Image Deblurring in the Presence of Impulsive Noise

Leah Bar Nir Sochen* Nahum Kiryati
School of Electrical Engineering
* Dept. of Applied Mathematics
Tel Aviv University
Tel Aviv, 69978, Israel

Abstract

Consider the problem of image deblurring in the presence of impulsive noise. Standard image deconvolution methods rely on the Gaussian noise model and do not perform well with impulsive noise. The main challenge is to deblur the image, recover its discontinuities and at the same time remove the impulse noise. Median-based approaches are inadequate, because at high noise levels they induce nonlinear distortion that hampers the deblurring process. Distinguishing outliers from edge elements is difficult in current gradient-based edge-preserving restoration methods. The suggested approach integrates and extends the robust statistics, line process (half quadratic) and anisotropic diffusion points of view. We present a unified variational approach to image deblurring and impulse noise removal. The objective functional consists of a fidelity term and a regularizer. Data fidelity is quantified using the robust modified L^1 norm, and elements from the Mumford-Shah functional are used for regularization. We show that the Mumford-Shah regularizer can be viewed as an extended line process. It reflects spatial organization properties of the image edges, that do not appear in the common line process or anisotropic diffusion. This allows to distinguish outliers from edges and leads to superior experimental results.

1 Introduction

Consider an image that has been blurred and contaminated by impulsive noise. Image deblurring is an ill-posed inverse problem, hence the noise characteristics play a major role in the mathematical analysis and in the eventual experimental outcome.

Most image deblurring methods rely on the standard model of a shift invariant kernel and additive noise $g = h * f + n$, that is applicable to a large variety of image degradation processes that are encountered in practice. Here h denotes a known space-invariant blur

Figure 1: Current image deblurring algorithms fail in the presence of salt and pepper noise. *Top-left:* Blurred image with Gaussian noise. *Top-right:* Restoration using the method of [45]. *Bottom-left:* Blurred image with salt and pepper noise. *Bottom-right:* Restoration using the method of [45].

kernel (point spread function), f is an ideal version of the observed image g and n is noise. Significant attention has been given to the case of Gaussian noise [7].

Image deblurring algorithms that were designed for Gaussian noise produce inadequate results with impulsive noise, see Fig. 1. The top-left image in Fig. 1 is the 256×256 *Lena* image, blurred by a pill-box kernel of radius 3 (7×7 kernel) and contaminated by Gaussian noise. Adopting the variational approach [10, 39, 40, 45], successful restoration is obtained using the Total Variation deblurring method of Vogel and Oman [45] (top-right). The bottom-left image in Fig. 1 is the same blurred Lena image, now contaminated by

salt and pepper noise of density 0.01. In this case restoration using the method of [45] is clearly inadequate (bottom-right). Note that due to the inadequacy of the noise model, the algorithm of [45] yields poor results even at lower salt and pepper noise density. The same regularization constant was used in Fig. 1 (top-right) and (bottom-right): 10^{-3}. Note that increasing the constant in the presence of salt and pepper noise effectively disables deblurring, while only reducing the amplitude of the noise.

In the absence of algorithms that concurrently deblur and remove impulse noise, the sequential approach is to first denoise the image, then to deblur it. This two-stage method is however prone to failure, especially at high noise density. Image denoising using median-type filtering [20, 28, 36] creates distortion that depends on the neighborhood size; this error can be strongly amplified by the deblurring process, even when using regularized methods. Consider the example shown in Fig. 2. The top-left 256×256 *Einstein* image, was blurred using a pill-box kernel of radius 4. The blurred image with added salt and pepper noise (noise density 0.11) is shown top-right. The outcome of 3×3 median filtering followed by deblurring using the algorithm of [45] is shown bottom-left. At this noise level, the 3×3 neighborhood size of the median filter is insufficient, the noise is not entirely removed, and the residual noise is greatly amplified by the deblurring process. If the neighborhood size of the median filter increases to 5×5, the noise is fully removed, but the distortion leads to inadequate deblurring (bottom-right).

In this paper we present a unified variational method for image deblurring and impulsive noise removal. Establishing a solid theoretical basis to [9], this study integrates and extends the robust statistics, line process (half quadratic) and anisotropic diffusion points of view [11, 12, 34]. The cost functional reflects the need to deblur the image, recover its discontinuities and at the same time remove the impulse noise.

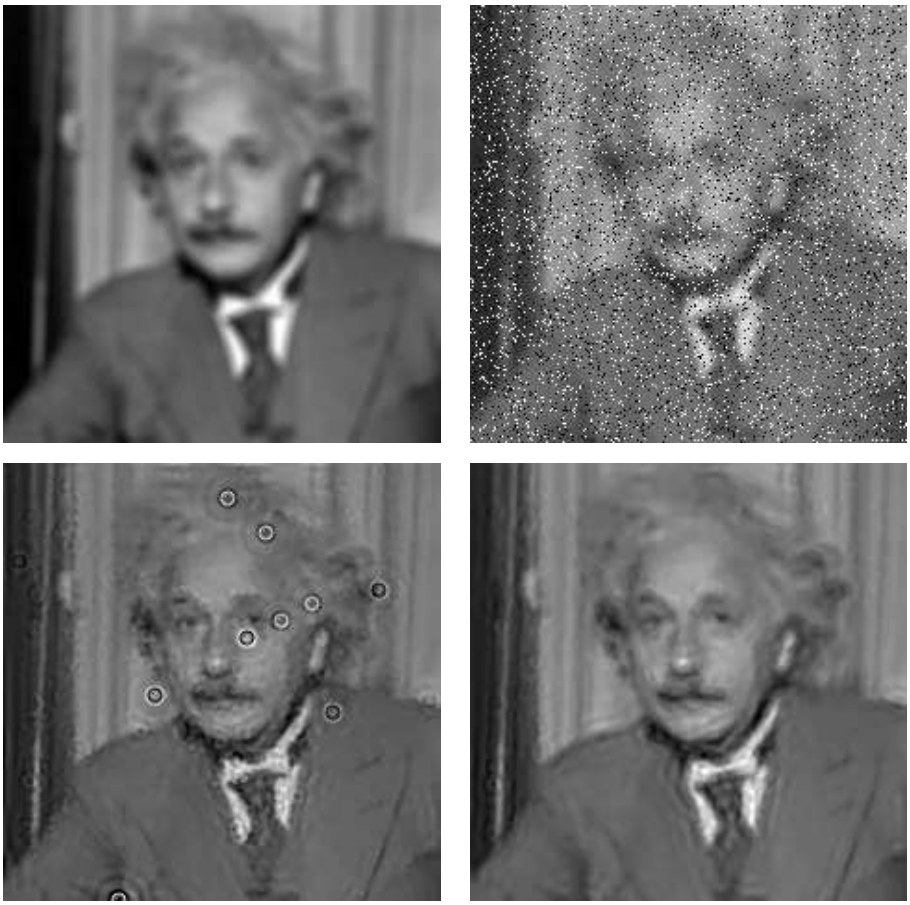

Figure 2: The failure of the two-stage approach to salt-and-pepper noise removal and image deblurring. *Top-left:* Blurred image. *Top-right:* Blurred image contaminated by salt and pepper noise. *Bottom-left:* The outcome of 3×3 median filtering, followed by deblurring. *Bottom-right:* The outcome of 5×5 median filtering, followed by deblurring.

The functional consists of a fidelity term and a regularizer. Data fidelity is quantified using the modified L^1 norm [15,32], that is robust to outliers, i.e., to impulse noise. Elements from the Mumford-Shah functional [31] are used for regularization, expressing compliance with the piecewise-smooth image model [26]. We show that the Mumford-Shah regularizer, in its Γ-convergence approximation [4], can be viewed as an extended line process. It reflects spatial organization properties of the image edges, that do not appear in the common line process [26] or anisotropic diffusion [35]. This allows to distinguish outliers from edges.

