ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 3, No. 2, 2008, *pp*. 104-110

A Bound Constrained Regularization for Total Variation-Based Image Denoising*

Weiguo Li +, Xiao Sun, Yanhong Wang, Jianbu Wang

School of Mathematics and Computational Sciences, Petroleum University of China
Dongying 257061, Shandong, P. R. China

(*Received October 14, 2007, accepted December 30, 2007*)

Abstract. TV minimizing based PDE models have been proved to be the most effective tool for image restoration. However for many applications, it is desirable to reduce or remove the staircasing effect of a TV solution to restore more realistic contrast, features, geometry and textures. In this paper, we propose and investing a bounded constrained regularization technique to improve the TV solution. A numerical method about the partial differential equations is attempted. Notable improvements are obtained by numerical experiments.

AMS subject classification: 68U10, 65K10

Keywords: Bound constrained, total variation, image denoising, PDE.

1. Introduction

Variational partial differential equations (PDEs) models are extremely studied in recent years for solving image restoration problems. Especially the total variation (TV) based PDEs [28] are extremely effective in edge preserving, as shown in the literature [1-3], [8-12], [15], [20-32].

In recent report [10], various attempts that have been made to improve the TV models are pointed out. The main purpose is to reduce or remove the staicasing effect associated with a TV solution. Sophisticated functional minimization problem have been proposed [10]. These models demand equally sophisticated computational algorithms which do not yet exist.

In this paper, a much simpler modification of the TV model is proposed, which results in reduction of the staircasing effect compared with TV. One idea of §2 is to introduce a BC into the TV model. (Moreover, we shall concentrate on deriving effective nonlocal schemes to solve the proposed model by its associated Euler-Lagrange equations. To this end, we generalize the Primal Dual idea of [11] to this case as discussed in §3. Finally in §4, we demonstrate the improvements obtained from our new BC model.

2. A Modified TV Model

Let Ω denote the image domain (for instance, the computer screen), and is usually a rectangle or a unit square in R^2 for the sake of simplification. We can use $f = f(x; y)$ to represent an observed image. Our image denosing problem is to find original image $u = u(x; y)$ s.t.

$$f = u + \eta \tag{2. 1}$$

Where η is an unknown Gaussian additive noise of variance σ^2.

As is known, the Rudin, Osher and Fatemi (ROF) model [28] proposed to solve u from

$$\inf_{u\in L^2(\Omega)} \int_\Omega |\nabla u| \, dxdy + \lambda \int_\Omega (u - f)^2 \, dxdy. \tag{2. 2}$$

However, such a solution will have the stairing effect e.g. when λ is relatively small, tower-shaped or dome-shaped features will not be visible in a TV solution [28]. We now consider a modified model.

Assume an upper bound, Σ, on the norm of the true solution, u, is known, and enforce regularization by

* This work is finished by the first author in the period of visiting the University of Liverpool
+ email: liwg20022004@yahoo.com.cn

computing a solution of the constrained minimization problem:

$$\begin{cases} \inf\limits_{u\in L^2(\Omega)} \int_\Omega |\nabla u|\, dxdy + \lambda \int_\Omega (u-f)^2 dxdy, \\ subject\ to\ \ ||u|| \le \Sigma \end{cases} \tag{2. 3}$$

or consider the following constrained minimization problem:

$$\begin{cases} \inf\limits_{u\in L^2(\Omega)} \int_\Omega |\nabla u|\, dxdy, \\ subject\ to \int_\Omega u dxdy = \int_\Omega f dxdy,\ \int_\Omega (u-f)^2 dxdy = \sigma\ and\ \ ||u|| \le \Sigma \end{cases} \tag{2. 4}$$

We note that in many applications it may not be possible to find a good upper bound Σ. But the numerical method considered in this paper may be appropriate in these situations. Especially, if a good bound is known, then knowledge of this information should be exploited. Similar considerations have been presented by many authors, such as Gander [18] and Dementiev and Nagy [17] and lots of LSS (Least Squares minimization over a Sphere) or constrained least squares problems. See [16] and its references.

