Nonlocal Mumford-Shah Regularizers for Color Image Restoration

Miyoun Jung, Xavier Bresson, Tony F. Chan, and Luminita A. Vese

***Abstract*—We propose here a class of restoration algorithms for color images, based upon the Mumford-Shah (MS) model and nonlocal image information. The Ambrosio-Tortorelli and Shah elliptic approximations are defined to work in a small local neighborhood, which are sufficient to denoise smooth regions with sharp boundaries. However, texture is nonlocal in nature and requires semilocal/non-local information for efficient image denoising and restoration. Inspired from recent works (nonlocal means of Buades, Coll, Morel, and nonlocal total variation of Gilboa, Osher), we extend the local Ambrosio-Tortorelli and Shah approximations to MS functional (MS) to novel nonlocal formulations, for better restoration of fine structures and texture. We present several applications of the proposed nonlocal MS regularizers in image processing such as color image denoising, color image deblurring in the presence of Gaussian or impulse noise, color image inpainting, color image super-resolution, and color filter array demosaicing. In all the applications, the proposed nonlocal regularizers produce superior results over the local ones, especially in image inpainting with large missing regions. We also prove several characterizations of minimizers based upon dual norm formulations.**

***Index Terms*—Ambrosio-Tortorelli elliptic approximations, deblurring, demosaicing, denoising, impulse noise, inpainting, Mumford-Shah (MS) regularizer, nonlocal operators, super-resolution.**

I. Introduction

We consider the degradation model of a color image

$$f = Hu+n, \qquad (f^i = H^i u^i + n^i,\ i = R,G,B) \quad (1)$$

where H is a linear operator accounting for some blurring, subsampling, or missing pixels (so that the observed data f loses some portion of the original image u we wish to recover), and n is additive noise. Problem (1) is highly ill-posed, thus, we formulate the restoration problem within the variational framework as: $\inf_u\{\Phi(f-Hu)+\Psi(\nabla u)\}$, where Φ defines a data-fidelity term, and Ψ defines the regularization that enforces a smoothness constraint on u, depending upon its gradient ∇u.

The regularization term Ψ alleviates the ill-posedness of the inverse problem by reflecting some a priori properties. Several edge-preserving regularization terms were suggested in the literature, including [2], [7], [14], [37], and [67]–[69]. These traditional regularization terms are based upon local image operators, which denoise and preserve edges and smooth regions very well, but may not deal well with fine structures like texture during the restoration process because texture is not local in nature.

In recent years, new image denoising models were proposed, based upon nonlocal image operators, to better deal with texture. Buades *et al.* [19] introduced the nonlocal means (NL-means) filter, producing excellent denoising and demosaicing results [21]. Kindermann *et al.* [52], and Gilboa-Osher [39], [40] formulated a variational framework of the NL-means filter by proposing nonlocal regularizing functionals. Lou *et al.* [53] used the nonlocal total variation (NL/TV) of Gilboa-Osher in grey-scale image deblurring in the presence of Gaussian noise. Moreover, Peyré *et al.* [64] used the total variation on nonlocal graphs for grey-scale image inpainting, super-resolution of a single image, and compressive sensing. Zhang *et al.* [78] also applied NL/TV to compressive sensing, and Protter *et al.* [66] generalized the NL-means filter to super-resolution.

Prior work using nonlocal methods has been done for the Gaussian noise model, but no study has been developed for the impulse noise model using nonlocal information. However, the impulse noise model was studied in the local case. Bar *et al.* [7] used the Ambrosio-Tortorelli and Shah approximations to Mumford-Shah (MS) regularizing functional for color image deblurring in the presence of impulse noise, producing better restoration than by the total variation regularization; moreover, the edge set is detected concurrently in the restoration process. We propose in this paper nonlocal versions (NL/MS) of Ambrosio-Tortorelli [2] and Shah [69] approximations to the MS regularizer for the multichannel case. We also propose a) several applications of NL/MS to color image denoising, deblurring in the presence of Gaussian or impulse noise, inpainting with larger missing regions, super-resolution of a single image, color filter array demosaicing, and b) an efficient preprocessing step to compute the weights w in the deblurring-denoising model in the presence of impulse noise. In all these applications, we show that the proposed nonlocal regularizers produce superior results over the local ones. We mention that preliminary results of this work have been presented in conference proceedings [47], [49] and technical report [16]. The work in [16] proposes also nonlocal level set approximations of the Mumford and Shah functional for curve evolution and boundary detection.

Manuscript received December 06, 2009; revised April 30, 2010; accepted September 20, 2010. Date of publication November 15, 2010; date of current version May 18, 2011. This work was supported by NSF-DMS Grants 031222 and 0714945, and by a UC Dissertation Year Fellowship. The associate editor coordinating the review of this manuscript and approving it for publication was Mr. Vishal Monga.

M. Jung was with the Department of Mathematics, University of California, Los Angeles, CA 90095 USA. She is now with CEREMADE, University of Paris IX Dauphine, Paris, France (e-mail: jung@ceremade.dauphine.fr).

L. A. Vese is with the Department of Mathematics, University of California, Los Angeles, CA 90095 USA (e-mail: lvese@math.ucla.edu).

X. Bresson is with the Department of Computer Science, City University of Hong Kong, Hong Kong (e-mail: xbresson@cityu.edu.hk).

T. F. Chan is with the Hong Kong University of Science and Technology, Hong Kong (e-mail: chan@math.ucla.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2010.2092433

II. Background

1) Local Regularizers: We recall two MS regularizing functionals [3], [61], [69] and their elliptic approximations [2], [3], [69] that have been used in several segmentation and restoration algorithms. The MSH^1 regularizer, depending upon the image $u:\Omega\longrightarrow\mathbb{R}$ and on its edge set $K\subset\Omega$, giving preference to piecewise smooth images, is

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

where $|\nabla u|=\sqrt{u_{x_1}^2+u_{x_2}^2}$ with $x=(x_1,x_2)$, $\mathcal{H}^1$ is the 1-D Hausdorff measure, and $\Omega\subset\mathbb{R}^2$ is the open image domain. 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 nonconvex MS functional.

There are several numerical approaches for minimizing the MS regularizing functional, one being the phase field approach using Γ-convergence [1], [2] (with applications to image deblurring and denoising [8], and to image inpainting [70]).

More specifically, 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^{MSH^1} by [2]

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

where $0\le v(x)\le 1$ represents the edges: $v(x)\approx 0$ if $x\in K=J_u$ (the jump set of u), and $v(x)\approx 1$ otherwise, ϵ is a small positive constant, α,β are positive weights. After incorporating the data fidelity term, a minimizer $u=u_\epsilon$ of $\Psi_\epsilon^{\mathrm{MSH}^1}$ approaches a minimizer u of Ψ^{MSH^1} as $\epsilon\longrightarrow 0$.

An alternative approach is the total variation regularization proposed in image restoration by Rudin, Osher, and Fatemi [67]: given a locally integrable function u, define

$$\Psi^{\mathrm{TV}}(u)=\mathrm{Sup}\left\{\int_\Omega u\nabla\cdot\phi dx:\phi\in C_c^1(\Omega,\mathbb{R}),\|\phi\|_{L^\infty(\Omega)}\le 1\right\}$$

which coincides with $\int_\Omega|\nabla u|dx$ when $u\in W^{1,1}(\Omega)$. Because of its benefit of preserving edges (which have high gradient levels) and convexity, the total variation has been widely used in image restoration.

Shah [69] suggested a modified version of the Ambrosio-Tortorelli approximation to the MS functional by replacing the quadratic term $|\nabla u|^2$ by $|\nabla u|$ in the first term

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

This functional Γ-converges to the Ψ^{MSTV} functional [3]

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

where u^+ and u^- denote the image values on two sides of the jump set $K=J_u$ of u, and D_cu is the Cantor part of the measure-valued derivative Du. Note that the nonconvex term $|u^+-u^-|/1+|u^+-u^-|$ is similar with the prior regularization by Geman-Reynolds [37]. We observe that this regularizing functional is also similar with the total variation of $u\in BV(\Omega)$ that can be written, for $K=J_u$, as

$$\Psi^{\mathrm{TV}}(u)=\int_{\Omega\backslash K}|\nabla u|dx+\int_K|u^+-u^-|d\mathcal{H}^1+|D_cu|(\Omega).$$

By comparing the second terms in Ψ^{TV} and Ψ^{MSTV}, we see that the MSTV regularizer does not penalize the jump part as much as the TV regularizer does.

In the case of color images, Blomgren and Chan [14] presented a color TV regularization by coupling the channels

$$\sqrt{\left[\int|\nabla u^R|\right]^2+\left[\int|\nabla u^G|\right]^2+\left[\int|\nabla u^B|\right]^2}.$$

The natural generalization of TV regularization to color images with coupled channels takes the form [5], [14], [75]

$$\Psi^{\mathrm{TV}}=\int_\Omega\|\nabla u\|dx=\int_\Omega\sqrt{|\nabla u^R|^2+|\nabla u^G|^2+|\nabla u^B|^2}dx$$

which was analyzed in [17]. Both channel coupling regularizers yield similar experimental results.

Bar *et al.* [7] used the color versions of MS regularizers for color image deblurring-denoising, by replacing the scalar-valued version of $|\nabla u|$ by the vector-valued version $\|\nabla u\|$ defined previously, such that

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

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

Note that, in both MS regularizers, the scalar-valued edge map v is common for the three channels and provides the necessary coupling between colors. We wish to refer to [18] for another very interesting work on color image restoration by regularization. We note that curve evolution level set approximations to MS functional have also been used for image deblurring in [50] and [48].

2) 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 texture. Nonlocal methods are generalized from the neighborhood filters (e.g., Yaroslavsky filter [77], bilateral filter [74], Susan filter [71]) and patch based methods [31], [76]. The idea of a neighborhood filter is to restore a pixel by averaging the values of neighboring pixels with a similar grey level value.

Buades *et al.* [19] generalized this idea by applying the patch-based method, and proposed the nonlocal means (or NL-means) filter for denoising a given noisy image f, given by

$$NLf(x) = \frac{1}{C(x)} \int_\Omega e^{-d_a(f(x),f(y))/h^2} f(y)dy, \quad \text{with}$$
$$d_a(f(x), f(y)) = \int_{R^2} G_a(t)\|f(x+t) - f(y+t)\|^2 dt$$

where $f(y)$ is the given color at pixel y, d_a is the patch distance, G_a is the Gaussian kernel with standard deviation a determining the patch size, $C(x) = \int_\Omega e^{-d_a(f(x),f(y))/h^2} dy$ 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 (color) value at a single point but the geometrical configuration in a whole neighborhood (patch).

