Image restoration via nonlocal Mumford-Shah regularizers

Miyoun Jung and Luminita Vese

University of California, Los Angeles
Department of Mathematics
Los Angeles, CA 90095-1555, USA
gomtaeng@math.ucla.edu, lvese@math.ucla.edu

Abstract. We wish to recover an image corrupted by blur and Gaussian or impulse noise in a variational framework. We use two data-fidelity terms depending on the type of noise, and several local and nonlocal regularizers. Inspired by Buades, Coll, Morel and Gilboa-Osher, we propose nonlocal versions of the Ambrosio-Tortorelli and Shah approximations to the Mumford-Shah regularizing functionals, with applications to image deblurring in the presence of noise. In the case of impulse noise model, we propose a necessary preprocessing step for the computation of the weight function. Experimental results show that these nonlocal MS regularizers yield better results than the corresponding local ones (proposed for deblurring by Bar et al.) for both noise models; moreover, these perform better than the nonlocal total variation in the presence of impulse noise. Characterization of minimizers is also given.

Keywords: image deblurring, image denoising, variational models, local and nonlocal regularizers, Gaussian noise, impulse noise.

1 Introduction

We consider the problem of the restoring an image blurred and then contaminated by Gaussian or impulse noise. Let $f, u : \Omega \to \mathbb{R}$ be image intensity functions, where $\Omega \subset \mathbb{R}^2$ is open and bounded. The standard linear degradation model that links the blurry-noisy image with the original image is given by $f = k * u + n$, where f is the observed degraded image, k is (known) space-invariant blurring kernel, u is the ideal image we want to recover, and n is the random additive noise independent of u. We approach the image restoration problem within the variational framework. Specifically, we consider the following minimization problem in the unknown u,

$$\inf_u \{ \Phi(f - k * u) + \Psi(\nabla u) \}$$

where $\Phi(\cdot)$ is a functional representing the data-fidelity term, and Ψ is the regularization functional which enforces a smoothness constraint on u, depending on its gradient ∇u.

First, two different fidelity terms can be considered based on the type of noise; in the case of Gaussian noise model, the L^2-fidelity term led by the maximum likelihood estimation is commonly used:

$$\Phi(f - K * u) = \int_\Omega |f - k * u|^2 dx.$$

However, the quadratic data fidelity term considers the impulse noise, which might be caused by bit errors in transmissions or wrong pixels, as an outlier. So, for the impulse noise model, the L^1-fidelity term is more appropriate, due to its robustness of removing outlier effects [7], [13]:

$$\Phi(f - K * u) = \int_\Omega |f - k * u| dx.$$

Image deblurring-denoising is an inverse problem, which is known to be ill-posed due to either the non-uniqueness of the solution or the numerical instability of the inversion of the blurring operator. The regularization term alleviates this problem by reflecting some a-priori properties. Several regularization terms were suggested in the literature, including [31], [27], [28], [25]. Here, we consider the total variation regularization [27], [28] and two approximations of the Mumford-Shah regularization [25], denoted MSH^1 and MSTV, proposed by Ambrosio-Tortorelli [2] and Shah [29] respectively and recently used for image deblurring in the presence of Gaussian and impulse noise by Bar et al [4], [5], [6], [7]. These traditional regularization terms are based on local image operators, which denoise and preserve edges very well, but may induce loss of fine structures like textures during the restoration process.

Recently, Buades et al [10] introduced the nonlocal means filter, which produces excellent denoising results. Kindermann et al [22] and Gilboa and Osher [19, 20] formulated the variational framework of NL-means by proposing nonlocal regularizing functionals and the nonlocal operators such as the nonlocal gradient and divergence. Lou et al [23] used the nonlocal total variation (NL/TV) of Gilboa-Osher in image deblurring in the presence of Gaussian noise with a preprocessing step for the computation of the weight function.

We propose here nonlocal versions of the approximated Mumford-Shah and Ambrosio-Tortorelli regularizing functionals, called NL/MSH^1 and NL/MSTV, by applying the nonlocal operators proposed by Gilboa-Osher to MSH^1 and MSTV respectively, for image restoration in the presence of blur and Gaussian or impulse noise. In addition, for the impulse noise model, we propose to use a preprocessed image to compute the weights w (the weights w defined in the NL-means filter are more appropriate for the additive Gaussian noise case). We note that the interesting parallel work [11] also proposed NL/MSH^1 regularizer for segmentation and denoising in the presence of Gaussian noise, but not for deblurring, nor for the impulse noise case.

Local regularizers In this section, we recall several regularization terms. The first one is the Mumford-Shah regularizing functional [25] which gives preference

to piecewise smooth images. The MS regularizer depends on the image u as well as on the set of edges $K \subset \Omega$, and is given by

$$\Psi^{MS}(u, K) = \beta \int_{\Omega \setminus K} |\nabla u|^2 dx + \alpha \int_K d\mathcal{H}^1$$

where $\mathcal{H}^1$ is the one-dimensional Hausdorff measure. The first term enforces smoothness of u everywhere except on the edge set K, and the second one minimizes the total length of edges. It is difficult to minimize in practice the non-convex MS functional.

Ambrosio and Tortorelli [2] approximated this functional by a sequence of regular functionals Ψ_ϵ using the Γ-convergence. The edge set K is represented by a smooth auxiliary function v. Thus we have an approximation to Ψ^{MS} as [2]

$$\Psi_\epsilon^{MSH^1}(u, v) = \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \Big(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \Big) dx,$$

where $0 \leq v(x) \leq 1$ represents the edges: $v(x) \approx 0$ if $x \in K$ and $v(x) \approx 1$ otherwise, ϵ is a small positive constant, α, β are positive weights. A minimizer $u = u_\epsilon$ of $\Psi_\epsilon^{MSH^1}$ approaches a minimizer u of Ψ^{MS} as $\epsilon \to 0$.

An alternative approach is the total variation [27] proposed by Rudin, Osher, and Fatemi, called TV regularizer:

$$\Psi^{TV}(u) = \int_\Omega |Du| \approx \int_\Omega |\nabla u| dx.$$

Because of its benefit of preserving edges (which have high gradient levels) and convexity, TV has been widely used in image restoration.

Shah [29] suggested a modified version of the approximation (1) to the MS functional by replacing the norm square of $|\nabla u|$ by the norm in the first term:

$$\Psi_\epsilon^{MSTV}(u, v) = \beta \int_\Omega v^2 |\nabla u| dx + \alpha \int_\Omega \Big(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \Big) dx.$$

This functional $\Gamma-$converges to the other Ψ^{MSTV} functional [1]:

$$\Psi^{MSTV}(u) = \beta \int_{\Omega \setminus K} |\nabla u| dx + \alpha \int_K \frac{|u^+ - u^-|}{1 + |u^+ - u^-|} d\mathcal{H}^1 + |D_c u|(\Omega)$$

where u^+ and u^- denote the image values on two sides of the jump set $K = J_u$ of u, and $D_c u$ is the Cantor part of the measure-valued derivative Du. Note that the non-convex term $\frac{|u^+ - u^-|}{1+|u^+ - u^-|}$ is similar with the prior regularization by Geman-Reynolds [18]. We observe that this regularizing functional is similar to the total variation of $u \in BV(\Omega)$ that can be written as

$$\int_\Omega |Du| = \int_{\Omega \setminus K} |\nabla u| dx + \int_K |u^+ - u^-| d\mathcal{H}^1 + |D_c u|(\Omega).$$

By comparing the second terms, we see that the MSTV regularizer does not penalize the jump part as much as the TV regularizer. In this paper, we consider the TV regularizer Ψ^{TV}, the MSH^1 regularizer $\Psi_\epsilon^{MSH^1}$, and the MSTV regularizer Ψ_ϵ^{MSTV}.

Nonlocal methods Nonlocal methods in image processing have been explored in many papers because they are well adapted to texture denoising while the standard denoising models working with local image information seem to consider texture as noise, which results in losing textures. Nonlocal methods are generalized from the neighborhood filters (e.g. Yaroslavsky filter, [32]) and patch based methods. The idea of neighborhood filter is to restore a pixel by averaging the values of neighboring pixels with a similar grey level value.

