7}$$

where $P = [P_1 \ldots P_m]$.

We can further use the second set of equations of (3.7) to eliminate δw:

$$\begin{bmatrix} -\nabla\cdot(\frac{1}{|\nabla u|}(1 - \frac{\nabla u\, w^T}{|\nabla u|})\nabla) + \sum_{i=1}^m \lambda_i P_i & P(u-z) \\ (P(u-z))^T & 0 \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\lambda \end{bmatrix} \tag{3.8}$$

$$= - \begin{bmatrix} -\nabla\cdot w + \sum_{i=1}^m \lambda_i P_i(u-z) \\ [\frac{1}{2}(\int_{\Omega_i}(u-z)^2\,dx - |\Omega_i|\sigma^2)] \end{bmatrix}, \tag{3.9}$$

$$\delta w = -w + \frac{\nabla u}{|\nabla u|} - (\frac{\nabla u\, w^T}{|\nabla u|^2} - 1)\nabla\delta u, \tag{3.10}$$

and make this into an iteration.

A fact that appears as crucial for the efficiency and stability of the iteration is that the bound $\|w\|_\infty < 1$ should be satisfied during the iteration. This can be achieved as follows: once computed δw^k from (3.10), the bound $\|w^{k+1}\|_\infty < 1$ can be enforced by setting

$$w^{k+1} = w^k + s\delta w^k, \tag{3.11}$$

where $s = min\{\rho\bar{s}, 1\}$, $\rho \in (0,1)$ and $\bar{s}$ satisfies $\|w^k + \bar{s}\delta w^k\|_\infty = 1$.

Another important fact for the stability of the iteration is that $\lambda_i > 0$, $i = 1, \ldots, m$. These bounds can be maintained during the iteration in a similar way as before, by setting

$$\lambda_i^{k+1} = \lambda_i^k + t\delta_i^k, \tag{3.12}$$

with $t = \min\{\rho\bar{t}, 1\}$ and $\bar{t}$ such that $\min_i\{\lambda_i^k + \bar{t}\delta_i^k\}$=0.

A final remark is that the obvious choice $u^0 = z$ would yield a singular coefficient matrix of the system (3.8), thus u^0 must be chosen in a somewhat random way. The complete algorithm appears in Fig. 3.1.

choose $\lambda_i > 0$, $i = 1, \ldots, m$
choose $u \neq z$
$w = 0$
for $k = 1, \ldots$
  compute right hand side of (3.8)
  if $\|$ right hand side $\| <$ tolerance
    stop
  end
  obtain δu^k, δw^k, $\delta \lambda^k$ from (3.8) and (3.10)
  $u^{k+1} = u^k + \delta u^k$
  compute s from (3.11)
  $w^{k+1} = w^k + s\delta w^k$
  compute t from (3.12)
  $\lambda^{k+1} = \lambda^k + t\delta\lambda^k$
end

FIG. 3.1. *Algorithm for solving the Total Variation denoising problem with local constraints.*

3.4. Numerical results. The algorithm shown in Fig. 3.1 is very sensitive to the initial guess for λ_i, in the sense that many iterations are usually spent until the typical Newton's method quadratic convergence appears. To alleviate this difficulty we propose a procedure to obtain a seemingly reasonable first guess for λ_i, $i = 1, \ldots, m$. It is based upon the heuristic reasoning that the Lagrange multiplier for the problem

$$\begin{aligned} &\min_u \int_{\Omega_i} \left(\alpha|\nabla u| + \frac{1}{2}(u-z)^2\right) dx, \\ &\text{s.t. } \frac{1}{2}\left(\int_{\Omega_i} (u-z)^2\, dx - |\Omega_i|\sigma^2\right) = 0, \end{aligned} \tag{3.13}$$

should be close to the Lagrange multiplier λ_i of the same constraint in problem (3.2), for

$$\int_\Omega |\nabla u|\, dx \approx \sum_i \int_{\Omega_i} |\nabla u|\, dx,$$

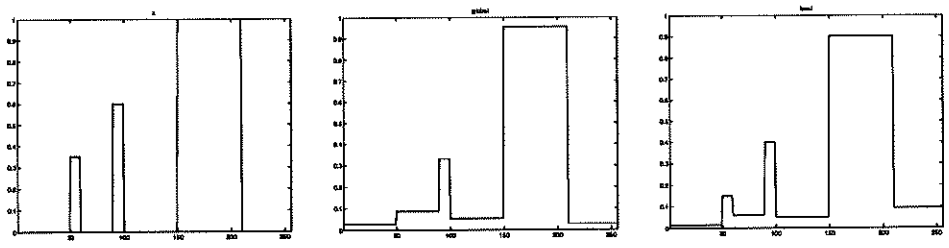
if the regions Ω_i do not overlap too much (in relation to this, it is worth noting that the solution of (3.1) satisfies Neumann boundary conditions on $\partial\Omega_i$, whereas the solution of (3.2) only satisfies them on $\partial\Omega \subsetneq \cup\partial\Omega_i$, if $i > 1$).