3) Nonlocal Regularizers: Nonlocal filtering can be understood as a quadratic regularization based upon a nonlocal graph, as detailed for instance in the geometric diffusion framework of Coifman *et al.* [27], which was applied to nonlocal image denoising by Szlam *et al.* [72]. Denoising using quadratic penalization on image graphs was studied by Gilboa and Osher for image restoration and segmentation [39]. These quadratic regularizations were extended to nonsmooth energies such as the total variation on graphs that was defined over the continuous domain by Gilboa *et al.* [38] and over the discrete domain by Zhou and Schölkopf [79], [80]. Elmoataz *et al.* [33] considered a larger class of nonsmooth energies involving a p-laplacian for $p < 2$. Peyré replaced these nonlinear flows on graphs by a non-iterative thresholding in a nonlocal spectral basis [63].

These graph-based regularizations were used to solve general inverse problems such as image deblurring [20], [52], inpainting of thin holes and removal of texture irregularities [40]. Moreover, Peyré *et al.* [64] extended the total variation on nonlocal graphs to solve arbitrary inverse problems such as inpainting, super-resolution and compressive sampling. The graph-based regularizations are adaptive since the graph depends upon the image, in other words, the graph is directly estimated from the measurements. Recently, for grey-scale image deblurring with Gaussian noise, Lou *et al.* [53] used a preprocessed image obtained by Wiener filter, instead of f, in order to construct the weights.

Let us review nonlocal differential operators over graphs and convex nonlocal functionals proposed by Gilboa and Osher [40]. Let $u : \Omega \to \mathbb{R}$ be a function, and $w : \Omega \times \Omega \to \mathbb{R}$ be a nonnegative and symmetric weight function. The nonlocal gradient vector $\nabla_w u : \Omega \times \Omega \to \mathbb{R}$ is $(\nabla_w u)(x, y) := (u(y) - u(x))\sqrt{w(x, y)}$. Hence, the norm of the nonlocal gradient of u at $x \in \Omega$ is defined by

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

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 upon these nonlocal operators, Gilboa and Osher introduced nonlocal regularizing functionals of the general form

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

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

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

which corresponds in the local 2-D case to $\Psi^{\mathrm{TV}}(u) = \int_\Omega |\nabla u|dx$.

Inspired by the previously mentioned work, we propose in the next section nonlocal versions of Ambrosio-Tortorelli and Shah approximations to the MS regularizer for color image restoration, such as deblurring-denoising in the presence of Gaussian noise or impulse noise, inpainting with large missing regions, superresolution of a single image, and image demosaicing, by extending the scalar nonlocal operators to the vector-valued case. This is also continuation or nonlocal extension of the work by Bar *et al.* [7], [8], first to propose the use of local MS-like approximations to color image deblurring-denoising in the presence of noise.

III. Proposed Nonlocal MS Regularizers

We propose the following nonlocal MS regularizers (NL/MS) by applying the nonlocal operators to the multichannel approximations of the MS regularizer

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

where $u : \Omega \rightarrow \mathbb{R}^3$, $v : \Omega \rightarrow [0,1]$, $\phi(s) = s$ or $\phi(s) = \sqrt{s}$ correspond to the nonlocal versions of MSH^1 and MSTV regularizers, called here $\mathrm{NL/MSH}^1$ and NL/MSTV, respectively

$$\begin{aligned}\Psi^{\mathrm{NL/MSH}^1}(u,v) &= \beta\int_\Omega v^2\|\nabla_w u\|^2 dx \\ &\quad + \alpha\int_\Omega\left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon}\right)dx \\ \Psi^{\mathrm{NL/MSTV}}(u,v) &= \beta\int_\Omega v^2\|\nabla_w u\| dx \\ &\quad + \alpha\int_\Omega\left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon}\right)dx\end{aligned}$$

where $\|\nabla_w u\| : \Omega \rightarrow \mathbb{R}$ is defined as

$$\begin{aligned}\|\nabla_w u\|(x) &:= \sqrt{\sum_{i=R,G,B}|\nabla u_w^i|^2(x)} \\ &:= \sqrt{\sum_{i=R,G,B}\int_\Omega (u^i(y)-u^i(x))^2 w(x,\,y)dy}.\end{aligned}$$

We apply these nonlocal MS regularizers to color image denoising, color image deblurring in the presence of Gaussian or impulse noise, color image inpainting, color image super-resolution, and to color filter array demosaicing, by incorporating proper fidelity terms. Furthermore, for deblurring in the presence of impulse noise, we propose a preprocessing step to evaluate the weights w based upon the preprocessed image. In practice, we use the standard weight function w at $(x,\,y) \in \Omega\times\Omega$ depending upon an image $q : \Omega \rightarrow \mathbb{R}$

$$w(x,\,y) = \exp\left(-\frac{d_a(q(x),q(y))}{h^2}\right)$$

which gives the similarity of the color values as well as of image features between two pixels x and y in the image q that will be defined in each section. For a fixed pixel $x \in \Omega$, we use a search window $S(x) = \{y \in \Omega : |x-y| \le r\}$, instead of Ω, to compute $w(x,\,y)$.

Note that the nonlocal and nonconvex continuous models proposed in the following sections have not been analyzed theoretically; however, these formulations become well-defined in the discrete, finite differences case, but we prefer to present them in the continuous setting for simplicity.

IV. Image Restoration With NL/MS Regularizers

A. Color Image Deblurring and Denoising

Image blur and noise are the most common problems in photography, which can be especially significant in light limited situations, resulting in a ruined photograph. Image deblurring (or deconvolution) is the process of recovering a sharp image from an input image corrupted by blurring and noise, where the blurring is due to convolution with a known or unknown kernel; see [7], [9], [17], [26], [44]. Recently, new image denoising models [19], [20], [39], [40], [55], [63], [72], based upon nonlocal image information, have been developed to better restore texture. The standard linear degradation model for color image deblurring-denoising (or denoising) is

$$f = k*u+n \quad (f^i = k*u^i+n^i,\ i=R,G,B)$$

where k is a (known) space-invariant blurring kernel, and n is additive Gaussian noise, additive Laplace noise, or impulse noise (salt-and-pepper noise or random-valued impulse noise; in this case, the relation between $k*u$ and n is no longer of the above form).

First, in the case of Gaussian noise model, the L^2-fidelity term led by the maximum likelihood estimation is commonly used

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

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 (or the additive Laplace noise model), the L^1-fidelity term is more appropriate, due to its robustness of removing outlier effects [4], [62]; moreover, we consider the case of independent channels noise [7]

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

Thus, we propose two types of total energies for color image deblurring-denoising (to be minimized with respect to u and v), depending upon the type of noise, as follows (Gau: Gaussian noise, Im: impulse noise)

$$E^{\mathrm{Gau}}(u,v) = \frac{1}{2}\int_\Omega\sum_i |f^i - k*u^i|^2 dx + \Psi^{\mathrm{NL/MS}}(u,v) \tag{2}$$

$$E^{\mathrm{Im}}(u,v) = \int_\Omega\sum_i |f^i - k*u^i| dx + \Psi^{\mathrm{NL/MS}}(u,v). \tag{3}$$

1) Preprocessing Step for the Impulse Noise Model: To extend the nonlocal methods to the impulse noise case, we need a preprocessing step for the weight function w 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 a preprocessed image $\bar{g}$, which removes the impulse noise (outliers) as well as preserves the texture as much as we can. Basically, we use the median filter, well-known for removing impulse noise. However, for the deblurring-denoising model, if we apply one-step of the median filter, then the output may be too smoothed out. In order to preserve fine structures as well as to remove noise properly, we define a preprocessing method for the deblurring-denoising model, inspired by the idea of Bregman iteration [15]. Thus, we propose the following algorithm to obtain a preprocessed image $\bar{g}$ that will be used only in the computation of the weight function w

Initialize: $r_0^i = 0,\ g_0^i = 0,\ i = R,G,B.$

do (iterate $m = 0,1,2,\ldots$)

$g_{m+1} = \mathrm{median}(f + r_m, [s\ s])$

$r_{m+1} = r_m + f - k*g_{m+1}$

while $\sum_i \|f^i - k*g_m^i\|_1 > \sum_i \|f^i - k*g_{m+1}^i\|_1$

Fig. 1. Preprocessed images $\bar{g}$ in the presence of Random-valued impulse noise. First–fourth columns: data $f, g_1, \bar{g}, \|f-k*g_m\|_1$ vs m. Data f: (top) motion blur kernel with length of 8 and orientation 0, noise density $d = 0.1$, (bottom) motion blur kernel with length of 4 and orientation 0, noise density $d = 0.2$. Preprocessed images: $\bar{g} = g_2$ (top), $\bar{g} = g_3$ (bottom) with 3×3 median filter.

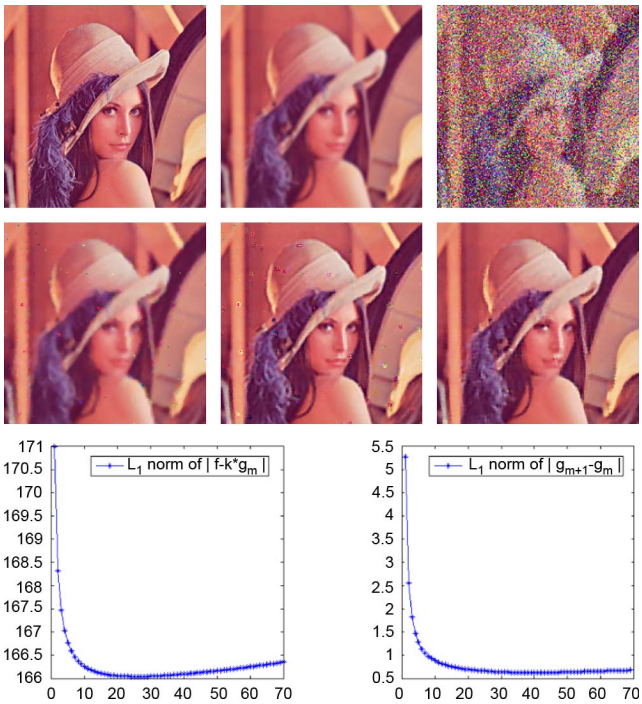

Fig. 2. Preprocessed image using iterative median filter. Top: original image, image blurred with Gaussian blur kernel with $\sigma_b = 2$, blurry-noisy data f contaminated by Salt-and-Pepper noise with noise density $d = 0.4$. Bottom: (first) recovered image using one-step median filter of size 5×5, recovered image using iterative median filter of size (second) 5×5, (third and fourth columns) 7×7 and its corresponding plots of energies $\|f-k*g_m\|_1$ and $\|g_m - g_{m+1}\|_1$ vs m.

where f is the given noisy-blurry data, and $\text{median}(f, [s\ s])$ is the median filter of size $s \times s$ with input f. The residual energy $\sum_i \|f^i - k*g_m^i\|_1$ has a minimum value at the lth iteration, thus, we obtain a preprocessed image $\bar{g} = g_l$. We show in Figs. 1 and 2 the residual norms $\sum_i \|f^i - k*g_m^i\|_1$ versus steps m. We also show in Fig. 2 the norm $\|g_{m+1} - g_m\|_1$, from which we can deduce that the sequence g_m does not converge to a limit. The preprocessed image $\bar{g}$ is a deblurred and denoised version of f, but it still includes some irregularities (or remaining impulse noise) that will be handled by constructing binary weights (values of 0 or 1) proposed for detecting and removing irregularities from texture in [40]: *in the case where the weights are computed with a Gaussian-like formula, the weights decay fast for distances above a certain threshold (usually related to the noise variance). This results in very weak connections (low weight values) for singular regions, thus such regions are essentially isolated from the rest of the image. This way rare features which also have a very large patch distance between them and any other patch in the image can be regularized as well.* The image $\bar{g}$ 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. We note that if we use the recent work [22] as our preprocessing step, then we might obtain better construction of weights. However, our proposed algorithm is simple, fast, and satisfactorily to construct the weights.