Buades et al. [10] generalized this idea by applying the patch-based methods proposed for texture synthesis [17], which is a famous neighborhood filter called nonlocal-means (or NL-means):

$$NLu(x) = \frac{1}{C(x)} \int_{\Omega} e^{-\frac{d_a(u(x),u(y))}{h^2}} u(y)dy$$
$$d_a(u(x), u(y)) = \int_R G_a(t)|u(x+t) - u(y+t)|^2 dt$$

where d_a is the patch distance, G_a is the Gaussian kernel with standard deviation a determining the patch size, $C(x) = \int_{\Omega} e^{-\frac{d_a(u(x),u(y))}{h^2}} dy$ is the normalization factor, and h is the filtering parameter which corresponds to the noise level; usually we set it to be the standard deviation of the noise. The NL-means not only compares the grey level at a single point but the geometrical configuration in a whole neighborhood (patch). Thus, to denoise a pixel, it is better to average the nearby pixels with similar structures rather than just with similar intensities.

In practice, we use the search window $\Omega_w = \{y \in \Omega : |y - x| \leq r\}$ instead of Ω (semi-local) and the weight function at $(x, y) \in \Omega \times \Omega$ depending on a function $u : \Omega \to \mathbb{R}$

$$w(x, y) = exp\Big(-\frac{d_a(u(x), u(y))}{h^2}\Big).$$

The weight function $w(x, y)$ gives the similarity of image features between two pixels x and y, which is normally computed based on the blurry noisy image f. Recently, for image deblurring with Gaussian noise, Lou et al [23] used a preprocessed image obtained by applying the Wiener filter, instead of f, to construct w. In our work, only for the impulse noise model, we propose a different preprocessing step and evaluate w using the preprocessed image.

Nonlocal regularizers In the variational framework, Kindermann et al [22] formulated the neighborhood filters and NL-means filters as nonlocal regularizing functionals which have the general form:

$$\Psi(u) = \int_{\Omega\times\Omega} \phi\left(\frac{|u(x) - u(y)|^2}{h^2}\right) w(|x - y|)dxdy$$

where $w(|x - y|)$ is a positive weight function. But, these functionals generally are not convex. Thus, Gilboa and Osher [19] formalized the convex nonlocal

functional inspired from graph theory:

$$\Psi(u) = \frac{1}{2}\int_{\Omega\times\Omega} \phi(|u(x) - u(y)|)w(x,y)dxdy$$

where ϕ is convex, positive, $\phi(0) = 0$ and further assume that $\lim_{s\to\infty}\phi(s)/s = 1$ for the L^1 type, and the weight function $w(x,y)$ is nonnegative and symmetric. Note that $\phi(s) = s^2$ is analogous to the standard H^1 semi-norm while $\phi(s) = s$ is analogous to the TV semi-norm.

Moreover, based on the gradient and divergence definitions on graphs in the context of machine learning, Gilboa and Osher [20] derived the nonlocal operators. Let $u : \Omega \to \mathbb{R}$ be a function, and $w : \Omega \times \Omega \to \mathbb{R}$ is a weight function assumed to be nonnegative and symmetric. The nonlocal gradient $\nabla_w u : \Omega\times\Omega \to \mathbb{R}$ is defined as the vector $(\nabla_w u)(x,y) := (u(y)-u(x))\sqrt{w(x,y)}$. Hence, the norm of the nonlocal gradient of u at $x \in \Omega$ is defined as

$$|\nabla_w u|(x) = \sqrt{\int_\Omega (u(y) - u(x))^2 w(x,y)dy}.$$

The nonlocal divergence $div_w \overrightarrow{v} : \Omega \to \mathbb{R}$ of the vector $\overrightarrow{v} : \Omega \times \Omega \to \mathbb{R}$ is defined as the adjoint of the nonlocal gradient

$$(div_w \overrightarrow{v})(x) := \int_\Omega (v(x,y) - v(y,x))\sqrt{w(x,y)}dy.$$

Based on these nonlocal operators, they introduced nonlocal regularizing functionals of the general form

$$\Psi(u) = \int_\Omega \phi(|\nabla_w u|^2)dx$$

where $\phi(s)$ is a positive function, convex in $\sqrt{s}$ with $\phi(0) = 0$. By taking $\phi(s) = \sqrt{s}$, they proposed the nonlocal TV regularizer (NL/TV) as

$$\Psi^{NL/TV}(u) = \int_\Omega |\nabla_w u|dx = \int_\Omega \sqrt{\int_\Omega (u(y) - u(x))^2 w(x,y)dy}dx$$

where this functional corresponds in the local two dimensional case to $\Psi^{TV}(u) = \int_\Omega |\nabla u|dx$. Inspired by these ideas, we propose in the next section nonlocal versions of Ambrosio-Tortorelli and Shah approximations to the MS regularizer for image denoising-deblurring. This is also continuation of work by Bar et al. [4], [5], [6], [7], first to propose the use of Mumford-Shah-like approximations to image deblurring.

2 Description of proposed models

We propose the following nonlocal approximated Mumford-Shah and Ambrosio-Tortorelli regularizing functionals (NL/MS) by applying the nonlocal operators

to the approximations of the MS regularizer,

$$\Psi^{NL/MS}(u,v) = \beta \int_\Omega v^2 \phi(|\nabla_w u|^2) dx + \alpha \int_\Omega \Big(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \Big) dx$$

where $\phi(s) = s$ and $\phi(s) = \sqrt{s}$ correspond to the nonlocal version of MSH[1] and MSTV regularizers, called here NL/MSH[1] and NL/MSTV, respectively:

$$\Psi^{NL/MSH^1}(u,v) = \beta \int_\Omega v^2 |\nabla_w u|^2 dx + \alpha \int_\Omega \Big(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \Big) dx$$

$$\Psi^{NL/MSTV}(u,v) = \beta \int_\Omega v^2 |\nabla_w u| dx + \alpha \int_\Omega \Big(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \Big) dx.$$

In addition, we use these nonlocal regularizers to deblur images in the presence of Gaussian or impulse noise. Thus, by incorporating the proper fidelity term depending on the noise model, we design two types of total energies as

Gaussian noise model:

$$E^G(u,v) = \int_\Omega (f - k * u)^2 dx + \Psi^{NL/MS}(u,v),$$

Impulse noise model:

$$E^{Im}(u,v) = \int_\Omega |f - k * u| dx + \Psi^{NL/MS}(u,v).$$

Minimizing these functionals in u and v, we obtain the Euler-Lagrange equations:

Gaussian noise model:

$$\begin{aligned} \frac{\partial E^G}{\partial v} &= 2\beta v \phi(|\nabla_w u|^2) - 2\epsilon\alpha \nabla^2 v + \alpha \left(\frac{v-1}{2\epsilon} \right) = 0, \\ \frac{\partial E^G}{\partial u} &= k^* * (k * u - f) + L^{NL/MS} u = 0 \end{aligned}$$

Impulse noise model:

$$\begin{aligned} \frac{\partial E^{Im}}{\partial v} &= 2\beta v \phi(|\nabla_w u|^2) - 2\epsilon\alpha \nabla^2 v + \alpha \left(\frac{v-1}{2\epsilon} \right) = 0, \\ \frac{\partial E^{Im}}{\partial u} &= k^* * \operatorname{sign}(k * u - f) + L^{NL/MS} u = 0 \end{aligned}$$

where $k^*(x) = k(-x)$ and (see Appendix I)

$$L^{NL/MS} u = -2 \int_\Omega (u(y) - u(x)) w(x,y) \left[(v^2(y) \phi'(|\nabla_w(u)|^2(y)) + v^2(x) \phi'(|\nabla_w(u)|^2(x)) \right] dy.$$

More specifically, the NL/MSH^1 and NL/MSTV regularizers give

$$L^{NL/MSH^1}u = -2\nabla_w\cdot\big(v^2(x)\nabla_w u(x)\big) = -2\int_\Omega (u(y)-u(x))w(x,y)\left[v^2(y)+v^2(x)\right]dy,$$

$$\begin{aligned} L^{NL/MSTV}u &= -\nabla_w\cdot\left(v^2(x)\frac{\nabla_w u(x)}{|\nabla_w u(x)|}\right) \\ &= -\int_\Omega (u(y)-u(x))w(x,y)\left[\frac{v^2(y)}{|\nabla_w u|(y)}+\frac{v^2(x)}{|\nabla_w u|(x)}\right]dy. \end{aligned}$$

The energy functionals $E^G(u,v)$ and $E^{Im}(u,v)$ are convex in each variable and bounded from below. Therefore, to solve two Euler-Lagrange equations simultaneously, the alternate minimization (AM) approach is applied: in each step of the iterative procedure, we minimize with respect to one function while keeping the other one fixed. Due to its simplicity, we use the explicit scheme for u based on the gradient descent method and the Gauss-Seidel scheme for v. Note that since both energy functionals are not convex in the joint variable, we may compute only a local minimizer. However, this is not a drawback in practice, since the initial guess for u in our algorithm is the data f.

