

A WELL-POSED MULTISCALE REGULARIZATION SCHEME FOR DIGITAL IMAGE DENOISING

V.B. SURYA PRASATH

Department of Mathematics
University of Coimbra, Apartado 3008, 3001–454 Coimbra, Portugal
e-mail: surya.iit@gmail.com

We propose an edge adaptive digital image denoising and restoration scheme based on space dependent regularization. Traditional gradient based schemes use an edge map computed from gradients alone to drive the regularization. This may lead to the oversmoothing of the input image, and noise along edges can be amplified. To avoid these drawbacks, we make use of a multiscale descriptor given by a contextual edge detector obtained from local variances. Using a smooth transition from the computed edges, the proposed scheme removes noise in flat regions and preserves edges without oscillations. By incorporating a space dependent adaptive regularization parameter, image smoothing is driven along probable edges and not across them. The well-posedness of the corresponding minimization problem is proved in the space of functions of bounded variation. The corresponding gradient descent scheme is implemented and further numerical results illustrate the advantages of using the adaptive parameter in the regularization scheme. Compared with similar edge preserving regularization schemes, the proposed adaptive weight based scheme provides a better multiscale edge map, which in turn produces better restoration.

Keywords: image restoration, edge-preserving, regularization, normalized local variance, adaptive parameter.

1. Introduction

Digital image restoration is an important task in image processing. It involves denoising and smoothing the input image which is often noisy and blurred. Due to the ill-posed nature of this inverse problem, regularization schemes are used to find an approximate solution effectively. To avoid the excessive smoothness of the classical quadratic penalty terms, edge preserving regularization methods (Aubert and Kornprobst, 2006) have been used in the past. For example, Total Variation (TV) based regularization (Rudin *et al.*, 1992) in the space of bounded variation (Giusti, 1984) gained popularity thanks to its good edge preservation qualities. Among other regularization based approaches for image denoising we should mention global models (You and Kaveh, 1999), and recently nonlocal models (Gilboa and Osher, 2007) are proven to be very effective in edge preserving noise removal. Half quadratic minimization schemes (Charbonnier *et al.*, 1997), which alternatively minimize for the estimated image and an edge variable, were initiated by Geman and Geman (1984) and further studied by many others (see the work of Aubert and Kornprobst (2006) for a re-

view). Due to the application of controlled regularization along edges, noise can still remain in these discontinuities (Santitissadeekorn and Bollt, 1996). Moreover, capturing the multiscale nature of images remains an open issue in these schemes.

To alleviate these drawbacks, in this paper we use an adaptive weight which is tuned by a contextual discontinuity detector based on variances computed over a window. Incorporating this adaptive regularization parameter into edge preserving regularization yields a scheme which is capable of removing noise and preserving multiscale boundaries. The well-posedness of the corresponding minimization problem is proved in the space of functions of bounded variation. The scheme proposed here is similar to the adaptive TV scheme studied by Chen and Wunderli (2002) (see also Prasath and Singh, 2010b) and a scheme proposed by Douiri *et al.* (2007) for diffusive optical tomography. The main differences are as follows. We prove the well-posedness of the scheme under the linear growth condition for the regularization function involved here. Hence, the adaptive TV scheme of Chen and Wunderli (2002) can be considered a special case. Moreover, in the work of Chen and Wunderli (2002), the adaptive

regularization parameter is defined via smoothed gradients, which inherits the drawbacks of the TV scheme near edges, namely noisy oscillations. The scheme proposed by Douiri *et al.* (2007) uses an incomplete edge prior as an adaptive weight. Motivated by the success of Douiri *et al.* (2007) in combining a well designed regularization function with an edge prior based weight, in this paper we propose a multiscale regularization for denoising digital images. In noisy images, we do not know *a priori* the edge map of the original image. Thus, to get an estimate for the edge map, we utilize a variance based local approach. Moreover, the scheme we propose here has flexibility in terms of tuning the contextual discontinuities and choosing the regularization function, which can be tuned according to the imaging modality.

The rest of the paper is organized as follows. In Section 2 we introduce the proposed regularization scheme and prove its well-posedness. Section 3 presents a comparison of numerical results with previous schemes. Finally, Section 4 concludes the paper.

2. Edge adaptive weighted regularization

The image restoration issue can be posed as an inverse problem in the following way. Find u given a noisy input image f which may be blurred, i.e.,

$$f = B(u) + n,$$

where n is additive Gaussian noise with variance σ_n and B is a linear map representing the blur, usually a convolution operator. To solve this ill-posed problem, the edge preserving regularization method can be used (Aubert and Kornprobst, 2006). A general formulation can be written in the form of a minimization problem:

$$\inf_{u \in BV(\Omega)} \mathcal{F}(u) := \int_{\Omega} |f(x) - B(u(x))|^2 dx + \alpha \int_{\Omega} \phi(|\nabla u|) dx$$

