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

April 18, 2005

Abstract

The problem of image deblurring in the presence of impulsive noise and in particular salt and pepper noise is considered. Standard image deconvolution algorithms, that are designed for Gaussian noise, do not perform well in this case. Median type filtering is a common method for salt and pepper noise removal. Deblurring an image that has been preprocessed by median-type filtering is however difficult, due to the amplification (in the deconvolution stage) of median-induced distortion. A unified variational approach to salt and pepper noise removal and image deblurring is presented. An objective functional that represents the goals of deblurring, noise-robustness and compliance with the piecewise-smooth image model is formulated. A modified L^1 data fidelity term integrates deblurring with robustness to outliers. Elements from the Mumford-Shah functional, that favor piecewise smooth images with simple edge-sets, are used for regularization. Promising experimental results together theoretical considerations are shown for several blur models.

1 Introduction

Consider an image that has been blurred and contaminated by salt and pepper noise. Typical sources of blur are defocus and motion [5]. Salt and pepper noise is a common model for the effects of bit errors in transmission, malfunctioning pixels and faulty memory locations [10].

Significant attention has been given to image deblurring in the presence of Gaussian noise [5]. We focus on variational methods, that have an important role in modern image deblurring research, see e.g. [25, 34, 35, 38]. Most methods rely on the standard model $g =$

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 [38]. *Bottom-left:* Blurred image with salt and pepper noise. *Bottom-right:* Restoration using the method of [38].

$h * f + n$, that is applicable to a large variety of image degradation processes that are encountered in practice. Here h represents a known space-invariant blur kernel (point spread function), f is an ideal version of the observed image g and n is (usually Gaussian) noise. In this research, we focus on the case of salt and pepper noise.

The assumption of Gaussian noise is a fundamental element of common image deblurring algorithms. It is therefore not surprising that those algorithms produce inadequate results in the presence of salt and pepper noise. This fact is illustrated in 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. Successful restoration is obtained using the state of the art deblurring method of [38] (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 [38] is clearly inadequate (bottom-right). Note that due to the inadequacy of the noise model, the algorithm of [38] yields poor results even at lower salt and pepper noise density. The regularization constants used to obtain Fig. 1 (top-right) and (bottom-right) are the same: 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.

Salt and pepper noise removal is considered in the literature by itself. It is commonly approached using median filters and their variants, see e.g [16, 24, 31]. Recently, a promising variational method for impulse denoising was proposed by [13, 28, 29].

In the absence of unified algorithms for deblurring and salt-and-pepper noise removal, the straightforward 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 creates distortion that depends on the neighborhood size; this error can be strongly amplified by the deblurring process, even in regularized methods. Consider the example shown in Fig. 2. The top-left image is the 256×256 *Einstein* image, 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 [38] 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 is increased to 5×5, the noise is fully removed, but the distortion leads to