2) Characterization of Minimizers: In this section, we characterize the minimizers of the functionals formulated with the nonlocal regularizers, using [59], [73]. The details of the proofs are given in the appendices. Assuming that a functional $\|\cdot\|$ on a subspace of $(L^2(\Omega))^3$ is a seminorm, we can define the dual norm (where $\langle \cdot, \cdot \rangle$ denotes the $(L^2(\Omega))^3$ inner product) of $f \in (L^2(\Omega))^3 \subset (L^1(\Omega))^3$ as $\|f\|_* := \sup_{\|\varphi\| \neq 0} \langle f, \varphi \rangle / \|\varphi\| \leq +\infty$, so that the usual duality $\langle f, \varphi \rangle \leq \|\varphi\| \|f\|_*$ holds for $\|\varphi\| \neq 0$. We define the functionals (here $Ku := k*u$)

$$F(u) = \lambda \int_\Omega \sum_i |f^i - Ku^i|^2 dx + |u|_{\text{NL/TV}}$$

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

$$E'(u,v) = \sum_i \int_\Omega \sqrt{|f^i - Ku^i|^2 + \eta^2} dx + \beta |u|_{\text{NL/MS}} + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx$$

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

$$|u|_{\text{NL/TV}} = \int_\Omega \|\nabla_w u\|(x) dx$$

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

$$|u|_{\text{NL/MSTV,v}} = \int_\Omega v^2(x) \|\nabla_w u\|(x) dx.$$

Note that the regularizers $|u|_{\text{NL/TV}}$, $|u|_{\text{NL/MSH}^1,v}$ and $|u|_{\text{NL/MSTV,v}}$ are seminorms. In addition, we modified the regularizing functional $|u|_{\text{NL/MSH}^1,v}$; 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 minimizers below, but the numerical calculations utilize the original formulations, which solve the same equivalent problems. The following characterizations of minimizers allow us to give conditions on the existence of minimizers (including the

case of trivial minimizers), and allow us to associate a dual "texture" norm on the residual and to quantify its size.

Proposition 1: Let $K : (L^2(\Omega))^3 \rightarrow (L^2(\Omega))^3$ be a linear and continuous blurring operator with adjoint K^* and let F be the associated functional. Then:

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

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

We omit the proof of Proposition 1, since it is similar with proofs given in [59], [73].

Proposition 2: Let $K : (L^2(\Omega))^3 \rightarrow (L^2(\Omega))^3$ be a linear and continuous blurring operator with adjoint K^* and let E be the associated functional. If (u, v) is a minimizer of E with $v \in [0, 1]$, then

$$\left\| K^* \frac{f - Ku}{\sqrt{\sum_i (f^i - Ku^i)^2 + \eta^2}} \right\|_* = \beta, \quad \text{and}$$

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

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

Note that, if we replace E by E' in Proposition 2, we obtain a similar result; the second formula in the conclusion must be replaced by

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

where $\langle \cdot, \cdot \rangle$ is the $L^2(\Omega)$ inner product.

B. Color Image Inpainting

Image inpainting, also known as image interpolation, is the process of reconstructing lost or corrupted parts of an image. This is an interesting and important inverse problem with many applications such as removal of scratches in old photographs, removal of overlaid text or graphics, and filling-in missing blocks in unreliably transmitted images. Nontexture image inpainting has received considerable interest since the pioneering paper by Masnou and Morel [57], [58], who proposed variational principles for image disocclusion. A recent wave of interest in inpainting has also started from Bertalmio *et al.* [11], where applications to the movie industry, video and art restoration were unified. These authors proposed nonlinear partial differential equations for nontexture inpainting. Moreover, many contributed works have been proposed for the solution of this interpolation task based upon (a) diffusion and transport PDE/variational principle [6], [13], [23]–[25], [34], [64], [70], (b) exemplar region fill-in [12], [28], [30], [31], [65], [76], (c) compressive sensing [32]. Inpainting corresponds to the operation H in (1) of losing pixels from an image, i.e., the observed data f is given by

$$f = u \quad \text{on} \quad \Omega - D$$

where $D = D^0$ is the region where the input data u has been damaged. Thus, inspired from [25], we propose the total energy functional for color image inpainting as

$$E^{\mathrm{Inp}}(u, v) = \frac{\lambda}{2} \int_\Omega \chi_{\Omega - D}(x) \sum_i |f^i - u^i|^2 dx + \Psi^{\mathrm{NL/MS}}(u, v) \quad (4)$$

where $\chi_{\Omega - D}$ is a characteristic function on Ω (i.e., $\chi_{\Omega - D}(x) = 1$ if $x \in \Omega - D$, 0 otherwise), and $\lambda > 0$ is a parameter. In addition, we update the weights w only in the damaged region D at every mth iteration for u using the patch distance

$$d_a^R(u(x), u(y)) = \int_{\mathbb{R}^2} \chi_{\Omega - R}(x + t) G_a(t) \|u(x + t) - u(y + t)\|^2 dt$$

where $\chi_{\Omega - R}$ is a characteristic function on Ω defined previously, and $R \subset D$ is an un-recovered region (still missing region). Therefore, the missing region $D = D^0$ is recovered by the following iterative algorithm, producing the un-recovered regions D^i, $i = 0, 1, 2, \ldots$, with $D^0 \supset D^1 \supset D^2 \supset \cdots$:

1) Compute weights w for $x \in \Omega$ s.t. $P(x) \bigcap (\Omega - D^0) \neq 0$ using $d_a^{D^0}(u^0(x), u^0(y))$ with $u^0 = f$ in $\Omega - D^0$ and ∞ in D^0, a patch $P(x)$ centered at x, and $y \in S(x) \cap (\Omega - D^0)$.
2) Iterate $n = 1, 2, \ldots$ to obtain a minimizer (u, v) starting with $u = u^0$:
   a) For fixed u, update v in Ω to obtain v^n.
   b) For fixed v, update u in Ω to obtain u^n with a recovered region $\Omega - D^n \supset \Omega - D^0$: at every mth iteration, update weights w only in $x \in D^0$ s.t. $P(x) \bigcap (\Omega - D^{n,m}) \neq 0$ with
   $$d_a^{D^{n,m}}(u(x), u(y))$$
   where $y \in S(x) \cap (\Omega - D^{n,m})$, $D^{n,m}$ is an un-recovered region in D^0 until mth iteration with $D^{n,m} \supset D^{n,2m} \supset \cdots \supset D^{n,n} = D^n$.

C. Color Image Super-Resolution

Super-resolution corresponds to the recovery of a high resolution image from a filtered and down-sampled image. It is usually applied to a sequence of images in video; see [29], [35], [36], [56], [66]. We consider here a simpler problem of increasing the resolution of a single still image, and the observed data f is given by

$$f^i = D_k(h * u^i), \qquad i \in R, G, B$$

where h is a low-pass filter, $D_k : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{p \times p}$ (with $p = [n/k^2]$ where $[q]$ is the integer part of q) is the down-sampling operator by factor k along each axis. We want to recover a high resolution image $u \in (\mathbb{R}^{n \times n})^3$ by minimizing

$$E^{\mathrm{Sup}}(u, v) = \frac{1}{2} \int_\Omega \sum_i |f^i - D_k(h * u^i)|^2 dxl + \Psi^{\mathrm{NL/MS}}(u, v). \quad (5)$$

In addition, we use a super-resolved image $\bar{g} \in (\mathbb{R}^{n \times n})^3$ obtained by bicubic interpolation of $f \in (\mathbb{R}^{p \times p})^3$ only for the computation of the weights w.

D. Color Filter Array Demosaicing

In a demosaicing algorithm we have to reconstruct a full color image from the incomplete color samples output from an image sensor overlaid with a color filter array (CFA). A color filter array is a mosaic of color filters in front of the image sensor, and we use here the Bayer filter [10] that has alternating green (G) and red (R) filters for odd rows and alternating blue (B) and green (G) filters for even rows. Since each pixel of the sensor is behind a color filter, the output is an array of pixel values, each indicating a raw intensity of one of the three filter colors. Thus, an algorithm is needed to estimate for each pixel the color levels for all color components, rather than a single component; see [21], [41]–[43], [45], [51], [54]. In this variational framework, we consider the observed data f as

$$f^i = H^i \cdot u^i, \quad i \in \{R, G, B\}$$

where $\cdot$ is the pointwise product, and H^i is the down-sampling operator; H^G has alternating 1 and 0 values for odd rows and alternating 0 and 1 values for even rows, H^R has alternating 0 and 1 values for odd rows and only 0 values for even rows, H^B has only 0 values for odd rows and alternating 1 and 0 values for even rows. We propose the following minimization problem to recover a full color image u

$$E^{\mathrm{Demo}}(u, v) = \frac{1}{2}\int_\Omega \sum_i |f^i - H^i \cdot u^i|^2 dx + \Psi^{\mathrm{NL/MS}}(u, v). \tag{6}$$

Moreover, we use the interpolated image obtained by applying Hamilton-Adams algorithm [43] for the green channel and bilinear interpolation for $R-G$ and $B-G$, to compute the initial weight function w. In the Hamilton-Adams method, the evaluation of the gradient at the missing green pixel is corrected by the second-order derivatives of the red or blue channels. In addition, as in [21], in order to gradually correct the erroneous structures and artifacts of the initial color image u_0, we also proceed by an iterative strategy, refining at each step the similarity search by reducing the value of parameter h in the weights as:

- Initialize $u = u_0$ with an interpolated image with Hamilton-Adams algorithm.
- Iterate for h (e.g., $h = \{16, 8, 4\}$):
  a) Construct the weight function $w = w_u$ using the image u.
  b) Compute a minimizer (u, v) by minimizing the functional $E^{\mathrm{Demo}}(u, v)$.

We refer to [21] using NL-means for prior relevant work that inspired us in this application.