where $\Omega \subset \mathbb{R}^2$ is the image domain (rectangle), ϕ is the regularization function defined on the gradient image $|\nabla u|$, and $\alpha > 0$. Note that, for $\phi(s) = s^2$, we get the classical Tikhonov regularization, which is known to give excessively smooth solutions. Half Quadratic (HQ) minimization introduced by Geman and Geman (1984) uses an auxiliary variable b which corresponds to edge strength,

$$b = \frac{\phi'(|\nabla u|)}{2|\nabla u|}, \tag{1}$$

and alternatively minimizes the following functional (Charbonnier *et al.*, 1997) with respect to u

and b :

$$\mathcal{F}(u, b) := \int_{\Omega} |f(x) - B(u(x))|^2 dx + \alpha \int_{\Omega} (b(|\nabla u|) dx + \psi(b)) dx,$$

see the works of Charbonnier *et al.* (1997) as well as Aubert and Kornprobst (2006) for more details regarding its well-posedness and derivation details. The implementation of the HQ algorithm is done via the Euler-Lagrange equation for \mathcal{F} and can be written as follows:

1. Solve for u :

$$(u^{t+1} - f) - \alpha \operatorname{div} (b^t \nabla u^{t+1}) = 0.$$

2. Solve for b :

$$b^{t+1} = \frac{\phi'(|\nabla u^{t+1}|)}{2|\nabla u^{t+1}|}.$$

Starting with the initial image, i.e., at time $t = 0$, $u^0 = f$, the above alternating minimization scheme is carried out for both the image u and the edge variable b .

It has been noted (Santitissadeekorn and Boltt, 1996), in the HQ scheme noise along edges can remain, since the variable b is computed from gradients which are prone to noise. One of the main reasons for this is the use of gradients alone to capture edge pixels, which may lead to false classifying noisy pixels as edge pixels. Computing gradients from noisy measurements is an ill-posed operation and, moreover, the gradient image is too localized to give a skeletal edge map, see Fig. 1(c). The auxiliary variable b (Eqn. (1)) in the HQ scheme is based on this gradient edge map $|\nabla u|$ alone and hence gives spurious edges, see Fig. 1(d). To avoid this, recently Santitissadeekorn and Boltt (1996) modified the HQ scheme by including an edge respecting term in the image minimization step. This approach yielded better results than the HQ scheme in terms of noise suppression along edges. Nevertheless, capturing the multiscale nature of images using gradients alone is not a feasible approach, as we have seen in this example. Thus, we note that, instead of trying to remedy the HQ scheme which is based solely on gradients, a better way is to use an explicit edge map computed from a multiscale descriptor.

2.1. Proposed scheme. To avoid the drawbacks of gradient based schemes and to keep the regularization property of the ϕ function formulation, we propose to make the parameter α adaptive. The parameter α balances the influence of the fidelity term and the *a priori* term in the regularization, and is an important factor to obtain good numerical results. Instead of using gradient based edge indicator functions as in the work of Strong and Chan

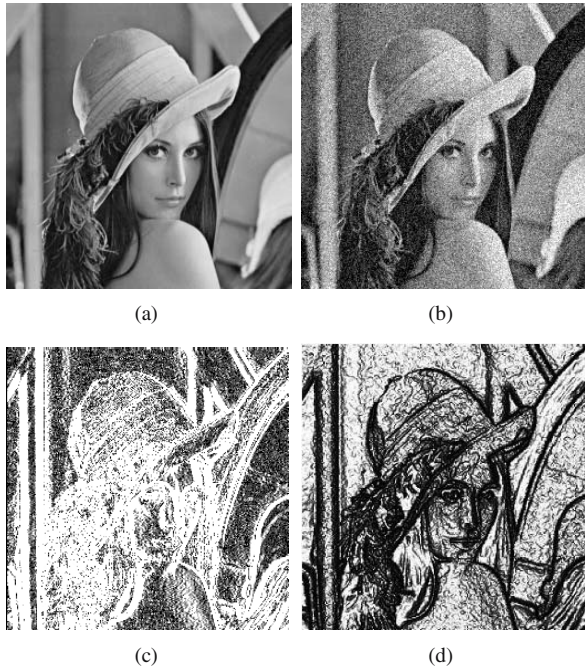


Fig. 1. Original image (a), noisy image with noise level $\sigma_n = 20$ (b), $|\nabla u|$ (c), variable b in (1) of the HQ algorithm (d).

(1996), we make use of the local information around each pixel (Szeliski *et al.*, 2008).

For a pixel $x \in \Omega$, consider the local variance computed in a neighborhood $\mathcal{N}_r(x)$ with radius r as follows:

$$\mathcal{V}_r(x) = \frac{1}{|\mathcal{N}_r(x)|} \sum_{y \in \mathcal{N}_r(x)} (f(y) - m_r(x))^2,$$

where $f(y)$ denotes the intensity (gray-scale) value at pixel $y \in \Omega$ and $m_r(x)$ is the mean value of the pixels situated in the neighborhood $\mathcal{N}_r(x)$. For the Gaussian noise case, $r = 1$ or $r = 2$ is well suited to capture the local variance around the pixel under consideration.