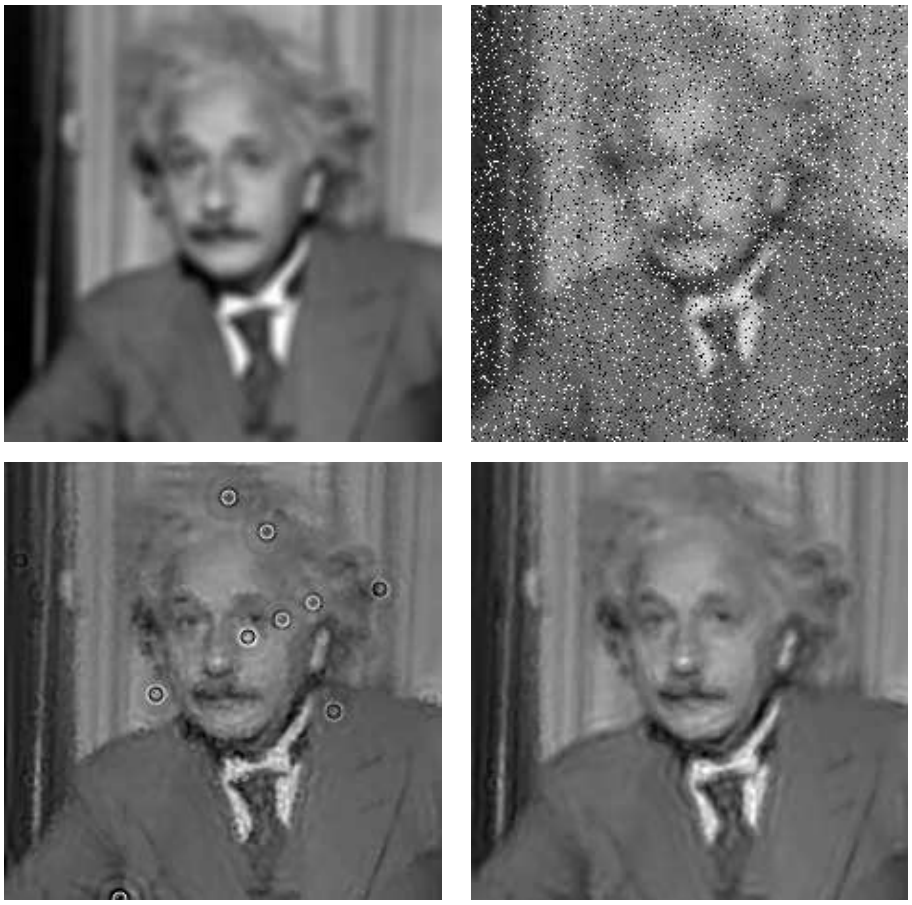

Figure 2: The failure of the two-stage approach to salt-and-pepper noise removal and image deblurring. *Top-left:* Blurred image. *Top-right:* Blurred image contaminated by salt and pepper noise. *Bottom-left:* The outcome of 3×3 median filtering, followed by deblurring. *Bottom-right:* The outcome of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

inadequate deblurring (bottom-right).

In this paper we present a unified method for image deblurring and salt-and-pepper noise removal. Using a variational technique, we introduce a cost functional that represents the goals of deblurring, robustness to salt and pepper noise, and compliance with a piecewise-smooth image model. In addition, we explore the cost functional from robust statistics point of view, and compare it to other methods both theoretically and experimentally. Finally, the experimental results exhibit effective image recovery, with various blur models and noise levels.

2 Unified Variational Framework

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 $f : \Omega \rightarrow [0, 1]$ 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 metric representing data-fidelity. In the case of Gaussian noise, a quadratic data-fidelity term is used:

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

The inverse problem represented by Eq. 1 is known to be ill-posed: small perturbations in the data may produce unbounded variations in the solution. 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 positive weighting scalar. Several regularization terms were suggested in the literature, for example Tikhonov [37] L^2 smoothness, Total variation (TV) L^1 norm [34, 35], modified L^1 norm [1] and recently an integrated TV and wavelet coefficients regularization [20, 21, 25, 26].

In the presence of salt and pepper noise, the quadratic data-fidelity term (2) is inadequate.

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 avoids the resulting PDE from being singular at zero. Brox et al [11] have recently used the modified L^1 norm as a fidelity term for precise optical flow estimation. Note that once η is increased, the modified L^1 tends to L^2 norm, since

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

which has a quadratic form.

2.1 Regularization

The regularization that we use represent preference for piecewise-smooth images with simple edge sets was chosen to be the Mumford-Shah [27] functional in its Γ-convergence approximation. This regularizer has been recently used in electrical impedance tomography [32] and in blind image restoration [6].

The following segmentation functional which was introduced by Mumford and Shah [27] models an image as 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 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 line integral form minimizes the total edge length. Due to the irregularity of this functional, classical calculus of variations methods are not applicable, therefore approximation approaches has to be used. The Mumford-Shah segmentation functional considered as a free-discontinuity problem could be approximated as regular functional in the framework of the Γ-convergence which was introduced by De Giorgi [19].

