XXI Congreso de Ecuaciones Diferenciales y Aplicaciones
XI Congreso de Matemática Aplicada
Ciudad Real, 21-25 septiembre 2009
(pp. 1–8)

Image denoising using some Total Variation minimization models

S. Amat[1], P. Pedregal[2]

[1] *Dpto. de Matemática Aplicada y Estadística, Univ. Politécnica de Cartagena. Cartagena. E-mail:* sergio.amat@upct.es.
[2] *Dpto. de Matemáticas, Univ. de Castilla-La Mancha, Campus de Ciudad Real. Ciudad Real. E-mail:* pablo.pedregal@uclm.es.

Palabras clave: Total Variation minimization, linearization, Helmholtz's equation, image denoising

Resumen

The Total Variation minimization model has been used in many applications related to image processing. It was introduced as a PDE-based algorithm for edge-preserving noise removal by Rudin, Osher and Fatemi [24]. The associated Euler-Lagrange equation includes the highly nonlinear term $\nabla \cdot (\frac{\nabla u}{|\nabla u|})$. The main difficulty in the numerical approximation of these equations is the "linearization.ºf this term. In this paper some approximation procedures are studied. The motivation of our approaches are based on both theoretical and practical aspects. Our models are closely related with a linearization proposed in [6]. We propose some changes in the approximation procedure that enable us to prove a strong convergence result in H^1. Finally, some numerical improvements, working in BV, are also presented.

1. Introduction

In the last two decades many authors have introduced and analyzed certain tools for the image denoising problem. The standard wavelet thresholding techniques, including linear (i.e. truncating the high frequency coefficients) and nonlinear thresholdings (i.e. retaining large wavelet coefficients), such as hard, soft wavelet thresholding, wavelet shrinkage, are well understood and have been widely used. These techniques are related to the following variational problem:

Given a positive parameter λ, and a noisy image $f(x, y)$, find a function u^* that minimizes, over all possible functions u in a given space Y, the functional

$$||f-u||^2_{L^2(I)} + 2\lambda||u||_Y$$

where

$$||f-u||^2_{L^2(I)} := \int_I |f(x,y)-u(x,y)|^2 dxdy$$

measures the mean-square error between f and u and $||\cdot||_Y$ is the norm in a smoothness space Y.

These techniques involve inter-resolution linear operators. The efficiency of these decompositions is generally limited by the presence of edges. It is well known that wavelet thresholdings may generate oscillations near discontinuities (the well-known Gibbs' phenomenon). Many methods have been proposed to reduce the severity of this problem. Alternatives using directly nonlinear operators can be found in [1]-[2]-[4]-[12]-[14]-[17]-[25].

In this paper, we focus our attention in alternatives which use Partial Differential Equation (PDE) techniques, specifically PDEs derived from variational principles, to reduce the oscillations near discontinuities. A particularly popular technique is the Total Variation (TV) restoration method, introduced by Rudin, Osher and Fatemi in [24]. The TV norm does not penalize discontinuities in u, and thus allows us to recover the edges of the original image.

The associated Euler-Lagrange equations are elliptic highly nonlinear. These equations, with Neumann's boundary conditions for u, are

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

where the positive parameter λ determines the relative importance of the smoothness of u and the quality of the approximation to the given signal f.

If there is not a good estimate of the variance of the noise, then we may consider the unconstrained optimization problem, and the Euler-Lagrange equation becomes

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

We refer to equation (1) as the nonlinear TV model. The classical linear model is governed by

$$-\triangle u+\lambda(u-f)=0, \tag{2}$$

that comes from the Euler-Lagrange equation of the corresponding unconstrained problem with the norm $||\nabla u||^2_{L^2(\Omega)}$.

In [24], the solution of this equation was originally obtained as the steady state solution of the related parabolic time dependent equation. Because of stability constraints, this algorithm can be slowly convergent. To improve the convergence behavior, Vogel and Oman proposed in [9]-[27] the use of a linearization technique for this nonlinear equation. This idea is quite commonly used in many PDE applications. Other numerical approaches can be found in [8]-[10]-[18]. For a review on the topic, see for instance [7]-[11].