The experimental results demonstrate effective image recovery, with various blur models

and noise levels. We also compare the suggested method to the variant of the Mumford-Shah functional described in [41].

2 Unified variational framework

2.1 Fidelity Term

Image deblurring is an inverse problem that can be formulated as a functional-minimization problem. Let Ω denote an open bounded set of $\mathbb{R}^2$, on which the image intensity function $g : \Omega \to [0, 1]$ is defined. Ideally, the recovered image $\hat{f}$ satisfies

$$\hat{f} = \arg\min_f \int_\Omega \Phi(h * f - g)dx, \tag{1}$$

where $\Phi(\cdot)$ is a norm representing data-fidelity. In the case of Gaussian noise, Maximum Likelihood considerations lead to a quadratic data-fidelity term:

$$\Phi(h * f - g) = (h * f - g)^2. \tag{2}$$

The inverse problem represented by (1) is known to be ill-posed due to either the non-uniqueness of the solution, or the numerical instability of the inverse kernel. To alleviate this difficulty, a regularization term, that reflects some a-priori preferences, is added. The functional to be minimized, thus, takes the form

$$\mathcal{F}(f) = \int_\Omega \Phi(h * f - g)dx \ + \alpha\mathcal{J}(f), \tag{3}$$

where $\mathcal{J}(f)$ is the regularization or smoothness operator and α is a positive weighting scalar.

Several regularization terms were suggested in the literature, for example the Tikhonov [44] L^2 norm of the gradient magnitude, the Total Variation (TV) L^1 norm [39, 40], the modified L^1 norm [1], the Beltrami regularization for color images [14, 42], and recently an integrated TV and wavelet coefficient regularization [10, 23, 24, 30].

The quadratic data fidelity term in (2) is incompatible with the salt and pepper noise model. The quadratic function assigns too much weight to distant points. We would like to minimize the effect of such outlier data. This is accomplished by using a robust ρ-function [29].

In this paper, we use a robust (modified L^1 norm) data-fidelity term

$$\Phi(h * f - g) = \sqrt{(h * f - g)^2 + \eta}\ , \tag{4}$$

where η is a small constant. The modified L^1 norm shares the robustness to outliers of the L^1 norm, but prevents the resulting PDE from being singular at zero. Brox *et al* [15] have recently used the modified L^1 norm as a fidelity term for precise optical flow estimation. Note that when $(f * h - g)^2 \ll \eta$, the modified L^1 norm tends to the L^2 norm, since

$$\sqrt{(f * h - g)^2 + \eta} = \sqrt{\eta}\ \sqrt{1 + \frac{(f * h - g)^2}{\eta}} \approx \sqrt{\eta}\ \left(1 + \frac{(f * h - g)^2}{2\eta}\right),$$

which has a quadratic form. The parameter η interpolates, therefore, between the L^1 and L^2 norms.

2.2 Regularization

The regularization that we use gives preference to piecewise-smooth images with simple edge sets. We employ the Mumford-Shah [31] functional in its Γ-convergence approximation. This regularizer has been recently used in electrical impedance tomography [37] and in blind image restoration [8].

The following segmentation functional which was introduced by Mumford and Shah [31], models an image as a set of piecewise smooth segments separated by well-behaved contours. Formally, the pair (f, K) is the minimizer of the following functional:

$$\mathcal{F}(f, K) = \frac{1}{2}\int_{\Omega}(f-g)^2 dx + \beta \int_{\Omega\setminus K} |\nabla f|^2 dx + \alpha \int_{K} d\sigma. \tag{5}$$

Here $f : \Omega \to [0, 1]$ and $g : \Omega \to [0, 1]$ are the model and the observed images, respectively. K denotes the edge set and α, β are positive scalars. The first term stands for data fidelity, piecewise smoothness is favored by the second term, and the third term minimizes the total edge length. Due to the irregularity of this functional, classical calculus of variations methods are not applicable, and approximation approaches have to be used. The Mumford-Shah segmentation functional, which is considered as a free-discontinuity problem [13], can be approximated by regular functionals in the framework of Γ-convergence, as introduced by De Giorgi [22].

The main idea of Γ-convergence is to approximate a functional $\mathcal{F}$ by a sequence $\mathcal{F}_\epsilon$ of regular functionals such that the minimizers of $\mathcal{F}_\epsilon$ approximate the minimizer of $\mathcal{F}$. Let (X, d) be a metric space. A sequence $F_\epsilon : X \to \mathbb{R}_+$ Γ-converges to $F : X \to \mathbb{R}_+$ as $\epsilon \to 0^+$

if for every $f \in X$

$$\begin{aligned} &1. \quad \forall f_\epsilon \to f, \quad \liminf_{\epsilon \to 0^+} F_\epsilon(f_\epsilon) \geq F(f) \\ &2. \quad \exists f_\epsilon \to f, \quad \limsup_{\epsilon \to 0^+} F_\epsilon(f_\epsilon) \leq F(f). \end{aligned}$$

The function F is called the Γ-limit of F_ϵ, denoted by $F = \Gamma\text{-lim}(F_\epsilon)$.

The fundamental theorem of Γ-convergence [13] states that if $F = \Gamma\text{-lim}(F_\epsilon)$ and there is a compact set $K \subset X$ such that $\inf_X F_\epsilon = \inf_K F_\epsilon$ for all ϵ, then there exists $\min_X F = \lim_{\epsilon \to 0^+} \inf_X F_\epsilon$. Moreover, if f_ϵ is the minimizer of F_ϵ and $f_\epsilon \to f$, then f is the minimizer of F. Another important property that we use below is *stability*. It is defined as follows: $\Gamma\text{-lim}(F_\epsilon + V) = F + V$ if $\Gamma\text{-lim}(F_\epsilon) = F$ and $V : X \to \mathbb{R}_+$ is continuous.

Ambrosio and Tortorelli [4] used the Γ-convergence framework to approximate the irregular Mumford-Shah functional by a sequence of regular functionals $\mathcal{F}_\epsilon$. The edge set K was represented by the characteristic function $(1 - \chi_K)$, which was approximated by a smooth auxiliary function v, where $v(x) \approx 0$ if $x \in K$ and $v(x) \approx 1$ otherwise. Thus,

$$\mathcal{F}_\epsilon(f, v) \;=\; \int_\Omega (f - g)^2 \, dx \;+\; G_\epsilon(f, v) \tag{6}$$

where $G_\epsilon(f, v)$ is the image regularization term defined as

$$G_\epsilon(f, v) = \beta \int_\Omega v^2 |\nabla f|^2 dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 \;+\; \frac{(v-1)^2}{4\epsilon} \right) dx. \tag{7}$$

The proof for the Γ-convergence of $G_\epsilon(f, v)$ to the regularization terms of the Mumford-Shah functional in (5) can be found in [13].

With the modified L^1 norm in the fidelity term, and the Mumford-Shah regularization

terms with the Γ-convergence approximation, the functional (3) takes the form

$$\mathcal{F}_\epsilon(f,v) \;=\; \int_\Omega \sqrt{(h*f-g)^2+\eta}\;dx \;+\; G_\epsilon(f,v) \tag{8}$$

with $G_\epsilon(f,v)$ defined as in (7). Due to the stability property of the Γ-convergence, and the continuity of the convolution operator

$$\Gamma\text{-}\lim(\mathcal{F}_\epsilon) = \int_\Omega \sqrt{(h*f-g)^2+\eta}\;dx + \beta\int_{\Omega\setminus K}|\nabla f|^2dx + \alpha\int_K d\sigma,$$

and by the fundamental theorem of the Γ-convergence, the existence of a minimizer is guaranteed.