V. Numerical Discretizations

Minimization of the proposed functionals (2)–(6): E^{Gau}, E^{Im}, E^{Inp}, E^{Sup}, E^{Demo} in u and v is carried out using the Euler–Lagrange equations

$$\frac{\partial E^{\mathrm{Gau,Im,Inp,Sup,Demo}}}{\partial v} = 2\beta v\phi(\|\nabla_w u\|^2) - 2\epsilon\alpha\triangle v + \alpha\left(\frac{v-1}{2\epsilon}\right) = 0$$

$$\frac{\partial E^{\mathrm{Gau}}}{\partial u} = \tilde{k} * (k * u - f) + L^{\mathrm{NL/MS}}u = 0$$
$$\frac{\partial E^{\mathrm{Im}}}{\partial u} = \tilde{k} * \mathrm{sign}(k * u - f) + L^{\mathrm{NL/MS}}u = 0$$
$$\frac{\partial E^{\mathrm{Inp}}}{\partial u} = \chi_{\Omega - D}(u - f) + L^{\mathrm{NL/MS}}u = 0$$
$$\frac{\partial E^{\mathrm{Sup}}}{\partial u} = \tilde{h} * (D_k^T(D_k(h * u) - f)) + L^{\mathrm{NL/MS}}u = 0$$
$$\frac{\partial E^{\mathrm{Demo}}}{\partial u} = H \cdot (H \cdot u - f) + L^{\mathrm{NL/MS}}u = 0$$

where $\tilde{k}(x) = k(-x)$, $\tilde{h}(x) = h(-x)$, $D_k^T : (\mathbb{R}^{p\times p})^3 \to (\mathbb{R}^{n\times n})^3$ is the transpose of D_k i.e., the up-sampling operator, and

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

To solve two Euler–Lagrange equations simultaneously, the alternate minimization approach is applied. Note that since the energy functionals are not convex in the joint variable (u, v), we may compute only a local minimizer. In our algorithm, we define the initial guess u^0 to be the data f (except for the super-resolution problem). As suggested by a referee, we have also tested other initial conditions (u^0 a constant, a random image, or the data f perturbed by a random component; v^0 constant equal to 0 or 1); we have observed that the final steady state result does not change, only the number of iterations needed to reach the same result changes; we noticed that, if the initial image u^0 is the input data f, then fewer iterations are needed to reach the steady state. Due to its simplicity, we use Gauss-Seidel scheme for v, and an explicit scheme for u using gradient descent method, leading to the following iterative algorithm:

- Initialization: $u^0 = f$, $v^0 = 1$.
- Iterate $n = 0, 1, 2, \ldots$, until $(\|u^{n+1} - u^n\|_2 < \eta\|u^n\|_2)$.
  1) Solve the equation for v^{n+1} using Gauss-Seidel scheme:
  $$\left(2\beta\phi(\|\nabla_w u^n\|^2) + \frac{\alpha}{2\epsilon} - 2\epsilon\alpha\triangle\right)v^{n+1} = \frac{\alpha}{2\epsilon}.$$
  2) Set $u^{n+1,0} = u^n$ and solve for u^{n+1} by iterating on l:
  $$u^{n+1,l+1} = u^{n+1,l} - dt \cdot \frac{\partial E}{\partial u}(u^{n+1,l}, v^{n+1}).$$

Here η is a small positive constant. The basic discretizations are explained next. Let u_k^i denote the value of a pixel k in the image $(1 \le k \le N)$ with channel i (i.e., the discretized version of $u^i(x)$ defined on Ω), and let $p_{k,l}^i$ be the discretized version of $p^i(x, y)$ with $x, y \in \Omega$. Also, $w_{k,l}$ is the sparsely discrete version of $w = w(x, y) : \Omega \times \Omega \to \mathbb{R}$. We use the neighbors set $l \in N_k$ defined by $l \in N_k := \{l : w_{k,l} > 0\}$. Then we have ∇_{wd} and div_{wd}, the discretizations of ∇_w and div_w, respectively given by [40]

$$\nabla_{wd}(u_k^i) := (u_l^i - u_k^i)\sqrt{w_{k,l}}, \quad l \in N_k$$
$$div_{wd}(p_{k,l}^i) := \sum_{l\in N_k}(p_{k,l}^i - p_{l,k}^i)\sqrt{w_{k,l}}.$$

TABLE I
PSNR VALUES OF RECOVERED LENA IMAGES IN FIG. 5 WITH SALT-AND-PEPPER NOISE

noisy density	$d = 0.3$	$d = 0.4$	$d = 0.5$
data f	10.3172	9.1047	8.1583
preprocessed $\bar{g}$	26.1987	25.7309	24.6567
TV	28.4720	28.0494	27.2286
MSH^1	28.4402	28.1353	27.1543
MSTV	28.8540	28.4862	27.5305
NL/TV	28.7825	28.5808	27.7406
NL/MSH^1	28.7334	28.2866	27.4064
NL/MSTV	**29.4641**	**28.7977**	**27.9640**

Moreover, the magnitude of $p^i_{k,l}$ at k is $|p^i|_k = \sqrt{\sum_l (p^i_{k,l})^2}$, thus, the discretization of $||\nabla u_w||^2(x)$ is done as

$$\begin{aligned} ||\nabla u_{wd}||^2_k &= \sum_{i\in\{R,G,B\}} |\nabla_{wd} u^i|^2_k \\ &= \sum_i \sum_l (u^i_l - u^i_k)^2 w_{k,l}. \end{aligned}$$

Basically, we construct the weight function $w_{k,l}$, following the algorithm in [39]: for each pixel k, (1) take a patch P_k around a pixel k, compute the distances $(d_a)_{k,l}$ (a discretization of d_a) to all the patches P_l in the search window $l \in S(k)$, and construct the neighbors set N_k by taking the m most similar and the four nearest neighbors of the pixel k, (2) compute the weights $w_{k,l}$, $l \in N_k$ and set to zero for all the other connections ($w_{k,l} = 0$, $l \notin N_k$), (3) set $w_{k,l} = w_{l,k}$, $l \in N_k$. For deblurring in the presence of impulse noise, we used $m = 5$, so a maximum of up to $2m + 4$ neighbors for each pixel is allowed, and 5×5 pixel patches with $a = 10$, a search window of size 11×11. The complexity of computing the weights using this algorithm is $N \times Window_{size} \times (Patch_{size} \times Channel_{size} + \log m)$. Thus, in this case, we need $121 \times (25 \times 3 + 2.5) \approx 9619$ operations per pixel. Note that, when we use a preprocessed image $\bar{g}$ to compute w in the impulse noise model, we construct the weight function w with the binary values of 0 or 1 [40]: in step (2) above, for $l \in N_k$, we assign the value 1 to $w_{k,l}$ and $w_{l,k}$.

VI. EXPERIMENTAL RESULTS AND COMPARISONS

The nonlocal MS regularizers proposed here, NL/MSH^1 and NL/MSTV, are tested on several color images corrupted by different blur kernels or different noise types, on color images with missing regions, on subsampled color images, as well as on incomplete color samples outputs. We mostly compare them with their local versions [7]. For deblurring in the presence of impulse noise, we also present results with TV [17] and NL/TV [53] models because our work is the first trial to extend nonlocal methods to impulse noise model. For denoising or deblurring in the presence of Gaussian noise and super-resolution, we only present the results of MSTV and NL/MSTV because NL/MSH^1 produces smoother images leading to similar PSNR values with MSTV. For more extensive results and comparisons obtained by other competitive methods, we would like to refer the reader to the webpage and work of Peyman Milanfar *et al.* [60] on image restoration.

First, in Figs. 1 and 3, we recover blurred images contaminated by random-valued impulse noise with noise density $d = 0.1$ or $d = 0.2$. Fig. 1 presents the preprocessed images $\bar{g}$ obtained by iterative median filter of size 3×3, and the corresponding residual energies $||f - k * g_m||_1$ vs m. As m increases, the image g_m gets deblurred to some extent where the residual energy has a minimum. Thus, the image having the minimum energy value is chosen as a preprocessed image $\bar{g}$, but this still contains some impulse noise or artifacts. However, the recovered images using nonlocal regularizers in Fig. 3 show that the artifacts on the preprocessed images are well handled when constructing weights, thus, these do not influence the final recovered images. Hence, by computing the weight function w based upon the preprocessed images $\bar{g}$, all nonlocal regularizers recover texture better and reduce the artifacts by blur kernel (especially on the face and hand), providing cleaner images as well as higher PSNR values (as seen in Table II). Fig. 4 provides the edge set v of local or nonlocal MS regularizers, concurrently obtained during the restoration process in Fig. 3.

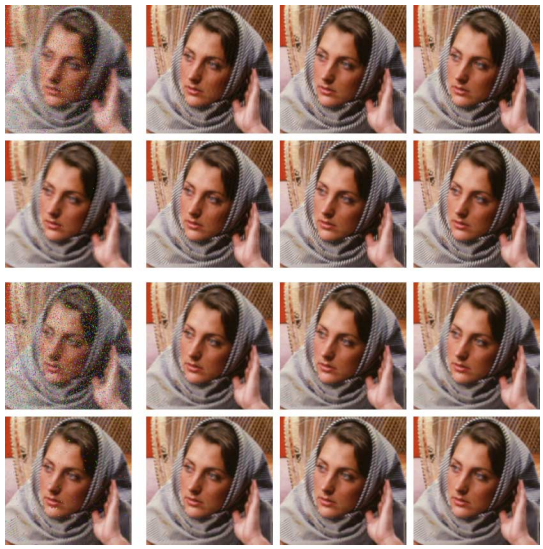

Fig. 3. DEBLURRING IN THE PRESENCE OF RANDOM-VALUED IMPULSE NOISE. Top two rows: data f given in Fig. 1 top row. Bottom two rows: data f given in Fig. 1 bottom row. First column: (top) data f, (bottom) preprocessed image $\bar{g}$. Second–fourth columns: recovered images using (top) local regularizers (MSH^1, MSTV, TV) and (bottom) nonlocal regularizers (NL/MSH^1, NL/MSTV, NL/TV).

Fig. 4. Edge set v obtained during the restoration process using MSH^1, NL/MSH^1, MSTV, NL/MSTV in Fig. 3 bottom two rows.