Let the minimum and maximum values of $\mathcal{V}_r(x)$ found in $\mathcal{N}_r(x)$ be $\mathcal{V}_r^{\min}(x)$, $\mathcal{V}_r^{\max}(x)$, respectively, and consider the normalized local variance of the pixel $x \in \Omega$:

$$\bar{\mathcal{V}}_r(x) = \frac{\mathcal{V}_r(x) - \mathcal{V}_r^{\min}(x)^2}{\mathcal{V}_r^{\max}(x)^2 - \mathcal{V}_r^{\min}(x)^2}.$$

Then use a simple threshold function to avoid too small values interfering with the contextual edges

$$\Theta(\bar{\mathcal{V}}_r^2(x), \theta) = \begin{cases} 0 & \text{if } \bar{\mathcal{V}}_r^2(x) < \theta, \\ \bar{\mathcal{V}}_r^2(x) & \text{if } \bar{\mathcal{V}}_r^2(x) \geq \theta. \end{cases}$$

Now, set the inhomogeneity term α using the variance controlled discontinuity measure Θ by the nonlinear function:

$$\alpha(x) = \exp\left(-\frac{1}{\delta}\Theta(\bar{\mathcal{V}}_r^2(x), \theta)\right). \quad (2)$$

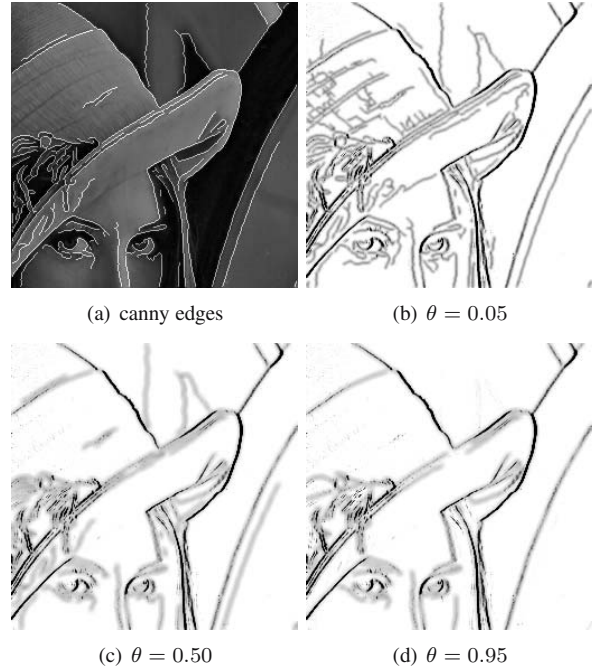


Fig. 2. Edge-map comparison showing the influence of the thresholding parameter θ in (2): edge map computed from the Canny edge detector applied to the original image (drawn on top of the image for better visualization) (a), adaptive parameter α given in Eqn. (2) for different values of θ (b)–(d).

The parameters θ and δ control the extent to which the variance induced discontinuities should be preserved. Typically they are in the range of θ from 0.05 to 0.95, and $\delta = 10$ can be fixed for a low to medium level of noise ($\sigma_n \leq 20$). Figure 2 shows an example of the effect of θ in a close-up shot of the noisy *Lena* image from Fig. 1(b). As the value of θ increases, we lose more small scale edge features, and it can be safely set at a medium level $\theta = 0.50$ for most of the noisy images (compare with edges from the original image given in Fig. 2(a). This image is computed using the MATLAB command `edge(f, 'canny', 0.2)`). We utilize the Canny edge detector for comparison as it is proven to be efficient among others (Basu, 2002).

Thus, the scheme we propose is of the following energy minimization form:

$$\inf_{u \in BV(\Omega)} \mathcal{E}(u) := \int_{\Omega} |f(x) - B(u(x))|^2 dx + \int_{\Omega} \alpha(x)\phi(|\nabla u|) dx. \quad (3)$$

Other choices for α are also possible (Stronh and Chan, 1996; 2010a; Prasath and Singh, 2010b), and here we chose the normalized local variance for its efficiency and success rate with respect to edge detection compared

with gradient based edge indicator functions as proposed by Strong and Chan (1996). Note that $\alpha \in C^\infty(\Omega)$ and $0 < \alpha \leq 1$.

2.2. Well-posedness. We now provide the well-posedness of the proposed scheme (3) in $BV(\Omega)$ using direct methods of the calculus of variations. We recall definitions and preliminaries about the space of functions of bounded variation $BV(\Omega)$. These results can be found e.g., in the work of Giusti (1984). We assume the following:

- H1: $B : L^2(\Omega) \rightarrow L^2(\Omega)$ is a linear continuous operator such that $B \cdot 1 \neq 0$.
- H2: $\phi : [0, \infty) \rightarrow [0, \infty)$ is a convex, even and nondecreasing function with $\phi(0) = 0$.
- H3: ϕ satisfies the linear growth condition, i.e., there exist $a \geq 0$ and $b > 0$ such that $as - b \leq \phi(s) \leq as + b$ for all $s \in \mathbb{R}$.