In recent years, a different variational framework has been studied [16]-[19]-[20]-[26]. It is based in the decomposition of an image f into $u+v$, with u a piecewise-smooth or *cartoon* component, and v an oscillatory component (texture or noise). Such new models separate better geometric structures from oscillatory ones, but it is difficult to handle them in practice [13]-[15].

After we thought about the approximation procedure we describe here, inspired by material in [21], we discover that very similar ideas had been previously pursued by A. Chambolle and P.L. Lions in [6]. In the present paper, we propose and analyze such a procedure to approximate the Euler-Lagrange equations of the TV denoising model. We have set to ourselves two aims. First to propose a simple PDE-type procedure for the denoising problem; and, secondly, only use the TV-norm in regions near discontinuities. In the oscillatory regions our approach mimics the linear denosing model associated with the norm $||\nabla u||^2_{L^2}$.

Our main result in this context is the strong convergence of a certain approximation sequence. This makes use of interesting ideas developed in [22].

2. The proposed regularization of the TV-model

Roughly speaking, the TV-model tries to solve the variational problem

$$\text{Minimize in } u: \quad \int_\Omega \left[|\nabla u(x)| + \lambda(u-f)^2\right]\,dx$$

where f is the original image. There are three main important issues which make the analysis of this problem really hard:

1. the lack of smoothness at the origin;
2. the lack of strict convexity;
3. the linear growth at infinity.

And yet, (some of) these ingredients are essential to the good performance of the model. So we would like to propose an approximation scheme to the problem that may retain these main aspects but at the same time, it could be easily implementable in practice.

The first issue above can be easily overcome by “rounded off.ª bit the corner at zero. That is no problem. Indeed, we will replace the TV-term $|\mu|$ by

$$\varphi_\epsilon(\mu) = \begin{cases} \frac{1}{2\epsilon}|\mu|^2, & |\mu| \le \epsilon, \\ |\mu| - \epsilon/2, & \epsilon \le |\mu|. \end{cases}$$

This term will cause a filtering behavior of small frequencies which will be very beneficial in practice.

The lack of strict convexity is more important. For this we follow here the strategy described in [21] where one expresses a certain density with problems of lack of convexity or strict convexity through a minimum of strictly convex functions, ideally, quadratics. This is easily done in our setting because

$$\varphi_\epsilon(\mu) = \min_\rho \left[\frac{1}{2\epsilon}|\mu-\rho|^2 + |\rho|\right].$$

For fixed μ, the optimal ρ is precisely given by the (thresholding) formula

$$\rho = \begin{cases} \left(1 - \frac{\epsilon}{|\mu|}\right)\mu, & |\mu| \geq \epsilon, \\ 0, & |\mu| \leq \epsilon. \end{cases}$$

Finally, the linear growth at infinity can also be avoided replacing φ_ϵ by

$$\varphi_{\epsilon,M}(\mu) = \begin{cases} \frac{1}{2\epsilon}|\mu|^2, & |\mu| \leq \epsilon, \\ |\mu| - \epsilon/2, & \epsilon \leq |\mu| \leq M + \epsilon, \\ \frac{1}{2\epsilon}\left(|\mu| - M\right)^2 + M, & M + \epsilon \leq |\mu|, \end{cases}$$

for M large. The linear behavior is kept until size M. From a practical point of view, this is fine as we will not destroy edges in the image with a jump less than a finite but large size M. Notice that

$$\varphi_{\epsilon,M}(\mu) = \min_{|\rho| \leq M} [\frac{1}{2\epsilon}|\mu - \rho|^2 + |\rho|],$$

and the thresholding formula now becomes

$$\rho(\lambda) = \begin{cases} 0, & |\lambda| \leq \epsilon, \\ \left(1 - \frac{\epsilon}{|\lambda|}\right)\lambda, & \epsilon \leq |\lambda| \leq M + \epsilon, \\ \frac{M}{|\lambda|}\lambda, & M + \epsilon \leq |\lambda|. \end{cases}$$