The solution of problem (3.1) for $i = 1, \ldots, m$ can be carried out efficiently in parallel, for they are independent. This embarrassingly parallel preprocessing stage has given a decrease by a factor of about 2 in the number of iterations of the algorithm of Fig. 3.1.

To illustrate the potential of this method we show the result of a 1-dimensional example. The signal of Fig. 3.2 (a),whose dynamic range is $[0, 1]$, is used as degraded image with a variance $\sigma^2 = 0.0077$ and a signal to noise ratio of approximately 30dB.

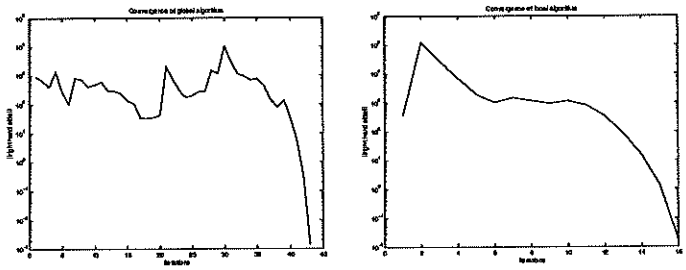
The output of the method (2.3), that is, of the algorithm of Fig. 3.1 with $m = 1$, is displayed in Fig. 3.2 (b), where we see that the smaller of the spikes of the original picture does not appear. The number of iterations has been 43, and the norm of the right hand side of (3.8) are plotted in Fig. 3.3 (a).

With a partition of $\Omega = [0, 1]$ into two equally sized domains with an overlap of 10 pixels, the algorithm of Fig. 3.1 ($m = 2$) gives the result shown in Fig. 3.2 (c). It is noteworthy to point out that the smaller of the spikes is recovered, although with the inherent contrast loss. The number of iterations is 16 and the norm of the right hand side of (3.8) are plotted in Fig. 3.3 (b).



(a) (b) (c)

FIG. 3.2. *(a) Original "noisy" image (z) for algorithm of Fig.3.1, (b) result of the algorithm with $m = 1$ (global constraint), (c) result of the algorithm with $m = 2$ (local constraints)*



(a) (b)

FIG. 3.3. *Plot of the norm of the right hand side of (3.7) versus iterations: (a) global constraint, (b) local constraints ($m = 2$)*

3.5. Final remarks. The dimension of the matrix C of the system (3.8) is the number of pixels in the image, hence it can be very large, but sparse. Iterative methods, as those based on Krylov subspaces, can be very advantageous in this situation. The matrix of (3.8) can be replaced by its symmetrization $\tilde{C}$ without losing the local quadratic convergence. However, in spite of C being symmetric, it is not positive definite, thus it is not clear how to solve the systems as efficiently as with the preconditioned conjugate gradient methods for symmetric positive definite matrices.

An important issue is finding a procedure for subdividing the domain Ω in such a way that the a priori advantages of the locally constrained methods could be exploited. This is an interesting open problem, especially for 2-d images.

4. The Staircase Effect.

4.1. Description. In one dimension, the Total Variation denoising method applied to an affine function degraded by noise will invariably yield a monotone piecewise constant solution. This is in part due to the fact that the TV-norm is not biased against discontinuous nor continuous functions. For instance, the minimum of the

Total Variation among all the functions whose values at 0 and 1 are a and b, respectively, is, precisely, $b - a$ and this minimum is attained by *any* monotone function taking those values.

On the other hand, the functional

$$H^1(u) = \int_\Omega |\nabla u|^2 \, dx,$$

has a strong bias against discontinuous functions and, contrary to the latter case, the problem of minimizing H^1 among all the functions whose values at 0 and 1 are a and b, respectively, has as unique solution the sole linear interpolatory in that set.