We consider the equivalent unconstrained optimization of (2.3), or (2.4) (the proof of equivalence is similar to [6])

$$\inf_{u\in L^2(\Omega)} \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega (u-f)^2 dx + \mu(\Sigma^2 - ||u||^2), \tag{2. 5}$$

where μ is another multiplier. This model retains the properties of ROF model—recover the edges of the original image and regularize the geometry of level sets without penalizing discontinuities. And it allows us to reduce the error of edges between original and restoration images. Consider $||\cdot||$ as Soblev norm $||\cdot||_{1,2} = \sqrt{|u|^2 + |\nabla u|^2}$. Its corresponding Euler-Lagrange equation is

$$g(u) \equiv -\alpha \nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}} \right) + u - f - \mu_1 (u - \nabla \cdot \nabla u) = 0, \tag{2. 6}$$

Where $\alpha = 1/(2\lambda)$, $\mu_1 = \mu / \lambda$. In this paper, we are concerned the development of efficient algorithms for computing solutions of (2.3), but the theoretical results characterizing properties of the solutions are similar to [6] or [14].

3. A Primal Dual Method

Before we discuss how to solve (2.6) we briefly review the commonly used algorithm for (2.2).

In their original paper [28], Rudin et al. proposed the use of artificial time marching to solve the Euler-Lagrange equations which is equivalent to the steepest descent of the energy function. More precisely, consider the images as a function of space and time and seek the steady state of the equation

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\beta} \right) - 2\lambda (u - f). \tag{3. 1}$$

Here, $|\nabla u|_\beta := \sqrt{|\nabla u|^2 + \beta}$ is a regularized version of $|\nabla u|$ to reduce degeneracies in flat regions where $|\nabla u| \approx 0$. In numerical implementation, an explicit time marching scheme with time step Δt and space step size Δx is used. Under this method, objective value of the ROF model is guaranteed to be decreasing and the solution will tend to the unique minimizer as time increases. However, the convergence is usually slow due to the Courant-Friedrichs-Lewy (CFL) condition, $\Delta t \le c \Delta x^2 / |\nabla u|$ for some constant $c > 0$, imposed on the size of the time step, especially in flat regions where $|\nabla u| \approx 0$.

In a later paper [1], Osher and Marquina considered an improved time marching method for

$$\frac{\partial u}{\partial t} = |\nabla u| \nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\beta} \right) - 2\lambda |\nabla u| (u - f) \tag{3. 2}$$

However, a different CFC condition must also be satisfied.

To completely get rid of CFL conditions, Vogel and Oman proposed in [32] a fixed point iteration scheme (FP) which solves the stationary Euler-Lagrange directly. The Euler-lagrange equation is linearized by lagging the diffusion coefficient and thus the $(i+1)-$th iterate is obtained by solving the sparse linear

equation:

$$\nabla \cdot (\frac{\nabla u^{i+1}}{|\nabla u^{i}|_{\beta}}) - \lambda(u^{i+1} - f) = 0. \tag{3. 3}$$

Although the linear solvers in [31-32] are fast, the NL steps are not generally fast.

In an attempt to improve the Newton menton, Chan, Golub and Mulet[11] proposed the Primal Dual idea (as in mixed FEM) by treating the term $\nabla u / |\nabla u|$ in the primal Euler-Lagrange equation as an independent variable g, leading to the system:

$$\begin{cases} -\nabla \cdot g + \lambda(u - f) = 0, \\ g\,|\nabla u|_{\beta} - \nabla u = 0. \end{cases} \tag{3. 4}$$

The above system of nonlinear equations is solved by Newton's method and quadratic convergence rate is almost always achieved. In the Newton updates, one may combine the two equations to eliminate the fixed point iteration (3.4). Empirically, this primal-dual method is much more robust than applying Newton's method directly to the primal problem in u only.