The main idea of Γ-convergence is to approximate a functional $\mathcal{F}$ by a regular sequence $\mathcal{F}_\epsilon$ such that the minimizers of $\mathcal{F}_\epsilon$ approximate the minimizers 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 $f \in X$

$$1. \quad \forall f_\epsilon \to f, \qquad \liminf_{\epsilon \to 0^+} F_\epsilon(f_\epsilon) \geq F(f)$$

$$2. \quad \exists f_\epsilon \to f, \qquad \limsup_{\epsilon \to 0^+} F_\epsilon(f_\epsilon) \leq F(f).$$

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, states that [9] if $F = \Gamma\text{-lim}(F_\epsilon)$ and that 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 a minimizer of F_ϵ and $f_\epsilon \to f$, then f is a minimizer of F. The *stability* property is formulated as the following: $\Gamma\text{-lim}(F_\epsilon + V) = F + V$ if $\Gamma\text{-lim}(F_\epsilon) = F$ and $V : X \to \mathbb{R}_+$ is continuous.

Ambrosio and Tortorelli [2] 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 characteristics 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 (5) could be found in [9].

Thus, the suggested unified functional (3) which consists of the modified L^1 norm with Mumford-Shah regularization terms 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 before (7). Due to the stability property and the continuity of the convolution operator

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

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

3 Minimization techniques

The objective functional (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 and lower bounded with respect to either of functions f and v if the other one is fixed. For example, given v, $\mathcal{F}_\epsilon$ is convex and lower bounded with respect to f. Therefore, following [14], 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. On the contrary, (11) is a nonlinear integro-differential equation. Linearization of this equation is carried out using the fixed point iteration scheme, as in [14, 38]. We set $f = f^l$ in the denominator, and $f = f^{l+1}$ elsewhere, where l is the current iteration number. Equation (9)

can thus be rewritten as

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

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

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

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 two E-L equations (9) and (11) have now become two linear PDE's, that can be represented by two systems of linear equations. These systems are solved in alternation. This leads 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. Steps 1 and 2 both call for a solution of a system of linear equations. Step 1 was implemented using the Minimal Residual algorithm [39]. As for step 2, following Vogel and Oman [38], Eq. (15) can be expressed in a quasi-Newton like form

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

where

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

and $\mathcal{H}(\cdot, \cdot)$ is the approximation of the Hessian operator. It can be shown that the Operator $\mathcal{H}(\cdot, \cdot)$ is self-adjoint and positive definite (see the proof in the appendix). Consequently $\mathcal{H}(\cdot, \cdot)^{-1} R(\cdot, \cdot)$ in (16) was computed via the Conjugate Gradients method.

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 denoted by $\Delta^x_{\pm} f_{ij} = \pm(f_{i\pm1,j} - f_{ij})$ and $\Delta^y_{\pm} f_{ij} = \pm(f_{i,j\pm1} - f_{ij})$. Hence, the discrete form of Eq. (14) is

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

and Eq. (15) is approximated by

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

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

In the discrete case, the Neumann boundary conditions were implemented as follows. The

observed image was extended by adding margins that are a few pixels wide. These margins were obtained by replicating the one-pixel thick outer frame of the image. The margins were then convolved with 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. All convolution procedures were performed in the Fourier Transform domain. The algorithm was implemented in the MATLAB environment.