The approach we propose is to consider regularizing functionals of the type

$$R(u) = \int_\Omega \Phi(|\nabla u|) \, dx, \tag{4.1}$$

for suitable real functions Φ. For instance, the TV-norm and the H^1 functionals are obtained with $\Phi(x) = x, x^2$, respectively. Consider the functionals (4.1) for $\Phi(x) = x^p$, for $p \in [1, 2]$

$$R(u) = \int_\Omega |\nabla u|^p \, dx. \tag{4.2}$$

The only functional among these for which problem (2.1) admits discontinuous solutions is, precisely, the TV-norm. Another important fact is that the penalization for large gradients increases with p.

Although problem (2.1) with regularizing functionals (4.2) for $p \in (1, 2]$ does not admit discontinuous solutions, this is not such an issue when problem (2.1) is discretized. A first possibility is to try an exponent p close to 1, for then we would expect a behavior close to the TV-norm when restoring edges. In Fig 4.1 (a), (b) we show the result of the denoising method (2.1) for the functionals (4.2) for several p, with original and noisy signals appearing at the bottom of each picture: we see in Fig. 4.1 (a) that sharp edges are obtained for $p = 1, 1.1$ and that those edges appear smeared for $p = 1.5, 2$. Fig 4.1 (b) shows the staircase effect for $p = 1$. This effect is alleviated for $p = 1.1, 1.5$ and, of course, $p = 2$.

4.2. Variable exponents. In view of the results shown in the previous subsection, a more sophisticated possibility would be the use of functions Φ that would behave like x for large values of x, i.e., obtain TV behavior at sharp gradients (edges) and like x^2 at smoother regions, H^1 behavior away from edges. Such a function could be, for instance, a "convex combination" of x and x^2, with variable weight $\alpha(x) \in [0, 1]$:

$$\Phi(x) = \alpha(x)x + (1 - \alpha(x))x^2, \tag{4.3}$$

with $\alpha(x) \to 1$ when $x \to \infty$ and $\alpha(x) \to 0$ when $x \to 0$.

Another possibility is

$$\Phi(x) = x^{p(x)}, \tag{4.4}$$

where $p(x) \in [1, 2]$ is a function that satisfies $\lim_{x \to 0} p(x) = 2$, $\lim_{x \to \infty} p(x) = 1$.

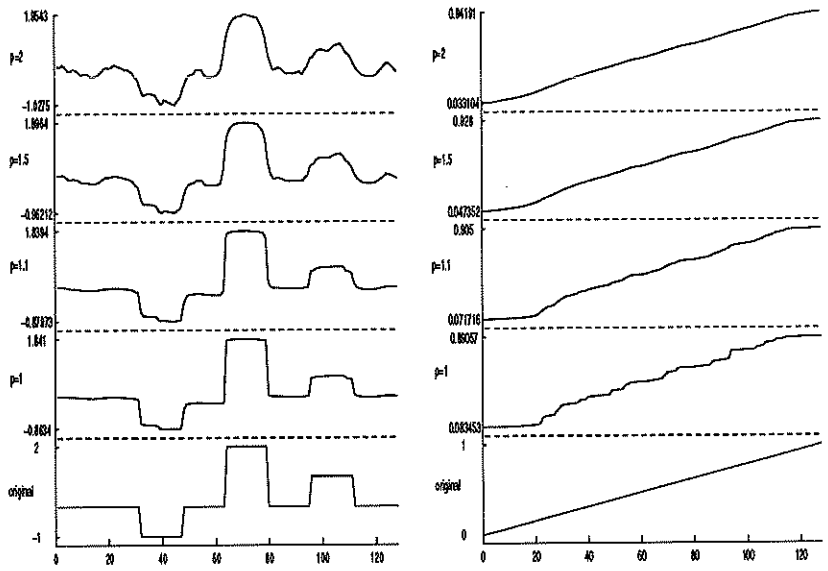


FIG. 4.1. *Denoising algorithm for functionals (4.2).*

4.3. Algorithms. The first and the second variations of the functional (4.1) are:

$$R'(u) = -\nabla \cdot \left(\frac{\Phi'(|\nabla u|)}{|\nabla u|} \nabla u \right) \tag{4.5}$$

$$R''(u)v = -\nabla \cdot \left(\frac{\Phi'(|\nabla u|)}{|\nabla u|} (\nabla v - \frac{(\nabla u, \nabla v)}{|\nabla u|^2} \nabla u) + \Phi''(|\nabla u|) \frac{(\nabla u, \nabla v)}{|\nabla u|^2} \nabla u \right). \tag{4.6}$$