3 Minimization techniques

The objective functional in (8) depends on the functions f (recovered image) and v (approximated edge map). Minimization with respect to both f and v is carried out using the Euler-Lagrange (E-L) equations (9) and (11), subject to the Neumann boundary conditions $\partial v/\partial N = 0$, $\partial f/\partial N = 0$, where N denotes the normal to the boundary.

$$\frac{\delta\mathcal{F}_\epsilon}{\delta v} = 2\,\beta\,v\,|\nabla f|^2 + \alpha\left(\frac{v-1}{2\epsilon}\right) - 2\,\epsilon\,\alpha\,\nabla^2 v = 0 \tag{9}$$

$$\frac{\delta\mathcal{F}_\epsilon}{\delta f} = \Phi'(h*f-g)*h(-x,-y) - 2\beta\,\mathrm{Div}(v^2\,\nabla f) = 0 \tag{10}$$

Substituting the modified L^1 norm (4) yields

$$\frac{\delta\mathcal{F}_\epsilon}{\delta f} = \frac{(h*f-g)}{\sqrt{(h*f-g)^2+\eta}} * h(-x,-y) - 2\beta\,\mathrm{Div}(v^2\,\nabla f) = 0 \tag{11}$$

Studying the objective functional (8), it can be seen that it is convex, lower bounded and coercive with respect to either function f or v if the other one is fixed. Therefore, following [18], the alternate minimization (AM) approach can be applied: in each step of the iterative procedure we minimize with respect to one function and keep the other one fixed.

Obviously, Eq. (9) is a linear partial differential equation with respect to v. In contrast, (11) is a nonlinear integro-differential equation. *Linearization* of this equation is carried out using the fixed point iteration scheme, as in [18, 45]. In this method, additional iteration index l serves as intermediate stage calculating f^{n+1}. We set $f = f^l$ in the denominator, and $f = f^{l+1}$ elsewhere, where l is the current iteration number. Equation (11) can thus be rewritten as

$$\mathcal{H}(v, f^l) f^{l+1} = G(f^l), \qquad l = 0, 1, ... \tag{12}$$

where $\mathcal{H}$ is the linear integro-differential operator

$$\mathcal{H}(v, f^l) f^{l+1} = \frac{h * f^{l+1}}{\sqrt{(h * f^l - g)^2 + \eta}} * h(-x, -y) - 2\beta \operatorname{Div}(v^2 \nabla f^{l+1})$$

and

$$G(f^l) = \frac{g}{\sqrt{(h * f^l - g)^2 + \eta}} * h(-x, -y). \tag{13}$$

Note that (12) is now a *linear* integro-differential equation in f^{l+1}.

The discretization of equations (9) and (12) yields two systems of linear algebraic equations. These systems are solved in alternation, leading to the following iterative algorithm:

`Initialization:` $f^0 = g, \quad v^0 = 1.$

1. Solve the Helmholtz equation for v^{n+1}

$$(2\beta\,|\nabla f^n|^2 + \frac{\alpha}{2\epsilon} - 2\,\alpha\,\epsilon\,\nabla^2)\,v^{n+1} = \frac{\alpha}{2\epsilon} \tag{14}$$

2. Set $f^{n+1,0} = f^n$ and solve for f^{n+1} (iterating on l)

$$\mathcal{H}(v^{n+1}, f^{n+1,l}) f^{n+1,l+1} = G(f^{n+1,l}) \tag{15}$$

3. if $(||f^{n+1} - f^n||_{L_2} < \varepsilon_1 ||f^n||_{L_2})$ stop.

Here ε_1 is a small positive constant. Both steps 1 and 2 call for a solution of a system of linear equations. Step 1 was implemented using the Minimal Residual algorithm [48]. For the solution of step 2, we followed the quasi-Newtom like method of Vogel and Oman [45]. $f^{n+1,l+1}$ is calculated incrementally, where

$$f^{n+1,l+1} = f^{n+1,l} + \Delta^l.$$

By Eq. (15),

$$\Delta^l = f^{n+1,l+1} - f^{n+1,l} = \left[\mathcal{H}(v^{n+1}, f^{n+1,l})\right]^{-1} G(f^{n+1,l}) - \left[\mathcal{H}(v^{n+1}, f^{n+1,l})\right]^{-1} \mathcal{H}(v^{n+1}, f^{n+1,l}) f^{n+1,l}.$$

Thus

$$\Delta^l = -\left[\mathcal{H}(v^{n+1}, f^{n+1,l})\right]^{-1} R(v^{n+1}, f^{n+1,l}) \tag{16}$$

where

$$R(v, f) = \frac{h * f - g}{\sqrt{(h * f - g)^2 + \eta}} * h(-x, -y) - 2\beta\,\mathrm{Div}(v^2\,\nabla f).$$

It can be shown that the operator $\mathcal{H}(\cdot,\cdot)$ is self-adjoint and positive definite (the proof is given in the appendix). Consequently, $\mathcal{H}(\cdot,\cdot)^{-1}R(\cdot,\cdot)$ in (16) was computed via the Conjugate Gradients method. In matrix presentation with column stack ordering of Δ^l, $\mathcal{H}(\cdot,\cdot)$ is five-diagonal, symmetric and positive definite.

Let f_{ij} denote the discretized image function. The forward and backward finite difference approximations of the derivatives $\partial f(x,y)/\partial x$ and $\partial f(x,y)/\partial y$ are respectively defined by $\Delta^x_{\pm} f_{ij} = \pm(f_{i\pm1,j} - f_{ij})$ and $\Delta^y_{\pm} f_{ij} = \pm(f_{i,j\pm1} - f_{ij})$, and the central finite differences are approximated by $\Delta^x_c f_{ij} = (f_{i+1,j} - f_{i-1,j})/2$ and $\Delta^y_c f_{ij} = (f_{i,j+1} - f_{i,j-1})/2$. Hence, the discrete form of Eq. (14) is

$$\left(2\beta\left[(\Delta^x_c f^n_{ij})^2 + (\Delta^y_c f^n_{ij})^2\right] + \frac{\alpha}{2\epsilon} - 2\alpha\epsilon\left[\Delta^x_+\Delta^x_- + \Delta^y_+\Delta^y_-\right]\right) v^{n+1}_{ij} = \frac{\alpha}{2\epsilon}$$

This equation can be expressed in matrix form as $\mathbf{Mv} = \mathbf{c}$, where the matrix $\mathbf{M}$ is five-diagonal, symmetric and positive definite, $\mathbf{v}$ is the column-stack vector representation of v^{n+1}_{ij}, and $\mathbf{c}$ is a vector of constants. Eq. (15) is approximated by

$$\frac{h * f^{n+1,l+1}}{\sqrt{(h * f^{n+1,l} - g)^2 + \eta}} * h(-x,-y) - 2\beta\left[\Delta^x_+((v^{n+1}_{ij})^2\Delta^x_-) + \Delta^y_+((v^{n+1}_{ij})^2\Delta^y_-)\right] f^{n+1,l+1}_{ij} = G(f^{n+1,l})$$

where $G(f^{n+1,l})$ is defined in Eq. (13).

Our implementation was restricted to rectangular domains. All convolution procedures were performed in the Fourier Transform domain. Special care should be taken to the Neumann boundary conditions. The observed image was extended by adding margins. Their width should be at least half of the kernel support. These margins were obtained by replicating the one-pixel thick outer frame of the image. The margins were then convolved with

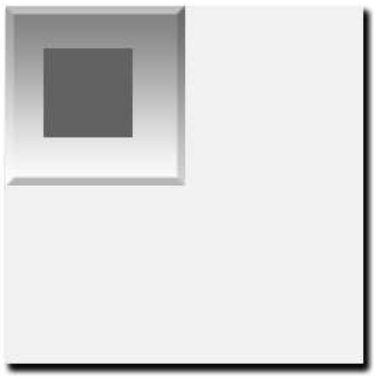

Figure 3: Extensions to the original image during the convolution procedure

the blur kernel. To avoid artifacts, in the presence of salt and pepper noise, care should be taken to ensure that the outer frame of the image is noise free. This limited task can easily be achieved using a median filter. According to the circular convolution theorem, the multiplication of the Discrete Fourier Transform of two signals corresponds to their circular convolution. Zero padding is therefore necessary to eliminate aliasing. Fig. 3 illustrates the image during the convolution procedure. The dark inner rectangle stands for the original image. It is surrounded by the margins approximating the Neumann boundary conditions, and zero-padded in the x and y directions. After the convolution process, the image is cropped back to its original support. The algorithm was implemented in the MATLAB environment.