Now we consider solving differential equation (2.6).

The idea to solve (2.6) is, similar to [11], to introduce new variable w:

$$\omega = \frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}, \tag{3. 5}$$

and replace (2.6) by the following equivalent system of nonlinear PDEs

$$\begin{cases} h_1(u, \omega) \equiv -\alpha \nabla \cdot \omega + (1 - \mu_1)u + \mu_1 \nabla \cdot \nabla u - f = 0, \\ h_2(u, \omega) \equiv \omega\sqrt{|\nabla u|^2 + \beta} - \nabla u = 0. \end{cases} \tag{3. 6}$$

Now we linearize this *(u;w)* system by Newton method.

$$\begin{pmatrix} (1-\mu)I_N + \mu_1 \nabla \cdot \nabla & -\alpha \nabla \cdot \\ -(I_{2N} - \frac{\omega \nabla u^T}{\sqrt{|\nabla u|^2 + \beta}})\nabla & \sqrt{|\nabla u|^2 + \beta} \end{pmatrix} \begin{pmatrix} \delta u \\ \delta \omega \end{pmatrix} = -\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \tag{3. 7}$$

Equation (3.7) can be solved by elimination $\delta\omega$, and solving the result equation for δu

$$[(1-\mu)I_N - \alpha \nabla \cdot (\frac{1}{\sqrt{|\nabla u|^2 + \beta}}(I_{2N} - \frac{\omega \nabla u^T}{\sqrt{|\nabla u|^2 + \beta}})\nabla + \mu_1 \nabla \cdot \nabla)]\delta u = -g(u) \tag{3. 8}$$

After δu is obtained we can compute $\delta\omega$ by the following equation

$$\delta\omega = \frac{1}{\sqrt{|\nabla u|^2 + \beta}}(I_{2N} - \frac{\omega \nabla u^T}{\sqrt{|\nabla u|^2 + \beta}})\nabla \delta u - \omega + \frac{\omega \nabla u^T}{\sqrt{|\nabla u|^2 + \beta}} \tag{3. 9}$$

In our experiments, μ_1 is nonnegative.

Algorithm 1

(i) Choose α and μ and make sure the predictor value u_1 of image u.

(ii) Use the formula (3.8), (3.9) to compute δu and $\delta\omega$ respectively.

(iii) If $\| g \|_{new} / \| g \|_{old} < tol$, stop, else go to (ii).

If we choose a simple form of $\| u \|, \Sigma^2 - \| u \|^2 = \Sigma^2 - u^2$, it is also an improvement. But Soblev norm seem to be more reasonable for our model. The numerical experiments show the validation. If we use the simple Euclid norm, the corresponding Euler-Lagrange equation is changed into

$$g(u) \equiv -\alpha \nabla \cdot (\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}) + u - f - \mu_1 u = 0 \tag{3. 10}$$

and it is easy to solve. Of course, if we set $\| u \| = \nabla u$, it is still an improvement. The corresponding Euler-Lagrange equation is changed into

$$g(u) \equiv -\alpha \nabla \cdot (\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}) + u - f - \mu_1 \nabla \cdot \nabla u = 0 \tag{3. 11}$$

Recently, in [19], a similar model are considered from the view of a bilaterally constrained optimization

problems and super-linear convergence is obtained. But it is different from our model. One is that the parameter $\mu > 0$ in our model but they let $-\mu = \alpha > 0$ in [19]. Another difference is that the norm $\| u \|$ are changeable in our model but the norm is fixed in their model.

4. Numerical Examples

In this section we present results of our denoising algorithms on some images. The noise image f is obtained by adding random noise n to the true image u_0. More precisely, we add random error to each pixel of the true images or synthetic images such that signal to noise ratio $snr = 1$, where

$$snr = \frac{\sum_{\Omega}(u_0(i,j) - \overline{u})^2}{\sum_{\Omega}(n(i,j))^2}$$

and $\overline{u}$ is the mean of the signal u_0 and n is the noise [28].