Our main result focuses on proving existence of a solution for the Euler Lagrange problem associated with the variational problem

$$\text{Minimize in } u \in H^1(\Omega): \quad \int_\Omega \left[\varphi_{\epsilon,M}(\nabla u(x)) + \lambda(u(x) - f(x))^2\right] dx,$$

through an approximation scheme that proceeds in each step by calculating the solution of a standard well-posed quadratic problem, and applying the above thresholding formula.

Inspired by these formal observations, we propose an iterative procedure as follows.

Let u^k be a certain "good", smooth approximation to our image f. We assume that it has associated a certain field $\rho^k : \Omega \to \mathbb{R}^2$, which in some sense is a measure of the "details-ecorded in u^k. From this pair (u^k, ρ^k), we find a new iterate by solving the diffusion problem

$$\begin{aligned} -\frac{1}{\epsilon}\mathrm{div}(\nabla u^{k+1} - \rho^k) + \lambda(u^{k+1} - f) &= 0 \text{ in } \Omega, \\ (\nabla u^{k+1} - \rho^k) \cdot n &= 0 \text{ on } \partial\Omega, \end{aligned}$$

and then applying the thresholding mechanism

$$\rho^{k+1} = \frac{M}{|\nabla u^{k+1}|}\nabla u^{k+1}$$

if $|\nabla u^{k+1}| \geq M + \epsilon$,

$$\rho^{k+1} = \left(1 - \frac{\epsilon}{|\nabla u^{k+1}|}\right)\nabla u^{k+1}$$

if $\epsilon \leq |\nabla u^{k+1}| \leq M+\epsilon$, and put $\rho^{k+1}=0$ if $|\nabla u^{k+1}| \leq \epsilon$. As initialization take any ρ^0 with the maximum L^∞-norm equal to M.

In regions near singularities the equation mimics the Euler-Lagrange equation (1) up to size M associated with the nonlinear TV-model. In regions, where only texture is important the equation mimics the equation (2) associated with the linear model. Moreover, in both cases we have to solve only a simple, standard Helmholtz's equation in each iteration.

2.1. Strong convergence in $H^1(\Omega)$

Consider the density

$$\varphi_{\epsilon,M}(\mu) = \begin{cases} \frac{1}{2\epsilon}|\mu|^2, & |\mu| \leq \epsilon, \\ |\mu| - \epsilon/2, & \epsilon \leq |\mu| \leq M+\epsilon, \\ \frac{1}{2\epsilon}\left(|\mu| - M\right)^2 + M, & M+\epsilon \leq |\mu|, \end{cases}$$

and the corresponding variational problem for the image f

$$\text{Minimize in } u \in H^1(\Omega): \quad \int_\Omega [\varphi_{\epsilon,M}(\nabla u(x)) + \frac{\lambda}{2}|u(x) - f(x)|^2]\, dx. \tag{3}$$

Notice that $\varphi_{\epsilon,M}$ is C^1, coercive (quadratic growth), and convex, although it is not strictly convex. Typical variational techniques allow to have the following.

Lemma 1 *There is at least one minimizer $u \in H^1(\Omega)$ which is a weak solution of the problem*

$$-div\,(\nabla\varphi_{\epsilon,M}(\nabla u(x))) + \lambda(u(x) - f(x)) = 0 \text{ in } \Omega, \quad \nabla\varphi_{\epsilon,M}(\nabla u(x)) \cdot n = 0 \text{ on } \partial\Omega, \tag{4}$$

where n is the outer, unit normal to $\partial\Omega$.

Our main theoretical result is the following.