In Figs. 2 and 5, we recover blurred images contaminated by salt-and-pepper noise with various noise densities $d = 0.3, 0.4, 0.5$. In Fig. 2, we present the preprocessed images $\bar{g}$ obtained by iterative median filter with different size, and the image obtained by one-step median filter for comparison. First, we observe that the images using iterative median filters are deblurred and denoised versions of noisy-blurry data. For the case of $d = 0.4$, we choose the preprocessed image with iterative median filter of size 7×7; the one with smaller median filter (5×5)

TABLE II
PSNR VALUES OF RECOVERED BARBARA AND GIRL IMAGES WITH RANDOM-VALUED IMPULSE NOISE

Image	Fig. 3		Fig. 6	
Data	$l=8$ $d=0.1$	$l=4$ $d=0.2$	$r=5$ $d=0.3$	$r=3$ $d=0.4$
data f	16.9528	14.9649	12.4583	11.2622
preprocessed $\bar{g}$	21.7047	22.3935	27.9720	25.7044
TV	25.2780	23.9149	30.7566	29.5530
MSH^1	25.2128	23.9964	30.9531	27,6346
MSTV	25.4572	24.1592	32.4722	29.8059
NL/TV	26.0389	24.1542	32.6614	29.9388
$\mathrm{NL/MSH}^1$	25.7009	24.2128	31.4532	28.1454
NL/MSTV	**26.2093**	**24.2962**	**33.0390**	**30.1133**

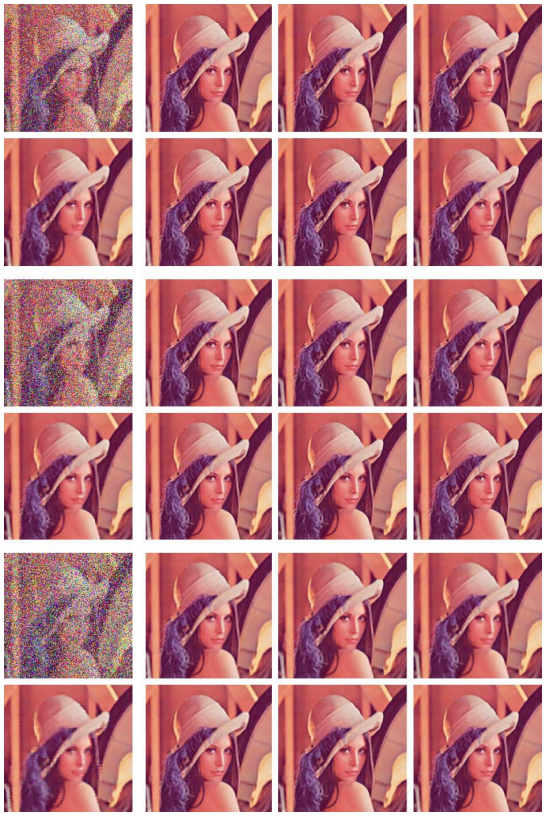

Fig. 5. DEBLURRING IN THE PRESENCE OF SALT-AND-PEPPER NOISE. Data f: Gaussian blur kernel with $\sigma_b = 2$, noise density $d = 0.3$ (Top two rows), $d = 0.4$ (Middle two rows), $d = 0.5$ (Bottom two rows). First column: (top) data f, (bottom) preprocessed image $\bar{g}$. Second–fourth columns: recovered images using (top) local regularizers (MSH^1, MSTV, TV) and (bottom) nonlocal regularizers ($\mathrm{NL/MSH}^1$, NL/MSTV, NL/TV).

TABLE III
PARAMETERS SELECTIONS $(\lambda,\ \beta)$ FOR FIGS. 3, 5, AND 6

	TV	NLTV	MSH^1	$\mathrm{NL/MSH}^1$	MSTV	NL/MSTV
Barbara	17.5	25	0.4	0.09	0.065	0.04
	3	6	2	0.6	0.3	0.2
Lena	20	30	0.17	0.07	0.04	0.02
	13	23	0.2	0.08	0.07	0.04
	5	12	0.8	0.2	0.2	0.08
Girl	14	33	0.3	0.12	0.05	0.03
	2.5	5	5	2	0.4	0.18

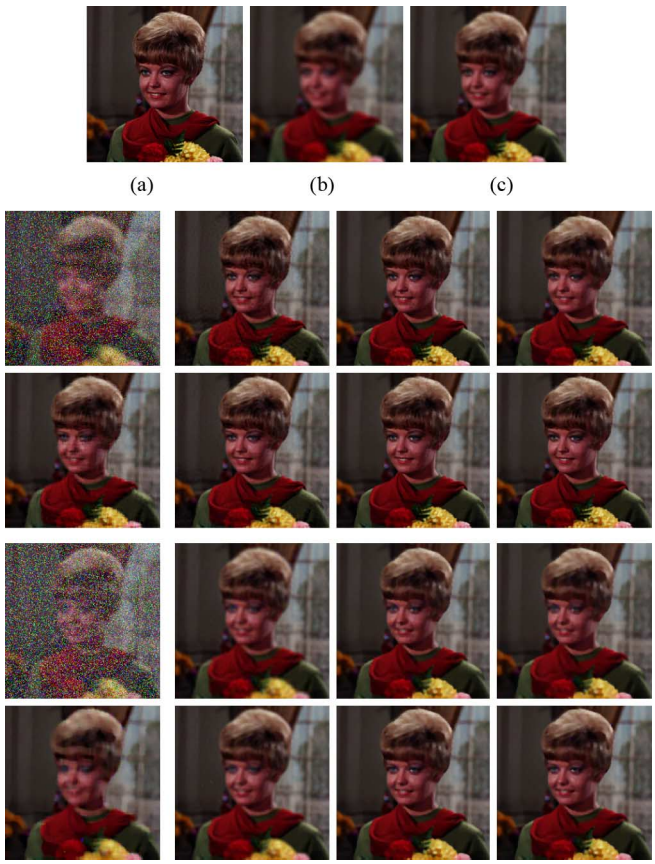

Fig. 6. DEBLURRING IN THE PRESENCE OF RANDOM-VALUED IMPULSE NOISE. (First row-a): original image. Data f: (second and third rows) out of focus blur kernel with radius $r = 5$ (first row-b) and noise density $d = 0.3$, (fourth and fifth rows) out of focus blur kernel with radius $r = 3$ (first row-c) and noise density $d = 0.4$. First column: (top) data f, (bottom) preprocessed image $\bar{g}$ with 9×9 (second and third rows) and 11×11 (fourth and fifth rows) median filters. Second–fourth columns: recovered images using (top) local regularizers (MSH^1, MSTV, TV) and (bottom) nonlocal regularizers ($\mathrm{NL/MSH}^1$, NL/MSTV, NL/TV).

produces more severe artifacts (that can influence the weights) tending to remain in the final recovered images. We show the recovered images using local and nonlocal regularizers in Fig. 5 and the corresponding PSNR values in Table I. Visually and according to PSNR values, the nonlocal regularizers recover the degraded images better than the local ones. Specifically, $\mathrm{NL/MSH}^1$ reduces ringing artifacts appeared in the recovered images with MSH^1; NL/MSTV and NL/TV reduce the staircase effect appeared in the results obtained by the corresponding local models, and suppress less the texture. Furthermore, we observe that the (local or nonlocal) MSTV regularizers are more robust than the (local or nonlocal) TV regularizers, resulting in clearer restoration, which might be caused by the fact that these do not penalize the edge part as much as TV regularizers do, and additionally yielding image edge sets v. Comparing with the (local or nonlocal) MSH^1 regularizers, the (local or nonlocal) MSTV regularizers produce cleaner images, despite the cartoon-like restored images especially in the case of high noise density.

Moreover, in Fig. 6, we use the "Girl" image corrupted by random-valued impulse noise; either high blur and noise with $d = 0.3$ (second, third rows) or low blur and higher noise with $d = 0.4$ (third, fourth rows). In the first case, $\mathrm{NL/MSH}^1$ reduces very much the ringing effect (especially appeared on the cloth part with MSH^1), providing cleaner image, and NL/TV gives much better restored image than by TV, leading to much

Fig. 7. Edge set v obtained during the restoration process using (left to right) MSH^1, $\mathrm{NL/MSH}^1$, MSTV, NL/MSTV in Fig. 6 Top two rows.

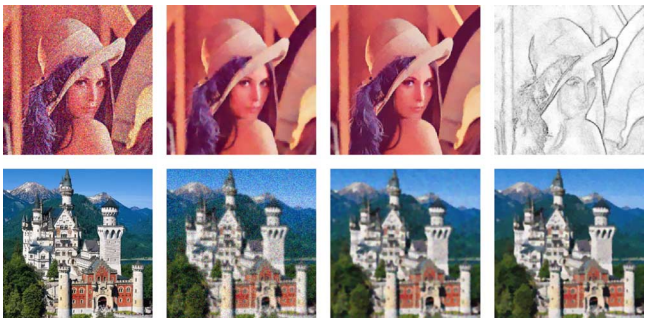

Fig. 8. DENOISING OR DEBLURRING IN THE PRESENCE OF GAUSSIAN NOISE. Top: (first) noisy image $f = u + n$ with noise variance $\sigma_n = 0.02$, recovered images using (second) MSTV and (third and fourth) NL/MSTV with edge set v. Bottom: (first) original, (second) blurry-noisy data $f = k * u + n$ with Gaussian blur kernel k with $\sigma_b = 1.2$ and noise variance $\sigma_n = 0.01$, recovered images using (third) MSTV and (fourth) NL/MSTV. PSNR: (top) f: 17.3719, MSTV: 26.1951, $\mathbf{NL/MSTV : 27.1599}$, (bottom) f: 17.2534, MSTV: 19.4849, $\mathbf{NL/MSTV : 19.8815}$. β: (top) MSTV: 0.17, NL/MSTV: 0.07, (bottom) MSTV: 0.06, NL/MSTV: 0.02.

Fig. 9. INPAINTING of 100×100 size image with 40×40 missing part. (First column) data f, (second–fourth columns) process of inpainting with $\mathrm{NL/MSH}^1$, recovered images using (fifth) $\mathrm{NL/MSH}^1$: $\mathrm{PSNR} = 35.6704$, (6th) $\mathbf{NL/MSTV : PSNR = 35.8024}$, with 41×41 search window and 9×9 patch.

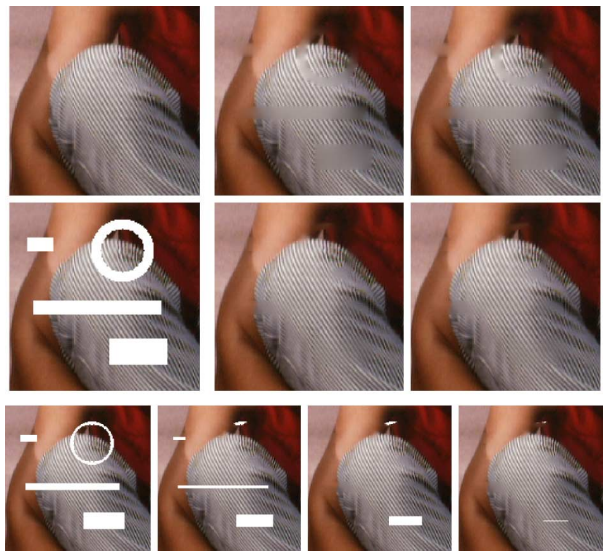

Fig. 10. INPAINTING of 150×150 size image. First column: (top row) original, (middle) data f. Second and third columns: recovered images using (top) MSH^1 : $\mathrm{PSNR} = 29.2797$, MSTV: 29.4205, (middle) $\mathbf{NL/MSH^1 : 34.4953}$, NL/MSTV: 34.2406, with 51×51 search window and 9×9 patch. Bottom row: process of inpainting with $\mathrm{NL/MSH}^1$ in 50th, 100th, 200th, and 350th iterations.

higher PSNR (as seen in Table II). Even though MSTV gives desired recovered image already, NL/MSTV additionally reduces the staircase effect (seen on the image obtained with MSTV), resulting in more realistic image as well as higher PSNR. Both NL/MSTV and NL/TV give very well recovered images visually and according to PSNR values. In the second case with less blur but more noise, NL/MSTV and NL/TV give sharper images than by corresponding local ones and the MSH^1 regularizers.