Our first group images are gray-scale images from 64×64 to 256×256 pixels and Gaussian white noise is added. The original and noisy images are shown below (Figure 1). We take values of 60 and 10^{-2} for α and β, and $tol = 10^{-4}$, $\mu = 8/n$. The processed images from ROF model and BC model are displayed in Figure 1 and the RMSE and SNR for comparison can be found in Table 1. We use here for the RMSE and the SNR the expressions

$$RMSE = \frac{\sqrt{\sum_{i-1}^{n}\sum_{j=1}^{n}(u^*(i,j) - u_0(i,j))^2}}{n^2} \tag{4. 1}$$

$$SNR = 10\log 10 \times \frac{\sum_{i-1}^{n}\sum_{j=1}^{n}[u^*(i,j) - \frac{\sum_{i-1}^{n}\sum_{j=1}^{n}u^*(i,j)}{n^2}]^2}{\sum_{i-1}^{n}\sum_{j=1}^{n}[u^*(i,j) - u_0(i,j) - \frac{\sum_{i-1}^{n}\sum_{j=1}^{n}(u^*(i,j) - u_0(i,j))}{n^2}]^2} \tag{4. 2}$$

where u^* is the computing solution of the true image u_0 [30]. The computing times are nearly same.

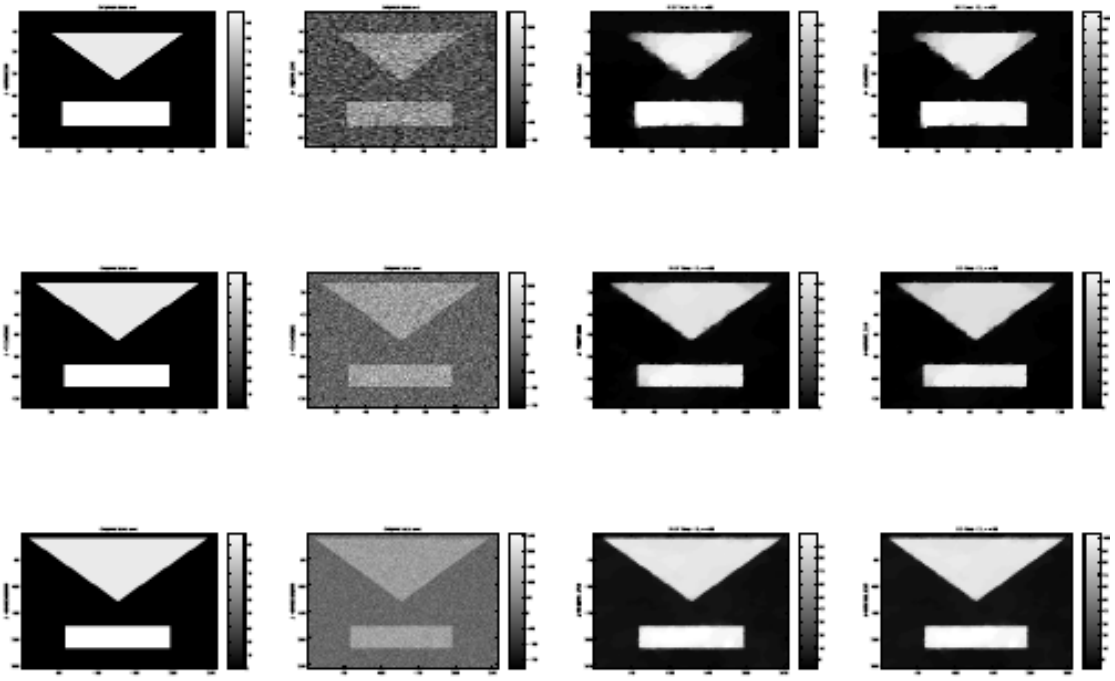

Figure 1 From first row to third row are 64×64, 128×128 and 256×256 separately. $\alpha = 60$, $\beta = 10^{-2}$ and $tol = 10^{-4}$. The far left is original image and the left second is noise image. The right second is from ROF and far right from BC.