Theorem 1 *The sequence $\{u^k\}$ constructed recursively in the previous section converges **strongly** in $H^1(\Omega)$ to a weak solution u of (4), and hence to a minimizer of (5).*

The main ingredient to claim such strong convergence is a certain compactness property ([22], [23]) of the operator $\mathbf{T}$ taking $u \in H^1(\Omega)$ into the solution v of the problem

$$\begin{aligned} -\frac{1}{\epsilon}\mathrm{div}(\nabla v - \rho(\nabla u)) + \lambda(v - f) &= 0 \text{ in } \Omega, \\ (\nabla v - \rho(\nabla u)) \cdot n &= 0 \text{ on } \partial\Omega, \end{aligned}$$

where $\rho(\lambda)$ is the thresholding operator defined above.

3. Numerical results

The introduced regularization has the same numerical properties as the original linearization proposed in [6].

In order to improve the denoising properties we propose in this section the following new approach:

$$\text{Minimize in } u \in BV(\Omega): \quad \int_\Omega [\varphi_\epsilon(\nabla u(x)) + \frac{\lambda}{2}|u(x) - f(x)|^2]\, dx. \tag{5}$$

where

$$\int_\Omega \varphi_\epsilon(\nabla u(x)) = \begin{cases} \frac{1}{2\epsilon}\int_\Omega |\nabla u(x)|^2, & |\nabla u(x)| \le \epsilon, \\ TV(u) - \epsilon/2|\Omega|, & otherwise. \end{cases}$$

In picture 1, we consider a step function perturbed by noise. We can see the advantage of using this new Total Variation minimization model.

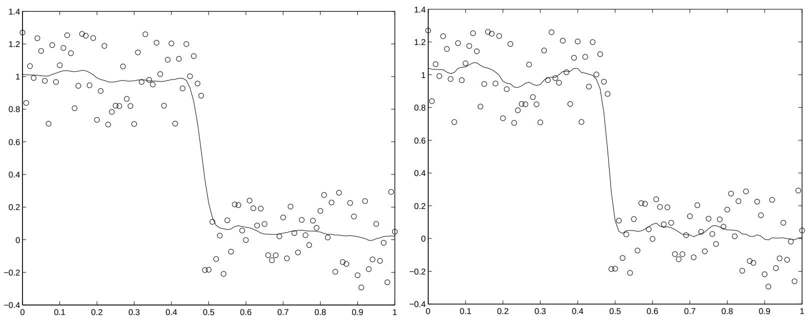

Figura 1: Left BV-minimization, right H^1-minimization. Noise level = 0.3

Finally, we consider an image in 2-d and we present a comparison with the original linear scheme.

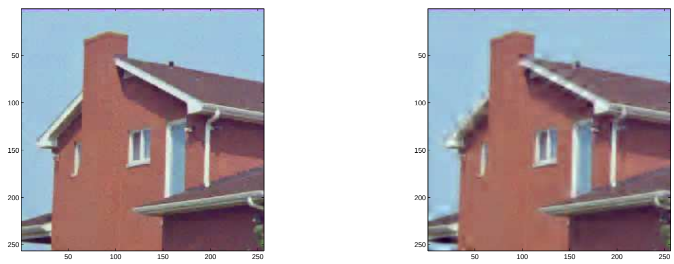

Figura 2: Left, BV-minimization, right, Linear model. Noise level = 15

Acknowledgements

Research supported in part by the Spanish grants MTM2007-62945, PCI08-0084-0424 and 08662/PI/08.

Referencias

[1] S. Amat, R. Donat, J. Liandrat and J.C. Trillo. *Analysis of a New Nonlinear Subdivision Scheme. Applications in Image Processing.*, Foundations of Computational Mathematics, **6** 2, (2000), 193-225.

[2] F. Aràndiga and R. Donat. *Nonlinear Multi-scale Decomposition: The Approach of A.Harten.*, Numerical Algorithms, **23**, (2000), 175-216.

[3] M. Bertalmío, V. Caselles, G. Haro and G. Sapiro. *PDE-based image and surface inpainting. Handbook of mathematical models in computer vision*, 33-61, Springer, New York, 2006.

