

Review Article

Total Variation Regularization Algorithms for Images Corrupted with Different Noise Models: A Review

Paul Rodríguez

Department of Electrical Engineering, Pontifical Catholic University of Peru, San Miguel, Lima 32, Peru

Correspondence should be addressed to Paul Rodríguez; prodrig@pucp.edu.pe

Received 26 October 2012; Revised 17 May 2013; Accepted 9 June 2013

Academic Editor: Florian Luisier

Copyright © 2013 Paul Rodríguez. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Total Variation (TV) regularization has evolved from an image denoising method for images corrupted with Gaussian noise into a more general technique for inverse problems such as deblurring, blind deconvolution, and inpainting, which also encompasses the Impulse, Poisson, Speckle, and mixed noise models. This paper focuses on giving a summary of the most relevant TV numerical algorithms for solving the restoration problem for grayscale/color images corrupted with several noise models, that is, Gaussian, Salt & Pepper, Poisson, and Speckle (Gamma) noise models as well as for the mixed noise scenarios, such the mixed Gaussian and impulse model. We also include the description of the maximum a posteriori (MAP) estimator for each model as well as a summary of general optimization procedures that are typically used to solve the TV problem.

1. Introduction

Acquired digital images are subject to different kinds of noise [1, Chapter 7] depending on the hardware used for their acquisition which may involve additional degradations due to transmission errors or other external factors. The more common image noise models include the Gaussian, Impulse (e.g., Salt & Pepper), Poisson, and Speckle (e.g., Gamma) and the mixed Gaussian and impulse noise models.

While there are several algorithms that can be considered state of the art for a particular noise model, typically the adaptation of such specialized algorithms to handle other noise models has been proven to be either severely difficult or just plain impossible. It is in this regard that regularization methods stand atop due to their flexibility to use any given noise model; while there are several examples of such methods (Tikhonov regularization, Wavelet image restoration, sparsity based denoising and inversion, etc.), here we focus on the Total Variation (TV) regularization [2] method. The original TV regularization method targeted image denoising under Gaussian noise [2]; nevertheless it has evolved into a more general technique for inverse problems (see [3] for more specific details) while retaining its edge preserving property ([4] gives an extended analysis of TV's properties).

The TV regularized solution of the inverse problem involving data \mathbf{b} and observation operator A (the identity in the case of denoising) is the minimum of the functional

$$T(\mathbf{u}) = F(\mathbf{b} | A\mathbf{u}) + \lambda \|\nabla \mathbf{u}\|_1, \quad (1)$$

where F is the data fidelity term, which depends on the noise model (for instance see (40), (42), (43), and (44)), the scalar λ is the regularization parameter, $\|\nabla \mathbf{u}\|_1$ represents the total variation of solution \mathbf{u} , and the ℓ^1 norm of vector $\mathbf{v} = \nabla \mathbf{u}$ is denoted by $\|\mathbf{v}\|_1 = \sum_k |v_k|$. Moreover, depending on the noise model a nonnegativity constraint may be imposed to (1).

While in Section 4 we will give a complete list of several types of numerical algorithms to solve (1), here we mention that such algorithms can be classified on those that need to solve a linear system of equations and those that do not. Usually the formers can handle the denoising as well as the deconvolution problems, whereas the latter can only handle the denoising problem. Moreover for the cases in which a nonnegativity constraint needs to be imposed to (1), such for the Poisson and Speckle noise models, numerical algorithms differ on how they handle the non-negativity constraint, which may add additional computational costs.

The outline of the paper is as follows: in Section 2 we describe the more common image noise models, where we

also include a brief description on how to find the best estimator of an observed data corrupted with a known noise model using the criterion based on the MAP (maximum a posteriori) estimator. Next in Section 3, we provide a general review of optimization procedures, with a particular focus on the TV problem, where we also include a more detailed description of some general algorithms that usually do not appear on well-known textbooks. In Section 4 we summarize several TV numerical algorithms for each noise model (Gaussian, Impulse, Poisson, Gamma, and mixed Gaussian plus Impulse) considered in the present paper; we also include some details that differentiate the Total Variation formulation for grayscale from color (vector-valued) images. Numerical simulations and examples are presented throughout Section 5. Finally, concluding remarks are listed in Section 6.

2. Image Noise Models

In this section we will describe the more common image noise models including the Gaussian, Impulse, Poisson, Speckle (Gamma), and the mixed Gaussian and impulse noise models; nevertheless we will start by succinctly describing the basics on how to find the best estimator of an observed data corrupted with a known noise model using the criterion based on the MAP (maximum a posteriori) estimator (for a more detailed introduction to this topic, we recommend [5] and the many references therein).

We will assume that \mathbf{u}^* , \mathbf{b} , and \mathbf{u} are a 1-dimensional (column) or 1D vector that represents the original, observed, and reconstructed 2D grayscale or color images, respectively, obtained via any ordering (although the most reasonable choices are row major or column major), that is, for the grayscale case, $\mathbf{b} = \{b_1, b_2, \dots, b_N\}$, whereas for the case of a 2D color image it will be represented by the 1D (column) vector $\mathbf{b} = [(\mathbf{b}_1)^T (\mathbf{b}_2)^T (\mathbf{b}_3)^T]^T$, where each \mathbf{b}_n with $n \in C = \{1, 2, 3\}$ represents a grayscale image. Additionally, we will also consider an observation operator A (see (1)) to be a $N \times N$ matrix, which we will assume to be either the identity or to represent an observation operator (e.g., blur); furthermore for the case of color images, the observation operator A is assumed to be decoupled; that is, A is a diagonal block matrix with elements A_n ; if A is coupled (interchannel blur) due to channel crosstalk, it is possible to reduced it to a diagonal block matrix via a similarity transformation [6].

The MAP estimator assumes that prior noise model (in the observed image \mathbf{b}) is known; then the joint probability distribution function for pixel k , given by $f(b_k | u_k^*)$, will be described by the particular noise model that degrades the observed image; furthermore, if we assume that the noise affecting the (observed) image is independent, then we can write

$$f_{\mathbf{B}|\mathbf{U}^*}(\mathbf{b} | \mathbf{u}^*) = \prod_{k=1}^N f(b_k | u_k^*). \quad (2)$$

Since the aim of any restoration procedure is to estimate $\mathbf{u} \approx \mathbf{u}^*$ conditioned on the knowledge of the observed data \mathbf{b} ,

Bayes' law directly leads to the description of the a posteriori conditional joint probability distribution function

$$f_{\mathbf{U}|\mathbf{B}}(\mathbf{u} | \mathbf{b}) = \frac{f_{\mathbf{B}|\mathbf{U}}(\mathbf{b} | \mathbf{u}) \cdot f(\mathbf{u})}{f(\mathbf{b})}, \quad (3)$$

where the maximization of (3) will lead us to the classical MAP estimator; moreover, this maximization is equivalent to the minimization of $-\ln(f_{\mathbf{U}|\mathbf{B}}(\mathbf{u} | \mathbf{b}))$. If we additionally assume that $f(\mathbf{u}) \approx f(\mathbf{u}^*)$ and $f(\mathbf{b})$ are constant, then the MAP estimator can be found by minimizing $-\ln(f_{\mathbf{B}|\mathbf{U}}(\mathbf{b} | \mathbf{u}))$.

In the next subsections we will use the previous description for each particular noise model to find its MAP estimator, which will be incorporated as the fidelity term in (1).

2.1. The Gaussian Noise Model. The additive Gaussian noise model, is the most frequently occurring noise model used to model thermal noise as well as the limiting behavior of other noise models (for details see the Central Limit theorem, [7, Chapter 8.4]); the model is given by

$$\mathbf{b} = \mathbf{A}\mathbf{u}^* + \boldsymbol{\eta}, \quad (4)$$

where the forward linear operator A models an observation operator which we will assume to be either the identity or to represent a blur operator and $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_N\}$ is a sample of a random vector of N independent Gaussian variables, each with probability density function given by $(1/\sqrt{2\pi\sigma^2})e^{-(\eta_k - \mu)^2/2\sigma^2}$.

Assuming that $\mathbf{u} \approx \mathbf{u}^*$ and $\mu = 0$, it is straightforward to check that

$$f_{\mathbf{B}|\mathbf{U}}(\mathbf{b} | \mathbf{u}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \cdot e^{-\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2/2\sigma^2}. \quad (5)$$

Using the MAP criterion (and since σ is a constant), then the estimator will be found by minimizing $-\ln(e^{-\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2/2\sigma^2})$. Typically this MAP estimator is referred as the solution of the optimization problem given by

$$\min_{\mathbf{u}} \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2. \quad (6)$$

This result is the classical and well-known direct approach for inverse filtering under additive Gaussian noise (see [8]).

We finish this subsection with a concisely list of general denoising methods under the additive Gaussian noise model, highlighting that [9] is a very good introduction to the classical methods for image restoration. Well-known restoration algorithms include the bilateral filter [10], Total Variation denoising [2], the SUSAN filter [11], Wavelet based denoising [12], nonlocal means [13], and so forth; several of the above mentioned algorithms are reviewed and compared in [13]. Finally we mention that patch based approaches, which group 2-D images fragments (or patches) with similar characteristics, are widely considered to be the current leaders in terms of reconstruction quality for the denoising problem; for instance [14] (and derived variants), known as the BM3D algorithm, is a good example of such approaches.