Condition H1 implies that the blur operator does not annihilate constant functions, which also guarantees that the functional $\mathcal{E}(u)$ is coercive in $BV(\Omega)$. A canonical choice of ϕ satisfying H2–H3 is the well-known TV function $\phi(|\nabla u|) = |\nabla u|$ studied by Rudin *et al.* (1992). We use it in our numerical experiments due to its edge preservation properties.

Lemma 1. *Let B be a linear operator satisfying H1, and the regularization function ϕ satisfy Assumptions H2–H3. Then the functional (3) is lower semicontinuous, i.e., if $\{u_n\} \rightarrow u$ in $L^1(\Omega)$, then*

$$\mathcal{E}(u) \leq \liminf_{n \rightarrow \infty} \mathcal{E}(u_n).$$

Proof. Consider $w \in C_0^1(\Omega)$ to be such that $|w(x)| \leq \alpha(x)$ for all $x \in \Omega$. Then

$$\begin{aligned} \int_{\Omega} u \operatorname{div} w \, dx &= \lim_{n \rightarrow \infty} \int_{\Omega} u_n \operatorname{div} w \, dx \\ &\leq \liminf_{n \rightarrow \infty} \int_{\Omega} \alpha |\nabla u_n| \, dx, \end{aligned}$$

where the inequality follows from the definition of the total variation of u . Now, taking the supremum over w gives

$$\int_{\Omega} \alpha \phi(\nabla u) \, dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} \alpha \phi(\nabla u_n) \, dx.$$

Since the L^2 norm is lower semicontinuous, we also have, for the fidelity term,

$$\begin{aligned} \int_{\Omega} |f(x) - B(u(x))|^2 \, dx \\ \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |f(x) - B(u_n(x))|^2 \, dx. \end{aligned}$$

Thus the proposed functional $\mathcal{E}(u)$ is lower semicontinuous in $BV(\Omega)$. ■

Theorem 1. *Let B and ϕ satisfy Assumptions H1–H3 and $f \in BV(\Omega) \cap L^2(\Omega)$. Then the proposed minimization problem*

$$\inf_{u \in BV(\Omega)} \mathcal{E}(u)$$

has a solution in $BV(\Omega) \cap L^2(\Omega)$. Moreover, the solution is unique if ϕ is strictly convex or B is injective.

Proof. The functional $\mathcal{E}(u)$ is coercive and convex, and by Lemma 1 the functional is lower semicontinuous in $BV(\Omega) \cap L^2(\Omega)$. Thus, by Theorem 3.2.2 of Aubert and Kornprobst (2006), the corresponding minimization problem (3) has a solution in $BV(\Omega)$ and it is unique if the regularization function ϕ is strictly convex or B is injective. ■

We note that the Euler–Lagrange equation of our functional (3) is given by a Partial Differential Equation (PDE) of the form

$$BB^*u - \operatorname{div} \left(\alpha(x) \frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right) = B^*f, \quad (4)$$

where B^* is the adjoint of B . Notice that this is a symbolic notation. There is a slight abuse of notation in writing (4) as an equality. Since the sub-differential of the functional given in (3) is multi-valued, the equality in (4) would be better written as $B^*f \in BB^*u - \operatorname{div}(\alpha \partial \phi(|\nabla u|))$. In spite of this, we use the equation as (4), understanding that the equality holds for an element of the sub-differential ∂ function. As a consequence of the convexity of the energy functional $\mathcal{E}(u)$, we also have the following theorem about the well-posedness of the evolution problem in Eqn. (4).

Theorem 2. *Let*

$$\mathcal{A}(u) = (BB^*u - B^*f) - \operatorname{div} \left(\alpha(x) \frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right).$$

Then the following time dependent evolution problem of (4) has a unique solution $u(t) : [0, \infty) \rightarrow L^2(\Omega)$:

$$\begin{cases} 0 \in \frac{\partial u}{\partial t} + \mathcal{A}(u) & \text{in } (0, \infty) \times \Omega, \\ u(0, x) = f(x) & \text{for } x \in \Omega. \end{cases}$$

Proof. Since \mathcal{A} is the derivative of a convex functional, it is a maximal monotone operator. The assertion follows from nonlinear semigroup theory, see, e.g., Theorem 3.1 of Brezis (1973). ■

2.3. Related schemes. Using the HQ algorithm given in Section 2.1, Santitissadeekorn and Boltt (1996) add an extra term in the image estimation step to remove noise along edges. This is carried out by the addition of a divergence based term in the corresponding gradient descent form in solving for u ,

$$\begin{aligned} (u^{t+1} - f) - \lambda \operatorname{div} (b^t \nabla u^{t+1}) \\ - \alpha \operatorname{div} ((1 - b^t) \nabla f) = 0. \end{aligned}$$

Note that the last term uses the noisy input image f , unlike the updated version employed in our adaptive parameter α induced PDE (4) above. Moreover, our edge adaptive scheme can be written in the HQ formulation with the edge variable b given by

$$b = \alpha(x) \frac{\phi'(|\nabla u|)}{|\nabla u|}. \tag{5}$$

Thus, we see that it uses a gradient weighted edge map via α and gives a better edge structure, see Fig. 4(a).