From (4.6) we deduce that:

$$(R''(u)v, v) = \int_\Omega (\frac{\Phi'(|\nabla u|)}{|\nabla u|} (|\nabla v|^2 - \frac{(\nabla u, \nabla v)^2}{|\nabla u|^2}) + \Phi''(|\nabla u|) \frac{(\nabla u, \nabla v)^2}{|\nabla u|^2}) \, dx. \tag{4.7}$$

The Cauchy-Schwartz inequality implies that

$$|\nabla v|^2 - \frac{(\nabla u, \nabla v)^2}{|\nabla u|^2} \geq 0,$$

therefore $\Phi'(x) \geq 0$ and $\Phi''(x) \geq 0$, $x \geq 0$, that is, Φ is an increasing convex function in $[0, \infty)$, are sufficient conditions for the functional R of (4.1) being convex.

A natural assumption on problem (2.1) is that no constant function u satisfies the inequality $\|u - z\|_{\mathcal{L}^2} \leq \sigma^2$, or, equivalently, that $\|\overline{z} - z\|_{\mathcal{L}^2} > \sigma^2$ ($\overline{z}$ denotes the average of z), for then z would be totally dominated by the noise and it would be hopeless to expect the image recovery. If Φ is convex and $\|\overline{z} - z\|_{\mathcal{L}^2} > \sigma^2$, then it can be proven (see Ref. [7]) that problem (2.1) has a unique solution, given by the solution of the Euler-Lagrange equation:

$$0 = -\nabla \cdot \left(\frac{\Phi'(|\nabla u|)}{|\nabla u|} \nabla u \right) + \lambda(u - z). \tag{4.8}$$

A natural approach for the solution of this non-linear PDE is to try a fixed point squeme, similar to the one introduced in Ref. [5], which consists in taking $u^0 = z$ and solving for u^{k+1} the linear PDE

$$-\nabla \cdot \left(\frac{\Phi'(|\nabla u^k|}{|\nabla u^k|} \nabla u^{k+1} \right) + \lambda(u^{k+1} - z) = 0. \tag{4.9}$$

It is shown in Ref. [8] that this scheme is robust (globally convergent) but only linearly convergent.

For these functionals, a faster convergent algorithm, Newton's method, would suffer similar problems to those mentioned in subsection 2.3 and a continuation procedure on the parameter β would be required if Φ is close to x (therefore R close to the TV-norm). A technique as the one described in subsection 3.3, consisting in the introduction of an auxiliary variable w that aims to transform the Euler-Lagrange equations (4.8) into a more *globally linear* system, can be applied here as well.

4.4. Final remarks. An aspect that needs further attention is the choice of the exponent p in (4.2) or (4.4). It is mathematically desirable to have a convex functional R, for then the solution to problem (2.1) has some kind of uniqueness. The functional in (4.2) is convex; however, it is not completely clear how to choose a function Φ such as in (4.3) or (4.4) verifying the sufficient conditions for convexity stated in subsection 4.3.

REFERENCES

[1] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, pp. 259–268, 1992.

[2] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley, New York, 1977.

[3] C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Integral Equations of the First Kind*, Pitman, Boston, 1984.

[4] R. Acar and C. R. Vogel, "Analysis of total variation penalty methods for ill-posed problems," *Inverse Problems* **10**, pp. 1217–1229, 1994.

[5] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM J. Sci. Statist. Comput.* **17**, pp. 227–238, 1996.

[6] 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.

[7] T. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," Tech. Rep. 95-43, UCLA, September 1995.

[8] T. F. Chan and P. Mulet, "On the convergence of the lagged diffusivity fixed point method in image restoration." preprint, 1997.

[9] C. R. Vogel, "A multigrid method for total variation-based image denoising." in Computation and Control IV, conference proceedings to be published by Birkhauser.

[10] D. Dobson and F. Santosa, "Recovery of blocky images from noisy and blured data," Tech. Rep. 94-7, Center for the Mathematics of Waves, University of Delaware, 1994.

[11] G. Golub and C. van Loan, *Matrix computations, 2^{nd} ed.*, The John Hopkins University Press, 1989.

[12] M. E. Oman, "Fast multigrid techniques in total variation-based image reconstruction." to appear in the Preliminary Proceedings of the 1995 Copper Mountain Conference on Multigrid Methods.

[13] A. R. Conn and M. L. Overton, "A primal-dual interior point method for minimizing a sum of euclidean norms." preprint, 1994.