2.2. The Impulse Noise Model. Degradations due to (wireless) transmission errors appearing during the acquisition of digital images are usually modeled as Impulse noise, which can be expressed as

$$\mathbf{b} = \mathcal{B} \cdot (\mathbf{A}\mathbf{u}^*) + (\mathbf{1} - \mathcal{B}) \cdot \mathbf{r}, \quad (7)$$

where arithmetic operations are to be considered as element-wise, \mathcal{B} is a sample drawn from an i.i.d. multivariate Bernoulli distribution with success probability $1 - p$, $\mathbf{1}$ represents a vector with all elements equal to one (both \mathcal{B} and $\mathbf{1}$ have the appropriate dimension), \mathbf{r} is either Salt & Pepper noise or random-valued impulse noise; for the former case, $r_k = c_{\min}$ or $r_k = c_{\max}$ with probability p_1 and p_2 , respectively, ($p = p_1 + p_2$), and for the latter case r_k is drawn from a uniform distribution in $[c_{\min}, c_{\max}]$.

Since Impulse noise (and Salt & Pepper noise in particular) is an example of heavy tailed noise (its probability density function or pdf approaches to zero slower than the Gaussian pdf) it is usually modeled as additive white Laplacian noise (in [15] it was noticed that Laplace noise has the heavy tail behavior associated with impulsive noise); in other words, the model described by (7) is approximated by $\mathbf{b} = \mathbf{A}\mathbf{u}^* + \mathbf{q}$, where $\mathbf{q} = \{q_1, q_2, \dots, q_N\}$ is a sample of a random vector of N independent Laplacian variables, each with probability density function given by $(1/2\sigma)e^{-|q_k - \mu|/\sigma}$; as in the previous sub-section, by assuming that $\mathbf{u} \approx \mathbf{u}^*$ and $\mu = 0$, then it is easy to verify that

$$f_{\mathbf{B}|U}(\mathbf{b} | \mathbf{u}) = \left(\frac{1}{2\sigma}\right)^N \cdot e^{-\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_1 / \sigma}. \quad (8)$$

Using the MAP criterion, the estimator can be found by solving the optimization problem given by

$$\min_{\mathbf{u}} \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_1. \quad (9)$$

The use of the ℓ^1 -norm (as in (9)) has attracted attention in the past decades [15–20] due to a number of advantages, including superior performance with impulse noise in scenarios ranging from detection to variational models.

There are several (nonvariational) approaches to denoise images corrupted with Salt & Pepper noise (as well as for the more general case of Impulse noise), here we succinctly mention the classical 2D median filter [21] and its variants, such the rank-order-based adaptive median filter [22] for the Salt & Pepper case (which uses two-phase approach: first detects the pixels corrupted with Salt & Pepper noise and then filters only those pixels), or the center-weighted median filter [23] for the random-valued impulse noise case. We also mention other methods that also use a two-phase approach: the progressive switching median filter [24], the filter based on the ranked over absolute difference statistics [25], the fuzzy impulse noise detection and reduction method [26], and the directional weighted median filter [27] among others.

2.3. The Poisson Noise Model. Fundamentally, almost every device for image acquisition is a photon counter: Positron Emission Tomography (PET), computed tomography (CT),

radiography, CCD cameras, and so forth. When the count level is low, the resulting image is noisy; such noise is usually modeled via the nonhomogeneous Poisson (NHP) model, with the resulting noise being nonadditive and pixel-intensity dependent. Given \mathbf{u}^* , the noise-free image, we consider each b_k (pixel of \mathbf{b} , observed image) to be an independent Poisson random variable such $E\{b_k | u_k^*\} = u_k^* \geq 0$ and $P(b_k | u_k^*) = u_k^{*b_k} e^{-u_k^*} / b_k!$; formally, the Poisson noise is defined as $v_k = b_k - E\{b_k | u_k^*\}$, with $E\{v_k | u_k^*\} = 0$ and $\text{var}\{v_k | u_k^*\} = u_k^*$, which stresses the fact that Poisson noise is signal dependent; furthermore, this characteristic of the Poisson noise (that Poisson noise's variance is signal dependent, and particularly for images, that its variance is pixel or spatial dependent) prevents from directly treating it as Gaussian noise (by invoking the Central Limit theorem, see [7, Chapter 8.4]) even when u_k^* (value that represents the “count”) is not low, since the variance of such approximation would be spatially dependent, and thus be in contradiction with the standard Gaussian noise model (which assumes a constant variance).

Following a similar approach as in the previous sub-sections, we will assume that each pixel of the observed image $\mathbf{b} = \{b_1, b_2, \dots, b_N\}$ is a sample of a random vector of N independent Poisson variables, with a joint pdf given by

$$P(\mathbf{b} | \boldsymbol{\lambda}) = \prod_{k=1}^N \frac{\lambda_k^{b_k} e^{-\lambda_k}}{b_k!}, \quad (10)$$

where $\boldsymbol{\lambda} = \mathbf{A}\mathbf{u}^* \geq 0$, A models an observation operator (identity or blur), and \mathbf{u}^* is the unknown noise-free image, which is usually assumed to be nonnegative.

Using the MAP criterion, the estimator can be found by solving the optimization problem given by

$$\begin{aligned} \min_{\mathbf{u}} \quad & \sum (\mathbf{A}\mathbf{u})_k - \mathbf{b}_k \cdot \log((\mathbf{A}\mathbf{u})_k) \\ \text{s.t.} \quad & \mathbf{u} \geq 0, \end{aligned} \quad (11)$$

where $(\mathbf{A}\mathbf{u})_k$ represents the k th element of $\mathbf{A}\mathbf{u}$ (i.e., if $\mathbf{v} = \mathbf{A}\mathbf{u}$, then $v_k = (\mathbf{A}\mathbf{u})_k$). The minimization problem summarized by (11) was first presented in [28, 29], where the (now classical) Richardson-Lucy algorithm was introduced.

It must be noted that the functional (Poisson likelihood) $T(\mathbf{u}) = \sum (\mathbf{A}\mathbf{u})_k - \mathbf{b}_k \cdot \log((\mathbf{A}\mathbf{u})_k)$ is not Lipschitz continuous when $(\mathbf{A}\mathbf{u})_k$ (for any k) is close to zero: that is, it is not possible to find a constant $0 < L < \infty$ such that $|T(\mathbf{u}^{(1)}) - T(\mathbf{u}^{(0)})| / \|\mathbf{u}^{(1)} - \mathbf{u}^{(0)}\| \leq L$ if either any element of $\mathbf{A}\mathbf{u}^{(0)}$ or $\mathbf{A}\mathbf{u}^{(1)}$ is (close to) zero; therefore, any given algorithm that optimizes a cost functional that includes that the Poisson likelihood must take this consideration into account.

Finally, we mention that there are several methods to restore (denoise/deconvolve) non-negative images corrupted with Poisson noise (a comprehensive review is given in [30]) which includes the Richardson-Lucy algorithm and its regularized variants as well as Wavelet methods, Bayesian methods, and so forth. Additionally we also notice that there are methods [31–33] based on Variance-Stabilizing Transformations or VST (the Anscombe [34] or Freeman-Tukey [35] VST for this particular case): given a set or

sequence of nonhomoskedastic random variables (random variables with different finite variances), with the same underlying distribution family, a VST is a specific distribution transformation that renders that set or sequence of random variables into an homoskedastic one (random variables with the same finite variance); moreover such transformation is chosen so the transformed random variables are Gaussian with unit variance. These methods ([31–33]) attain very high quality reconstruction results by using state-of-the-art denoising methods for the Gaussian noise model and then applying the inverse VST.

2.4. The Speckle (Gamma) Noise Model. For images acquired via ultrasound, SAR (synthetic aperture radar) or coherent light imaging systems (e.g., Fabry-Perot interferometer, etc.), Speckle noise is a critical factor that limits their visual perception and processing (see [36, 37] among others). Typically, the multiplicative model

$$\mathbf{b} = (\mathbf{A}\mathbf{u}^*) \cdot \boldsymbol{\zeta}, \quad (12)$$

is considered, where (as in previous sub-sections) A is the observation operator, \mathbf{u}^* is the noise-free data, and $\boldsymbol{\zeta}$ is the noise; also, it is assumed that $\mathbf{b}, \mathbf{u}^* > 0$. Moreover, in the Speckle noise model, it is considered that $\boldsymbol{\zeta} = \{\zeta_1, \zeta_2, \dots, \zeta_N\}$ is a sample of a random vector of N independent Gamma variables, each with pdf given by $f_Z(\zeta) = (L^L/\Gamma(L))\zeta^{L-1}e^{-L\zeta}$, with $E\{\zeta_k\} = 1$ and $\text{var}\{\zeta_k\} = 1/L$, where $\Gamma(L) = (L-1)!$ for a positive value of L .

The derivation of the MAP estimator is not as straightforward as for the previous noise models; on what follows we summarized the derivation of the MAP estimator as originally presented in [38], which leads to a nonconvex optimization problem.