Strong and Chan (1996) as well as Chen and Wunderli (2002) studied the adaptive TV case for the denoising case $B = I$, i.e.,

$$\begin{aligned} & \inf_{u \in BV(\Omega)} \mathcal{F}(u) \\ &= \int_{\Omega} |f(x) - u(x)|^2 dx + \int_{\Omega} \alpha(x) |\nabla u| dx, \end{aligned}$$

with $\alpha(x) = (1 + K |G_{\sigma} \star \nabla f|^2)^{-1}$, where $K > 0$ is a parameter. This particular choice of α includes the same problems associated with other schemes which make use of $|\nabla u|$. Although pre-smoothing with the Gaussian G_{σ} makes the gradient image of f less oscillatory, the corresponding results suffer from the same artefacts associated with gradient based regularization schemes, such as the lack of multiscale edges capturing and localization of weak edges.

3. Numerical results

The parameter θ of the adaptive function (2) is set to $\theta = 0.50$ and the neighborhood size $r = 2$ in the following experiments. Increasing the value of θ results in the reduction of high frequency edges and hence the contrast, see Fig. 2(b)–(d). The scale parameter δ is fixed at $\delta = 10$ in what follows and it works well for images corrupted noise with variance of up to $\sigma_n \leq 20$. If the noise level is high ($\sigma_n > 20$), the parameter δ needs to be set higher to avoid noisy oscillations around edges. All the images are normalized to be in the range $[0, 1]$. We implement the gradient descent version of the Euler–Lagrange equation (4) of the proposed convex functional (3). We employ a standard finite difference scheme as follows.

Let h be the grid size and U_{ij}^t the intensity value $u(i, j)$ at iteration t . Instead of the classical explicit scheme, which severely restricts the step size, we make use of an unconditionally stable semi-implicit scheme. In 1D with matrix–vector notation, it reads as

$$U^{t+1} = [1 - \tau A(U^t)]^{-1} U^t,$$

where τ is the time step, $A(U^t) = [a_{ij}(U^t)]$, and

$$a_{ij}(U^t) := \begin{cases} \frac{b_i^t + b_j^t}{2h^2} & \text{if } j \in \mathcal{N}_i, \\ - \sum_{k \in \mathcal{N}_i} \frac{b_i^t + b_k^t}{2h^2} & \text{if } j = i, \\ 0 & \text{otherwise,} \end{cases}$$

with $b_i = \alpha_i g_i$, where g_i is the discrete version of the diffusion function $g = \phi'(|\nabla u|) / |\nabla u|$, see Eqn. (4). For n D images, the semi-implicit scheme is written as

$$U^{t+1} = \left[1 - \tau \sum_{l=1}^n A_l(U^t) \right]^{-1} U^t. \tag{6}$$

The matrix $A_l = (a_{ijl})_{ij}$ corresponds to derivatives along the l -th coordinate axis.

We consider the denoising case first, i.e., $B = I$ in (3). The first example, given in Fig. 3, shows the restoration of the noisy *Lena* image from Fig. 1(b). The following schemes are compared: Perona and Malik’s (PM) scheme (Perona and Malik, 1990), Total Variation (TV) regularization (Rudin *et al.*, 1992), Adaptive TV (ATV) (Chen and Wunderli, 2002), the HQ method (Charbonnier *et al.*, 1997), and its modification proposed by Santitissadeekorn and Bollt (1996) (SB). To compare the image quantitatively and to gauge the parameters associated with each of the schemes compared here, we used the following error metric (Santitissadeekorn and Bollt, 1996):

$$E(u) = \frac{\|u - f_0\|}{\|f_0\|} + \left\| \frac{|\nabla u|}{\|\nabla u\|} - \frac{|\nabla f_0|}{\|\nabla f_0\|} \right\|,$$

where f_0 is the true noiseless image and $\|\cdot\|$ denotes the l^2 -norm. This error metric evaluates the error in data fidelity as well as the steepness error of edges.

Further, the usual image error metrics such as the Peak Signal to Noise Ratio (PSNR) and the Mean Absolute Error (MAE) were calculated, and comparison results for the USC-SIPI database and other standard test images corrupted by a higher noise level $\sigma_n = 35$ are given in Table 1. The PSNR and MAE are given by the following formula:

$$\text{PSNR} := 20 \log_{10} \left(\frac{u_{\max}}{\sqrt{\text{MSE}}} \right) [\text{dB}],$$

where $\text{MSE} = (mn)^{-1} \sum \sum (u - f_0)$, $m \times n$ denotes the image size, u_{\max} denotes the maximum value, e.g., in 8-bit images $u_{\max} = 255$, and

$$\text{MAE} := \text{Mean} |f_0 - u|.$$

The PM scheme (Fig. 3(a)) removes small scale edges, which leads to diffusion leakage and reduction in



Fig. 3. Denoised images of various schemes: PM scheme with $K = 30, t = 50$ error $E = 0.78$ (a), TV scheme with $\epsilon = 10^{-4}, t = 100$, error $E = 0.85$ (b), adaptive TV scheme with $\sigma = 2, K = 10$, error $E = 0.80$ (c), half-quadratic method with $\phi(s) = (1 + (s/K)^2)^{-1/2}, K = 10, t = 120$, error $E = 0.45$ (d), SB method with $\lambda = 10, \alpha = 50$, error $E = 0.22$ (e), proposed scheme (3) with $\delta = 10, \theta = 0.5$, error $E = 0.25$ (f).