In Fig. 8, we only test MSTV and NL/MSTV models for the Gaussian noise model. For the noisy Lena image corrupted by Gaussian noise with noise variance $\sigma_n = 0.02$, NL/MSTV recovers texture much better (hat part) and provides cleaner edges, while MSTV smoothes out many details and has noisy edges. In addition, for the noisy-blurry castle image, NL/MSTV gives cleaner and sharper restored images, leading to higher PSNR values.

In Figs. 9 and 10, we use the NL/MS regularizers to recover images with texture and large missing regions. In Fig. 9, we present the process of inpainting, and final recovered images using NL/MS regularizers. We can easily see that both nonlocal regularizers recover the missing regions very well, and moreover NL/MSTV gives slightly better result than $\mathrm{NL/MSH}^1$ according to PSNR even though these visually seem to produce very similar results. However, in Fig. 10 with a real image, $\mathrm{NL/MSH}^1$ gives slightly higher PSNR values, especially recovering better the part damaged by the rectangle in the bottom. Both NL/MS regularizers gradually recover the missing regions, as seen in Fig. 10 bottom row, while local MS regularizers fail to recover them.

In Figs. 11 and 12, we recover an image filtered with a low-pass filter and then subsampled, using MSTV and NL/MSTV. As seen in both recovered images and edge sets v, NL/MSTV (incorporating weights based upon the interpolated image $\bar{g}$) provides clearer edges, leading to much higher PSNR, while MSTV produces some artifacts especially on the edges. For comparison, we apply Lucy-Richardson deconvolution iterative algorithm to the blurry image $\bar{g}$, but this produces severe artifacts and provides an even lower PSNR than for $\bar{g}$ itself and lower than the values of recovered images using MSTV and NL/MSTV.

In Figs. 13–15, we reconstruct full color images from the incomplete color samples outputs by Bayer color filter, using NL/MSTV regularizer with the decreasing sequences of h. In all the examples, we use the initial color image u_0 obtained by applying Hamilton-Adams algorithm to the green channel (G) and bilinear interpolation to both $R - G$ and $B - G$. Figs. 13 and 14 show that, with NL/MSTV regularizer, the color artifacts of u_0 are much reduced even after one iteration, and these are gradually further reduced by using decreasing sequences $h = \{16, 8, 4\}$ or $h = \{16, 8, 2\}$. Even in the case with the well-interpolated initial u_0 as in Fig. 15, NL/MS regularizer improves u_0 in one iteration, by recovering the window parts, which can be seen in the residual with the original image. Note that NL/MSTV produces better reconstructed images than $\mathrm{NL/MSH}^1$, with less color artifacts in the final images as well as higher PSNR values; the PSNR values of $\mathrm{NL/MSH}^1$ for the

Fig. 11. SUPER-RESOLUTION of a still image. Top: (first) original image of size 272×272, (second) down-sampled data $f = D_k(h * u)$ of size 68×68 with Gaussian blur kernel h with $\sigma_b = 2$ and $k = 4$, (third) preprocessed (up-sampled) image $\bar{g}$ using bicubic interpolation, (fourth) deblurred image by applying Lucy-Richarson deconvolution algorithm to $\bar{g}$ with 10 iteration number. Bottom: recovered images using (left) MSTV, (right) NL/MSTV and corresponding edge set v. PSNR: $h * u$: 21.9423, $\bar{g}$: 18.3970, MSTV: 23.4551, $\mathbf{NL/MSTV}$: $\mathbf{24.3336}$.

Fig. 12. SUPER-RESOLUTION of a still image. Top: original image of size 272×272, blurred image $h * u$ with out of focus blur kernel h with radius $r = 3$, down-sampled data $f = D_k(h * u)$ of size 68×68 with $k = 4$, preprocessed (up-sampled) image $\bar{g}$ using bicubic interpolation. Bottom: recovered images using (left) MSTV, (right) NL/MSTV. PSNR: $h * u$: 21.7459, $\bar{g}$: 19.9989, MSTV: 21.8869, $\mathbf{NL/MSTV}$: $\mathbf{22.1116}$.

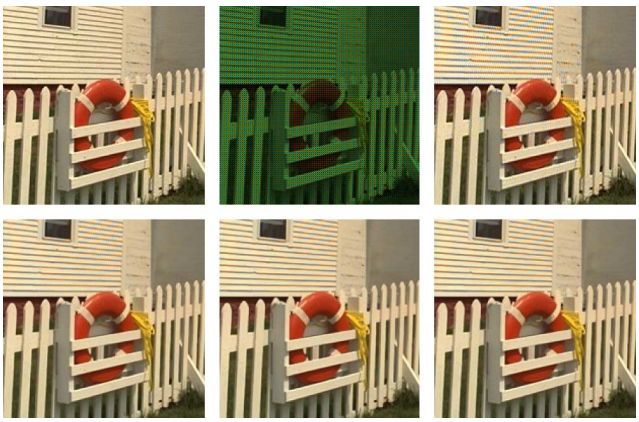

Fig. 13. DEMOSAICING using NL/MSTV with iterative algorithm. Top: original image, data f, interpolated image ($\mathrm{PSNR} = 27.3076$) using Hamilton-Adams for green and bilinear interpolation for $R-G$ and $B-G$. Bottom: demosaiced images with decreasing sequence of $h = \{16, 8, 4\}$ and corresponding PSNR values: $(\mathbf{first})$ $\mathbf{29.5101}$, $(\mathbf{second})$ $\mathbf{29.7029}$, $(\mathbf{third})$ $\mathbf{29.7128}$. Note that the PSNR values of $\mathrm{NL/MSH}^1$ for this case are (first) 29.3584, (second) 29.5858, (third) 29.5903.

roof image (in Fig. 14) are (first) 28.0111, (second) 28.0933, (third) 28.1232. Furthermore, we note that the reconstructed images using local MS regularizers are slightly improved from

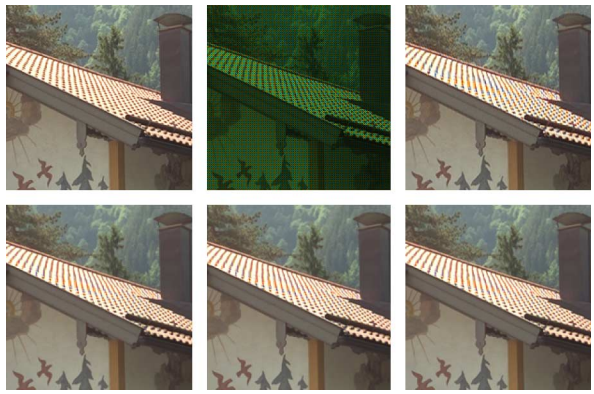

Fig. 14. DEMOSAICING using NL/MSTV with $h = \{16, 8, 2\}$. PSNR values: H-A (top-third column): 26.5008, $\mathbf{NL/MSTV}$ $(\mathbf{bottom})$: $(\mathbf{first})$ $\mathbf{28.1480}$ $(\mathbf{second})$ $\mathbf{28.2220}$, $(\mathbf{third})$ $\mathbf{28.2615}$.

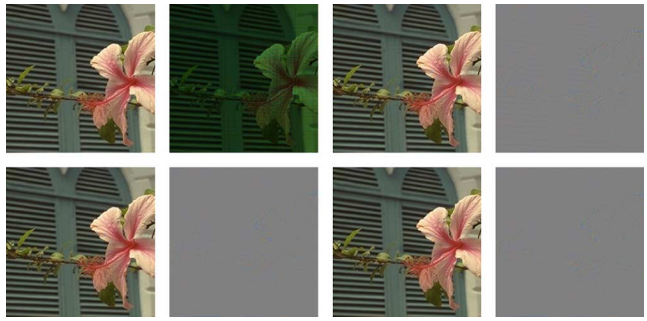

Fig. 15. DEMOSAICED IMAGES using NL/MS regularizers with $h = \{16\}$, and corresponding residuals with the original image. Top: original, data f, interpolated image with Hamilton-Adams (PSNR=36.5672) and residual. Bottom: recovered images using (left) $\mathrm{NL/MSH}^1$ (37.3598) and (right) $\mathbf{NL/MSTV(37.3606)}$.

Fig. 16. Edge sets v of $\mathrm{NL/MSH}^1$ (first, third) and NL/MSTV (second, fourth columns) obtained in Figs. 13 and 15.

the initial u_0, but in the case where the artifacts in u_0 are severe such as the Fence and the Roof images, these do not look different from the initial u_0.

Finally, we note that the parameters α, β and ϵ were selected manually to provide the best PSNR results. The smoothness parameter β increases with the noise level, while the other parameters α, ϵ are approximately fixed, $\alpha = 0.001$, $\epsilon = 0.00000001$ for deblurring-denoising, inpainting, super-resolution, demosaicing, and $\alpha = 0.1$, $\epsilon = 0.001$ for denoising (although in theory $\epsilon \longrightarrow 0$, it is common in practice to work with a small fixed ϵ). For the weights w, we use 11×11 search window with a 5×5 patch for the deblurring-denoising model (or denoising), 21×21 search window with a 9×9 patch for super-resolution, and 15×15 search window with a 3×3 patch for demosaicing, while larger search windows are needed for inpainting. 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 2 GHz processor and 2 GB memory. The minimization for the (local or nonlocal) MS regularizers in the deblurring-denoising model takes about 150 s for the computations of both u using an explicit scheme based upon the gradient descent method and v using a semiimplicit scheme with the total iterations $5 \times (2 + 100)$ (without including the computation of the weight function $w(x,\ y)$). For the inpainting model with 150×150 size image, it takes about 20 minutes with total iteration numbers $5 \times (2+100)$ since we update the weight function at every 50th iteration for u. For both super-resolution and demosaicing, $10 \times (2 + 200)$ iteration numbers are needed for all regularizers.

VII. Summary and Conclusions

The proposed nonlocal MS regularizers, $\mathrm{NL/MSH}^1$ and NL/MSTV, outperform the local ones on all the applications. For deblurring in the presence of impulse noise, the nonlocal regularizers (including NL/TV) incorporating an efficient preprocessing step perform very well and provide better recovered images than by the local ones; $\mathrm{NL/MSH}^1$ reduces ringing artifacts appeared in MSH^1, and both NL/MSTV and NL/TV reduce the staircase effect appeared in images obtained by local models, resulting in more realistic images, and better recovery of details. For denoising or deblurring in the presence of Gaussian noise, NL/MSTV recovers fine scales better, and gives cleaner and shaper images than by local regularizers or by $\mathrm{NL/MSH}^1$. Even for super-resolution, NL/MSTV provides the cleanest and sharpest results, and for color filter array demosaicing, NL/MSTV reconstructs images best. For inpainting, both $\mathrm{NL/MSH}^1$ and NL/MSTV provide superior results to those by local models by better recovering texture and large missing regions. To sum up, in all the experiments (except inpainting), NL/MSTV produces superior results to local regularizers and to $\mathrm{NL/MSH}^1$. Moreover, NL/MSTV provides edge sets concurrently obtained in the restoration process, which is another stronger point over the NL/TV model, while NL/TV model is computationally faster.