Furthermore, to extend the nonlocal methods to the impulse noise case, we need a preprocessing step for the weight function $w(x,y)$ since we cannot directly use the data f to compute w. In other words, in the presence of impulse noise, the noisy pixels tend to have larger weights than the other neighboring points, so it is likely to keep the noise value at such pixel. Thus, we propose a simple algorithm to obtain first a preprocessed image g, which removes the impulse noise (outliers) as well as preserves the textures as much as possible. Basically, we use the median filter, well-known for removing impulse noise. However, if we apply one-step of the median filter, then the output may be too smoothed out. In order to preserve the fine structures as well as to remove the noise properly, we use the idea of Bregman iteration [8], [26], and we propose the following algorithm to obtain a preprocessed image g that will be used only in the computation of the weight function:

Initialize : $r_0 = 0$, $g_0 = 0$.
do (iterate $n = 0, 1, 2, \ldots$)
$\quad g_{n+1} = median(f + r_n, [a\ a])$
$\quad r_{n+1} = r_n + f - k * g_{n+1}$
while $\|f - k * g_n\|_1 > \|f - k * g_{n+1}\|_1$
[Optional] $g_m = median(g_m, [b\ b])$

where f is the given noisy blurry data, $median(u, [a\ a])$ is the median filter of size $a \times a$ with input u; the optional step is needed in the case when the final g_m still has some salt-and-pepper-like noise. This algorithm is simple and requires a few iterations only, so it takes less than 1 second for a 256×256 size image. As seen in Fig. 9, the energy functional $\|f - k * g_n\|_1$ has a minimizer at mth iteration, and the preprocessed image g_m is a deblurred and denoised version of f; it will be used only in the computation of the weights w, while keeping f in the data fidelity term, thus artifacts are not introduced by the median filter.

Characterization of minimizers In this section, we characterize the minimizers of the functionals formulated with the nonlocal regularizers using [24, 30, 3]. Assuming that a functional $\|\cdot\|$ is a semi-norm, we can define its dual with respect to the $L^2(\Omega)$ scalar product $\langle\cdot,\cdot\rangle$ as $\|f\|_* := \sup_{\|\varphi\|\neq 0} \frac{\langle f,\varphi\rangle}{\|\varphi\|} \leq +\infty$, so that the usual duality $\langle f,\varphi\rangle \leq \|\varphi\|\|f\|_*$ holds for $\|\varphi\| \neq 0$. We define two functionals (here $Ku := k * u$),

$$F_K(u) = \lambda \int_\Omega (f-Ku)^2 dx + |u|_{NL/TV},$$
$$G_K(u,v) = \int_\Omega \sqrt{(f-Ku)^2+\eta^2}dx + \beta|u|_{NL/MS} + \alpha\left(\int_\Omega \epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon}dx\right)$$

where $\lambda > 0$ and $|u|_{NL/MS} \in \{|u|_{NL/MSTV,v}, |u|_{NL/MSH^1,v}\}$ with

$$|u|_{NLTV} = \int_\Omega |\nabla_w u|(x)dx, \quad |u|_{NL/MSTV,v} = \int_\Omega v^2(x)|\nabla_w u|(x)dx,$$
$$|u|_{NL/MSH^1,v} = \sqrt{\int_\Omega v^2(x)|\nabla_w u|^2(x)dx}$$

Note that the regularizing functionals $|u|_{NL/TV}$, $|u|_{NL/MSTV,v}$ and $|u|_{NL/MSH^1,v}$ are semi-norms (Appendix II). In addition, we modified the regularizing functional $|u|_{NL/MSH^1,v}$; the square-root term replaces the original term,

$$\int_\Omega v^2(x)|\nabla_w u|^2(x)dx,$$

of our model. It is introduced here to enable the characterization of the minimizers, but the numerical calculations will still utilize the original formulation.

Proposition 1. *Let $K : L^2(\Omega) \to L^2(\Omega)$ be a linear bounded blurring operator with adjoint K^* and let F_K be the associated functional. Then*

(1) $\|K^*f\|_* \leq \frac{1}{2\lambda}$ if and only if $u \equiv 0$ is a minimizer of F_K.
(2) Assume that $\frac{1}{2\lambda} < \|K^*f\|_* < \infty$. Then u is a minimizer of F_K if and only if $\|K^*(f-Ku)\|_* = \frac{1}{2\lambda}$ and $\langle u, K^*(f-Ku)\rangle = \frac{1}{2\lambda}|u|_{NL/TV}$

where $\|\cdot\|_*$ is the corresponding dual norm of $|\cdot|_{NL/TV}$.

We omit the proof of Proposition 1, because it is the same with the proof of Corollary 2.5 and 2.7 from [30].

Proposition 2. *Let $K : L^2(\Omega) \to L^2(\Omega)$ be a linear bounded blurring operator with adjoint K^* and let G_K be the associated functional. If (u,v) is a minimizer of G_K with $v \in [0,1]$, then*

$$\left\| K^* \frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}} \right\|_* = \beta, \quad \text{and} \quad \langle K^* \frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}, u\rangle = \beta|u|_{NL/MS}$$

where $\|\cdot\|_*$ is the corresponding dual norm of $|\cdot|_{NL/MS}$.

Proof. Let $[u, v]$ be a minimizing pair. Considering the variation of F only with respect to u, we find that for any $\varphi \in NL/MS(\Omega) = \{u \in L^2(\Omega) : |u|_{NL/MS} < \infty\}$, we have

$$\int_\Omega \sqrt{(f-Ku)^2+\eta^2}dx+\beta|u|_{NL/MS} \le \int_\Omega \sqrt{(f-K(u+\epsilon\varphi))^2+\eta^2}dx+\beta|u+\epsilon\varphi|_{NL/MS}.$$

Let

$$g(\epsilon) := \sqrt{(f-K(u+\epsilon\varphi))^2+\eta^2}.$$

Taylor's expansion gives $g(\epsilon) = \sqrt{(f-Ku)^2+\eta^2} - \epsilon\frac{(f-Ku)(K^\varphi)}{\sqrt{(f-Ku)^2+\eta^2}} + \frac{\epsilon^2}{2}g''(\epsilon_\xi)$ and hence

$$\begin{aligned}&\int_\Omega \sqrt{(f-K(u+\epsilon\varphi))^2+\eta^2}dx\\ &\le \int_\Omega \sqrt{(f-Ku)^2+\eta^2}dx - \epsilon\left\langle \frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}, K\varphi\right\rangle + \frac{\epsilon^2}{2}\max_x|g''(x)|.\end{aligned}$$

Then, the first inequality implies that

$$\epsilon\left\langle \frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}, K\varphi\right\rangle \le \epsilon\beta|\varphi|_{NL/MS} + \frac{\epsilon^2}{2}\max_x|g''(x)|.$$

Dividing by $\epsilon > 0$ and letting $\epsilon \downarrow 0_+$ (while noticing that $\lim_{\epsilon\to 0}\frac{\epsilon^2}{2}\max_x|g''(x)| = 0$) yield that for any $\varphi \in NL/MS(\Omega)$,

$$\left\langle K^*\frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}, \varphi\right\rangle \le \beta|\varphi|_{NL/MS},$$

thus,

$$\left\|K^*\frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}\right\|_* \le \beta.$$

Now let $\varphi = u$. Then, dividing by $\epsilon < 0$, and letting $\epsilon \uparrow 0_-$, we obtain

$$\left\langle K^*\frac{f-Ku}{\sqrt{(f-Ku)^2+\eta^2}}, u\right\rangle \ge \beta|u|_{NL/MS}.$$

Combining the last two inequalities concludes the proof.