4 Relation to robust statistics, anisotropic diffusion and line process

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. Consider the image denoising problem with a smoothing ρ-function

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

Minimization of this functional results in a smoother (denoised) image, where careful selection of $\rho(\cdot)$ would eliminate the effect of gross errors or outliers. In this case $\rho(\cdot)$ is considered as a robust function. In the least-squares approach $\rho(s) = s^2$ which is obviously sensitive to outliers and therefore is non-robust function. The significance of the robust smoothness in this context is that the edges (which have relatively high gradient levels) are preserved. Other $\rho-$functions such as L^1 norm, Huber's MiniMax, Hampel and so on [8] satisfy the robustness property to outliers and some of them are non convex functions. Eq. (17) can be

optimized by the gradient descent

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

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

An alternative formulation of the denoising problem can be expressed by the heat or isotropic diffusion equation,

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

This equation is equivalent to the convolution of the image with a Gaussian kernel, which is an over-smoothing operator, meaning that the edges are not well preserved. Perona and Malik [30] modified this equation to anisotropic diffusion 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 relationship between robust smoothness and anisotropic diffusion was presented by Black et al. [7]. The equivalence relation can be shown by comparing Eq. (18) and (19), hence

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

The third aspect of the edge preserving regularization is the *line process* which was first

introduced by Geman and Geman [22] followed by the *half-quadratic* [15] formulation. The idea was to transform the robust smoothness term into an enhanced functional. This is accomplished by introducing a dual variable b which represents the image edges such that $b \to 0$ in the presence of an edge and $b \to 1$ otherwise. In addition, a convex and decreasing penalty function $\Psi(b)$ is defined where $\Psi(b) \to 1$ as $b \to 0$ and $\Psi(b) \to 0$ otherwise. Thus, the denoising equation with its half-quadratic regularization takes the form

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

This functional is quadratic with respect to f when b is fixed and convex with respect to b when f is fixed.

The Relation between robust estimation and line processes with detailed mechanism of transformations between them was shown by Black and Rangarajan [8]. By comparing Eq. (17) and (20), the connection is given by

$$\rho(s) = \min_{b}(s^2 b + \Psi(b))$$

and $\rho(\sqrt{s})$ is a concave function. Black and Rangarajan [8] also showed that a line process could be extended such that it embodies spatial organization constraints of the outliers for example edges connectivity. In this case, another term which reflects the preference of continuous edges is added to the cost functional:

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

Experimental results in edge preserving denoising application validate the superiority of

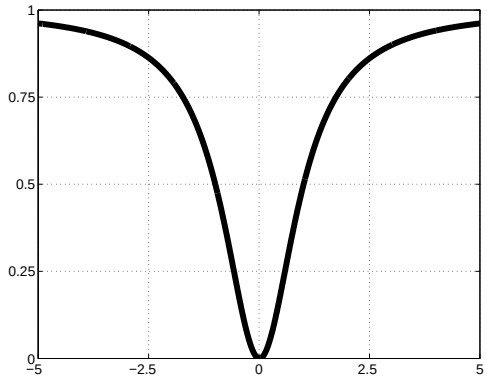

Figure 3: Geman and McClure robust function

this extension [8]. In addition, Teboul et al. [36] were also integrated the spatial interaction constraint as a robust function $\phi_b(|\nabla b|)$ within the half-quadratic regularization and obtained an efficient edge-preserving denoising outcome.

Consider for instance the Geman and McClure [23] robust function (Fig. 3)

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

thus, the corresponding half-quadratic penalty function of this function takes the form

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

and the edge stopping function in the anisotropic diffusion representation is

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

This function is one of non-convex functions family which might suffer from some mathematical difficulties. In the following, two aspects of these problems will be presented. Catte

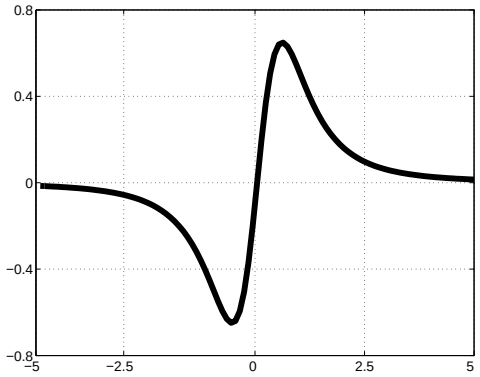

Figure 4: sA(s) for Geman and McClure robust function

et al. [12] noted that if $sA(s)$ in the anisotropic diffusion form is non increasing function, then Perona and Malik's equation is an ill-posed problem. This can be easily verified in the one dimensional case:

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

if $sA(s)$ decreases at some point, it has a negative derivative which results in the inverse heat equation or backward diffusion process. For the Geman and McClure robust function (21),

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

which is partially decreasing as shown in Fig. 4. The solution to this problem suggested by Catte et al. [12] was by selective smoothing approach

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

where G_σ is the Gaussian function and D is a derivative operator. The proof of the regularity of this equation is given in their work. Another difficulty aspect is exemplified in the following

denoising functional:

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

f and g are the model and observed images respectively, and the regularizer is again the Geman and McClure robust function (21). Chipot et al. [17] proved that $\mathcal{F}(f)$ fails to attain a minimum. They also showed that if the functional was perturbed in the form

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

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

4.1 Relation to Mumford-Shah functional

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

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

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

$$\mathcal{F}(f,v) = \frac{1}{2}\int_{\Omega}(f-g)^2 dx + \lambda \int_{\Omega} \left(v^2|\nabla f|^2 + \gamma(v-1)^2 \right) dx. \tag{24}$$

Note the relation between (24) and (6): if $\beta = \lambda$, and $\alpha/4\varepsilon = \lambda\gamma$, then both equations are identical except for the $|\nabla v|^2$ term in (6).

This relation had been also shown by Teboul et al. [36] where interaction constraint was added to the line process, and by Rosati [33] who showed the Γ-convergence of a discrete version of Geman and McClure function to the Mumford-Shah functional. As it shows, the Mumford-Shah functional in its Γ-convergence approximation, is actually a denoising problem with an *extended* line process where the penalty function $\Psi(b)$ corresponds to the robust Geman and McClure function. It better explains the advantages of this regularization. First, the robustness to gradient outliers is well compatible with the impulsive noise model. The smoothing operation is not influenced by these outliers and therefore it is appropriate for the denoising and deconvolution tasks. In addition, the edge smoothing constraint efficiently removes the impulse noise. Second, from mathematical point of view, the problem is well posed and minimizers existence is guaranteed due to the Ambrosio and Tortorelli's [2] proof, together with the fact that the fidelity term (both in denoising and deconvolution cases) is continuous with respect to the image f. Therefore, the Mumford-Shah regularization terms cannot be described as simple anisotropic diffusion or robust smoothness process, but as a more complex structure.

5 Experimental Results

The performance of the algorithm are presented in Figs. 5, 6, 7 and 8. Fig. 5 (left) is a blurred and noisy version of the *Einstein* image. The blur kernel was a pill-box of radius 4; the noise density was 0.11. Fig. 5 (right) is the outcome of the suggested method. The parameters were $\beta = 0.5, \alpha = 0.5, \epsilon = 0.1$. The superiority of the proposed method, with respect to the sequential one (Fig. 2), is clear.

In all the examples in this section, the convergence tolerance of $\varepsilon_1 = 10^{-4}$ was reached