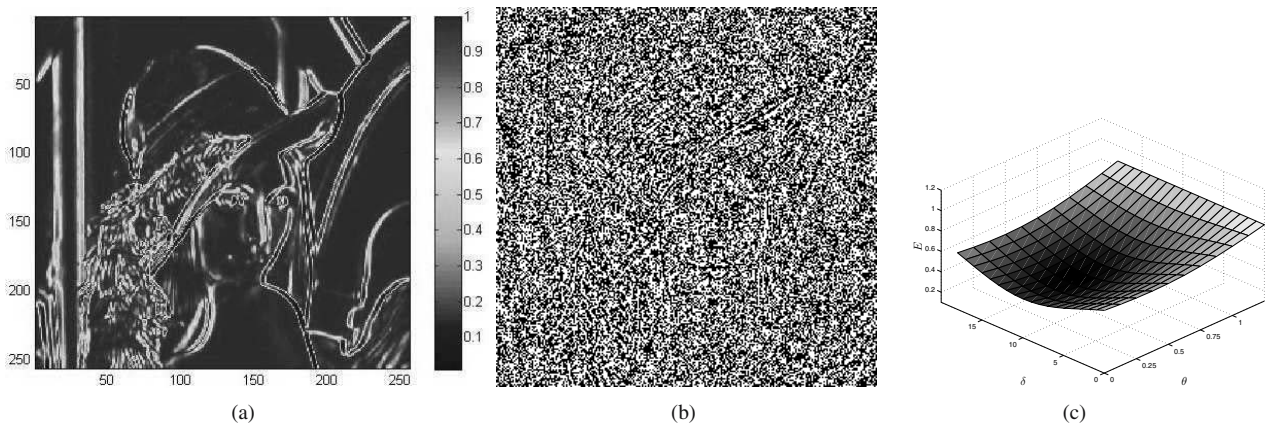


Fig. 4. Adaptive variance based edge indicator function $\alpha(x)\phi'(|\nabla u|)/|\nabla u|$ produces a weighted edge map (a), method noise $(|u - f|)$ for our result given in Fig. 3(f), multiplied by a factor of 2 for visualization (b), error metric $E(u)$ as a function of parameters θ and δ (c).

Table 1. Comparison of different error metrics for USC-SIPI standard test images with a high noise level $\sigma_n = 35$. PSNR [dB]/MAE are given for each scheme. A higher value for the PSNR and a lower value for the MAE indicate better results. The best result in the PSNR is indicated in boldface for each image.

Image/method	PM	TV	ATV	HQ	SB	Our scheme
Lena	19.85/16.22	19.02/20.92	20.31/16.99	22.67/13.87	24.55/11.47	26.20 /8.66
Cameraman	21.97/10.93	24.29/11.52	25.60/08.19	25.40/07.82	25.06/07.87	26.67 /07.87
House	21.99/11.39	28.53/06.59	28.56/05.92	27.77/06.28	28.70 /06.71	28.34/10.17
Peppers	20.52/16.06	24.92/11.11	27.17/07.94	27.75 /08.32	27.45/08.10	27.46/07.78
Tree	19.55/17.69	24.04/12.25	25.12/09.80	24.29/10.30	25.67/10.87	27.07 /10.45
Barbara	19.85/18.25	22.62/13.78	22.74/12.20	22.38/12.53	21.91/12.21	23.25 /13.58
Couple	21.69/14.19	24.89/10.98	26.59/08.20	25.50/10.69	25.97/08.53	27.10 /08.42
Man	21.13/15.14	24.92/11.08	26.67/08.44	25.45/10.87	26.22/08.58	26.75 /08.17
Boat	21.23/14.34	24.93/11.04	26.67/08.24	25.67/08.54	25.97/110.87	26.80 /08.59
Tank	23.86/10.25	25.73/10.24	28.75/07.08	28.46 /07.17	28.21/07.00	27.88/07.54
F-16	21.63/12.24	25.25/10.60	27.10/07.09	27.07/07.05	27.24/07.83	27.69 /08.28
Car	20.67/15.46	24.66/11.42	25.09/08.87	25.33/09.26	25.79/09.35	26.42 /08.80
Bridge	19.85/18.99	23.75/12.78	24.24/11.62	23.46/12.66	23.54/12.40	24.62 /11.20
Baboon	19.14/20.53	22.30 /14.75	21.96/14.47	21.11/15.87	21.77/15.10	21.67/14.85
Hill	22.04/13.67	25.30/10.70	27.56/07.84	27.04/08.14	27.49/08.25	27.80 /08.18