We first reproduce Proposition 3.1 from [38] (for the proof, see the mentioned reference): assume that \mathbf{u} and $\boldsymbol{\zeta}$ are independent random variables with independent continuous pdfs f_U and f_Z ; then for $\mathbf{b} = \mathbf{u} \cdot \boldsymbol{\zeta}$ and $\mathbf{u} > 0$

$$f_Z\left(\frac{\mathbf{b}}{\mathbf{u}}\right) \cdot \frac{1}{\mathbf{u}} = f_{\mathbf{B}|\mathbf{U}}(\mathbf{b} | \mathbf{u}), \quad (13)$$

where arithmetic operations are to be understood as element-wise; using this result, then MAP estimator can be computed by maximizing

$$f_{\mathbf{B}|\mathbf{U}}(\mathbf{b} | \mathbf{u}) = \left(\frac{L^L}{\Gamma(L)}\right)^N \prod_{k=1}^N \left(\frac{b_k}{(\mathbf{A}\mathbf{u})_k}\right)^{L-1} \frac{1}{(\mathbf{A}\mathbf{u})_k} e^{-L(b_k/(\mathbf{A}\mathbf{u})_k)}, \quad (14)$$

such maximization derives in the following non-convex optimization problem

$$\begin{aligned} \min_{\mathbf{u}} \quad & \sum_k \frac{b_k}{(\mathbf{A}\mathbf{u})_k} + \log((\mathbf{A}\mathbf{u})_k) \\ \text{s.t.} \quad & \mathbf{u} \geq 0. \end{aligned} \quad (15)$$

As for the Poisson noise model, it must be noted that the functional (Gamma likelihood) $T(\mathbf{u}_k) = \sum_k (\mathbf{b}_k/(\mathbf{A}\mathbf{u})_k) + \log((\mathbf{A}\mathbf{u})_k)$ is not Lipschitz continuous when $(\mathbf{A}\mathbf{u})_k$ (for any k) is close to zero, and thus any given algorithm that optimizes a cost functional that includes the Gamma likelihood must take this consideration into account. In [38] the authors give a detailed proof of the existence and the uniqueness for this MAP estimator (15) when used within a variational model.

While in [37] an extensive list of despeckle filtering algorithms is provided, which includes methods based on linear filtering (first-order statistics, local statistics, etc.), nonlinear filtering (median, linear scaling, etc.), diffusion filtering (anisotropic diffusion, etc.), and wavelet filtering, we also mention that methods following a VST-like approach [39] can be also an effective alternative.

2.5. The Mixed Gaussian-Impulse Noise Model. As mentioned before, acquired digital images are frequently subject to additive Gaussian noise; if we then add the degradation due to transmission errors, the observed image will be corrupted by Gaussian with salt-and-pepper (constant-valued impulse) noise or random-valued impulse noise. The mixed Gaussian and impulse noise model is summarized by

$$\mathbf{b} = \mathcal{B} \cdot (\mathbf{A}\mathbf{u}^* + \boldsymbol{\eta}) + (\mathbf{1} - \mathcal{B}) \cdot \mathbf{r}, \quad (16)$$

where we have the same considerations as for the Impulse noise model (see Section 2.2, in particular the considerations for (7)), and additionally $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_N\}$ is a sample of the random vector of N independent Gaussian variables, each with probability density function given by $(1/\sqrt{2\pi\sigma^2})e^{-(\eta_k-\mu)^2/2\sigma^2}$.

The key idea of the methods that target this mixed noise model scenario is to use a two-phase approach: detect the outlier pixels before proceeding with the filtering phase, where the detected outliers (pixels corrupted with Impulse noise) and the rest of the pixels are filtered using a suitable (and sometimes independent) method for the particular noise model. For example (in the case of non-TV based methods), [25] introduced an universal noise removal filter that first detects the impulse corrupted pixels and estimates its local statistics and incorporate them into the bilateral filter [10], resulting in the trilateral filter. In [40] somewhat similar ideas were used to develop a similarity principle which in turn drives the weights of the “mixed noise filter”; the reconstruction performance (as reported in [40]) outperformed that of the trilateral filter.

Although all the methods that target (16) assume that \mathcal{N} , the set of outliers (pixels corrupted with impulse noise), is known; the fact is that \mathcal{N} must be estimated; for this purpose there are several options (for a performance comparison between several of the listed methods we refer the reader to [41]): the rank-order-based adaptive median filter [22], the progressive switching median filter [24], the filter based on the ranked over absolute difference statistics [25], the fuzzy impulse noise detection and reduction method [26], the center weighted median filter [23], and the directional weighted median filter [27] among others.

3. Optimization Algorithms: A Review from the Total Variation's Perspective

Optimization theory deals with methods to find the argument of a cost function that minimizes (maximizes) it, where the cost function is a quantitative measure of the performance of a given system. In this Section we focus on listing general optimization procedures that are usually used to solve the TV problem, giving references to well-known textbooks or succinctly describing such general optimization procedures. For a general introduction to the topic of optimization theory we recommend [42, 43] among other textbooks.

We first start this Section by reminding of the general TV problem (as described in (1))

$$T(\mathbf{u}) = F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda \|\nabla \mathbf{u}\|_1, \quad (17)$$

and by stressing that a non-negativity constraint ($\mathbf{u} \geq 0$) may be imposed to (17); also, throughout this paper we will consider that the discrete version of $\|\nabla \mathbf{u}\|_1$ is given by

$$\left\| \sqrt{\sum_{n \in C} (D_x \mathbf{u}_n)^2 + (D_y \mathbf{u}_n)^2} \right\|_1, \quad (18)$$

where $n \in C = \{1\}$ (grayscale images case) or $n \in C = \{1, 2, 3\}$ (or “red,” “green,” and “blue,” for the color images case; also note that C could represent an arbitrary number of channels), the horizontal and vertical discrete derivative operators are denoted by D_x and D_y , respectively, and that the scalar operations are considered to be applied element-wise, so that, for example, $\mathbf{u} = \sqrt{v} \Rightarrow u[k] = \sqrt{v[k]}$. We also note that (18) is the generalization of TV regularization to color images (for other choices see [44, Chapter 4.8.2]) with coupled channels (see [45, Section 9]), typically used within the color or vector-valued TV framework (first introduced in [46]), which coincides with the discrete version of $\|\nabla \mathbf{u}\|_1$ for grayscale images ($n \in C = \{1\}$ in (18)), that is $\left\| \sqrt{(D_x \mathbf{u})^2 + (D_y \mathbf{u})^2} \right\|_1$.

Due to the nondifferentiability of $\|\nabla \mathbf{u}\|_1$, it is sometimes replaced by the smooth approximation

$$\frac{1}{2} \sum_k \tau_\epsilon(v_k), \quad (19)$$

where $v^2 = \sum_{n \in C} (D_x \mathbf{u}_n)^2 + (D_y \mathbf{u}_n)^2$ and some sensible choice of function $\tau_\epsilon(\cdot)$ such $\tau_\epsilon(v_k) = 2\sqrt{v_k^2 + \epsilon^2}$ or such the Huber function [47]

$$\tau_\epsilon(v_k) = \begin{cases} \epsilon^{-1} v_k^2 & \text{if } |v_k| \leq \epsilon \\ 2|v_k| - \epsilon & \text{if } |v_k| > \epsilon. \end{cases} \quad (20)$$

We also mention that in other cases an anisotropic separable approximation of $\|\nabla \mathbf{u}\|_1$ is preferred, that is $\|\nabla \mathbf{u}\| \approx |D_x \mathbf{u}| + |D_y \mathbf{u}|$ (for instance, see [48]).

Since its original formulation [2], the development of numerical algorithms for solving (17), as well as its non-negativity version, has attracted considerable interest; while in Section 4 we will give a complete list of such algorithms,

here we focus on briefly describing or listing (whenever already described in well-known textbooks) general optimization procedures that are usually used to solve (17). For instance in [49, Chapter 8] a complete description of several general optimization procedures for TV are given, such as the Steepest Descent (also referred as the artificial time marching approach, see [44, Chapter 4.5.5]), the Newton method, the lagged fixed point iteration (originally described in [50, 51]), and the Primal-dual Newton method (originally described in [52]). Although all of these algorithms focus on the ℓ^2 -TV case (Gaussian noise model), several of them have been used for other noise models (see Section 4). Similarly in [44], several numerical algorithms for TV are described; of particular interest is the dual formulation of the TV problem described in [44, Chapter 4.5.6], which is a different approach than the Primal-dual Newton method (and was used in [53] to derive an effective algorithm).

The enforcement of a nonnegativity constraint, that is: $\mathbf{u} \geq 0$, for the solution of (17) is not only physically meaningful in most of the cases: images acquired by digital cameras, CT, and so forth; it also improves the quality of the reconstruction (see [54]). Nevertheless, the non-negativity constraint is seldom considered in the practice (unless explicitly needed by the noise model), since it makes a hard problem even harder. There are several well-known methods to attack this problem: for instance in [49, Chapter 9] very good descriptions are given for the gradient projection method (which can be understood as a generalization of the Steepest Descent method) and for the Projected Newton method (as well as its variants).

We finalize this Section by succinctly describing three general optimization procedures that have been used to solve the TV problem: the Iterative Reweighted Least Squares (IRLS), the Alternating Direction method of multipliers (ADMM) or Split-Bregman (SB), and the Nonnegative Quadratic Programming (NQP) method.