3 Experimental results and comparisons

The nonlocal MS regularizers proposed here, NL/MSTV and NL/MSH^1, are tested on several images with different blur kernels and noise types. We compare them with their traditional (local) versions, such as MSTV and MSH^1 and with

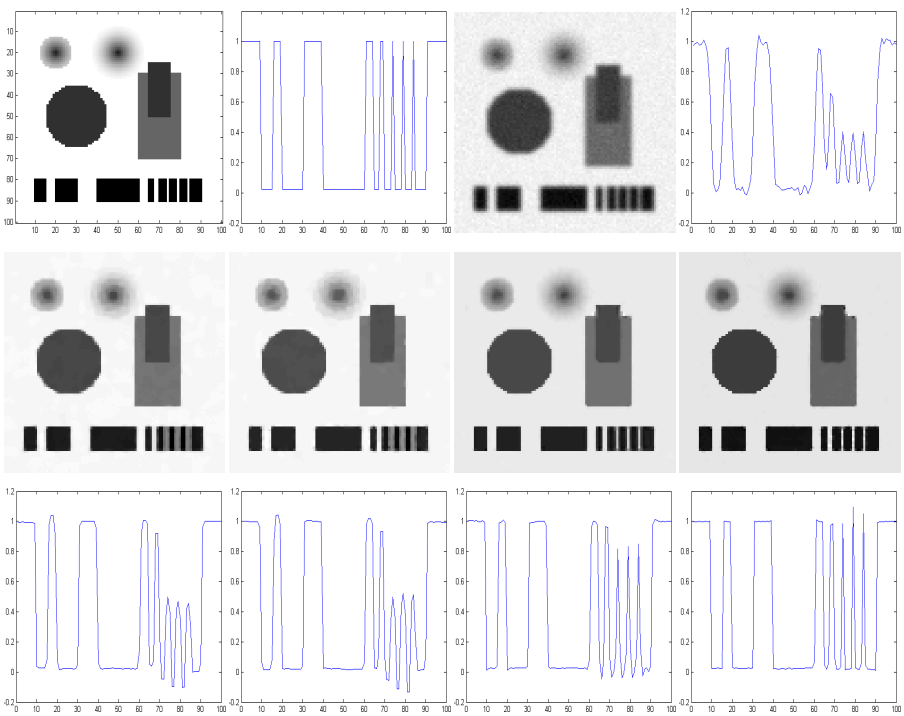

Fig. 1. Image recovery and cross sections: Gaussian blur kernel with $\sigma_b = 1$ and Gaussian noise with $\sigma_n = 5$. Top row: original image and its cross section, noisy blurry image and its cross section. Middle, bottom rows: recovered images (middle) and recovered cross sections (bottom) using TV, MSTV, NL/TV, NL/MSTV. SNR for recovered results: TV=32.9485, MSTV=33.5629, NL/TV=45.1943, NL/MSTV=50.6618. $\beta = 0.0045$ (MSTV), 0.001 (NL/MSTV), $\alpha = 0.00000015$, $\epsilon = 0.000001$.

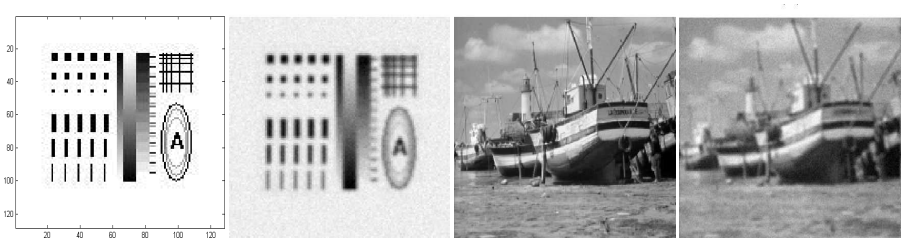

Fig. 2. Original and noisy blurry images: (second image) noisy blurry image using Gaussian kernel with $\sigma_b = 1$ and Gaussian noise with $\sigma_n = 5$; (forth image) noisy blurry image using the pill-box kernel of radius 2 and Gaussian noise with $\sigma_n = 5$.

the nonlocal total variation (NL/TV) proposed by Osher and Gilboa [20]. In addition, we experiment the nonlocal regularizers in the impulse noise model with a preconditioning step for the weight function. We use the signal-to-noise

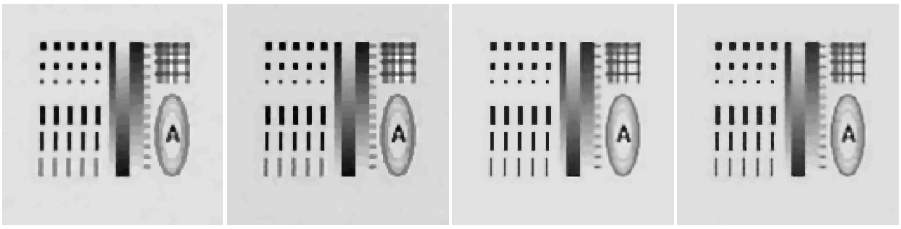

Fig. 3. Recovery of noisy blurry image from Fig. 2 using TV (SNR=14.4240), MSTV (SNR=14.4693), NL/TV (SNR=17.4165), NL/MSTV (SNR=16.5776). $\beta = 0.007$, $\alpha = 0.00000015$ (MSTV), $\beta = 0.0025$, $\alpha = 0.00000025$ (NL/MSTV), $\epsilon = 0.0000005$

ratio (SNR) as a measure of performance

$$SNR(g,u) = 40\log_{10}\left\{\frac{\|u-\overline{u}\|_{L^2}}{\|u-g\|_{L^2}}\right\}$$

where g is the original true image, u is the recovered image and $\overline{u}$ is its mean.

First, we test the Gaussian noise model in Figures 1-4. As expected, NL/MSTV and NL/MSH^1 perform better than MSTV and MSH^1 respectively in the sense that not only they recover the fine scales such as textures better, but also in the case of NL/MSTV, it does not produce any staircase effects appeared in MSTV. Furthermore, comparing the nonlocal MS regularizers with NL/TV, NL/MSTV and NL/TV seem to lead to similar results visually and according to SNR, while NL/MSH^1 gives a smoother image and lower SNR. Specifically, in Fig. 1, we use a simple image and its one-dimensional cross section. In this example, we use 11×11 size search window for NL/MSTV which is sufficient to obtain the best result, while NL/TV needs a 31×31 size. Moreover, NL/MSTV recovers the signals much better than NL/TV, which might be caused by the fact that MSTV regularizer does not suppress the jump part as much as TV. On the other hand, in Fig. 3, NL/TV produces clearer edges leading to higher SNR, while NL/MSTV has some artifacts near the edges of especially small black boxes. However, in the other real boat image, there is no significant difference between them visually and according to SNR (see Fig. 4). Fig. 4 also justifies the result that the nonlocal regularizers preserve edges and details better than the traditional local ones because we see less textures in the residuals $f - k * u$. Before ending the discussion about the Gaussian noise model, we note that we can also improve the results of the nonlocal MS regularizers by using a preprocessed image for the weight function, as Lou et al [23] did for the NL/TV.

Next, we recover a blurred image contaminated by impulse noise (salt-and-pepper noise or random-valued impulse noise). First, we test all the nonlocal regularizers and the corresponding local ones on two images, Barbara (Fig. 5) and Lenna (Fig. 6), contaminated by salt-and-pepper noise with the noise density $d = 0.2$ and $d = 0.3$ respectively, and then we test MSH^1 and NL/MSH^1 on the Einstein image (Fig. 7 and Fig. 8) with different blur kernels and both impulse