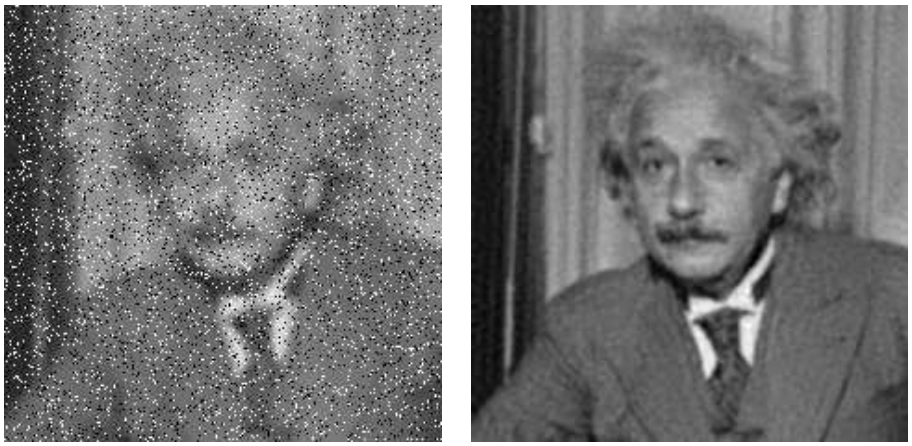

Figure 5: Deblurring in the presence of salt and pepper noise. *Left:* Source image, blurred with a pill-box kernel of radius 4, and contaminated by noise of density 0.11. *Right:* Recovered image, using the suggested algorithm.

with 3-5 external iterations (over n). The number of internal iterations (over l) was set to 5. The constant η (Eq. 4) was set to 10^{-4}.

The examples presented in Fig. 6 demonstrate the performance of the algorithm at a variety of noise levels. The images in the left column were all blurred by a pill-box kernel of radius 3. The noise densities were, 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 for the three cases respectively. Clearly, the image smoothness weight β should be increased with noise level.

Recovery of motion blur in the presence of salt and pepper noise is demonstrated in Fig. 7. The 256×256 *cameraman* image was 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 contaminated 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 [38] is demonstrated in the bottom row. The left image in that row is the outcome of

Figure 6: *Left column:* The *Lena* image, blurred with a pill-box kernel of radius 3, and contaminated by salt and pepper noise. The noise densities are (top to bottom) 0.01, 0.1 and 0.3. *Right column:* The corresponding recovered images.

Figure 7: 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 [38] restoration. *Bottom-right:* The outcome of 5×5 median filtering followed by TV [38] restoration.

3×3 median filtering followed by the TV restoration. The right image in the bottom row was obtained in a similar way, but with a 5×5 median filter.

Finally, the suggested algorithm was tested in the presence of a *random* impulsive noise. In this case, shown in Fig. 8 the image was blurred by pill box kernel of radius 3 and the noisy pixels were set to random values in the range [0,1]. It is clear that the suggested algorithm performs well, since the outliers model is adequate and is even more general than the salt-and-pepper case. In this example the noise density level was 0.1 and the parameters set was $\beta = 0.5, \alpha = 0.1, \epsilon = 0.1$ as in the 10% noise case in Fig. 6.

Figure 8: 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.

6 Comparison to Other Methods

The suggested method of image deconvolution in the presence of salt and pepper noise was tested against two other approaches. In the first experiment, the image was denoised by the Mumford-Shah functional (6) and then restored by the TV method of Vogel and Oman [38]. Fig. 9 top-left is the observed image blurred by a pill-box kernel of radius 3 with 0.1 noise density. Top-right image is the outcome of the denoising procedure ($\beta = 0.9, \alpha = 0.5, \epsilon = 0.1$) and bottom-left is the restored image with weighted factor of 10^{-4}. The recovered image using the suggested method is shown bottom-right. It can be noticed that the noise that had not been entirely removed in the first stage, was amplified in the deconvolution process. On the contrary, the unified approach performs both tasks simultaneously and therefore yields better recovery.

We also compared our method to a unified functional with L^1 regularization instead of

Figure 9: Deblurring in two stages. *Top-left:* Source image, blurred with a pill-box kernel of radius 3, and contaminated by noise of density 0.1. *Top-right:* Denoised image by Mumford-Shah method *Bottom-left:* TV Deconvolution of the denoised image *Bottom-right:* Recovered image using the suggested algorithm

the Mumford-Shah terms. The following cost functional was optimized

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

The results are presented in Fig. 10. In this case, the original image was blurred by a pill-box of radius 3 with 0.3 noise density. Top-left is the outcome of the functional with $\lambda = 0.1$ and $t = 10^{-10}$. Top-right is the recovered image using the suggested method. The bottom row is a magnification of the top one. It can be seen that the suggested algorithm produces a cleaner image.