3.1. Iterative Reweighted Least Squares (IRLS). Since its introduction [55] the IRLS method has been applied to a variety of optimization problems. Originally, the IRLS method was used to solve the ℓ^p minimization problem $T(\mathbf{u}) = (1/p)\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_p^p$ by iteratively approximating it by a weighted ℓ^2 norm. At iteration k the solution $\mathbf{u}^{(k+1)}$ is the minimizer of $(1/2)\|W^{(k)1/2}(\mathbf{A}\mathbf{u} - \mathbf{b})\|_2^2$, with weighting matrix $W^{(k)} = \text{diag}(|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}|^{p-2})$. For $2 \leq p \leq 3$ this algorithm converges to the global minimizer [56], while for $1 \leq p < 2$ the definition of the weighting matrix $W^{(k)}$ must be modified to avoid numerical instability due to division by zero or by a very small value. A standard approach is to threshold elements of $\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}$ in constructing the corresponding elements of $W^{(k)}$, but other choices are also possible [57, 58]. For $0 < p < 1$ it has been shown [58], within a sparse representation framework, that the IRLS algorithm not only converges but increases its convergence rate as p goes to zero.

Without loss of generality, we will focus on the ℓ^2 -TV [2] case for grayscale images: $T(\mathbf{u}) = (1/2)\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2 + \lambda \left\| \sqrt{(D_x \mathbf{u})^2 + (D_y \mathbf{u})^2} \right\|_1$, where A is the forward operator, \mathbf{b}

is the observed noisy data, λ is a weighting factor controlling the relative importance of the data fidelity and regularization terms, and \mathbf{u} is the restored image data.

The key idea [59] is to express the regularization term by the quadratic approximation $Q_R^{(k)}(\mathbf{u}) = (1/2)\|W_R^{(k)1/2} D\mathbf{u}\|_2^2$, where

$$W_R^{(k)} = I_2 \otimes \Omega_R^{(k)}, \quad D = [D_x^T D_y^T]^T, \quad (21)$$

$$\Omega_R^{(k)} = \text{diag} \left(\left((D_x \mathbf{u}^{(k)})^2 + (D_y \mathbf{u}^{(k)})^2 \right)^{-0.5} \right),$$

and I_N is a $N \times N$ identity matrix, and \otimes is the Kronecker product. The resulting iterations can be expressed in the form of the standard IRLS problem:

$$T^{(k)}(\mathbf{u}) = \frac{1}{2} \left\| \begin{bmatrix} 1 & 0 \\ 0 & W_R^{(k)} \end{bmatrix}^{1/2} \begin{bmatrix} A \\ \sqrt{\lambda} D \end{bmatrix} \mathbf{u} - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2. \quad (22)$$

For a given current solution $\mathbf{u}^{(k)}$, the weighting matrix $W_R^{(k)}$ can be easily computed, and the threshold (to avoid numerical instability due to division by zero or by a very small value) may be automatically adapted to the input image [59, Section 4.7]. Finally, the resulting algorithm has to iteratively solve the linear system

$$(A^T A + \lambda D^T W_R^{(k)} D) \mathbf{u}^{(k+1)} = A^T \mathbf{b}, \quad (23)$$

which is its most computationally demanding part. The same strategy can be used to solve all other noise models within the TV framework, including the vector-valued (color) TV.

Moreover, for the particular case of denoising ($A = I$ in (22)) the solution of the linear system described by (23) is given by

$$(I + D^T \widehat{W}_R^{(k)} D) \mathbf{u}^{(k+1)} = \mathbf{b}, \quad (24)$$

where $\widehat{W}_R^{(k)} = \lambda W_R^{(k)}$; if we now apply the (well-known) matrix inversion lemma to (24), we get

$$\mathbf{u}^{(k+1)} = \mathbf{b} - D^T \left((\widehat{W}_R^{(k)})^{-1} + DD^T \right)^{-1} D \mathbf{b}. \quad (25)$$

It is important to notice that $(\widehat{W}_R^{(k)})^{-1}$ can be computed directly (no division by zero); furthermore, by solving

$$\left((\widehat{W}_R^{(k)})^{-1} + DD^T \right) \mathbf{z}^{(k)} = D \mathbf{b}. \quad (26)$$

We can compute $\mathbf{u}^{(k+1)}$ via

$$\mathbf{u}^{(k+1)} = \mathbf{b} - D^T \mathbf{z}^{(k)}. \quad (27)$$

This approach ((26) and (27)) was first proposed, within the Total Variation framework (particularly for the ℓ^2 -TV denoising case), in [60].

3.2. Alternating Direction Method of Multipliers (ADMMs) or Split Bregman (SB). Alternating minimization methods have become popular in the past few years due to their ability to

solve ℓ^1 regularized problems, that is, $F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda R(\mathbf{u})$, where R is the regularization term (being (17) a particular example with $R(\mathbf{u}) = \|\nabla \mathbf{u}\|_1$), in a simple and computationally efficient fashion. Although there are several incarnations of these methods [61], we focus on the Split-Bregman (SB) [62, 63] algorithm, while noting that it is now recognized that the SB algorithm is equivalent to the older Alternating Direction Method of Multipliers (ADMM) [64, 65] algorithm.

The key idea of the SB method [62, 63] is to introduce an auxiliary variable to modify the original cost functional (17) so that it can be iteratively minimized using simple steps per iteration; this is done by first considering the following constrain optimization problem

$$F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda R(\mathbf{z}) \quad \text{s.t. } \mathbf{z} = \mathbf{u}; \quad (28)$$

the standard Lagrangian of (28) will lead us to $L(\mathbf{u}, \mathbf{z}, \mathbf{w}) = F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda R(\mathbf{z}) + \mathbf{w}^T(\mathbf{u} - \mathbf{z})$; nevertheless the augmented Lagrangian $L_\rho(\mathbf{u}, \mathbf{z}, \mathbf{w})$,

$$F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda R(\mathbf{z}) + \mathbf{w}^T(\mathbf{u} - \mathbf{z}) + \frac{\rho}{2} \|\mathbf{u} - \mathbf{z}\|_2^2, \quad (29)$$

is preferred since it gives a more robust framework (see [61, Section 2.3]). Furthermore, by setting $\mathbf{r} = \mathbf{u} - \mathbf{z}$ and noting that $\mathbf{w}^T \mathbf{r} + (\rho/2) \|\mathbf{r}\|_2^2 = (\rho/2) \|\mathbf{v} + \mathbf{r}\|_2^2 + (\rho/2) \|\mathbf{v}\|_2^2$, with $\mathbf{w} = \mathbf{v}/\rho$, we can write (29) as

$$F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \lambda R(\mathbf{z}) + \frac{\rho}{2} \|\mathbf{v} + \mathbf{u} - \mathbf{z}\|_2^2 + \frac{\rho}{2} \|\mathbf{v}\|_2^2, \quad (30)$$

which can be iteratively solved by considering the following alternating optimization problems:

$$\mathbf{u}^{(k+1)} : \arg \min_{\mathbf{u}} F(\mathbf{b} | \mathbf{A}\mathbf{u}) + \frac{\rho}{2} \|\mathbf{v}^{(k)} + \mathbf{u} - \mathbf{z}^{(k)}\|_2^2, \quad (31)$$

$$\mathbf{z}^{(k+1)} : \arg \min_{\mathbf{z}} \lambda R(\mathbf{z}) + \frac{\rho}{2} \|\mathbf{v}^{(k)} + \mathbf{u}^{(k+1)} - \mathbf{z}\|_2^2, \quad (32)$$

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \mathbf{u}^{(k+1)} - \mathbf{z}^{(k+1)}. \quad (33)$$

It is important to highlight some observations regarding the previous optimization problems: assuming that $F(\mathbf{b} | \mathbf{A}\mathbf{u})$ is quadratic then (31) is a generalized Tikhonov problem step, which is simple to solve; if $R(\mathbf{z})$ can be expressed as the ℓ^1 norm of \mathbf{z} , then (32) can be solved via shrinkage (in some cases the generalized shrinkage is needed, see [61, Section 4.1]).

It is also important to note that the SB (or ADMM) method can also be used to solve the non-negative quadratic optimization (see (35)) or to solve the optimization problems derived from the Poisson and Speckle noise models (see Sections 2.3 and 2.4) using a similar approach (as described in the previous paragraphs); for a complete description we refer the reader to [61, Section 5.2], [66, 67], respectively.