4 Relation to robust statistics, anisotropic diffusion and line processes

In this section we explore the relations of the regularization terms of the suggested functional (8) to robust statistics, anisotropic diffusion, and line process (or half quadratic formulations).

Images are modelled by piecewise smooth functions. One way to regularize images while satisfying this structure is to consider the gradient image as a smooth function and to regard the gradient values at discontinuities (edges) as outliers. This can be accomplished, for example, in the spirit of M-estimation, by using a robust ρ-function of the gradient magnitude

$$\mathcal{F}(f) = \int_{\Omega} \rho(|\nabla f|)dx. \tag{17}$$

Decreasing this regularization term amounts to edge preserving image smoothing (denoising). In the least-squares approach $\rho(s) = s^2$, which obviously leads to sensitivity to outliers and is not robust. The significance of robust smoothness in this context is that the edges (which have high gradient levels) are preserved. Examples of commonly used ρ-functions are the L^1 norm, Huber's MiniMax and Hampel (see [12] and references therein). It is important for the subsequent analysis to note that some of these functions are not convex. The minimization of (17) can be accomplished by gradient descent

$$\frac{\partial f(x,t)}{\partial t} = \text{Div}\left(\frac{\rho'(|\nabla f|)\nabla f}{|\nabla f|}\right) , \tag{18}$$

where $f(x,0)$ is the given (noisy) image, and t is an artificial time variable.

An alternative denoising process is described by the heat or isotropic diffusion equation,

$$\frac{\partial f(x,t)}{\partial t} = \mathrm{Div}(\nabla f) \ .$$

The solution of this equation is via the convolution of the image with scaled Gaussian kernels. This is a shift invariant operation that ignores the nature of the processed data. It results, therefore, in an over-smoothed image, meaning that the edges are not well preserved. Perona and Malik [35] modified this equation to the so called anisotropic diffusion[1] with

$$\frac{\partial f(x,t)}{\partial t} = \mathrm{Div}(A(|\nabla f|)\nabla f), \tag{19}$$

where $A(|\nabla f|)$ is a smooth and non increasing "edge stopping" function satisfying

$$A(0) = 1, \qquad \lim_{s\to\infty} A(s) = 0,$$

so diffusion is low across the edges (high gradients).

The relation between robust smoothness and anisotropic diffusion was presented by Nordstrom [34] and Black *et al* [11]. The equivalence can be shown by comparing Eq. (18) and (19), with

$$A(s) = \frac{\rho'(s)}{s}.$$

The third approach to edge preserving regularization is the *line process* which was first introduced by Geman and Geman [26], Geman and Reynolds [25] and followed by Charbonnier *et al* [19]. The idea is to express the robust smoothness term in a different but equivalent

[1] This is actually an isotropic and non-homogeneous process. True anisotropic diffusion equations are the Beltrami flow [42], edge-enhancing diffusion [46] and coherence-enhancing diffusion [47].

way, in which the dependence on $|\nabla f|$ is quadratic. This is accomplished by introducing a dual variable $b(x)$ which represents the image edges, such that $b(x) \to 0$ in the presence of an edge and $b(x) \to 1$ otherwise, and a strictly convex and decreasing penalty function $\Psi(b)$. This leads to the *half-quadratic* formulation

$$\mathcal{L}(f) = \int_{\Omega} \min_{0 \le b \le 1} (|\nabla f|^2 b + \Psi(b)) dx. \tag{20}$$

The function $\Psi(b)$ satisfies

$$\varphi(|\nabla f|) = \min_{0 \le b \le 1} (|\nabla f|^2 b + \Psi(b)) \ ,$$

where $\varphi(s)$ is a non-decreasing function, $\theta(s) = \varphi(\sqrt{s})$ is a concave function [25], $\Psi(b) = \theta[(\theta')^{-1}(b)] - b(\theta')^{-1}(b)$ and $b(x)$ for which the minimum of $\varphi(|\nabla f|)$ is reached is unique and given by

$$b(x) = \frac{\varphi'(|\nabla f|)}{2|\nabla f|}.$$

The mechanism that relates robust estimation and line processes was described by Black and Rangarajan [12], thus comparing Eq. (17) and (20) yields the connection

$$\rho(s) = \min_{0 \le b \le 1} (s^2 b + \Psi(b)).$$

Black and Rangarajan [12] also showed that a line process could be extended to embody spatial organization constraints on the outliers. For example, they presented two kinds of spatial interaction terms which enforce edge connectivity: the *hysteresis* term attracts the edges towards unbroken contours while the *non-maximum suppression* term increases the

penalty for parallel multiple edges. The preference for continuous edges is thus reflected by the additional spatial organization term in the cost functional:

$$\mathcal{L}(s) = \int_\Omega \min_{0 \le b \le 1} (s^2 b + \Psi(b) + \text{spatial}(b)) dx \ .$$

Edge preserving denoising experiments validate the superiority of this extension [12]. Teboul *et al* [43] have also integrated the spatial interaction constraint as a robust function $\phi_b(|\nabla b|)$ within the half-quadratic regularization and obtained efficient edge-preserving denoising.

Consider for instance the Geman and McClure [27] robust function (Fig. 4)

$$\rho(s) = \frac{s^2}{1+s^2} \ . \tag{21}$$

The corresponding half-quadratic penalty function takes the form

$$\Psi(b) = (\sqrt{b} - 1)^2 \ ,$$

and the corresponding edge stopping function in the anisotropic diffusion equation is

$$A(s) = \frac{2}{(1+s^2)^2} \ .$$

The denoising process based on this non convex ρ-function is prone to mathematical difficulties. Catté et al. [16] showed that if $sA(s)$ is a non increasing function then, Perona and Malik's equation is ill-posed. This can be easily verified in the one dimensional case:

$$\frac{\partial f}{\partial t} = (A(f')f')' = (A'(f')f' + A(f'))\, f'' = \frac{d}{ds} sA(s)\bigg|_{s=f'} f'' \ .$$

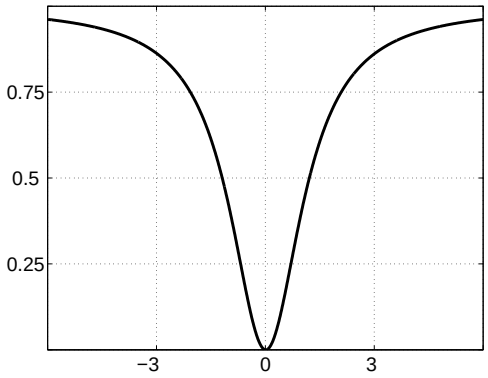

Figure 4: Geman and McClure robust function

If $sA(s)$ decreases at some point, the expression that multiplies the second derivative is negative. This results in inverse heat equation or backward diffusion process. The robust function of Geman and McClure (21) corresponds to

$$sA(s) = \frac{2s}{(1+s^2)^2} \ .$$

This expression is decreasing in most of its domain (see Fig. 5), and the anisotropic flow is therefore unstable. The solution suggested by Catté *et al* was to employ a selective smoothing term

$$\frac{\partial f(x,t)}{\partial t} = \mathrm{Div}\left(A(|DG_\sigma * f|)\nabla f\right) \ ,$$

where G_σ is the Gaussian function and D is the derivative operator. A regularity proof for this equation is given in [16]. Another facet of the inherent instability of the process based on the Geman-McClure robust function is revealed in the associated denoising functional:

$$\mathcal{F}(f) = \frac{1}{2}\int_\Omega (f-g)^2 dx + \lambda \int_\Omega \frac{|\nabla f|^2}{1+|\nabla f|^2} dx \ . \tag{22}$$