We apply both methods to three sizes of the image shown in Figure 1 and compare the performance of ROF and our model for the images. The results for comparison can be found in Table 1 clearly.

Next we test our method on a more realistic image, the satellite image with SNR=4 and SNR=1. Both images are 256×256 in size and have range[0,255]. $\beta=10^{-2}$ in two cases, $\alpha=50$. We choose (3.11) as our model and RMSE and SNR still use (4.1) and (4.2). The results can be observed from Figure 2 and Table 2.

n	RMSE			SNR		
	ROF	BC	improvement(%)	ROF	BC	improvement(%)
64	0.13246	0.09942	24.940	328.73	969.05	194.78
128	0.04697	0.03953	15.836	901.88	1573.3	74.446
256	0.01763	0.01638	7.0886	1848.4	2397.0	29.679

Table 1 The comparison of the denoising results in various size by ROF and BC.

snr	RMSE			SNR		
	ROF	BC	improvement(%)	ROF	BC	improvement(%)
4 ($\mu=48/n$)	0.04031	0.03703	8.1390	489.84	605.81	23.675
1 ($\mu=38/n$)	0.04386	0.04149	5.3847	424.33	489.49	15.357

Table 2. The comparison of the denoising results for Satellite by ROF and BC.

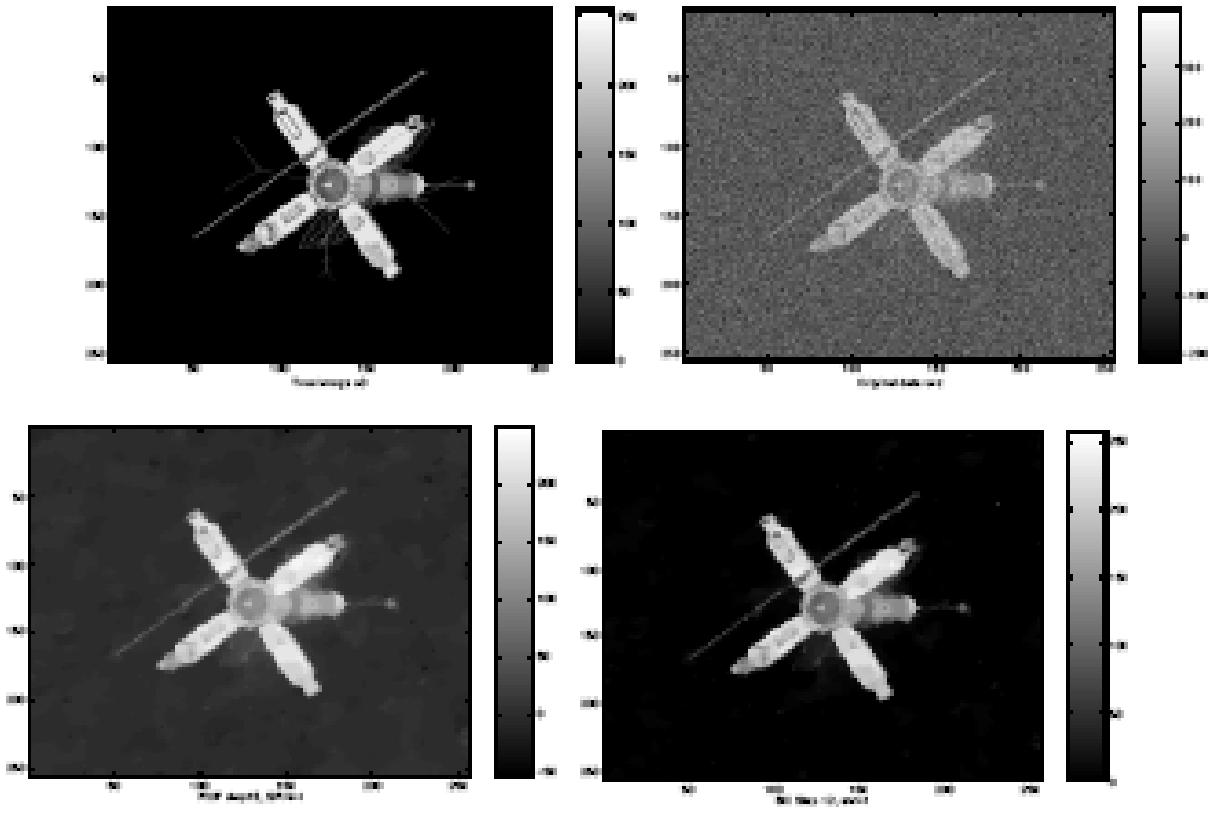

Figure 2 (Satellite) Top left is original image and top right is noise image.

$\alpha=50, \beta=10^{-2}$ and $tol=10^{-4}$. Bottom left is from ROF model and bottom right from BC model.

5. Conclusions

The standard ROF model worked well in denoising problems but we found that a modified ROF model can improve its performance in reducing its stairing effect. With our recommended Bound Constrained Regularization technique we observe that it gives over 5% improvement for RSEM and we can get more than

15% for SNR. So our method is efficient and robust. Here we applied some projection methods that is elementary. We suggest that the continuation Newton method can be used directly if small β is desired (such as $\beta = 10^{-4}$), or the truncated versions of Newton (continuation) algorithm based on the conjugate gradient method with incomplete Cholesky as preconditioner. As an alternative solution method, we may consider the multi-grid method with the use of the Krylov acceleration procedure made the convergence fast [13].