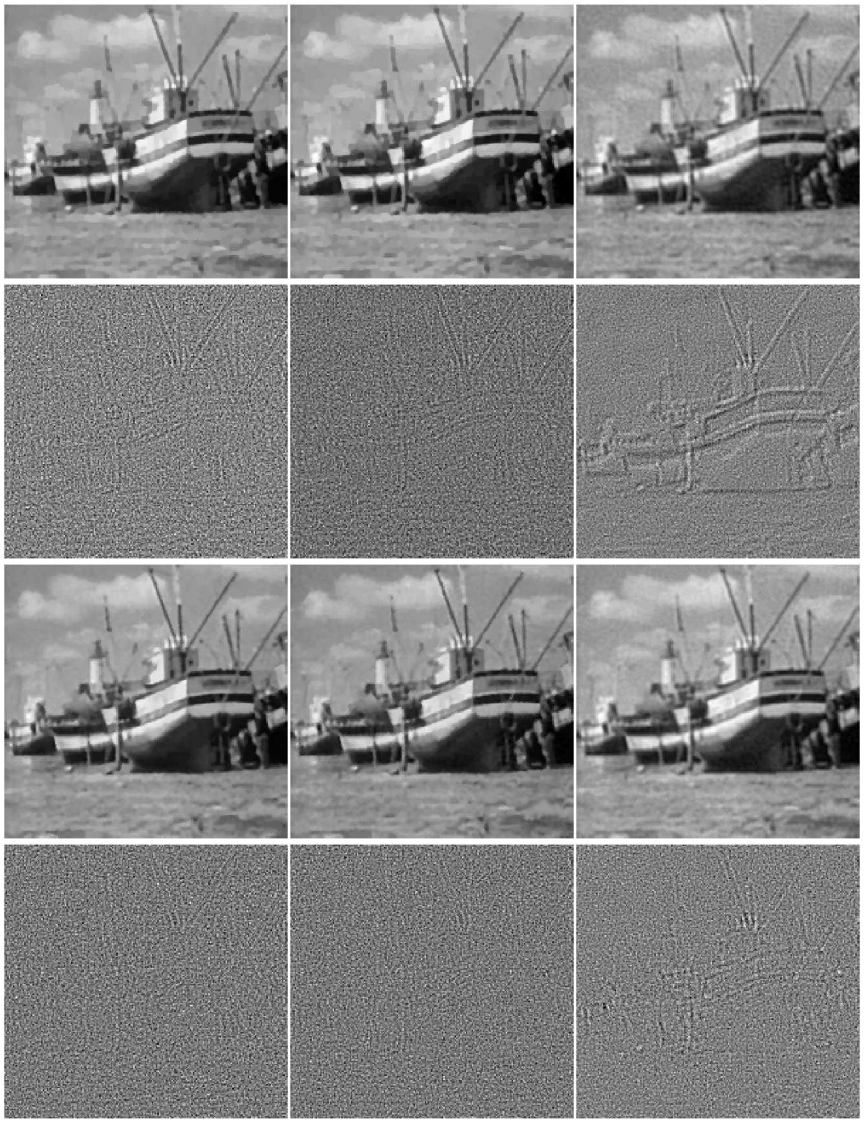

Fig. 4. Recovery of noisy blurry image from Fig. 2. Top row: recovered image u using TV (SNR=25.0230), MSTV (SNR=25.1968), MSH^1 (SNR=23.1324). Third row: recovered image u using NL/TV (SNR=26.4554), NL/MSTV (SNR=26.4696), NL/MSH^1 (SNR=24.7164). Second, bottom rows: corresponding residuals $f - k * u$. $\beta = 0.0045$ (MSTV), 0.001 (NL/MSTV), 0.06 (MSH^1), 0.006 (NL/MSH^1), $\alpha = 0.00000001$, $\epsilon = 0.00002$.

noise models, salt-and-pepper noise (Fig. 7) and random-valued impulse noise (Fig. 8), with the same noise density $d = 0.4$. By using a preconditioned image

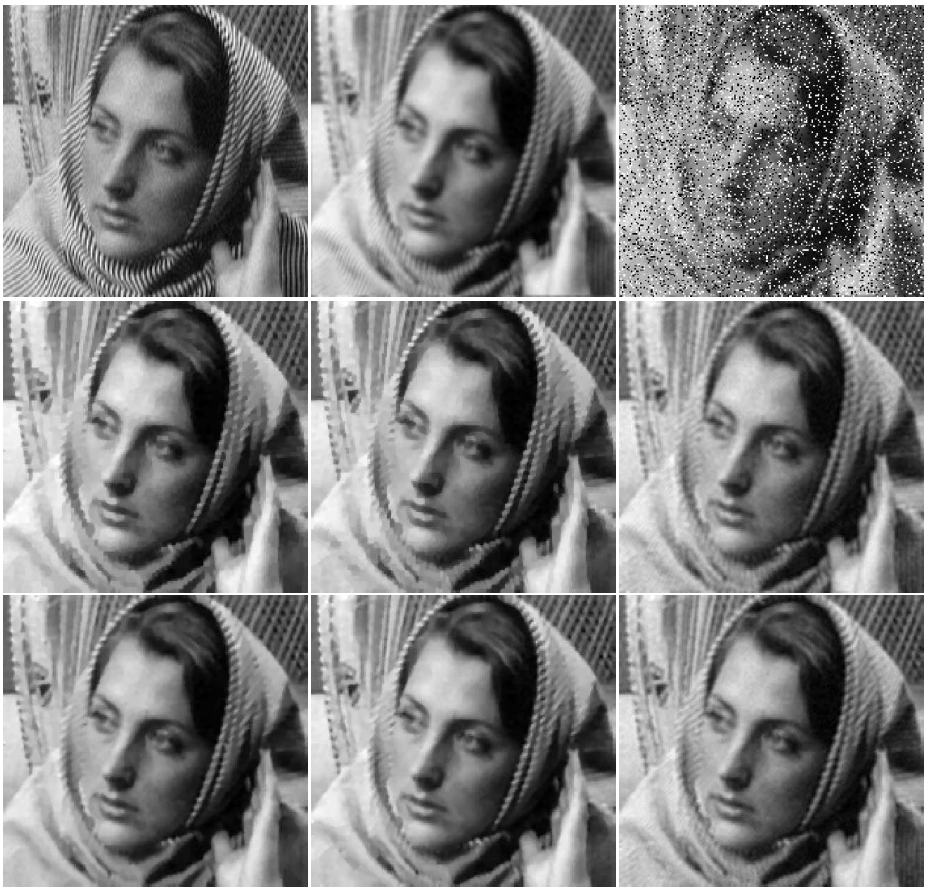

Fig. 5. Recovery of image blurred with a pill-box kernel of radius 2 and contaminated by salt-and-pepper noise with $d = 0.2$. Top row: original image, a blurry image, noisy-blurry image. Middle row: recovered images using TV (SNR=18.2909), MSTV (SNR=18.8140), MSH^1 (SNR=17.9379). Bottom row: recovered images using NL/TV (SNR=18.5693), NL/MSTV (SNR=18.8551), NL/MSH^1 (SNR=18.9862). $\beta = 0.18$ (MSTV), 0.09 (NL/MSTV), 2.2 (MSH^1), 0.35 (NL/MSH^1), $\alpha = 0.001$, $\epsilon = 0.0002$.

for the weight function, all the nonlocal regularizers outperform the traditional local ones by reducing the staircase effect and recovering the details better. And, comparing the nonlocal regularizers, NL/TV and NL/MSTV sometimes seem to give better results than NL/MSH^1 in the sense of SNR, but visually NL/MSH^1 looks much natural by preserving textures or details better especially in the case of high noise density (see Fig. 6). Moreover, in the presence of high noise density, MSH^1 suffers from restoring images well especially blurred with Gaussian kernel, while it works satisfactorily with the other blur kernels such as motion blur. But, NL/MSH^1 performs very well with Gaussian blur as well as it

Fig. 6. Recovery of noisy blurry image with Gaussian kernel with $\sigma = 1$ and salt-and-pepper noise with $d = 0.3$. Top row: original image, blurry image, noisy-blurry image. Middle row: recovered images using TV (SNR=26.9251), MSTV (SNR=27.8336), MSH^1 (SNR=23.2052). Bottom row: recovered images using NL/TV (SNR=29.2403), NL/MSTV (SNR=29.3503), NL/MSH^1 (SNR=27.1477). Second column: $\beta = 0.25$ (MSTV), 0.1 (NL/MSTV), $\alpha = 0.01$, $\epsilon = 0.002$. Third column: $\beta = 2$ (MSH^1), 0.55 (NL/MSH^1), $\alpha = 0.001$, $\epsilon = 0.0001$.

produces better results with the other blur kernels. This can be seen in Figures 6, 7, and 8. More specifically, in Fig. 5 with the noise density $d = 0.2$, NL/MSH^1 not only gives better result than MSH^1 visually and in the sense of SNR, but recovers the texture parts well while NL/TV and NL/MSTV still smooth out part of the texture. In Fig. 6 with Gaussian blur and high noise density $d = 0.3$, MSH^1 suffers from some artifacts induced by noise, while MSTV and TV give cleaner results. On the other hand, NL/MSH^1 provides better result than the