Here f and g are the model and observed images, respectively, and the regularizer is of the

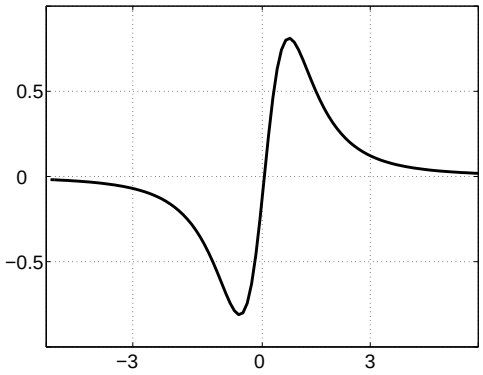

Figure 5: $sA(s)$ that corresponds to the robust function of Geman and McClure.

form (21). Chipot *et al* [21] proved that $\mathcal{F}(f)$ does not have a minimum. However, they showed that if the functional (22) is perturbed as follows

$$\mathcal{F}_\eta(f) = \frac{1}{2}\int_\Omega (f-g)^2 dx + \lambda \int_\Omega \left(\frac{|\nabla f|^2}{1+|\nabla f|^2} + \eta |\nabla f|^2 \right) dx, \tag{23}$$

then $\mathcal{F}_\eta(f)$ has a minimizer.

5 Relation to the Mumford-Shah functional

Careful examination of Eq. (22) with the Geman and McClure robust function, shows a relation to the Γ-convergence version of the Mumford-Shah functional (6). If the $\rho(|\nabla f|)$ function is replaced by its corresponding line process,

$$\frac{|\nabla f|^2}{1+|\nabla f|^2/\gamma} \longleftrightarrow b|\nabla f|^2 + \gamma(\sqrt{b}-1)^2, \tag{24}$$

then by substituting $b = v^2$, Eq. (22) can be rewritten as

$$\mathcal{F}(f, v) = \frac{1}{2}\int_{\Omega}(f-g)^2 dx + \lambda \int_{\Omega}\left(v^2|\nabla f|^2 + \gamma(v-1)^2\right) dx. \tag{25}$$

Note the relation between (25) and (6,7): if $\beta = \lambda$, and $\alpha/4\varepsilon = \lambda\gamma$, then both equations are identical except for the $|\nabla v|^2$ term in (7).

This relation was also shown by Teboul *et al* [43] where an interaction constraint was added to the line process, and by Rosati [38] who showed the relation of the discrete version of the Geman and McClure function to the Mumford-Shah functional. As it was shown, the Mumford-Shah functional in its Γ-convergence approximation, is actually an *extended* line process where the penalty function $\Psi(b)$ corresponds to the robust Geman and McClure function. This observation explains the advantages of Mumford-Shah regularization. It applies a robust function for the detection of edges while demanding that these edges are smooth and continuous. This combination does not admit impulse noise as an edge and smoothes it. The robust fidelity term regards the impulse noise point as an outlier and allows the necessary large change of value without penalty. From a mathematical point of view, the problem is well posed and a minimizer exists. This follows from Ambrosio and Tortorelli's proof [4], together with the fact that the fidelity term (both in the denoising and deconvolution cases) is continuous with respect to the image f. The stability criterion completes the proof. Therefore, the Mumford-Shah regularizer is more general and advantageous with respect to anisotropic diffusion or robust smoothing.

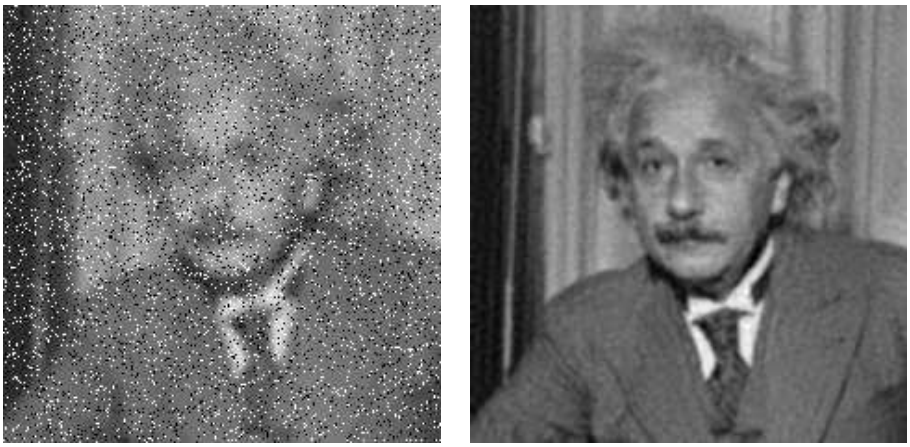

Figure 6: Deblurring in the presence of salt and pepper noise. *Left:* Source image, blurred with a pill-box kernel of radius 4, and degraded by noise of density 0.11. *Right:* Recovered image, using the proposed algorithm.

6 Experimental results

Consider the blurred and noisy version of the *Einstein* image, shown in Fig. 6 (left). The blur kernel is a pill-box of radius 4; the noise density is 0.11. Fig. 6 (right) is the outcome of the suggested method. The parameters are $\beta = 0.5, \alpha = 0.5, \epsilon = 0.1$. The superiority of the proposed method with respect to the sequential one (Fig. 2) is evident.

In all examples in this section the convergence tolerance of $\varepsilon_1 = 10^{-4}$ is reached with 3-5 external iterations (over n). The number of internal iterations (over l) is set to 5. The constant η (Eq. 4) is set to 10^{-4}.

The examples presented in Fig. 7 demonstrate the performance of the algorithm at several noise levels. The images in the left column are all blurred by a pill-box kernel of radius 3. The noise densities are, from top to bottom, 0.01, 0.1 and 0.3. The corresponding recovered images are shown in the right column. In all three cases $\alpha = 0.5$ and $\epsilon = 0.1$, while $\beta = 0.05,\ 0.1$ and 0.5 respectively. Clearly, the image smoothness weight β should be increased with noise level.

Figure 7: *Left column:* The *Lena* image blurred with a pill-box kernel of radius 3, and contaminated by salt and pepper noise. The noise density is (top to bottom) 0.01, 0.1 and 0.3. *Right column:* The corresponding recovered images.

Figure 8: The case of motion blur. *Top-left:* Blurred and noisy image. *Top-right:* Restoration using the proposed method. *Bottom-left:* The outcome of 3×3 median filtering followed by TV [45] restoration. *Bottom-right:* The outcome of 5×5 median filtering followed by TV [45] restoration.

Recovery of motion blur in the presence of salt and pepper noise is demonstrated in Fig. 8. The 256×256 *cameraman* image is blurred by a motion blur kernel of length 8, oriented at an angle $\theta = 25^o$ with respect to the horizon. The blurred image was further degraded by salt and pepper noise of density 0.1 (top-left). The outcome of the method suggested in this paper (with $\beta = 0.6, \alpha = 0.01, \epsilon = 0.1$) is shown top-right. The inadequacy of the sequential strategy, of median filtering followed by the Total Variation (TV) deconvolution [45] is demonstrated in the bottom row. The left image in that row is the outcome of 3×3 median filtering followed by the TV restoration. The bottom-right image was obtained in a similar

Figure 9: Deblurring in the presence of random impulsive noise. *Left:* Source image, blurred with a pill-box kernel and contaminated by random impulse noise of 0.1 density. *Right:* Recovered image, using the suggested algorithm.

way, but with a 5×5 median filter.