The general SB (or ADMM) algorithm can be easily adapted to handle isotropic TV ($R(\mathbf{u}) = \|\nabla \mathbf{u}\|_1$, that should be replaced by its appropriate discrete version for grayscale and vector-valued images); as an example we succinctly focus on the ℓ^2 -TV [2] case for grayscale images: $T(\mathbf{u}) = (1/2)\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2 + \lambda \sqrt{(D_x \mathbf{u})^2 + (D_y \mathbf{u})^2}$, where the auxiliary

variable $\mathbf{z} = [\mathbf{z}_x, \mathbf{z}_y]^T$ is used along with the constrains $\mathbf{z}_x = D_x \mathbf{u}$ and $\mathbf{z}_y = D_y \mathbf{u}$; among other operations, such as generalized shrinkage (see [63, Section 4.1]) and auxiliary vector updates (which are not computationally demanding) use to solve the particular optimization problems (related to the general ones, i.e., (32) and (33)), the SB-TV algorithm has to solve the linear system

$$(A^T A - \rho D^T D) \mathbf{u}^{(k+1)} = A^T \mathbf{b} + \rho D^T D (\mathbf{z}^{(k)} - \mathbf{v}^{(k)}). \quad (34)$$

Note that the left-hand side (LHS) of (34) is constant across different iterations while its right-hand side (RHS) changes at each iteration; the opposite is true for the resulting linear system in (23). Furthermore, since LHS of (34) is constant across iterations (and strictly diagonally dominant in the practice), it seems natural (as suggested in [63]) to use the Gauss-Seidel method to solve (34); furthermore, when $A = I$ or A is a circulant (diagonalizable by the Fourier transform) matrix, (34) may also be solved in the Fourier or DCT domain.

One key computational difference between the SB-TV and the IRLS based algorithm is that even though the number of floating-point operations for each SB-TV iteration is slightly smaller than that of each IRLS-TV iteration, typically the number of global iterations to attain good reconstruction results for the latter algorithm is less than for the former (see [68] for a detailed analysis).

3.3. Nonnegative Quadratic Programming (NQP). Recently [69] an interesting and quite simple algorithm has been proposed to solve the NQP problem:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \frac{1}{2} \mathbf{u}^T \Phi \mathbf{u} + \mathbf{c}^T \mathbf{u} \\ \text{s.t.} \quad & 0 \leq \mathbf{u} \leq u \max, \end{aligned} \quad (35)$$

where the matrix Φ is assumed to be symmetric and positive defined, and $u \max$ is some positive constant. The multiplicative updates for the NQP are summarized as follows (see [69] for details on derivation and convergence):

$$\Phi_{nl}^+ = \begin{cases} \Phi_{nl} & \text{if } \Phi_{nl} > 0 \\ 0 & \text{otherwise} \end{cases}, \quad \Phi_{nl}^- = \begin{cases} |\Phi_{nl}| & \text{if } \Phi_{nl} < 0 \\ 0 & \text{otherwise} \end{cases}, \quad (36)$$

$$\mathbf{u}^{(k+1)} = \min \left\{ \mathbf{u}^{(k)} \left[\frac{-\mathbf{c} + \sqrt{\mathbf{c}^2 + \mathbf{v}^{(k)} \boldsymbol{\nu}^{(k)}}}{2\mathbf{v}^{(k)}} \right], u \max \right\}, \quad (37)$$

where $\mathbf{v}^{(k)} = \Phi^+ \mathbf{u}^{(k)}$, $\boldsymbol{\nu}^{(k)} = \Phi^- \mathbf{u}^{(k)}$, and all algebraic operations in (37) are to be carried out element-wise. One key observation to consider is that this algorithm cannot be initialized with zeros. Interestingly, once an element has been zeroed by the multiplicative updates it remains zero under successive updates; this property is specially appealing when this algorithm is applied to sparse representation problems. Furthermore, the NQP is quite efficient and has been used to solve interesting problems such as statistical learning [69] and compressive sensing [70] among others.

We finalize this sub-section by noting that the constraint problem,

$$T(\mathbf{u}) = \frac{1}{p} \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_p^p \quad \text{s.t. } 0 \leq \mathbf{u} \leq u \max, \quad (38)$$

could be iteratively solved using an IRLS type approach that benefits from the NQP algorithm, by approximating (38) with

$$T(\mathbf{u}) = \frac{1}{2} \left\| W^{(k)1/2} (\mathbf{A}\mathbf{u} - \mathbf{b}) \right\|_2^2 \quad \text{s.t. } 0 \leq \mathbf{u} \leq u \max, \quad (39)$$

and by setting $\Phi = A^T W^{(k)} A$ and $\mathbf{c} = -A^T W^{(k)1/2} \mathbf{b}$ in (37).

4. Numerical Algorithms for TV

In the following sub-sections we will summarize a complete list of TV numerical algorithms for each noise model; we will also include the particular cost functional that is targeted in each case, highlighting particular cases such the methods that adapt the regularization parameter, either in a global or local fashion.

4.1. TV Numerical Algorithms for the Gaussian Noise Model.

The minimization of the cost functional,

$$T(\mathbf{u}) = \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2 + \lambda \|\nabla \mathbf{u}\|_1, \quad (40)$$

is usually referred to as the ℓ^2 -TV solution. Since its original formulation [2] a large number of algorithms have been proposed to specifically solve (40); generally speaking these algorithms can be classified based on the necessity (or not) to solve a linear system of equations.

For instance algorithms based on a dual formulation of (40) do not need to solve a linear system of equations and are usually computational efficient, although they lack the ability to handle a nontrivial forward (observation) operator A in (40); the Chambolle [53] and the Aujol's ℓ^2 -TV approximation (or A^2 BC model) [71] algorithms are examples of these type of methods.

On the other hand, methods that are based on the Euler-Lagrange equations (derived from the direct optimization of (40)) usually use a smooth approximation of $\|\nabla \mathbf{u}\|_1$ (see Section 3) and need to solve a linear system of equations, which is usually carried out via conjugate gradient (CG), the preconditioned CG (PCG), or via the Fourier transform (to solve the linear system when the forward operator A is equivalent to a circular convolution). For example, we mention that in [2] the authors used the steepest descent (artificial time marching) algorithm along with a line-search to solve (40) (the computational performance of the resulting method is very slow compare to all others); more elaborated algorithms include the primal-dual method [52] (which although uses a dual formulations, it does need to solve a set of linear equations), the lagged diffusivity algorithm [50, 51], the majorization-minimization or MM method [72, 73] (which can be considered a generalization of the Expectation-Maximization, or EM algorithm [74]), as well as its variant

[60] for the denoising case that uses the matrix inversion lemma (see (26) and (27)), the TwIST method [75] (which is similar to MM method, previously mentioned and uses a predetermined PCG splitting method), and the Iteratively Reweighted Norm (IRN) [59] (which can be understood as an IRLS method). We also mention the FTVd method [62, 76] (which is a FFT based algorithm) and the linear programming via interior point method [77] (for which the linear system to be solved has twice as many variables as the original problem) and the Newton's method [49].

4.1.1. Vector-Valued ℓ^2 -TV Algorithms. The TV minimization scheme for deblurring color images was first introduced in [46], and since then several alternatives have been proposed: in [78] the authors proposed an algorithm that considered the HSV color model, in [79] an algorithm based on a dual formulation (similar to Chambolle's [53]) was proposed (only the denoising case was considered), [80] presented an extension to [79] with a dual/alternate algorithm, in [81] the IRN algorithm [59] was modified to handle vector-valued images (this algorithm is related to the IRLS algorithm for vector-valued datasets, see [6]), and in [82] the FTVd algorithm [62, 76] was extended to handle vector-valued images. Finally we also mention that a direct extension of [50], called vectorial lagged diffusivity, was used in [83] for comparison purposes.

4.1.2. Nonnegative ℓ^2 -TV Algorithms. For scalar (grayscale) images, only a handful of numerical algorithms, for example, [49, Chapter 9] and more recently [84] (non-negativity is handled via the NQP algorithm, see Section 3.3), [54] (non-negativity is handled via a variant of the Primal-dual Newton method, [49, 52, Chapter 8]), [85] (non-negativity is handled via combination of a gradient based approach along with an iterative shrinkage/thresholding algorithm), and [86] (non-negativity is handled via the SB algorithm, see Section 3.2), include a non-negativity constraint on the solution of the ℓ^2 -TV problem (40), and to the best of our knowledge, for vector-valued (color) images only [87] (non-negativity is handled via the NQP algorithm, see Section 3.3) explicitly includes the non-negativity constraint within the ℓ^2 -TV framework.

4.1.3. Adaptation of the Regularization Parameter for ℓ^2 -TV. Several methods have been proposed to adapt the regularization parameter for ℓ^2 -TV; they differ on how the amount of noise is estimated and on the nature of the regularization parameter which can be a scalar (global) value or a spatially dependent (diagonal matrix) set of regularization parameters. For the latter case, the regularization parameter is moved from the regularization term into the fidelity term; moreover if we consider $\Lambda = \text{diag}(1/\lambda)$, then (40) can be written (for the denoising case) as

$$T(\mathbf{u}) = \frac{1}{2} \left\| \Lambda^{1/2} (\mathbf{u} - \mathbf{b}) \right\|_2^2 + \|\nabla \mathbf{u}\|_1, \quad (41)$$

where Λ will be updated for each pixel (independently) in an iteratively fashion.

Succinctly we mention that for the ℓ^2 -TV case in [88] two adaptive regularization schemes were presented to spatially update the regularization parameters. Also in [89] the Stein's Unbiased Risk Estimate (SURE) was used to estimate the mean square error between the observed image and the denoised one; a global regularization parameter was used. In [90] an extension to the Unbiased Predictive Risk Estimator (UPRE) was used to estimate the global regularization parameter as well.