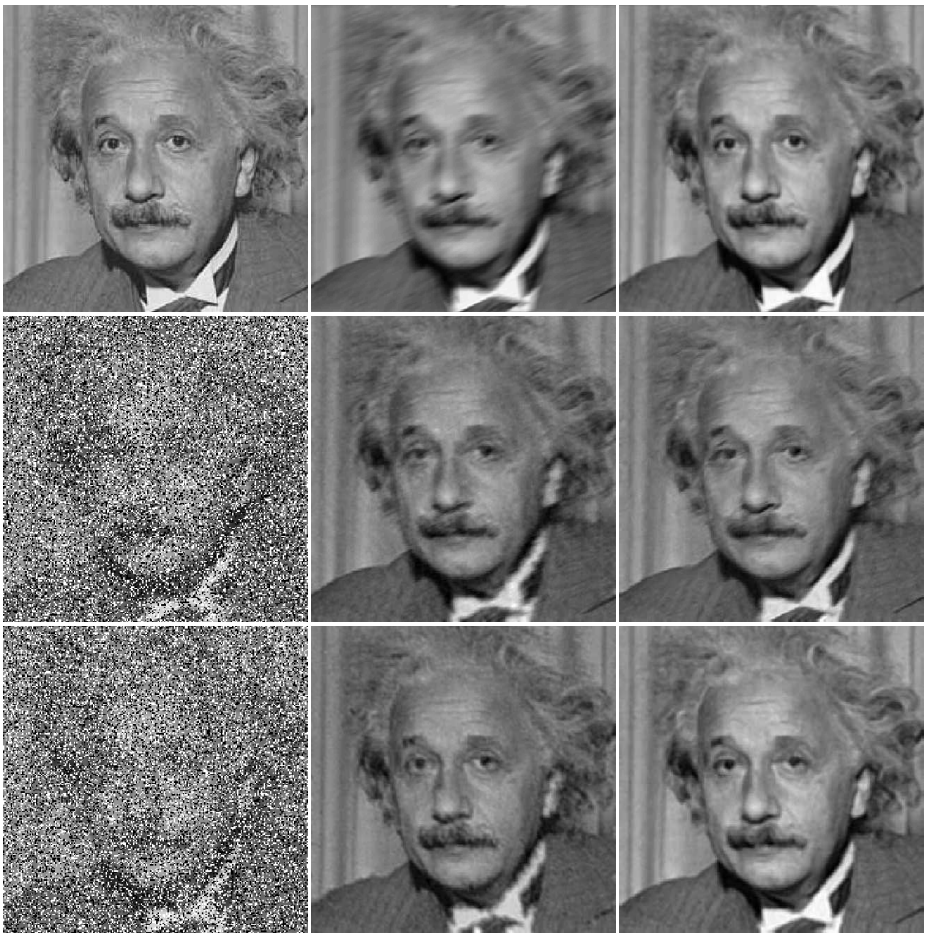

Fig. 7. Comparison between MSH^1 and NL/MSH^1 with the image blurred and contaminated by high density ($d = 0.4$) of salt-and-pepper noise. Top row: original, image blurred with motion blur kernel of length=10, oriented at angle $\theta = 25^\circ$ w.r.t. the horizon, image blurred with Gaussian kernel with $\sigma_b = 1$. Middle row: noisy blurry image blurred with the above motion blur, recovered images using MSH^1 (left, SNR=17.1106) and NL/MSH^1 (right, SNR=21.2464). Bottom row: noisy blurry image blurred with the above Gaussian blur, recovered images using MSH^1 (left, SNR=15.2017) and NL/MSH^1 (right, SNR=23.1998). Middle row: $\beta = 2$ (MSH^1), 0.4 (NL/MSH^1), bottom row: $\beta = 2$ (MSH^1), 1 (NL/MSH^1), $\alpha = 0.001$, $\epsilon = 0.0002$.

other nonlocal methods visually by preserving fine structures, and the result looks more natural. Even though NL/MSTV gives the highest SNR, it still looks more like cartoon by suppressing the texture parts especially in the hat part. So we choose NL/MSH^1 as the best result. Based on the above results, in Fig. 7 and Fig. 8, we only compare MSH^1 and NL/MSH^1 with the different blur kernels for

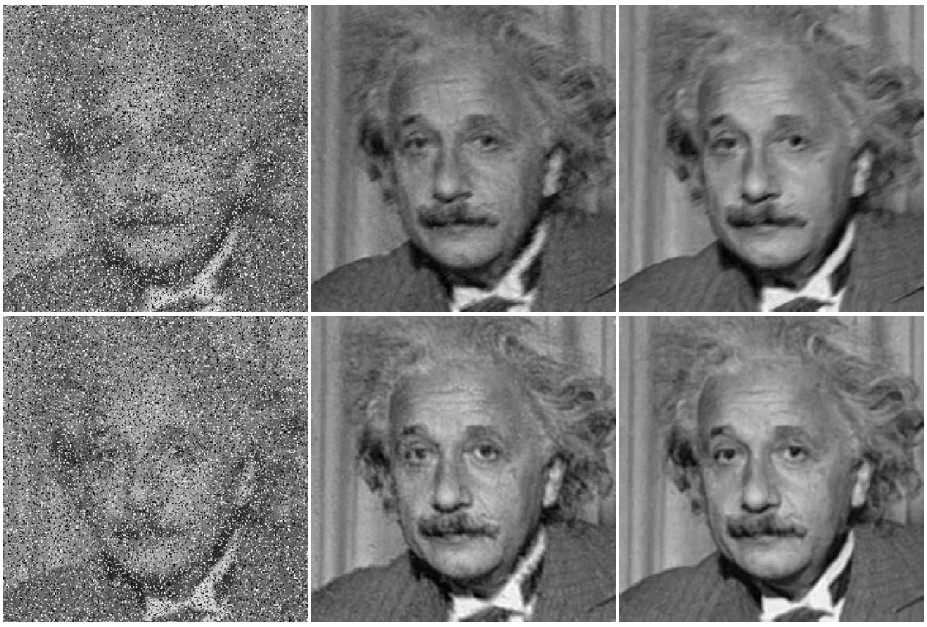

Fig. 8. Comparison between MSH^1 and NL/MSH^1 with the image blurred and contaminated by high density ($d = 0.4$) of random-valued impulse noise. Top: noisy blurry image blurred with the motion blur in the Fig.7, recovered images using MSH^1 (left, SNR=17.9608) and NL/MSH^1 (right, SNR=20.7563). Bottom: noisy blurry image blurred with the Gaussian blur in the Fig.7, recovered images using MSH^1 (left, SNR=16.6960) and NL/MSH^1 (right, SNR=24.2500). Top: $\beta = 1.5$ (MSH^1), 0.5 (NL/MSH^1), $\alpha = 0.0001$, $\epsilon = 0.002$. Bottom: $\beta = 2.5$ (MSH^1), 0.65 (NL/MSH^1), $\alpha = 0.000001$, $\epsilon = 0.002$.

both impulsive noise models, salt-and-pepper noise and random-valued impulse noise, with higher noise density $d = 0.4$. As expected, NL/MSH^1 produces better results than MSH^1 for both blur cases, and especially in the Gaussian blur case, the results do not have any artifacts (by contrast with MSH^1). Additionally, Fig. 10 shows the approximated edge sets v obtained in the restoration process. We observe that the model with the nonlocal regularizers provide smoother and more continuous edges.

Finally we note that in the MS regularizers, the parameters α, β and ϵ were selected manually to provide the best SNR results. The smoothness parameter β increases with noise level while the other parameters α, ϵ are approximately fixed. For the computational time, it takes about 5 minutes for constructing the weight function of a 256×256 image with the 11×11 search window and 5×5 patch in MATLAB on a dual core laptop with 2GHz processor and 2GB memory. The minimization for the (local or nonlocal) MS regularizers takes about 60 seconds for the computations of both u using an explicit scheme based on the gradient descent method and v using a semi-implicit scheme with the total iterations