In Fig. 9, the suggested algorithm is tested in the presence of impulsive noise with *random* intensity. In this case, the image is blurred by a pill-box kernel of radius 3 and the noisy pixels are set to random intensity values in the range [0,1]. It is clear that the suggested algorithm removes outliers of random intensity, and is not limited to the white and black salt-and-pepper case. In this example the noise density is 0.1 and the parameters are set to $\beta = 0.5, \alpha = 0.1, \epsilon = 0.1$ as in the case of 10% noise in Fig. 7.

We proceed to compare the suggested method to a sophisticated two-stage restoration process. Fig. 10 top-left is the observed image, blurred by a pill-box kernel of radius 3 with salt and pepper noise of density 0.1. The top-right image is the outcome of denoising with the Mumford-Shah functional (6) ($\beta = 0.9, \alpha = 0.5, \epsilon = 0.1$). At the bottom-left is the restored image using the TV method of Vogel and Oman [45], with a weight factor of 10^{-4}. It is apparent that the noise, that had not been entirely removed in the first stage, is amplified in the deconvolution process. The image recovered using the suggested method is shown bottom-right. The unified approach performs the noise and restoration

Figure 10: Two-stage deblurring. *Top-left:* Source image, blurred with a pill-box kernel of radius 3, and contaminated by noise of density 0.1. *Top-right:* Denoised by the Mumford-Shah method. *Bottom-left:* TV Deconvolution of the denoised image. *Bottom-right:* The image recovered using the algorithm suggested in this paper.

tasks simultaneously, and therefore yields better recovery.

We next compare our method to an alternative unified denoising and restoration method, employing modified L^1 regularization instead of the Mumford-Shah terms. The competing functional is thus

$$\mathcal{F}(f) = \int_\Omega \sqrt{(f * h - g)^2 + \eta}\, dx + \lambda \int_\Omega \sqrt{|\nabla f|^2 + t}\, dx. \tag{26}$$

The results are presented in Fig. 11. In this case, the original image is blurred by a pill-box of radius 3 with salt and pepper noise density of 0.3. Top-left is the outcome of minimizing (26)

with $\lambda = 0.1$ and $t = 10^{-10}$. Top-right is the image recovered using the method developed in this paper. The bottom row shows magnifications of the images in the top one. It can be seen that the suggested algorithm produces a cleaner image.

A possible explanation to the better performance of the proposed method is the robustness of the regularizer. The Mumford-Shah terms are associated with a function that is more robust than the L^1 norm (see Fig. 12) and therefore better deals with edges. It demands edges to be continuous and, unlike the L^1 method, does not confuse impulse noise with edges. Edges are therefore better preserved and noise is better removed. Additional support to this observation can be found in Aubert *et al* [6]. They have recently shown that the Geman and McClure robust smoothing function, embedded in a supervised classification functional, yields better results than convex smoothing functions.

A variant of the Mumford-Shah functional in its Γ-convergence approximation was suggested by Shah [41]. In this version the L^2 norm of $|\nabla f|$ in (7) was replaced by its L^1 norm in the first term of G_ϵ

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

The Γ-convergence of this form was later proved by Alicandro *et al* [2]. The Mumford-Shah and Shah regularizers are compared in Fig. 13. The 256×256 *Window* image was blurred by a pill-box kernel of radius 3 with noise densities 0.01, 0.1 and 0.2 (left column top to bottom). The results of the restoration using the Mumford-Shah stabilizer are presented in the middle column and the images recovered using the Shah regularizer are shown in the right column. The recovery using both methods is satisfactory, but it can be clearly seen that while the Mumford-Shah restoration performs better in the high-frequency image

Figure 11: *Left column:* Restoration using L^1 regularization. *Right column:* Restoration using the proposed method.

content (see the shades for instance), the Shah restoration attracts the image towards the piecewise constant or cartoon limit which yields much cleaner images. This can be explained by the fact that the Shah regularizer is more robust to image gradients and hence eliminates high-frequency contributions.

A promising variational method for pure impulse denoising (no blurring) was proposed by Nikolova *et al* [17, 32, 33]. We degenerated our method to this special case and compared it to [33]. The salt-and-pepper noise density was 0.02. The images in the left column of Fig. 14 show the outcome of the algorithm of [33] with L^1 norm for both the fidelity and regularization, and with a weight factor of 0.01. The recovery using the suggested method is shown in the right column ($\beta = 0.5, \alpha = 0.5, \epsilon = 0.3$). It can be observed that the better

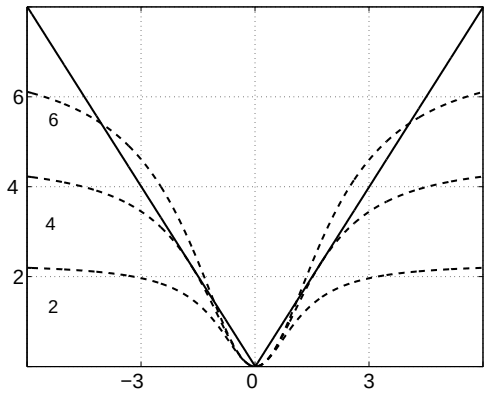

Figure 12: The Geman-McClure function (24) for several γ values (dashed) vs. the L^1 robust function. The superior performance of Mumford-Shah regularization can be explained by the better robustness of the Geman-McClure function with respect to the L^1 function.

robustness of the suggested algorithm leads to better performance in the presence of salt and pepper noise.

Another useful outcome of the suggested method is the auxiliary function v – an approximated edge map of the image. For example, Fig. 15 shows the v-maps obtained as part of the processing of the blurred and noisy *Lena* (pill-box blur, Fig. 7) and *Cameraman* (motion-blur, Fig. 8) images.

7 Discussion

We present in this paper a method for image deblurring in the presence of impulsive (e.g., salt and pepper) noise. Our unified approach to deblurring and outlier removal is novel and unique. Experimental results demonstrate the superiority of the suggested method with respect to sequential approaches, in which noise removal and image deconvolution are separate steps.

The algorithm is fast, robust and stable. Computation time for 256×256 images is about 3 minutes, using interpreted MATLAB on a 2GHz PC. The robustness of the algorithm is

Figure 13: *Left column:* The *Window* image blurred with a pill-box kernel of radius 3, and contaminated by salt and pepper noise. The noise density is (top to bottom) 0.01, 0.1 and 0.2. *Middle column:* The corresponding recovered images with Mumford-Shah regularization. *Right column:* The corresponding recovered images with Shah regularization.

Figure 14: *Left column:* Restoration using the L^1 regularization [33]. *Right column:* Restoration using the proposed method.

demonstrated by the fact that similar parameters can be used in the processing of different images with the same noise level. The numerical convergence is fast.

In the variational approach, image deblurring in the presence of noise is expressed as a functional minimization problem. The functional consists of a data fidelity term and a regularization term, that stabilizes the inherent ill-posedness of the image deconvolution problem. The data fidelity term used in this study is the modified L^1 norm. It is more robust than the common L^2 norm, and is thus more suitable for images contaminated by outliers. Yet, it is differentiable and convex.

Elements from the Mumford-Shah segmentation functional, in the Γ-convergence formulation, serve as the regularization term. They reflect the underlying piecewise-smooth image

Figure 15: Approximated edge maps obtained as a by-product of the restoration process. *Left:* The v-function that corresponds to the deblurring of the *Lena* image with a pill-box kernel and noise density 0.1. *Right:* The v-function that corresponds to the deconvolution of the *Cameraman* image with motion-blur and noise density 0.1.

model, and in addition, guarantee the existence of a minimizer to the problem. Exploring these terms from the robust statistics point of view shows that this regularization is an extended line process derived from the Geman and McClure robust function. This has the theoretical and mathematical advantages of being robust to large gradients (and noise), while preferring structured or smooth edges. The alternative edge-preserving stabilizer, the L^1 norm in the Total Variation approach, is less robust to outliers than the Geman and McClure function (Figs. 11,14). Another advantage of the proposed regularization terms is that they do not induce nonlinearity. The only non-linearity in our approach comes from the fidelity term. Finally, the extraction of the edge map is a useful by-product of the suggested restoration method.