Appendix A

We show that the following regularizing functionals are seminorms, a necessary step in the proofs of Propositions 1 and 2. We let

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

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

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

with $u : \Omega \longrightarrow \mathbb{R}^3$, $v : \Omega \longrightarrow \mathbb{R}$ and $w : \Omega \times \Omega \longrightarrow \mathbb{R}$ is nonnegative and symmetric. We only need to show that these functionals satisfy the triangle inequality.

Define

$$|u| = \int_\Omega g(x) \left(\sqrt{\sum_i \int_\Omega (u^i(y) - u^i(x))^2 w(x,y)dy} \right) dx$$

for any nonnegative function $g : \Omega \longrightarrow \mathbb{R}$, and show that $|u+v| \leq |u| + |v|$ for $u, v : \Omega \longrightarrow \mathbb{R}^3$. First, we have the equality

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

Using Cauchy-Schwarz inequality, we have

$$\begin{aligned} &\int_\Omega (u^i(y) - u^i(x))(v^i(y) - v^i(x)) w(x,y)dy \\ &\quad \leq \left(\int_\Omega \left((u^i(y) - u^i(x)) \sqrt{w(x,y)} \right)^2 dy \right)^{1/2} \\ &\qquad \cdot \left(\int_\Omega \left((v^i(y) - v^i(x)) \sqrt{w(x,y)} \right)^2 dy \right)^{1/2}. \end{aligned}$$

Denote

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

Using $\sum_i \sqrt{a^i}\sqrt{b^i} \leq \sqrt{\sum_i a^i}\sqrt{\sum_i b^i}$, we obtain

$$\begin{aligned} &\sum_i \int_\Omega ((u^i + v^i)(y) - (u^i + v^i)(x))^2 w(x,y)dy \\ &\quad \leq \sum_i (a^i + b^i + 2\sqrt{a^i}\sqrt{b^i}) \\ &\quad = \sum_i a^i + \sum_i b^i + 2\sum_i \sqrt{a^i}\sqrt{b^i} \\ &\quad \leq \sum_i a^i + \sum_i b^i + 2\sqrt{\sum_i a^i}\sqrt{\sum_i b^i} \\ &\quad = \left(\sqrt{\sum_i a^i} + \sqrt{\sum_i b^i} \right)^2 \end{aligned}$$

which finally leads to

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

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

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

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

Similarly, we can also show that $|u|_{\mathrm{NL/MSH^1,v}}$ is a seminorm using Cauchy-Schwarz inequality and the inequality $\sum_i \sqrt{a^i}\sqrt{b^i} \le \sqrt{\sum_i a^i}\sqrt{\sum_i b^i}$:

$$\begin{aligned}\sum_i \int_\Omega \int_\Omega & v^2(x)((u^i+\varphi^i)(y)-(u^i+\varphi^i)(x))^2 w(x,y)dydx \\ &= \sum_i \int_\Omega \int_\Omega v^2(x)(u^i(y)-u^i(x))^2 w(x,y)dydx \\ &\quad + \sum_i \int_\Omega \int_\Omega v^2(x)(\varphi^i(y)-\varphi^i(x))^2 w(x,y)dydx \\ &\quad + 2\sum_i \int_\Omega \int_\Omega v^2(x)(u^i(y)-u^i(x))(\varphi^i(y)-\varphi^i(x)) \\ &\qquad \times w(x,y)dydx.\end{aligned}$$

Using Cauchy-Schwarz inequality

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

and the inequality $\sum_i \sqrt{a^i}\sqrt{b^i} \le \sqrt{\sum_i a^i}\sqrt{\sum_i b^i}$, we have

$$\begin{aligned}&\sqrt{\sum_i \int_\Omega \int_\Omega v^2(x)((u^i+\varphi^i)(y)-(u^i+\varphi^i)(x))^2 w(x,y)dydx} \\ &\le \sqrt{\sum_i \int_\Omega \int_\Omega v^2(x)(u^i(y)-u^i(x))^2 w(x,y)dydx} \\ &\quad + \sqrt{\sum_i \int_\Omega \int_\Omega v^2(x)(\varphi^i(y)-\varphi^i(x))^2 w(x,y)dydx}.\end{aligned}$$

Hence, $|u|_{\mathrm{NL/MSH^1,v}}$ satisfies the triangle inequality as well, so $|u|_{\mathrm{NL/MSH^1,v}}$ is a seminorm.

APPENDIX B

Proof of Proposition 2: Let $[u,v]$ be a minimizing pair. Considering the variation of F only with respect to u, we find that for any $\varphi \in NL/MS(\Omega) = \{u \in (L^2(\Omega))^3 : |u|_{\mathrm{NL/MS}} < \infty\}$, we have

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

Let

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

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

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

Then, the first inequality implies that

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

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

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

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

thus

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

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

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

Combining the last two inequalities concludes the proof.

Appendix C

We compute the Gâteaux derivative of J with respect to a minimizing function $u : \Omega \longrightarrow \mathbb{R}^3$, with $g : \Omega \longrightarrow \mathbb{R}$ and $\phi : \mathbb{R} \longrightarrow \mathbb{R}$, necessary in the computation of the Euler–Lagrange equations, where

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

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 differentiating G w.r.t ϵ, we obtain

$$
\begin{aligned}
G'(\epsilon) &= J'(u+\epsilon h)h \\
&= \int_\Omega g(x)\phi'(\|\nabla_w(u+\epsilon h)\|^2(x))2 \\
&\quad \times \int_\Omega ((u(y)-u(x)) + \epsilon(h(y)-h(x))) \\
&\quad \times (h(y)-h(x))w(x,y)dydx.
\end{aligned}
$$

Taking $\epsilon = 0$, we obtain the variation of J with respect to u i.e., See equation at the top of the page. where $\phi'(s)$ is the derivative of ϕ with respect to s and $w(x,y) = w(y,x)$. Hence, we obtain

$$
\begin{aligned}
Lu = &-2\int_\Omega (u(y)-u(x))w(x,y) \\
&\times \Big[(g(y)\phi'(\|\nabla_w(u)\|^2(y)) + g(x)\phi'(\|\nabla_w(u)\|^2(x))\Big]dy
\end{aligned}
$$

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

$$
\begin{aligned}
J^{\mathrm{NL/MSH}^1}(u) &= \int_\Omega v^2(x)\|\nabla_w u\|^2(x)dx : \\
L^{\mathrm{NL/MSH}^1}u &= -2\nabla_w \cdot \left(v^2(x)\nabla_w u(x)\right) \\
&= -2\int_\Omega (u(y)-u(x))w(x,y) \\
&\quad \times \left[v^2(y) + v^2(x)\right]dy \\
J^{\mathrm{NL/MSTV}}(u) &= \int_\Omega v^2(x)\|\nabla_w u\|(x)dx : \\
L^{\mathrm{NL/MSTV}}u &= -\nabla_w \cdot \left(v^2(x)\frac{\nabla_w u(x)}{\|\nabla_w u(x)\|}\right) \\
&= -\int_\Omega (u(y)-u(x))w(x,y) \\
&\quad \times \left[\frac{v^2(y)}{\|\nabla_w u\|(y)} + \frac{v^2(x)}{\|\nabla_w u\|(x)}\right].
\end{aligned}
$$

References

[1] L. Ambrosio and V. M. Tortorelli, "Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence," *Commun. Pure Appl. Math.*, vol. 43, pp. 999–1036, 1990.

[2] L. Ambrosio and V. M. Tortorelli, "On the approximation of free discontinuity problems," *Bollettino dell'Unione Matematica Italiana*, vol. B(7), no. 6, pp. 105–123, 1992.

[3] R. Alicandro, A. Braides, and J. Shah, "Free-discontinuity problems via functionals involving the L^1-norm of the gradient and their approximation," *Interfaces and Free Boundaries*, vol. 1, no. 1, pp. 17–37, 1999.

[4] S. Alliney, "Digital filters as absolute norm regularizers," *IEEE Trans. Signal Process.*, vol. 40, no. 6, pp. 1548–1562, Jun. 1992.

[5] G. Aubert and P. Kornprobst, "Mathematical problems in image processing: Partial differential equations and the calculus of variations," in *Applied Mathematical Sciences*. New York: Springer-Verlag, 2006, vol. 147.

[6] J. F. Aujol, S. Ladjal, and S. Masnou, "Exemplar-based inpainting from a variational point of view," *SIAM J. Math. Anal.*, vol. 42, no. 3, pp. 1246–1285, 2010.

[7] L. Bar, A. Brook, N. Sochen, and N. Kiryati, "Deblurring of color images corrupted by impulsive noise," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1101–1111, Apr. 2007.

[8] L. Bar, N. Sochen, and N. Kiryati, "Semi-blind image restoration via Mumford-Shah regularization," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 483–493, Feb. 2006.

[9] L. Bar, N. Sochen, and N. Kiryati, "Image deblurring in the presence of impulsive noise," *Int. J. Comput. Vis.*, vol. 70, no. 3, pp. 279–298, 2006.

[10] B. E. Bayer, "Color Imaging Array," U.S. Patent 3 971 065, Jul. 20, 1976.

[11] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. SIGGRPAPH*, 2000, pp. 417–424.

[12] M. Bertalmio, L. A. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.

[13] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," *Proc. IEEE Comput. Vis. Pattern Recognit.*, vol. 1, pp. I–355, 2001.

[14] P. Blomgren and T. F. Chan, "Color TV: Total variation methods for restoration of vector-valued images," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 304–309, Mar. 1998.

[15] L. M. Bregman, "The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming," *USSR Comput. Math. Math. Phys.*, vol. 7, no. 3, pp. 200–217, 1967.

[16] X. Bresson and T. F. Chan, "Non-local unsupervised variational image segmentation models," Univ. California, Los Angeles, UCLA C.A.M. Rep. 08-67, 2008.

[17] X. Bresson and T. F. Chan, "Fast minimization of the vectorial total variation norm and applications to color image processing," *Inv. Problems Imag.*, vol. 2, no. 4, pp. 455–484, 2008.

[18] A. Brook, R. Kimmel, and N. A. Sochen, "Variational restoration and edge detection for color images," *JMIV*, vol. 18, no. 3, pp. 247–268, 2003.

[19] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *SIAM Multiscale Model. Sim.*, vol. 4, no. 2, pp. 490–530, 2005.

[20] A. Buades, B. Coll, and J. M. Morel, *Image Enhancement by Non-Local Reverse Heat Equation, Preprint CMLA 2006-22*. : , 2006.

[21] A. Buades, B. Coll, J. M. Morel, and C. Sbert, "Self similarity driven demosaicking," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1192–1202, Jun. 2009.