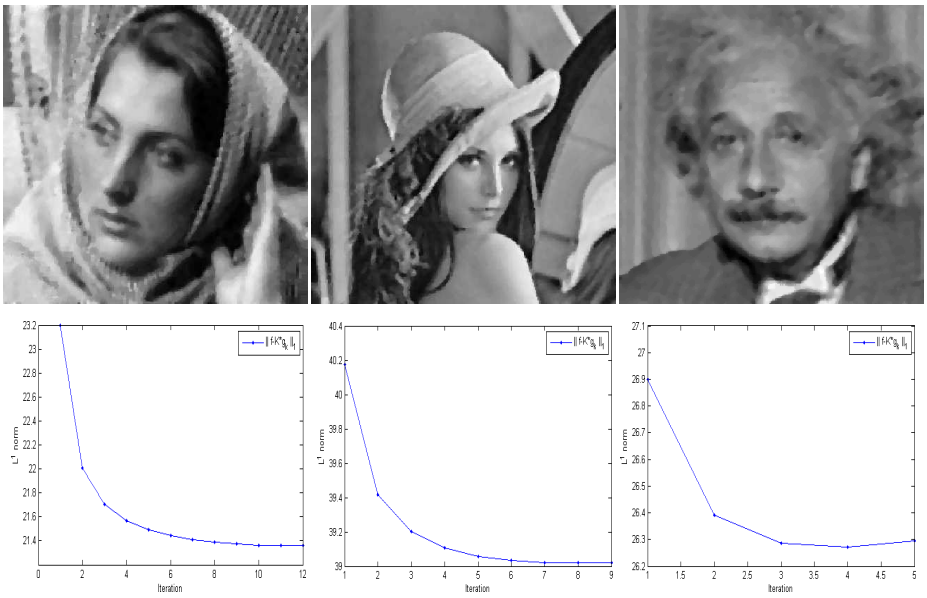

Fig. 9. The final g_m and plot of $\|f - k * g_n\|_1$ in the preprocessing step for the impulse noise model. Top: g_m of (1) Barbara (m=11) in Fig. 5, $a = 5$, $b = 3$, (2) Lena (m=8) in Fig. 6, $a = 5$, $b = 3$, (3) Einstein (m=4) with motion blur in Fig. 8, $a = 5$, $b = 5$. SNR of g_m: Barbara=16.4790, Lena=24.6619, Einstein=15.1443. Bottom: plot of $\|f - k * g_n\|_1$ of the case (1), (2), (3).

$5 \times (100 + 5)$, while the (local or nonlocal) TV regularizer using gradient descent with an explicit scheme takes less than 55 seconds with 500 iterations.

4 Summary and Conclusions

In this work, we proposed the nonlocal versions of Ambrosio-Tortorelli and Shah approximations to the MS regularizer, NL/MSH^1 and NL/MSTV, with applications to image deblurring in the presence of Gaussian or impulse noise. In the case of impulse noise model, we proposed a preprocessing step using the median filter and Bregman refinement for the computation of the weight function. Moreover, we compare the results of the proposed nonlocal MS regularizers with the ones obtained by their local versions, MSH^1 and MSTV, as well as the nonlocal total variation (NL/TV). For both noise models, the nonlocal MS regularizers perform better than the traditional local ones as expected. Comparing them with NL/TV, for the Gaussian noise model, NL/MSTV produces similar results with NL/TV in the visual sense and according to SNR, while NL/MSH^1 gives smoother images leading to lower SNRs. On the other hand, for the impulse noise case, NL/MSH^1 produces better results than the other nonlocal regularizers by

Fig. 10. Edge map v using the MS regularizers in the recovery of the Lena image blurred with Gaussian blur kernel with $\sigma_b = 1$ and contaminated by salt-and-pepper noise with density $d = 0.3$. Top: (left) MSTV, (right) NL/MSTV. Bottom: (left) MSH^1, (right) NL/MSH^1.

preserving details well, especially in the presence of high density of noise. Furthermore, the nonlocal MS regularizers provide more continuous and smoother edges, and the edges are detected concurrently with the restoration process.

Appendix I

We derive the Euler-Lagrange equation of the functional J with respect to the function $u : \Omega \to \mathbb{R}$, where $g\Omega \to \mathbb{R}$ and $\phi : \mathbb{R} \to \mathbb{R}$ are given, with

$$J(u) = \int_{\Omega} g(x)\phi(|\nabla_w u|^2(x))dx.$$

We assume that u is a minimizer of J and define $G(\epsilon) = J(u + \epsilon h)$ for $\epsilon \in \mathbb{R}$ and a test function h. Then

$$G(\epsilon) = \int_{\Omega} g(x)\phi(|\nabla_w (u + \epsilon h)|^2(x))dx.$$

By differenciating G w.r.t ϵ, we obtain

$$G'(\epsilon) = J'(u + \epsilon h)h$$

$$= \int_\Omega g(x)\phi'(|\nabla_w(u+\epsilon h)|^2(x)) \cdot \left[2\int_\Omega ((u(y)-u(x))+\epsilon(h(y)-h(x)))(h(y)-h(x))w(x,y)dy\right]dx$$

Taking $\epsilon = 0$, we obtain the variation of J with respect to u i.e.

$$\begin{aligned}
J'(u)h &= G'(0) \\
&= 2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))(h(y)-h(x))w(x,y)dy\right]dx \\
&= 2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))h(y)w(x,y)dy\right]dx \\
&\quad -2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))h(x)w(x,y)dy\right]dx \\
&= 2\int_\Omega \left[\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))(u(y)-u(x))w(x,y)dx\right]h(y)dy \\
&\quad -2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))w(x,y)dy\right]h(x)dx \\
&= 2\int_\Omega \left[\int_\Omega g(y)\phi'(|\nabla_w(u)|^2(y))(u(x)-u(y))w(y,x)dy\right]h(x)dx \\
&\quad -2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))w(x,y)dy\right]h(x)dx \\
&= -2\int_\Omega \left[\int_\Omega g(y)\phi'(|\nabla_w(u)|^2(y))(u(y)-u(x))w(x,y)dy\right]h(x)dx \\
&\quad -2\int_\Omega g(x)\phi'(|\nabla_w(u)|^2(x))\left[\int_\Omega (u(y)-u(x))w(x,y)dy\right]h(x)dx
\end{aligned}$$

where $\phi'(s)$ is the derivative of ϕ with respect to s and $w(x,y) = w(y,x)$.

Hence, we obtain

$$Lu = -2\int_\Omega (u(y)-u(x))w(x,y)\left[(g(y)\phi'(|\nabla_w(u)|^2(y)) + g(x)\phi'(|\nabla_w(u)|^2(x))\right]dy$$

where the operator L is the gradient flow corresponding to the functional J.

Specifically, by taking $g(x) = v^2(x)$ and $\phi(s) = s$ or $\phi(s) = \sqrt{s}$, we obtain two functionals and the corresponding gradient flows:

$J^{NL/MSH^1}(u) = \int_\Omega v^2(x)|\nabla_w u|^2(x)dx$:

$$L^{NL/MSH^1}u = -2\nabla_w \cdot \left(v^2(x)\nabla_w u(x)\right)$$

where $\nabla_w \cdot \left(v^2(x)\nabla_w u(x)\right) = \int_\Omega (u(y)-u(x))w(x,y)\left[v^2(y)+v^2(x)\right]dy$,

$J^{NL/MSTV}(u) = \int_\Omega v^2(x)|\nabla_w u|(x)dx$:

$$L^{NL/MSTV}u = -\nabla_w \cdot \left(v^2(x)\frac{\nabla_w u(x)}{|\nabla_w u(x)|} \right)$$

where $\nabla_w \cdot \left(v^2(x)\frac{\nabla_w u(x)}{|\nabla_w u(x)|} \right) = \int_\Omega (u(y)-u(x))w(x,y)\left[\frac{v^2(y)}{|\nabla_w u|(y)} + \frac{v^2(x)}{|\nabla_w u|(x)} \right] dy.$

Appendix II

We show that the following regularizing functionals are semi-norms (necessary for showing the characterization of minimizers):

$$|u|_{NLTV} = \int_\Omega \sqrt{\int_\Omega (u(y)-u(x))^2 w(x,y)dy}dx,$$

$$|u|_{NL/MSTV,v} = \int_\Omega v^2(x) \left(\sqrt{\int_\Omega (u(y)-u(x))^2 w(x,y)dy} \right) dx,$$

$$|u|_{NL/MSH^1,v} = \sqrt{\int_\Omega v^2(x) \left(\int_\Omega (u(y)-u(x))^2 w(x,y)dy \right) dx}$$

with $u, v : \Omega \to \mathbb{R}$ and $w : \Omega \times \Omega \to \mathbb{R}$ is nonnegative and symmetric. We only need to show that the functionals satisfy triangle inequality, i.e. $|u+v| \le |u|+|v|$.