Acknowledgment

This research was supported by MUSCLE: Multimedia Understanding through Semantics, Computation and Learning, a European Network of Excellence funded by the EC 6th Frame-

work IST Programme. It was supported also by the Israel Science Foundation.

Appendix

Theorem. *The operator $A(v)$ defined as*

$$A(v)f = h(-x,-y) * \left(\frac{h(x,y) * f(x,y)}{C(x,y)}\right) - 2\beta Div(v^2 \nabla f(x,y)) \tag{27}$$

with

$$C(x,y) = \sqrt{(h * \tilde{f} - g)^2 + \eta^2} \tag{28}$$

is self adjoint and positive definite.

Note that $\tilde{f}$ in $C(x,y)$ refers to $f(x,y)$ in the previous iteration.

Proof. Let $A(v) = A^I + A^{II}$, where

$$A^I f = h(-x,-y) * \left(\frac{h(x,y) * f(x,y)}{C(x,y)}\right) \tag{29}$$

and

$$A^{II} f = -2\beta \mathrm{Div}(v^2 \nabla u) \ .$$

Let $u, v \in H$ be arbitrary real functions in Hilbert space. $\mathcal{H}$ is a continuous linear operator, and $< \cdot, \cdot >$ denotes the inner product. The adjoint operator $\mathcal{H}^*$ satisfies

$$< u, \mathcal{H}^* v > = < \mathcal{H} u, v > \tag{30}$$

or

$$\int u\mathcal{H}^* v dx = \int (\mathcal{H}u) v dx. \tag{31}$$

$\mathcal{H}$ is a *self adjoint* operator if $\mathcal{H} = \mathcal{H}^*$.

Let $\mathcal{H}u = h(x) * u(x)$, then

$$< \mathcal{H}^* u, v > \;=\; < u, \mathcal{H}v > \;=\; \int u \cdot (h * v) dx = \int [h(-x) * u(x)] \cdot v(x) dx.$$

hence $\mathcal{H}^* u(x) = h(-x) * u(x)$. In the same manner, for $\mathcal{M}u = M(x)u(x)$,

$$< \mathcal{M}^* u, v > \;=\; < u, \mathcal{M}v > \;=\; \int u(x) \cdot [M(x)v(x)]\, dx = \int [M(x)u(x)] \cdot v(x) dx$$

so $\mathcal{M}^* u(x) = M(x)u(x)$, and $\mathcal{M}$ is a self-adjoint operator. Equation (29) can now be rewritten as

$$A^I f = \mathcal{H}^* \mathcal{M} \mathcal{H} f. \tag{32}$$

Since $M(x) = 1/C(x)$ is a real positive function, it can be decomposed as $M(x) = D(x) \cdot D(x)$, where $D(x) = 1/\sqrt{C(x)}$. Together with the fact that multiplication operator $\mathcal{M}$ is self adjoint, Eq. (32) can be replaced by

$$A^I = \mathcal{H}^* \mathcal{D}^* \mathcal{D} \mathcal{H}.$$

Recall that

$$(ABCD)^* = D^* C^* B^* A^*,$$

thus

$$A^{I*} = (\mathcal{H}^*\mathcal{D}^*\mathcal{D}\mathcal{H})^* = \mathcal{H}^*\mathcal{D}^*\mathcal{D}\mathcal{H} = A^I,$$

verifying that A^I is a self adjoint operator. Now,

$$\begin{aligned} < g, A^I g > &= \int g\mathcal{H}^*\mathcal{D}^*\mathcal{D}\mathcal{H}g \; dx = \int g(\mathcal{D}\mathcal{H})^*\mathcal{D}\mathcal{H}g \; dx = \\ &= < g, (\mathcal{D}\mathcal{H})^*\mathcal{D}\mathcal{H}g > = < \mathcal{D}\mathcal{H}g, \mathcal{D}\mathcal{H}g > = \int |\mathcal{D}\mathcal{H}g|^2 \; dx > 0 \end{aligned} \quad (33)$$

for all functions g that are not identically zero, which proves that A^I is positive definite.

For the second part of the operator,

$$< g, A^{II} f >= -2\beta \int_\Omega g \left[\nabla \cdot \left(v^2 \nabla f\right)\right] dx \; .$$

Recall that for a scalar function g and a vector field ϕ,

$$g\nabla \cdot \phi = \nabla \cdot (g\phi) - \nabla g \cdot \phi, \quad (34)$$

thus

$$< g, A^{II} f > = -2\beta \int_\Omega \nabla \cdot (gv^2 \nabla f) dx + 2\beta \int_\Omega \nabla g \cdot v^2 \nabla f dx.$$

Applying the divergence theorem,

$$< g, A^{II} f > = -2\beta \int_{\partial\Omega} \left(gv^2 \nabla f\right) \cdot dn + 2\beta \int_\Omega \nabla g \cdot v^2 \nabla f dx \quad (35)$$

Using the Neumann boundary condition, the first term vanishes, hence

$$< g, A^{II} f > \;=\; 2\beta \int_\Omega \nabla g \cdot v^2 \nabla f dx \;=\; 2\beta \int_\Omega \nabla f \cdot v^2 \nabla g \, dx \tag{36}$$

Substituting Eq. (34) in (36) and using the divergence theorem we obtain

$$< g, A^{II} f > \;=\; 2\beta \int_{\partial\Omega} \left(g v^2 \nabla f\right) \cdot dn \;-\; 2\beta \int_\Omega f \nabla \cdot \left(v^2 \nabla g\right) dx$$

Applying again the Neumann boundary condition, the first term vanishes and we observe that

$$< g, A^{II} f > \;=\; -\,2\beta \int_\Omega \nabla \cdot \left(v^2 \nabla g\right) f \, dx \;=\; < A^{II} g, f >$$

which proves that the operator A^{II} is self adjoint. We proceed to show that A^{II} is positive definite.

$$< g, A^{II} g > \;=\; -\,2\beta \int_\Omega g \nabla \cdot \left(v^2 \nabla g\right)$$

Applying Eq. (34) and using the divergence theorem yields

$$< g, A^{II} g > \;=\; -\,2\beta \int_{\partial\Omega} \left(g v^2 \nabla g\right) \cdot dn \;+\; 2\beta \int_\Omega \nabla g \cdot \left(v^2 \nabla g\right) dx$$

The first term vanishes due to the Neumann boundary condition, thus

$$< g, A^{II} g > \;=\; 2\beta \int_\Omega v^2 \left|\nabla g\right|^2 dx \;>\; 0$$

We conclude that A^{II} is positive definite. Since both A^I and A^{II} are self adjoint and positive definite, their sum $A(v)$ is also self adjoint and positive definite. □

References

[1] R. Acar and C.R. Vogel, "Analysis of Total Variation Penalty Methods", *Inverse Problems*, Vol. 10, pp. 1217-1229, 1994.

[2] 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, pp. 17-37, 1999.

[3] G.R. Arce, J.L. Paredes and J. Mullan, "Nonlinear Filtering for Image Analysis and Enhancement", in A.L. Bovik (Ed.), *Handbook of Image & Video Processing*, Academic Press, 2000.

[4] L. Ambrosio and V.M. Tortorelli, "Approximation of Functionals Depending on Jumps by Elliptic Functionals via Γ-Convergence", *Communications on Pure and Applied Mathematics*, Vol. XLIII, pp. 999-1036, 1990.

[5] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer, New York, 2002.

[6] G. Aubert, L. Blanc-Féraud and R. March, "Γ-Convergence of Discrete Functionals with Nonconvex Perturbation for Image Classification", *SIAM Journal of Numerical Analysis*, Vol. 42, pp. 1128-1145, 2004.

[7] M. Banham and A. Katsaggelos, "Digital Image Restoration", *IEEE Signal Processing Mag.*, Vol. 14, pp. 24-41, 1997.