6. References

[1] Antonio Marquina, Stanley Osher, Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for Nonlinear Deblurring and Noise Removal. SIAM Journal on Scientific Computing. 22: 387-405, 2000.

[2] P. Blomgren, Total Variation Methods for Restoration of Vector Valued Images. PhD thesis, Dept. of Math., Univ. of Califonia, Los Angeles, June 1998.

[3] P. Blomgren, T. Chan and P. Mulet, Extensions to total variation denoising. Advanced Signal Processing Algorithms, Architectures and Implementations (Proc. SPIE vol 3162) Vol 7, ed F. Luk (Bellingham, WA: SPIE Optical Engineering Press), 1997.

[4] J. Carter. Dual Methods for Total Variation-based Image Retoration. PhD thesis, UCLA, Los Angeles, CA, 2001.

[5] A. Chambolle. An Algorithm for Total Variation Minimization and Applications. J. Math. Imaging Vision, 20: 89-97, 2004.

[6] A. Chambolle and P. Lions. Image Recovery via Total Variation Minimization and Related Problems. Numer. Math., 76: 167-188, 1997.

[7] R. H. Chan, T. F. Chan, and H. M. Zhou, Advanced Signal Processing Algorithms, in Proceedings of the International Society of Photo-Optical Instrumentation Engeers, F. T. Luk, ed., SPIE, 1995, pp. 314-325.

[8] T. Chan, S. Esedoglu, and F. Park, A Fourth Order Dual Method for Staircase Reduction in Texture Extraction and Image Restoration Problems. UCLA CAM Reports, 05-28.

[9] T. Chan, S. Esedoglu and F. Park, Image Decomposition Combining Staircase Reduction and Texture Extraction. UCLA CAM Reports, 05-18.

[10] T. Chan, S. Esedoglu, F. Park and A. Yip, Recent Developments in Total Variation Image Restoration. UCLA Reports, 05-01.