4.2. TV Numerical Algorithms for the Salt & Pepper Noise Model. TV for the Salt & Pepper noise model is usually referred as ℓ^1 -TV; in this case the minimizer of the cost functional,

$$T(\mathbf{u}) = \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_1 + \lambda \|\nabla \mathbf{u}\|_1, \quad (42)$$

is the solution of the ℓ^1 -TV problem. The use of the ℓ^1 for the fidelity term in (42) was a significant development [16, 17, 19, 20] for TV based restoration methods and attracted attention due to a number of advantages, including superior performance with impulse noise [91].

The numerical algorithms that solve (42) are based on a variety of methods (although they can also be loosely classified between methods that need to solve a linear system and methods that need not): for instance in [20, 91] a smooth approximation of the ℓ^1 norm was used, along with the steepest descent algorithm (the computational performance of this algorithm is slow compared to all other options); in [92] a Markov random field model was used, where the resulting algorithm, that used the anisotropic separable approximation, operated over integer-valued images and did not need to solve a linear system of equations; moreover, similar ideas were used in [93–95] (which are related to an earlier algorithm [96]) which yielded to a computational efficient algorithm (specially [93, 94]), [71] also uses the anisotropic separable approximation, and does not need to solve a linear system of equations nevertheless; it introduces an auxiliary variable (similar to the core idea of ADDM/SB, see Section 3.2) resulting in an alternating algorithm; a second-order cone programming approach was proposed in [97], although it had high memory requirements and needed to solve a linear system of equations; in [77] the ℓ^1 -TV problem was solved using (linear programming) interior point method; finally we mention that the IRN algorithm [59] (IRLS based) can also handle the ℓ^1 -TV problem (both denoising and deconvolution) and was a computational efficient alternative to methods based on the Markov random field model for the denoising case.

4.2.1. Adaptation of the Regularization Parameter for ℓ^1 -TV. The original ℓ^1 -TV problem (42) features a single regularization parameter (λ), which influences the entire pixel set, and has a direct impact on the quality of the reconstructed data. Ideally, for the salt-and-pepper noise model, noise-free pixels should preserve their values in the reconstructed (denoised) image. However, the use of a global parameter forces the entire pixel set to be penalized, which results in

an inaccurate reconstruction. To the best of our knowledge, [41, 98–100] are the only published papers that tackle the above-mentioned shortcomings for the ℓ^1 -TV problem.

All the above-mentioned algorithms (in a way or another) minimize the cost functional $T(\mathbf{u}) = \|\Lambda(\mathbf{u} - \mathbf{b})\|_1 + \|\nabla\mathbf{u}\|_1$, where Λ represents a diagonal matrix, whose entries are set to zero if the pixel is declared noise-free. Nevertheless, whereas in [98] a set of corrupted pixel candidates were estimated (via [22]) to then proceed to solve the ℓ^1 -TV problem only for the corrupted candidates, being the regularization parameter hand-picked, in [99] and in [41, 100] a scheme was proposed to spatially adapt the regularization parameter; for the former, such adaptation was based only on local statistics and on the estimation of the noise level (although all pixels are still regularized); for the latter, which also estimated the set of corrupted pixel candidates via [22], the adaptation was initially based on similar ideas [100], however it was then proved [41] that the scheme that obtains the best results (reconstruction quality) is such that the nonzero entries of Λ should be increased at each iteration.

4.3. TV Numerical Algorithms for the Poisson Noise Model.

The minimization of the TV regularized Poisson log-likelihood cost functional [66, 101–108], summarized by $\min T(\mathbf{u})$ s.t. $\mathbf{u} \geq 0$, where

$$T(\mathbf{u}) = \sum_k (A\mathbf{u})_k - \mathbf{b}_k \cdot \log((A\mathbf{u})_k) + \lambda \|\nabla\mathbf{u}\|_1, \quad (43)$$

has been successfully employed to restore non-negative images corrupted with Poisson noise from a number of applications, such as PET images [101], confocal microscopy images [102], and others.

Within the TV framework, numerical algorithms that solve (43) can be loosely classified as those that use a second-order Taylor approximation of the data fidelity term $F(\mathbf{u})$ to make (43) a tractable problem and those that use other alternative approaches.

For the latter case, we mention that in [101] an EM algorithm along with a smooth approximation of $\|\nabla\mathbf{u}\|_1$ was proposed to solve the Poisson-TV problem; in [102] a multiplicative gradient based algorithm (equivalent to the penalized EM algorithm) was used; a multilevel algorithm was used in [103] to solve a modified version of (43); in [66] the authors used the ℓ^1 regularized loss minimization (a particular case of the ADMM algorithm, see [61, Section 6.3]) to minimize (43); in [107] two alternative variational methods were used (called L2-L2Log and TV-log) which can be understood as a change of variable ($z_k = \log((A\mathbf{u})_k)$) in (43).

Algorithms that do use a second-order Taylor approximation of the data fidelity term $F(\mathbf{u})$ need to explicitly address the problem of it ($F(\mathbf{u})$) being not Lipschitz continuous when any element of \mathbf{u} is (close to) zero (see Section 2.3). This issue is typically addressed by using a modified version of the second order Taylor approximation or by adding a constant to the observed data, and although the exact original problem is not solved (due to the approximations involved) the reconstruction (numerical) results have proven to be competitive. Moreover, this type of algorithms need

solve a linear system of equations and typically differ in the constrained optimization algorithm used to carry out the actual minimization and in whether the true Hessian of $F(\mathbf{u})$ or an approximation is used. For instance, in [109] a linear approximation of the Hessian is used along with a quasi-Newton optimization method, in [104] the minimization was carried out via a nonnegatively constrained, projected quasi-Newton minimization algorithm; $\nabla^2 F(\mathbf{u})$ was used. In [106] a constrained TV algorithm, described in [85], plus the approximation $\nabla^2 F(\mathbf{u}) \approx \eta_k I$, $\eta_k > 0$ was used. In [105] a Expectation-Maximization TV approach was used and only the denoising problem was addressed. In [108] the NQP algorithm was used to optimize (43) for both grayscale and color images; $\nabla^2 F(\mathbf{u})$ was used.

4.4. TV Numerical Algorithms for the Speckle (Gamma) Noise Model. The minimization of the non-convex TV multiplicative model (introduced in [38]) summarized by $\min T(\mathbf{u})$ s.t. $\mathbf{u} \geq 0$, where

$$T(\mathbf{u}) = \sum_k \frac{\mathbf{b}_k}{(A\mathbf{u})_k} + \log((A\mathbf{u})_k) + \lambda \|\nabla\mathbf{u}\|_1, \quad (44)$$

is not the only alternative within the TV framework to restore images corrupted with Speckle noise; here we mention that [110] (the first method within the TV framework), used a constrained optimization approach with two Lagrange multipliers; the denoising and deconvolution problems were addressed; additionally in [39] the multiplicative model was converted into an additive one and used a multigrid algorithm to solve the resulting Euler-Lagrange equation; Also in [111] the multiplicative model was converted into an additive one and used the SB (or ADMM) algorithm to solve the optimization problem; only the denoising problem was addressed. Moreover, a framework based on MRF with levelable priors for restoration of images corrupted by Gaussian or Speckle (Rayleigh) was proposed in [112] where only the denoising problem was addressed, and in [113] a hard thresholding of the curvelet transform of the log-image followed by a ℓ^1 -TV in the log-image domain was used; only the denoising problem was addressed.

Regarding the methods that use (44) within the TV framework, we first mention that [38] introduced such data fidelity term, which was derived using the MAP criterion; moreover, a detailed mathematical study of (44) was also carried in [38, Section 4] to address several issues, including the problem of (44) being not Lipschitz continuous (see Section 2.4): the unique solution to (44) is such that all its elements are strictly greater than zero. This result was reached without any assumption on the underlying optimization procedure to solve (44); furthermore, as to give numerical evidence on such result, in [38] an artificial time marching approach (steepest descent) was used to numerically solve the resulting Euler-Lagrange equation from (44); the denoising and deconvolution problems were addressed. A general TV formulation was proposed in [114] which included several models ([38, 110, 112]) as special cases; it also replaced the regularizer $\text{TV}(u)$ by $\text{TV}(\log u)$; only the denoising problem was addressed. In [115] the non-convex model introduced

in [38] was augmented with the Weberized TV (a term of the form $\|\nabla \mathbf{u}\|_1 / \mathbf{u}$, see also [116]) as an extra regularizer and solved the Euler-Lagrange equation via a fixed-point iteration; only the denoising problem was addressed. In [117] a second order Taylor approximation of the data fidelity term $F(\mathbf{u})$ in (44) was used along with the NQP algorithm.

4.5. TV Numerical Algorithms for the Mixed Gaussian-Impulse Noise Model. A number of algorithms have recently been proposed for denoising of images subject to the mixed noise model (16), reproduced here for convenience

$$\mathbf{b} = \mathcal{B} \cdot (\mathbf{A}\mathbf{u}^* + \boldsymbol{\eta}) + (\mathbf{I} - \mathcal{B}) \cdot \mathbf{r}; \quad (45)$$

we also recall that the key idea of the non-TV based methods is to use a two-phase approach: detect the outlier pixels before proceeding with the filtering phase.