the overall contrast. The TV scheme (Fig. 3(b)) yields edge preservation but contains staircasing artefacts in flat regions. Adaptive TV produces a better result (Fig. 3(c)) than the TV scheme but contains noise along edges. Similar artefacts appear in the HQ and SB methods (Figs. 3(d) and (e), respectively), removes noise along edges and improves the result of HQ. As can be seen from comparing our scheme's output (Fig. 3(f)), we obtain a cleaner image with strong edges and without many of the aforementioned artefacts. Figure 4(a) shows the weighted edge map computed from the normalized local variance based parameter (2) used in our scheme (3). Note that the details from different scales are included, and a better edge map of the input image is obtained. Figure 4(b) shows the method noise ($|u - f|$) for the result shown in Fig. 3(f). Figure 4(c) shows the effect of the parameters θ and δ against the error metric E . The multiscale edge controlling is taken care by the parameter θ and the parameter δ is similar to the contrast parameter in traditional Perona–Malik type diffusion PDEs.

We next show the result of using our scheme for a Gaussian blurred image (Fig. 5(a)) with additional noise level $\sigma_n = 10$,

$$B(u(x)) = (G_{\sigma_b} \star u)(x) = \int_{\Omega} \exp\left(-\frac{|x-y|^2}{2\sigma_b}\right) u(y) dy.$$

We took the blurring kernel of size $\sigma_b = 2$, and Fig. 5(b) shows the result of our scheme with a TV regularization function. The scheme can be used to remove small scale features along with noise, which can be useful in denoising scanned documents. Figure 5(c) shows a denoising example where we increased the variance threshold parameter to $\theta = 0.80$ to avoid the small vertical ink streaks.

Finally, Fig. 6 shows a small line (in the 50-th row) taken across the *Lena* image from the original image given in Fig. 1(a) and its restoration by our scheme given in Fig. 3(f). As can be seen, our scheme results in strong smoothing in flat regions with edge preservation. Note the removal of wiggleness, which corresponds to small scale textures when compared with the original signal. Hence, our scheme based on normalized local variance removes textures and other small scale features along with noise. To preserve such textures, we have to include a texture detector (e.g., Gabor filters) based measure into the value of α which defines a further improvement of the proposed scheme.

The scheme we proposed here is general in the sense that many of the existing edge preserving regularization and PDE schemes can be used along with (2). Computations are done on a desktop computer with a Pentium IV 2.14 GHz processor and MATLAB 7.4 used for visualization. For a 512×512 image it takes about 1 min to denoise (100 iterations) and this can be further reduced if we make use of advanced numerical techniques such as operator splitting. The main computational bottleneck is the evaluation of the adaptive parameter $\alpha(x)$ which uses a window based approach for each pixel. The computational time is significantly higher than in the traditional HQ schemes (Charbonnier *et al.*, 1997; Santitissadeekorn and Bollt, 1996).

4. Conclusion

In this paper we considered edge preserving regularization methods in the context of image restoration and denoising. Half quadratic schemes separate the edge detection part using an auxiliary variable and can leave noise along edges. By using an explicit edge map adaptively com-

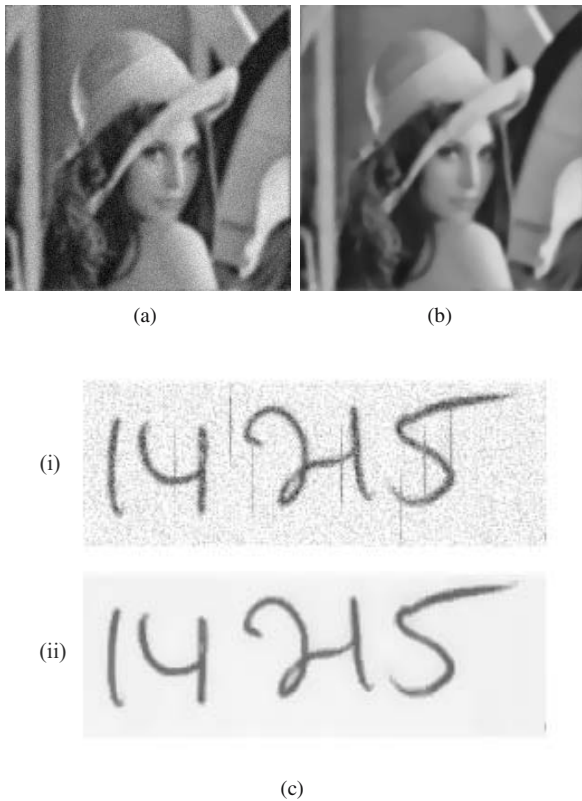


Fig. 5. Blurred and noisy image (a), proposed scheme applied to the blurred and noise image (b), denoising document images: (i) input image contains noise as well as some vertical streaks of ink, (ii) denoising result using our scheme (c).