[11] T. Chan, G. Golub, and P. Mulet, A Nonlinear Primal-Dual Method for Total Variation-based Image Restoration. SIAM J. Sci. Comp., 20: 1964-1977, 1999. 8

[12] T. Chan, A. Marquina and P. Mulet, Second Order Differential Functionals in Total Variation-Based Image Restoration, UCLA CAM Reports, 98-35.

[13] Q. Chang and I-Lian Chern, Acceleration Method for Total Variation-Based Image Denoising. SIAM J. Sci. Comput., 25: 982-994, 2003.

[14] G. Chavent and K. Kunisch, Regularation of linear least squares problems by total bounded variation. ESAIM Control Optim. Cal. Var., 2:359-376, 1997.

[15] Y. Chen, S. Levine and M. Rao, Functional with p(x)-growth in image restoration, SIAM Journal on Applied Mathematics, (to appear).

[16] A. J. Cox, Stability of algrothms for solving weighted and constrained least squares problems, PhD thesis, University of Manchester, Manchester, England, Oct., 1997.

[17] D. Dementiev and J. G. Nagy. Bound constrained regularization for ill-posed problems. Technical Report TR-2003-0007-A, Mathematics and Computer Science, Emory University.

[18] W. Gander. Least squares with a quadratric constraint. Numer. Math., 36:291-307, 1981.

[19] M. Hintermuller and K. Kunisch, Total bounded variation regulation as a bilaterally constrained optimization problems. SIAM J. Appl. Math., 64:1311-1333,2004.

[20] K. Joo and S. Kim, PDE-based image restoration, I: Anti-staircasing and anti-diffusion. Technical Report, 2003-07, Dept. of maths, University of Kentucky, Lexington, USA.

[21] T. Karkkainen and K. Majava, SAC-methods for image restoration recent advances in Applied and Theoretical Mathematics. ed N. Mastorakis (Greece: World Scitific and Engineering Society). 2000.

[22] S. Kim and K. Joo, Efficient and reliable procedures for PDE-based denoising. In: Proceedings of Hawaii International Conference on Statistics, Mathematics and Related Fields, pp. 1126-1141, 2004.

[23] Suk-Ho Lee and Jin Keun Seo, Noise removal with Gauss curvature-driven diffusion. Image Processing, IEEE Transactions. 14: 904- 909, 2005.

[24] S. Levine, An Adaptive Variational Model for Image Decomposition, to appear in Energy Minimization Methods

in Computer Vision and Pattern Recognition (2005). 9

[25] S. Levine, J. Stanich, and Y. Chen, Image Restoration via Nonstandard Diffusion. Technical Reports: 04-01,Dept. of Mathematics and Computer Science, Duquesne University.

[26] M. Lysaker, A. Lundervold, Xue-cheng Tai, Noise Removal Using Fourth-order Partial Differential Equations with Applications to Medical Magnetic Resonance Images in Space and Time. IEEE Trans. Image Processing, 12: 1579-1590.

[27] M. Lysaker and X.-C. Tai, Iterative image restoration combining total variation minimization and a second order functional. International Journal of Computer Vision (to appear) 2005.

[28] L. Rudin, S. Osher, and E. Fatemi. Nonlinear Total Variation Based Noised Removal Algorithms. Physica D, 60:259-268, 1992.

[29] D. Strong and T. Chan. Edge-preserving and Scale-dependent Properties of Total Variation Regularization. Inv. Probl., 19: 165-187, 2003.

[30] L. Vese and S. Osher, Image denoising and decomposition with total variation minimization and oscillatory functions. J. Mathematical Imaging and Vision, 20: 7-18, 2004.

[31] C. R. Vogel, Computational Methods for Inverse Problems, SIAM, Philadelphia, 2002.

[32] C. Vogel and M. Oman. Iteration Methods for Total Variation Denoising. SIAM J. Sci. Comp., 17:227-238, 1996. 10.