[8] L. Bar, N. Sochen and N. Kiryati, "Variational Pairing of Image Segmentation and Blind Restoration", Proc. ECCV'2004, Prague, Czech Republic, Part II: *LNCS* #3022, pp. 166-177, Springer, 2004.

[9] L. Bar, N. Sochen and N. Kiryati, "Image Deblurring in the Presence of Salt and Pepper Noise", Proc. Scale-Space 2005, Hofgeismar, Germany: *LNCS* #3459, pp. 107-118, Springer, 2005.

[10] J. Bect, L. Blanc-Féraud, G. Aubert and A. Chambolle, "A l^1-Unified Variational Framework for Image Restoration", Proc. ECCV'2004, Prague, Czech Republic, Part IV: *LNCS* #3024, pp. 1-13, Springer, 2004.

[11] M.J. Black, G. Sapiro, D. Marimont, and D. Heeger, "Robust Anisotropic Diffusion", *IEEE Trans. Image Processing*, Vol. 7, pp. 421-432, 1998.

[12] M.J. Black and A. Rangarajan, "On the Unification of Line Processes, Outlier Rejection, and Robust Statistics with Applications in Early Vision", *International Journal of Computer Vision*, Vol. 19, pp. 57-92, 1996.

[13] Braides A., *Approximation of Free-Discontinuity Problems*, Lecture Notes in Mathematics, Vol. 1694, Springer-Verlag, 1998.

[14] A. Brook, R. Kimmel and N. Sochen, "Variational Segmentation for Color Images", *International Journal of Computer Vision*, Vol. 18, pp. 247-268, 2003.

[15] T. Brox, A. Bruhn, N. Papenberg and J. Weickert, "High Accuracy Optical Flow Estimation Based on a Theory for Warping", Proc. ECCV'2004, Prague, Czech Republic, Part IV: *LNCS* #3024, pp. 25-36, Springer, 2004.

[16] F. Catté, P.L. Lions, J.M. Morel, and T. Coll, "Image Selective Smoothing and Edge Detection by Nonlinear Diffusion", *SIAM Journal of Numerical Analysis*, Vol. 29, pp. 182-193, 1992.

[17] R.H. Chan, C. Ho and M. Nikolova, "Salt-and-Pepper Noise Removal by Median-type Noise Detectors and Detail-preserving Regularization", to appear.

[18] T.F. Chan and C. Wong, "Total Variation Blind Deconvolution", *IEEE Trans. Image Processing*, Vol. 7, pp. 370-375, 1998.

[19] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic Edge-Preserving Regularization in Computed Imaging", *IEEE Trans. Image Processing*, Vol. 6, pp. 298-311, 1997.

[20] T. Chen and H.R. Wu, "Space Variant Median Filters for the Restoration of Impulse Noise Corrupted Images", *IEEE Trans. Circuits and Systems II*, Vol. 48, pp. 784-789, 2001.

[21] M. Chipot, R. March, M. Rosati, and G. Vergara Caffarelli, "Analysis of a Nonconvex Problem Related to Signal Selective Smoothing", *Mathematical Models and Methods in Applied Science*, Vol. 7, pp. 313-328, 1997.

[22] Dal Maso G., *An Introduction to* Γ*-Convergence*, Progress in Nonlinear Differential Equations and their Applications, Birkhauser, 1993.

[23] S. Durand and J. Froment, "Reconstruction of Wavelet Coefficients Using Total Variation Minimization", *SIAM Journal of Scientific Computing*, Vol. 24, pp. 1754-1767, 2003.

[24] S. Durand and M. Nikolova, "Restoration of Wavelet Coefficients by Minimizing a Specially Designed Objective Function", Proc. IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision, pp. 145-152, 2003.

[25] D. Geman and G. Reynolds, "Constrained Restoration and the Recovery of Discontinuities", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, pp. 367-383, 1992.

[26] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 721-741, 1984.

[27] S. Geman and D.E. McClure, "Statistical Methods for Tomographic Image Reconstruction", *Bulletin of the International Statistical Institute*, LII-4, pp. 5-21, 1987.

[28] H. Hwang and R. A. Haddad, "Adaptive Median Filters: New Algorithms and Results", *IEEE Trans. Image Processing*, Vol. 4, pp. 499-502, 1995.

[29] P.J. Huber, *Robust Statistics*, John Wiley and Sons, New York, 1981.

[30] F. Malgouyres, "Minimizing the Total Variation Under a General Convex Constraint", *IEEE Trans. Image Processing*, Vol. 11, pp. 1450-1456, 2002.

[31] D. Mumford and J. Shah, "Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems", *Communications on Pure and Applied Mathematics*, Vol. 42, pp. 577-684, 1989.

[32] M. Nikolova, "Minimizers of Cost-Functions Involving Nonsmooth Data-Fidelity Terms: Application to the Processing of Outliers", *SIAM Journal on Numerical Analysis*, Vol. 40, pp. 965-994, 2002.

[33] M. Nikolova, "A Variational Approach to Remove Outliers and Impulse Noise", *Journal of Mathematical Imaging and Vision*, Vol. 20, pp. 99-120, 2004.

[34] N. Nordstrom, "Biased Anisotropic Diffusion - A Unified Regularization and Diffusion Approach to Edge Detection", Proc. 1st European Conference on Computer Vision, pp. 18-27, Antibes, France, 1990.

[35] P. Perona and J. Malik, "Scale Space and Edge Detection Using Anisotropic Diffusion", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 629-639, 1990.

[36] G. Pok, J.-C. Liu and A.S. Nair, "Selective Removal of Impulse Noise based on Homogeneity Level Information", *IEEE Trans. Image Processing*, Vol. 12, pp. 85-92, 2003.

[37] L. Rondi and F. Santosa, "Enhanced Electrical Impedance Tomography via the Mumford-shah Functional", *ESAIM: Control, Optimization and Calculus of Variations*, Vol. 6, pp. 517-538, 2001.

[38] M. Rosati, "Asymptotic Behavior of a Geman and McClure Discrete Model", *Applied Math. Optim.*, Vol. 41, pp. 51-85, 2000.

[39] L. Rudin, S. Osher and E. Fatemi, "Non Linear Total Variation Based Noise Removal Algorithms", *Physica D*, Vol. 60, pp. 259-268, 1992.

[40] L. Rudin and S. Osher, "Total Variation Based Image Restoration with Free Local Constraints", *Proc. IEEE ICIP*, Vol. 1, pp. 31-35, Austin TX, USA, 1994.

[41] J. Shah, "A Common Framework for Curve Evolution, Segmentation and Anisotropic Diffusion", Proc. IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, pp. 136-142, 1996.

[42] N. Sochen, R. Kimmel and R. Malladi, "A General Framework for Low level Vision" *IEEE Trans. Image Processing*, Vol. 7, pp. 310-318, 1998.

[43] S. Teboul, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Variational Approach for Edge-Preserving Regularization Using Coupled PDE's", *IEEE Trans. Image Processing*, Vol. 7, pp. 387-397, 1998.

[44] A. Tikhonov and V. Arsenin, *Solutions of Ill-posed Problems*, New York, 1977.

[45] C. Vogel and M. Oman, "Fast, Robust Total Variation-based Reconstruction of Noisy, Blurred Images", *IEEE Trans. Image Processing*, Vol. 7, pp. 813-824, 1998.

[46] J. Weickert, "Anisotropic Diffusion Filters for Image Processing Based Quality Control", Proc. Seventh European Conference on Mathematics in Industry, Teubner, Stuttgart, pp. 355-362, 1994.

[47] J. Weickert, "Coherence-Enhancing Diffusion Filtering", *International Journal of Computer Vision*, Vol. 31, pp. 111-127, 1999.

[48] E.W. Weisstein *et al*, "Minimal Residual Method", from *MathWorld–A Wolfram Web Resource*. `http://mathworld.wolfram.com/MinimalResidualMethod.html`