[4] E.J. Candès and D.L. Donoho. *Recovering edges in ill-posed inverse problems: optimality of curvelet frames.* Dedicated to the memory of Lucien Le Cam. Ann. Statist., **30** (3), (2002), 784-842.

[5] V. Caselles. *Total variation based image denoising and restoration.* International Congress of Mathematicians. Vol. III, 1453-1472, Eur. Math. Soc., Zürich, 2006.

[6] A. Chambolle and P.L. Lions, *Image recovery via total variation minimization and related problems.* Numer. Math. **76** 2, 167-188, (1997).

[7] T.F. Chan, S. Esedoglu, F. Park, and A. Yip. *Total variation image restoration: overview and recent developments.* Handbook of mathematical models in computer vision, 17–31, Springer, New York, 2006.

[8] 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** 6, (1999), 1964-1977.

[9] 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** 2, (1999), 354-367.

[10] T.F. Chan and P. Mulet. *Iterative methods for total variation image restoration.* Iterative methods in scientific computing (Hong Kong, 1995), 359-381, Springer, Singapore, 1997.

[11] T.F. Chan, J. Shen and L. Vese. *Variational PDE models in image processing.* Notices Amer. Math. Soc., **50** 1, (2003), 14-26.

[12] A. Cohen, R. DeVore, P. Petrushev and H. Xu. *Nonlinear approximation and the space* $\mathrm{BV}(R^2)$. Amer. J. Math., **121** 3, (1999),587-628.

[13] I. Daubechies and G. Teschke. *Variational image restoration by means of wavelets: simultaneous decomposition, deblurring, and denoising.* Appl. Comput. Harmon. Anal., **19** (1), (2005), 1-16.

[14] D.L. Donoho. *Ridge functions and orthonormal ridgelets.* J. Approx. Theory, **111** (2), (2001), 143-179.

[15] J.B. Garnett, T.M. Le, Y. Meyer and L.A. Vese. *Image decompositions using bounded variation and generalized homogeneous Besov spaces.* Appl. Comput. Harmon. Anal., **23** 1, (2007), 25-56.

[16] A. Haddad and Y. Meyer. *An improvement of Rudin-Osher-Fatemi model.* Appl. Comput. Harmon. Anal., **22** 3, (2007), 319-334.

[17] E. Le Pennec and S. Mallat. *Image compression with geometrical wavelets.* IEEE Conference on Image Processing(ICIP), Vancouver, September, 2000.

[18] 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** 2, (2000), 387-405.

[19] Y. Meyer. *Oscillating patterns in some nonlinear evolution equations.* Mathematical foundation of turbulent viscous flows, 101-187, Lecture Notes in Math., 1871, Springer, Berlin, 2006.

[20] S. Osher, A. Solé and L. Vese. *Image decomposition and restoration using total variation minimization and the* H^{-1} *norm.* Multiscale Model. Simul., **1** 3, (2003), 349-370.

[21] P. Pedregal. *Variational principles with integrands defined through a minimum*, Non-lin. Anal. Th. Meth. Appl., **66** 12, (2007), 2777-2793.

[22] P. Pedregal. *On a generalization of compact operators, and its application to the existence of critical points without convexity*, (submitted).

[23] P. Pedregal, H. Serrano. *A main example of NFO operator in higher dimension*, (in preparation).

[24] L. Rudin, S. Osher and E. Fatemi. *Nonlinear total variation based noise removal algorithms.* Physica D, **60**, (1992), 259-268.

[25] J-L. Starck, E.J. Candès and D.L. Donoho. *The curvelet transform for image denoising.* IEEE Trans. Image Process, **11** (6), (2002), 670-684.

[26] L.A. Vese and S.J. Osher. *Modeling textures with total variation minimization and oscillating patterns in image processing.* J. Sci. Comput., **19**, 1-3, (2003), 553-572.

[27] C.R. Vogel and M.E. Oman. *Iterative methods for total variation denoising.* SIAM J. Sci. Statist. Comput., **17**, (1996), 227-238.