Define $|u| = \int_\Omega g(x) \left(\sqrt{\int_\Omega (u(y)-u(x))^2 w(x,y)dy} \right) dx$ for any positive function g defined on Ω and show that $|u+v| \le |u|+|v|$ for u, v. First, we have the equality

$$\begin{aligned}
&\int_\Omega ((u+v)(y)-(u+v)(x))^2 w(x,y)dy \\
&= \int_\Omega (u(y)-u(x))^2 w(x,y)dy + \int_\Omega (v(y)-v(x))^2 w(x,y)dy \\
&+2\int_\Omega (u(y)-u(x))(v(y)-v(x))w(x,y)dy.
\end{aligned}$$

Using Cauchy-Schwarz inequality,

$$\begin{aligned}
&\int_\Omega (u(y)-u(x))(v(y)-v(x))w(x,y)dy \\
&\le \left(\int_\Omega \left((u(y)-u(x))\sqrt{w(x,y)} \right)^2 dy \right)^{1/2} \left(\int_\Omega \left((v(y)-v(x))\sqrt{w(x,y)} \right)^2 dy \right)^{1/2},
\end{aligned}$$

we obtain the quadratic formula on the right side,

$$\begin{aligned}
&\int_\Omega ((u+v)(y)-(u+v)(x))^2 w(x,y)dy \\
&\le \left\{ \left(\int_\Omega (u(y)-u(x))^2 w(x,y)dy \right)^{1/2} + \left(\int_\Omega (v(y)-v(x))^2 w(x,y)dy \right)^{1/2} \right\}^2
\end{aligned}$$

which leads to the following inequality

$$\sqrt{\int_\Omega ((u+v)(y)-(u+v)(x))^2 w(x,y)dy}$$
$$\leq \sqrt{\int_\Omega (u(y)-u(x))^2 w(x,y)dy} + \sqrt{\int_\Omega (v(y)-v(x))^2 w(x,y)dy}.$$

Multiplying by $g(x)$ and integrating both sides w.r.t x, we obtain

$$\int_\Omega g(x)\sqrt{\int_\Omega ((u+v)(y)-(u+v)(x))^2 w(x,y)dydx}$$
$$\leq \int_\Omega g(x)\sqrt{\int_\Omega (u(y)-u(x))^2 w(x,y)dydx} + \int_\Omega g(x)\sqrt{\int_\Omega (v(y)-v(x))^2 w(x,y)dydx}.$$

Thus, $|u|$ satisfies triangle inequality, so we conclude that $|u|$ is a semi-norm. Specifically, by taking $g(x)=1$ or $g(x)=v^2(x)$, $|u|_{NL/TV}$ and $|u|_{NL/MSTV}$ are semi-norms.

Similarly, we can also show that $|u|_{NL/MSH^1}$ is a semi-norm using Cauchy-Schwarz inequality:

$$\int_\Omega\int_\Omega v^2(x)((u+\varphi)(y)-(u+\varphi)(x))^2 w(x,y)dydx$$
$$= \int_\Omega\int_\Omega v^2(x)(u(y)-u(x))^2 w(x,y)dydx + \int_\Omega\int_\Omega v^2(x)(\varphi(y)-\varphi(x))^2 w(x,y)dydx$$
$$+2\int_\Omega\int_\Omega v^2(x)(x)(u(y)-u(x))(\varphi(y)-\varphi(x))w(x,y)dydx.$$

Using Cauchy-Schwarz inequality,

$$\int_\Omega\int_\Omega v^2(x)(x)(u(y)-u(x))(\varphi(y)-\varphi(x))w(x,y)dydx$$
$$\leq \left(\int_\Omega\int_\Omega (v(x)(x)(u(y)-u(x))\sqrt{w(x,y)})^2 dydx\right)^{1/2}$$
$$\cdot \left(\int_\Omega\int_\Omega (v(x)(\varphi(y)-\varphi(x))\sqrt{w(x,y)})^2 dydx\right)^{1/2},$$

the righthand side becomes a quadratic formula, hence we have

$$\sqrt{\int_\Omega\int_\Omega v^2(x)((u+\varphi)(y)-(u+\varphi)(x))^2 w(x,y)dydx}$$
$$\leq \sqrt{\int_\Omega\int_\Omega v^2(x)(u(y)-u(x))^2 w(x,y)dydx} + \sqrt{\int_\Omega\int_\Omega v^2(x)(\varphi(y)-\varphi(x))^2 w(x,y)dydx}.$$

Hence, $|u|_{NL/MSH^1}$ satisfies triangle inequality, so $|u|_{NL/MSH^1}$ is a semi-norm.

References

1. Alicandro, R., Braides, A., Shah, J.: Free-discontinuity problmes via functionals involving the L^1-norm of the gradient and their approximation. Interfaces Free Bound. 1, 17–37 (1999)
2. Ambrosio, L., Tortorelli, V.M.: On the approximation of free discontinuity problems. Boll. Un. Mat. Ital. 6-B, 105–123 (1992)
3. Andreu-Vaillo, F., Ballester, C., Caselles, V., Mazon, J.M.: Minimizing total variation flow. C.R. Acad. Sci. Paris Ser. I Math. 331, 867–872 (2000)
4. Bar, L., Sochen, N., Kiryati, N.: Variational pairing of image segmentation and blind restoration. LNCS 3022, 166–177 (2004)
5. Bar, L., Sochen, N., Kiryati, N.: Semi-Blind Image Restoration via Mumford-Shah Regularization. IEEE TIP 15(2), 483–493 (2006)
6. Bar, L., Sochen, N., Kiryati, N.: Image deblurring in the presence of salt-and-pepper noise. LNCS 3439, 107–118 (2005)
7. Bar, L., Sochen, N., Kiryati, N.: Image deblurring in the presence of impulsive noise. IJCV 70, 279–298 (2006)
8. Bregman, L.M. The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics 7, 200-217, 1967.
9. Cheon, E., Paranjpye, A., Osher, S., Vese, L.: Image denoising with total variation minimization and different L^p norms for the fidelity term. UCLA Summer REU project, 2002.
10. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. SIAM MMS 4(2), 490–530 (2005)
11. Bresson, X., Chan, T.F.: Non-local unsupervised variational image segmentation models. UCLA CAM Report 08–67 (2008)
12. Chan, R.H., Ho, C.W., Nikolova, M.: Salt-and-pepper noise removal by median-type noise detector and egde preserving regularization. IEEE TIP 14, 1479–1485 (2005)
13. Chan, R.H., Hu, C., Nikolova, M.: An Iterative procedure for removing random-valued impulse noise. IEEE SPL 11, 921-924 (2004)
14. Cai, J.F., Chan, R.H., Nikolova, M.: Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise. CMLA ENS Cachan Teport 2008-09 (2008)
15. Cai, J.F., Chan, R.H., Nikolova, M.: Fast two-phase image deblurring under impulse noise. CLMA ENS Chachan Report 2008-12 (2008)
16. Chan, T., Esedoglu, S.: Aspects of total variation regularized l^1 function approximation. UCLA CAM Report 04-07 (2004)
17. Efros, A.A., Leung, T.K.: Texture synthesis by non-parametric sampling. ICCV vol. 2, 1033–1038 (1999)
18. Geman, D, Reynolds, G.: Constrained Restoration and the Recovery of Discontinuities, IEEE TPAMI 14(3), 367-383 (1992)
19. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. SIAM MMS 6(2), 595–630 (2007)
20. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. UCLA CAM report 07-23 (2007)
21. Hwang, H., Haddad, R.A.: Adaptive median filters :new algorithms and results. IEEE Trans. Image Processing 4, 492–502 (1995)

22. Kindermann, S., Osher, S., Jones, P.W.: Deblurring and denoising of images by nonlocal functionals. SIAM MMS 4(4), 1091–1115 (2005)
23. Lou, Y., Zhang, X., Osher, S., Bertozzi, A.: Image recovery via nonlocal operators. UCLA CAM report 08-35 (2008)
24. Meyer, Y.: Oscillating Patterns in Image Processing and Nonlinear Evolution Equations. Univ. Lecture Ser. 22., Am. Math. Soc., Providence (2002)
25. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. 42, 577–685 (1989).
26. Osher, S., Burger, M., Goldfarb, D., Xu, J., Yin, W.: An iterative regularization method for total variation based image restoration. SIAM MMS 4, 460–489 (2005)
27. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D 60, 227–238 (1996)
28. Rudin, L., Osher, S.: Total variation based image restoration with free local constraints. Proc. IEEE ICIP, vol. 1, 31–35, Austin TX, USA (1994)
29. Shah, J.: A common framework for curve evolution, segmentation and anisotropic diffusion. Proc. IEEE CVPR, 136–142 (1996)
30. Tadmor, E., Nezzar, S., Vese, L.: Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation: Commun. Math. Sci. 6(2), 281–307 (2008)
31. Tichonov A., Arsensin, V.: Solution of ill-posed problems. New York: Wiley (1977)
32. Yaroslavsky, L.P.: Digital image processing: An Introduction. Springer (1985)