Figure 10: *Left column:* Restoration using the L^1 regularization. *Right column:* Restoration using the proposed method.

Probable explanation to this difference relies in the regularizer robustness property. The Mumford-Shah terms are basically derived from a function which is more robust than the L^1 norm and that is why the noise is better removed. Fig. 11 illustrates the Geman and McClure robust function with several γ values (dashed line) versus the L^1 function (solid line) which is clearly less robust. Additional support to this observation can be found by Aubert et al. [4]. They have recently shown that the Geman and McClure robust smoothing function embedded in a supervised classification functional, yields better results than other convex smoothing functions.

In the last example we compare the special case of the denoising problem (with no blur process) to the method of Nikolova [29]. Here the salt-and-pepper noise density was 0.02.

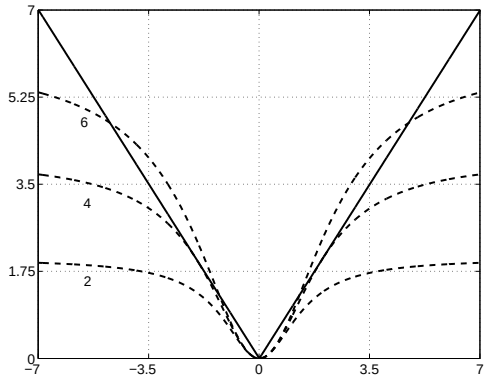

Figure 11: Geman and McClure vs. L^1 robust functions

The images on the left column in Fig. 12 are the outcome of the algorithm of [29] with L^1 norm for both the fidelity and regularization with a weight factor of 0.01. The recovered image in the suggested method in shown on the right column ($\beta = 0.5, \alpha = 0.5, \epsilon = 0.3$). It can be noticed that the suggested algorithm performs better due to the robustness to the impulsive noise.

7 Discussion

We presented a method for image deblurring in the presence of impulsive noise and in particular salt and pepper noise. Our unified approach to deblurring and outliers 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 demonstrated by the fact that similar parameters can be used in the processing of different images in the same noise level. Furthermore, note the fast numerical convergence in our

Figure 12: *Left column:* Restoration using the L^1 regularization [29]. *Right column:* Restoration using the proposed method.

experiments.

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 for images contaminated by outliers, and yet it is still differentiable and convex.

Elements from the Mumford-Shah segmentation functional, in the Γ-convergence formulation, served as the regularization term. They reflect the profound piecewise-smooth image model, and in addition, guarantees the existence of minimizers 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 in which the functional is robust to gross gradients (noise) and in the same time prefers organized or smooth edges. Investigating the total variation, the alternative edge-preserving stabilizer, supports this relation - the L^1 norm in the total variation approach is less robust to outliers then the Geman and McClure function (Figs. 10,12). Another advantage of the proposed regularization terms is that they do not induce nonlinearity beyond that of the fidelity term.

The relations which had been shown between the regularizing terms and the Geman and McClure function have a major significance in parameters selection. The parameter β reflects the smoothness of the recovered image. This parameter is increased with noise level. The parameters α and ϵ are related as $\alpha/4\epsilon = \beta\gamma$ while γ represents the robustness level of the Geman and McClure function where L^1 serves as a reference (Fig. 11). The parameter γ should also be increased with noise level, so if for example α and ϵ are fixed, then β is increased with noise level and γ increases as well (see parameters of Fig. 6).

Finally, an additional advantage of this method is the production of the auxiliary function v, that is an approximated edge map corresponding to the image. For example, Fig. 13 shows the v-maps obtained during the processing of the blurred and noisy *Lena* (pill-box blur, Fig. 6) and *Cameraman* (motion-blur, Fig. 7) images.

Figure 13: 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.