Within the TV framework, most of the algorithms also start with an outlier detection pre-processing phase followed by a denoising phase. The approach in [119] was based on two augmented cost functionals ([119, equations (15)-(16)]) which had to be chosen depending on the noise characteristics; the reconstruction performance was competitive for grayscale images, but its computational performance was extremely poor. In [120] the computational performance of [119] was improved, although it still was highly dependent on the noise level, specially for high noise levels; In [121] an $\ell^1 - \ell^0$ minimization approach was proposed for grayscale and color images, resulting in a three-phase algorithm; reconstruction quality results and computational performance were quite good when compared with published works (and could be considered state-of-the-art), but the proposed cost functional ([121, Equation (5)]) is complicated and has several regularization parameters which have to be hand-picked, and the use of a dictionary learning second phase affects the overall computational performance. In [118] a spatially adaptive algorithm was proposed, along with the cost functional $T(\mathbf{u}) = \|\Lambda_O \cdot (\mathbf{u} - \mathbf{b})\|_1 + (1/2)\|\Lambda_G \cdot (\mathbf{u} - \mathbf{b})\|_2^2 + \|\nabla \mathbf{u}\|_1$ (minimized via an IRLS based algorithm) to address the denoising problem; the computational performance of this algorithm greatly exceeded that of [120, 121] while its reconstruction quality results can be placed somewhere in between those reported by [120, 121].

5. Numerical Simulations and Examples

The aim of this Section is to provide practical examples of restoration results that can be obtained with TV for different noise models, for which we have used several (grayscale and color) images from [122] as test images, which are scaled between 0 and 1 before blurring (by a 7×7 out-of-focus kernel); afterwards, they are corrupted with noise (models described in Section 2). To assess the reconstruction quality of the restored images the SNR (all cases) and SSIM (only for grayscale images) [123] metrics are reported.

Results presented throughout this Section may be reproduced using the [124], a Matlab-only implementation of IRLS based algorithms for TV with different fidelity terms. Since it is not the aim of this paper to present a computational

performance comparison between the many TV numerical algorithms (for all the considered different noise models) here we will only report the time (computational) performance of [124] for completeness' sake and to offer the reader a rough idea of the processing time for each case (elapsed for Matlab, version R2011a, running on a 2.20 GHz Intel core i7 CPU laptop, with 6144 K of L2 memory and 6 G of RAM). For the readers interested in a thoroughly performance comparison between the many TV numerical algorithms we recommend to review the performance reports listed in the bibliography referenced throughout Section 4.

For the ℓ^2 -TV case we choose to present simulations where we compare the reconstruction performance results of the (standard) ℓ^2 -TV (Section 4.1) with the nonnegative ℓ^2 -TV (Section 4.1.2) for the deconvolution case (see Figure 1). Reconstruction SNR values and computation times are compared in Table 1; The nonnegative ℓ^2 -TV method has a better results, in terms of visual quality although SNR/SSIM metrics are about the same, than the (standard) ℓ^2 -TV method, although the computational performance for the latter is about 2 to 3 times faster.

For the ℓ^1 -TV case we choose to present simulations where we compare the reconstruction performance results of the (standard) ℓ^1 -TV (Section 4.2) with the adaptive ℓ^1 -TV (Section 4.2.1) for the denoising case (see Figure 2). Reconstruction SNR and SSIM values and computation times are compared in Table 2; The adaptive ℓ^1 -TV method has significantly better reconstruction results, both in terms of SNR, SSIM and visual quality than the (standard) ℓ^1 -TV; there is, however, a trade-off to be made for the superior reconstruction quality: the first step of the adaptive ℓ^1 -TV, that is the estimation of the set of pixels corrupted with Salt & Pepper noise, is very slow compared to the processing time needed to solve the ℓ^1 -TV problem. Here we want to highlight that although we only report the performance for the grayscale "Boats" and color "Barbara" images, these results are representative for all other test images.

For the Poisson-TV (Section 4.3) method, we choose to present simulations for the denoising and deconvolution cases. We use the grayscale "Cameraman" and color "Peppers" images (both 512×512) which were first scaled to a maximum value $M \in \{5, 30, 100, 255\}$ (the lower the value of M is, the noisier the image will be perceived), and then were blurred ("Peppers" only), and finally the Poisson noise was added (this matches typical experimental setups, such [66, Section 6.3] and [108, Section 4]). In Table 3 (see also Figure 3) we summarize the restoration results for the Poisson-TV method.

Similarly, for the Gamma-TV (Section 4.4) method, we choose to present simulations for the denoising and deconvolution cases. We use the grayscale "Tank" and color "Lena" images (both 512×512) which were corrupted with Speckle noise, generated according to (12) with $L = 5, 10, 33$ (note that the lower the value of L is, the noisier the image will be perceived, only "Lena" was blurred). This matches typical experimental setups, such as [111, 115, 117]. In Table 4 (see also Figure 4) we summarize the restoration results for the Gamma-TV method.



FIGURE 1: Deconvolution with the ℓ^2 -TV and the nonnegative ℓ^2 -TV (via [81, 87] resp.) methods for the grayscale “Goldhill” (720×576) image and the color “Lena” (512×512) image.

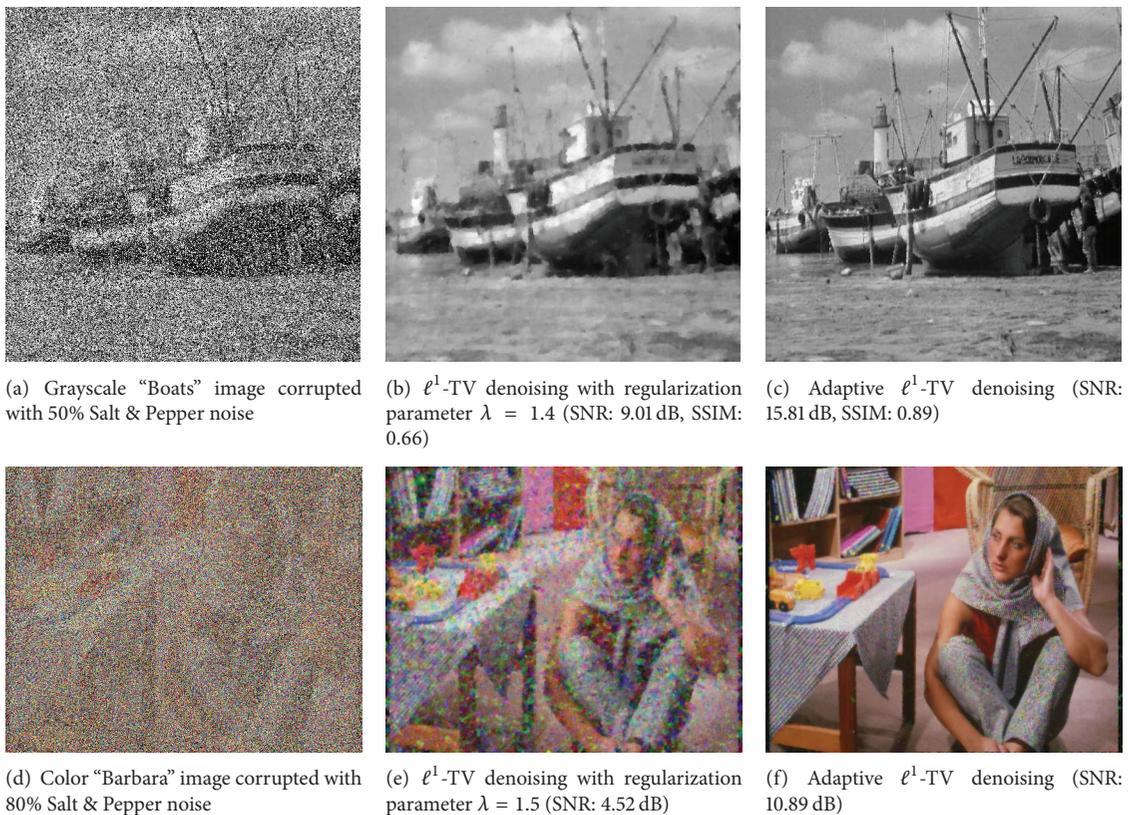


FIGURE 2: Denoising with the ℓ^1 -TV and the adaptive ℓ^1 -TV methods (via [81] and via [41, 100], resp.) for the grayscale “Boats” (512×512) image and the color “Barbara” (720×576) image.