[22] J. F. Cai, R. H. Chan, and M. Nikolova, "Fast two-phase image deblurring under impulse noise," *JMIV*, vol. 36, no. 1, pp. 46–53, 2010.

[23] V. Caselles, G. Sapiro, C. Ballester, M. Bertalmio, and J. Verdera, "Filling-in by joint interpolation of vector fields and grey levels," *IEEE Trans. Image Process*, vol. 10, no. 8, pp. 1200–1211, Aug. 2001.

[24] T. F. Chan, S. H. Kang, and J. Shen, "Euler's elastica and curvature based inpainting," *SIAM J. Appl. Math.*, vol. 63, no. 2, pp. 564–592, 2002.

[25] T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, 2002.

[26] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.

[27] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proc. Nat. Acad. Sci.*, vol. 102, no. 21, pp. 7426–7431, 2005.

[28] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.

[29] D. Datsenko and M. Elad, "Example-based single image super-resolution: A global map approach with outlier rejection," *J. Mult. Syst. Signal. Process.*, vol. 18, no. 2–3, pp. 103–121, 2007.

[30] L. Demanet, B. Song, and T. Chan, "Image inpainting by correspondence maps: A deterministic approach," in *Proc. 2nd IEEE Workshop Variational, Geomet. Level Set Meth. Comput. Vis.*, Nice, France, 2003.

[31] A. A. Effros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, 1999, vol. 2, pp. 1033–1038.

[32] M. Elad, J. L. Starck, D. Donoho, and P. Querre, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, 2005.

[33] A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1047–1060, Jul. 2008.

[34] S. Esedoglu and J. Shen, "Digital inpainting based on the Mumford-Shah-Euler image model," *Eur. J. Appl. Math.*, vol. 13, pp. 353–370, 2002.

[35] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in superresolution," *Int. J. Imag. Syst. Tech.*, vol. 14, no. 2, pp. 47–57, 2004.

[36] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based superresolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, Mar./Apr. 2002.

[37] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

[38] G. Gilboa, J. Darbon, S. Osher, and T. Chan, "Nonlocal convex functionals for image regularization," Univ. California, Los Angeles, UCLA CAM Rep. 06-57, 2006.

[39] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *SIAM Multiscale Model. Sim.*, vol. 6, no. 2, pp. 595–630, 2007.

[40] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *SIAM Multiscale Model. Sim.*, vol. 7, no. 3, pp. 1005–1028, 2008.

[41] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 997–1013, Sep. 2002.

[42] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicing: Color filter array interpolation," *IEEE Signal Process. Mag.*, vol. 22, no. 1, pp. 44–54, 2005.

[43] J. F. Hamilton and J. E. Adams, "Adaptive Color Plan Interpolation in Single Sensor Color Electronic Camera," U.S. Patent 5506619, Jul. 29, 1997.

[44] L. He, A. Marquina, and S. Osher, "Blind deconvolution using TV regularization and Bregman iteration," *Int. J. Imag. Syst. Technol.*, vol. 15, no. 1, pp. 74–83, 2005.

[45] K. Hirakawa and T. W. Paks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 360–369, Mar. 2005.

[46] K. Hirakawa and T. W. Paks, "Joint demosaicing and denoising," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2146–2157, Aug. 2006.

[47] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese, "Color image restoration using nonlocal Mumford-Shah regularizers," *LNCS*, vol. 5681, pp. 373–387, 2009.

[48] M. Jung, G. Chung, G. Sundaramoorthi, L. A. Vese, and A. L. Yuille, "Sobolev gradients and joint variational image segmentation, denoising and deblurring," *Electron. Imag.*, vol. 7246, pp. 72460I–1, 2009.

[49] M. Jung and L. A. Vese, "Nonlocal variational image deblurring models in the presence of Gaussian or impulse noise," *LNCS*, vol. 5567, pp. 402–413, 2009.

[50] D. Kim, A. Tsai, M. Cetin, and A. Willsky, "A curve evolution based variational approach to simultaneous image restoration and segmentation," in *Proc. IEEE Int. Conf. Image Process.*, 2002, vol. 1, pp. 109–112.

[51] R. Kimmel, "Demosaicing: Image reconstruction from color CCD samples," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp. 1221–1228, Sep. 1999.

[52] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *SIAM Multiscale Model. Sim.*, vol. 4, no. 4, pp. 1091–1115, 2005.

[53] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, "Image recovery via nonlocal operators," *J. Sci. Comput.*, vol. 42, no. 2, pp. 185–197, 2010.

[54] W. Lu and Y.-P. Tan, "Color filter array demosaicing: New method and performance measures," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1194–1210, Oct. 2003.

[55] M. Mahmoudi and G. Sapiro, "Fast image and video denoising via nonlocal means of similar neighborhoods," *IEEE Signal Process. Lett.*, vol. 12, no. 12, p. 839, Dec. 2005.
[56] F. Malgouyres and F. Guichard, "Edge direction preserving image zooming: A mathematical and numerical analysis," *SIAM Numer. Anal.*, vol. 39, no. 1, pp. 1–37, 2001.
[57] S. Masnou and J.-M. Morel, "Level lines-based disocclusion," in *Proc. IEEE Int. Conf. Image Process.*, 1998, vol. 3, pp. 259–263.
[58] S. Masnou and J.-M. Morel, "Disocclusion: A variational approach using level lines," *IEEE Trans. Image Process.*, vol. 11, no. 1, pp. 68–76, Jan. 2002.
[59] Y. Meyer, "Oscillating patterns in image processing and nonlinear evolution equations," *Univ. Lecture Ser. 22., Am. Math. Soc., Providence*, 2002.
[60] P. Milanfar [Online]. Available: http://users.soe.ucsc.edu/~milanfar/software/
[61] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *CPAM*, vol. XLII, pp. 577–685, 1989.
[62] M. Nikolova, "Minimizers of cost-functions involving non-smooth data-fidelity terms. Application to the processing of outliers," *SIAM Numer. Anal.*, vol. 40, no. 3, pp. 965–994, 2002.
[63] G. Peyré, "Image processing with non-local spectral bases," *SIAM Multiscale Model. Sim.*, vol. 7, no. 2, pp. 703–730, 2008.
[64] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," *LNCS*, vol. 5304, pp. 57–68, 2008.
[65] P. Pérez, M. Gangnet, and A. Blake, "PatchWorks: Example-based region tiling for image editing," Microsoft Corporation, Redmond, WA, Tech Rep. Microsoft Research, MSR-TR-2004-04, 2004.
[66] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 36–51, Jan. 2009.
[67] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
[68] L. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Proc. IEEE Int. Conf. Image Process.*, 1994, vol. 1, pp. 31–35.
[69] J. Shah, "A common framework for curve evolution, segmentation and anisotropic diffusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 1996, pp. 136–142.
[70] J. Shen and T. F. Chan, "Variational image inpainting," *CPAM*, vol. 58, no. 5, pp. 579–619, 2005.
[71] S. M. Smith and J. M. Brady, "SUSAN—A new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, 1997.
[72] A. Szlam, M. Maggioni, and R. R. Coifman, "Regularization on graphs with function-adapted diffusion processes," *J. Mach. Learn. Res.*, vol. 9, pp. 1711–1739, 2008.
[73] E. Tadmor, S. Nezzar, and L. Vese, "Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation," *Commun. Math. Sci.*, vol. 6, no. 2, pp. 281–307, 2008.
[74] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.
[75] D. Tschumperlé and R. Deriche, "Vector-valued image regularization with PDE's: A common framework for different applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 506–517, Apr. 2005.
[76] L. Y. Wei and M. Levoy, "Fast texture synthesis using tree-structured vector quantization," *Proc. SIGGRAPH 2000*, pp. 479–488, 2000.
[77] L. P. Yaroslavsky, *Digital Image Processing: An Introduction*. New York: Springer-Verlag, 1985.
[78] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.
[79] D. Zhou and B. Scholkopf, "A regularization framework for learning from graph data," in *Proc. Int. Conf. Mach. Learn.*, Banff, Canada, 2004, pp. 132–137.
[80] D. Zhou and B. Scholkopf, "Regularization on discrete spaces," *LNCS*, vol. 3663, pp. 361–368, 2005.

Miyoun Jung received the B.S. and M.S. degrees in mathematics from Seoul National University, Seoul, Korea, in 2002 and 2004, respectively, and the Ph.D. degree from the University of California, Los Angeles, in 2009.

She is currently a Postdoctoral Fellow at CEREMADE, University of Paris IX Dauphine. Her research interests are in mathematical image processing, and she has worked on image restoration and segmentation problems in a variational framework.

Xavier Bresson was born in France, in 1976. He received the M.Sc. degree in theoretical physics from University of Marseille, France, in 1998, the M.Sc. degree in electrical engineering from Ecole Superieure d'Electricite, Paris, France, in 2000, and in automatism and signal Processing from the University of Paris XI, and the Ph.D. degree in computer vision at the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 2005.

He is currently an Assistant Professor in the Department of Computer Science, City University of Hong Kong. In 2006–2010, he was a Postdoctoral Scholar in the Department of Mathematics, University of California, Los Angeles (UCLA). His main research activities are focused on computational modeling (fast and generic optimization algorithms) applied to computer vision, image processing, medical imaging, and machine learning.

Tony F. Chan received the B.S. and M.S. degrees in engineering from CalTech, Pasadena, CA, and the Ph.D. degree in computer science from Stanford University, Stanford, CA.

He was with CalTech (applied math) as a Research Fellow and taught computer science at Yale University, New Haven, CT, before joining the faculty at the University of California, Los Angeles (UCLA), in 1986 as Professor of mathematics. He served as a Chair of the Department of Mathematics, UCLA, from 1997–2000, and as Director of the Institute for Pure and Applied Mathematics (IPAM) from 2000 to 2001. From July 2001 to June 2006, he served as Dean of Physical Sciences, UCLA, and became an Assistant Director, Directorate for Mathematics and Physical Sciences at National Science Foundation (2006–2009). He has been the President at the Hong Kong University of Science and Technology since October 2009, holding positions in the Mathematics and Computer Science Departments. His current research interests include mathematical image processing, computer vision, and computer graphics, computational brain mapping, VLSI physical design optimization, multiscale computational methods, multigrid and domain decomposition algorithms, iterative methods, Krylov subspace methods, parallel algorithms, and computational linear algebra.

Luminita A. Vese received the M.S. degree in mathematics from West University of Timisoara, Romania, in 1993, and the M.S. and Ph.D. degrees in applied mathematics from University of Nice, Sophia, Antipolis, France, in 1992 and 1997, respectively.

She is currently a Professor of Mathematics at the University of California, Los Angeles (UCLA). Before joining the UCLA faculty, she held postdoctoral research and teaching positions at the University of Nice, the University of Paris IX Dauphine, and UCLA. Her research interests include variational methods and partial differential equations, inverse problems, image analysis, and computer vision.