Appendix

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

$$A(v)f = h(-x,-y) * \left(\frac{h(x,y) * f(x,y)}{C(x,y)}\right) - 2\beta Div(v^2 \nabla f(x,y)) \tag{25}$$

with

$$C(x,y) = \sqrt{(h * \tilde{f} - g)^2 + \eta^2} \tag{26}$$

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

$$A^I f = h(-x,-y) * \left(\frac{h(x,y) * f(x,y)}{C(x,y)}\right) \tag{27}$$

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 $\mathcal{H}^*$ is one which satisfies

$$< u, \mathcal{H}^* v > = < \mathcal{H} u, v > \tag{28}$$

or

$$\int u \mathcal{H}^* v dx = \int (\mathcal{H} u) v dx. \tag{29}$$

$\mathcal{H}$ is referred as *self adjoint* operator in the case of $\mathcal{H} = \mathcal{H}^*$.

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 $\mathcal{M} u = M(x) u(x)$,

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

hence $\mathcal{M}^* u(x) = M(x) u(x)$, which is a self-adjoint operator.

Equation (27) could then be rewritten as

$$A^I f = \mathcal{H}^* \mathcal{M} \mathcal{H} f. \tag{30}$$

Since $M(x) = 1/C(x)$ is a real positive function, it could 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. (30) could 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,

$$< g, A^I g >= \int g \mathcal{H}^* \mathcal{D}^* \mathcal{D} \mathcal{H} g \; dx = \int g (\mathcal{D}\mathcal{H})^* \mathcal{D} \mathcal{H} g \; dx = \tag{31}$$
$$< g, (\mathcal{D}\mathcal{H})^* \mathcal{D} \mathcal{H} g >=< \mathcal{D}\mathcal{H} g, \mathcal{D}\mathcal{H} g >= \int |\mathcal{D}\mathcal{H} g|^2 \; dx \quad > \quad 0$$

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,

$$g \nabla \cdot \phi = \nabla \cdot (g\phi) - \nabla g \cdot \phi, \tag{32}$$

thus

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

Applying the divergence theorem,

$$< g, A^{II} f > \;\; = \;\; -2\beta \int_{\partial\Omega} \left(g v^2 \nabla f \right) \cdot dn \;\; + \;\; 2\beta \int_\Omega \nabla g \cdot v^2 \nabla f dx \tag{33}$$

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

$$< g, A^{II} f > \;\; = \;\; 2\beta \int_\Omega \nabla g \cdot v^2 \nabla f dx \;\; = \;\; 2\beta \int_\Omega \nabla f \cdot v^2 \nabla g \, dx \tag{34}$$

Substituting Eq. (32) in (34) 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. (32) 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. □

Acknowledgment

This research was supported by MUSCLE: Multimedia Understanding through Semantics, Computation and Learning, a European Network of Excellence funded by the EC 6th Framework IST Programme. It was supported also by the Israel Science Foundation. Leah Bar was supported by the Weinstein Center for Signal Processing Research at Tel Aviv University.

References

[1] R. Acar and C.R. Vogel, "Analysis of total variation penalty methods", *Inverse Problems*, Vol.10, pp. 1217-1229, 1994.

[2] 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.

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

[4] 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.

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

[6] 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.

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

[8] 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.

[9] Braides A., *Approximation of free-discontinuity problems*, volume 1694 of Lecture Notes in Mathematics. Springer-Verlag, 1998.

[10] 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.

[11] 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.

[12] F. Catte, 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(1), pp. 182-193, 1992.

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

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

[15] 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.

[16] 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.

[17] 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.

[18] Dacorogna B., *Direct Methods in the Calculus of Variations*, Number 78 in Applied Mathematics Sciences. Springer-Verlag, 1989

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

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

[21] 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.

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

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

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

[25] 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.

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

[27] 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.

[28] 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.

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

[30] 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.

[31] 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.

[32] 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.

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

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

[35] 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.

[36] 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.

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

[38] 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.

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