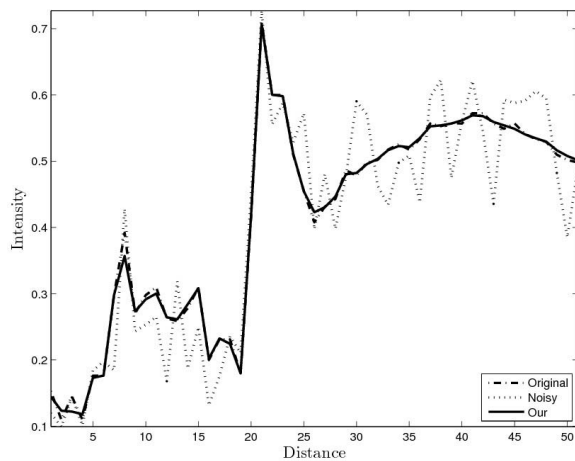


Fig. 6. 1D slice taken across Fig. 3(f) and from the noisy image.

puted from a multiscale edge detector, the regularization scheme proposed here restores the images better and removes spurious oscillations. The well-posedness of the minimization scheme is proved in the space of functions of bounded variation. Related schemes are compared in

terms of detail preservation, and numerical results indicate that addition of a pixel-wise adaptive parameter improves denoising and restoration. Future works in this direction include extending the scheme to color images, using other contextual edge detectors in the edge variable process, and using a median based adaptive parameter (Lukac and Smolka, 2003) to handle multiplicative noise.

Acknowledgment

This work was supported by an FCTUC/CMUC fellowship at the University of Coimbra, Portugal, and partially by the FCT and UT Austin|Portugal project: UTAustin/MAT/0009/2008. The author would like to thank the four reviewers for their valuable comments and suggestions that helped him to improve both the technical content and the presentation quality of this paper.

References

Aubert, G. and Kornprobst, P. (2006). *Mathematical Problems in Image Processing: Partial Differential Equation and Calculus of Variations*, Springer-Verlag, New York, NY.

Basu, M. (2002). Gaussian based edge-detection methods—A survey, *IEEE Transactions on System, Man, and Cybernetics, Part C: Applications and Reviews* **32**(3): 252–260.

Brezis, H. (1973). *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Company, Amsterdam.

Charbonnier, P., Blanc-Feraud, L., Aubert, G. and Barlaud, M. (1997). Deterministic edge-preserving regularization in computed imaging, *IEEE Transactions on Image Processing* **6**(2): 298–311.

Chen, Y. and Wunderli, T. (2002). Adaptive total variation for image restoration in BV space, *Journal of Mathematical Analysis and Applications* **272**(3): 117–137.

Douiri, A., Schweiger, M., Riley, J. and Arridge, S.R. (2007). Anisotropic diffusion regularization methods for diffuse optical tomography using edge prior information, *Measurement Science and Technology* **18**(1): 87–95.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6): 721–741.

Gilboa, G. and Osher, S. (2007). Nonlocal linear image regularization and supervised segmentation, *SIAM Journal on Multiscale Modelling and Simulation* **6**(2): 595–630.

Giusti, F. (1984). *Minimal Surfaces and Functions of Bounded Variation*, Birkhauser, Basel.

Lukac, R. and Smolka, B. (2003). Application of the adaptive center-weighted vector median framework for the enhancement of cDNA microarray images, *International Journal of Applied Mathematics and Computer Science* **13**(3): 369–383.

- Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(7): 629–639.
- Prasath, V.B.S. and Singh, A. (2010a). A hybrid convex variational model for image restoration, *Applied Mathematics and Computation* **215**(10): 3655–3664.
- Prasath, V.B.S. and Singh, A. (2010b). Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising, *Journal of Applied Mathematics*, Article ID 763847.
- Rudin, L., Osher, S. and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms, *Physica D* **60** (1–4): 259–268.
- Santitissadeekorn, N. and Bollt, E.M. (1996). Image edge respecting denoising with edge denoising by a designer non-isotropic structure tensor method, *Computational Methods in Applied Mathematics* **9**(3): 309–318.
- Strong, D.M. and Chan, T.F. (1996). Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing, *Technical Report 96-46*, Computational and Applied Mathematics, University of California Los Angeles, CA, <ftp://ftp.math.ucla.edu/pub/camreport/cam96-46.ps.gz>.
- Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tapen, M. and Rother, C. (2008). A comparative study of energy minimization methods for Markov random fields with smoothness based priors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(6): 1068–1080.
- You, Y.-L. and Kaveh, M. (1999). Blind image restoration by anisotropic regularization, *IEEE Transactions on Image Processing* **8**(3): 396–407.



V.B. Surya Prasath is currently a postdoctoral fellow at the Department of Mathematics at the University of Coimbra, Portugal. He obtained his Ph.D. degree from the Indian Institute of Technology Madras, India, in 2009. His research interests include mathematical image processing, color image processing, anisotropic diffusion models, variational and regularization methods for inverse problems. He was granted the US Nair Young Statistician Award by the Indian Society for Probability and Statistics in 2008. He won the Best Paper Award at the 6th Indian Conference on *Computer Vision, Graphics and Image Processing (ICVGIP'08)* held in Bhubaneswar, India.

Received: 4 November 2010

Revised: 19 May 2011

Re-revised: 12 June 2011