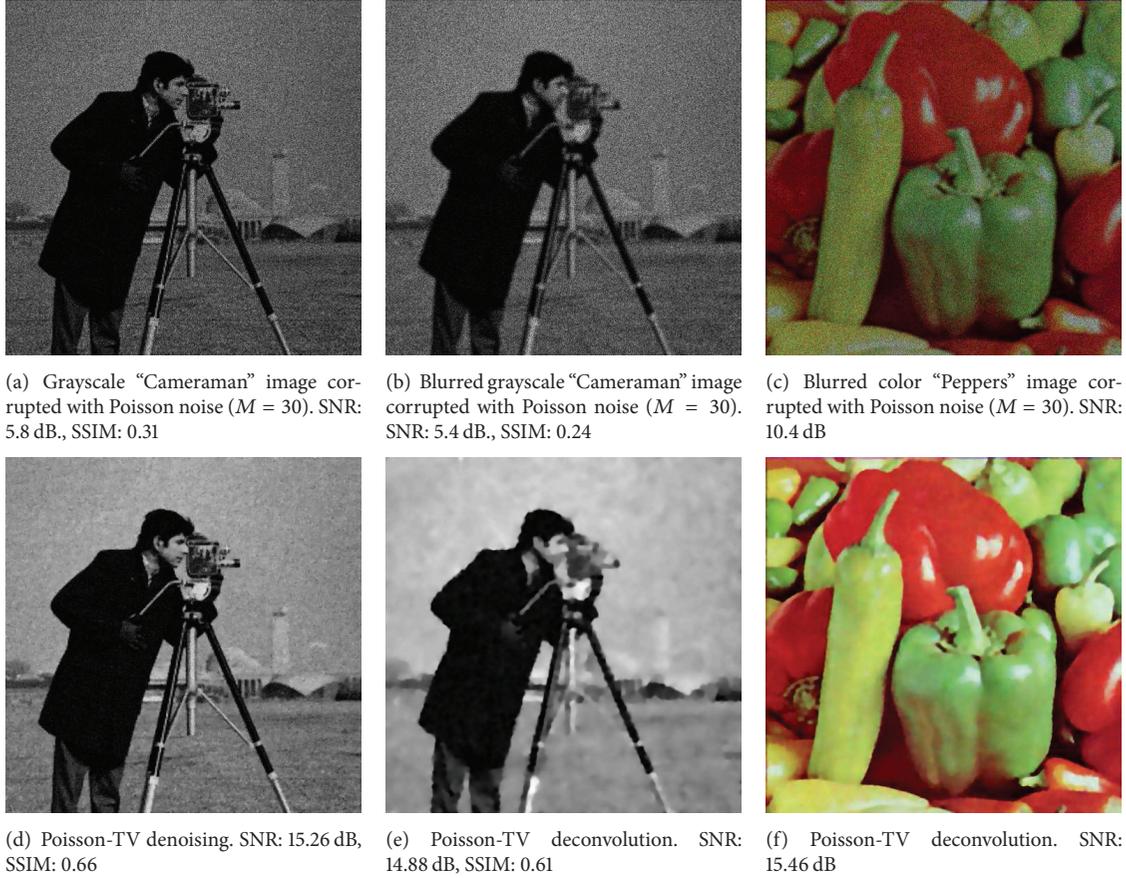


FIGURE 3: Denoising and deconvolution with the Poisson TV via [108] (a spatially adaptive algorithm).

TABLE 1: Deconvolution performance results for the standard ℓ^2 -TV and for the nonnegative ℓ^2 -TV methods for the grayscale “Goldhill” (720×576) image and the color “Lena” (512×512) image.

Image	Noise (σ^2)	SNR (dB)		SSIM		Time (s)	
		ℓ^2 -TV	Nonneg.	ℓ^2 -TV	Nonneg.	ℓ^2 -TV	Nonneg.
Goldhill (gray)	10^{-2}	13.39	13.89	0.59	0.60	6.69	11.99
	$4 \cdot 10^{-2}$	11.59	12.12	0.55	0.54	7.61	12.05
Lena (color)	10^{-2}	14.98	15.17	—	—	11.21	23.34
	$4 \cdot 10^{-2}$	13.02	13.15	—	—	12.50	23.19

TABLE 2: Denoising performance results for the ℓ^1 -TV and the adaptive ℓ^1 -TV methods for the grayscale “Boats” (512×512) image and the color “Barbara” (720×576) image. Time in parenthesis (adaptive case) corresponds to time elapsed to estimate the pixels corrupted with Salt & Pepper noise.

Image	Noise	SNR (dB)		SSIM		Time (s)	
		ℓ^1 -TV	Adaptive	ℓ^1 -TV	Adaptive	ℓ^1 -TV	Adaptive
Boats (gray)	30%	11.41	18.40	0.76	0.93	2.07	0.44 (5.11)
	50%	9.01	15.81	0.66	0.89	2.30	0.40 (5.11)
	80%	4.02	10.81	0.44	0.74	2.41	0.52 (8.57)
Barbara (color)	30%	12.23	17.08	—	—	11.40	2.89 (23.08)
	50%	10.22	14.71	—	—	12.46	2.46 (26.01)
	80%	4.52	10.89	—	—	13.66	2.84 (40.25)

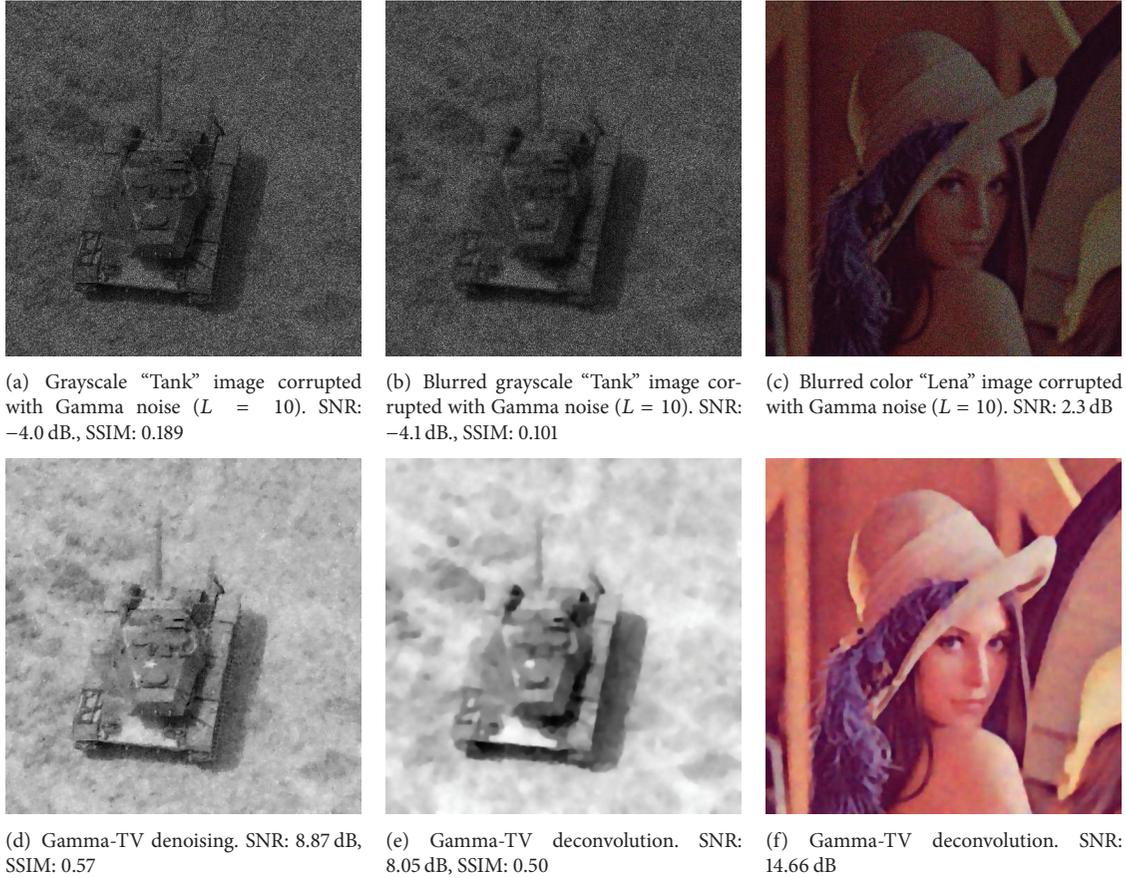


FIGURE 4: Denoising and deconvolution with the Gamma TV via [117].

TABLE 3: Denoising and deconvolution performance results for the Poisson-TV method for the grayscale “Cameraman” (512×512) image and the color “Peppers” (512×512) image.

Image	M	SNR (dB)		SSIM		Time (s)	
		Denoising	Deconv.	Denoising	Deconv.	Denoising	Deconv.
Cameraman (gray)	5	7.48	6.89	0.34	0.35	3.19	3.78
	30	15.26	14.88	0.66	0.61	3.10	3.77
	100	18.91	16.47	0.80	0.84	3.13	3.75
	255	19.55	17.70	0.78	0.86	3.10	3.98
Peppers (color)	5	13.59	12.44	—	—	10.85	12.53
	30	16.31	15.46	—	—	10.47	12.49
	100	17.40	16.46	—	—	10.89	12.36
	255	17.72	17.15	—	—	10.54	12.43

TABLE 4: Denoising and deconvolution performance results for the Gamma-TV method for the grayscale “Tank” (512×512) image and the color “Lena” (512×512) image.

Image	M	SNR (dB)		SSIM		Time (s)	
		Denoising	Deconv.	Denoising	Deconv.	Denoising	Deconv.
Tank (gray)	5	7.36	7.50	0.519	0.490	27.13	32.18
	10	8.87	8.05	0.573	0.502	27.02	32.27
	33	10.39	8.50	0.624	0.517	26.60	32.42
Lena (color)	5	12.93	13.41	—	—	85.88	96.36
	10	15.72	14.66	—	—	86.47	89.82
	33	18.16	15.54	—	—	56.83	77.87

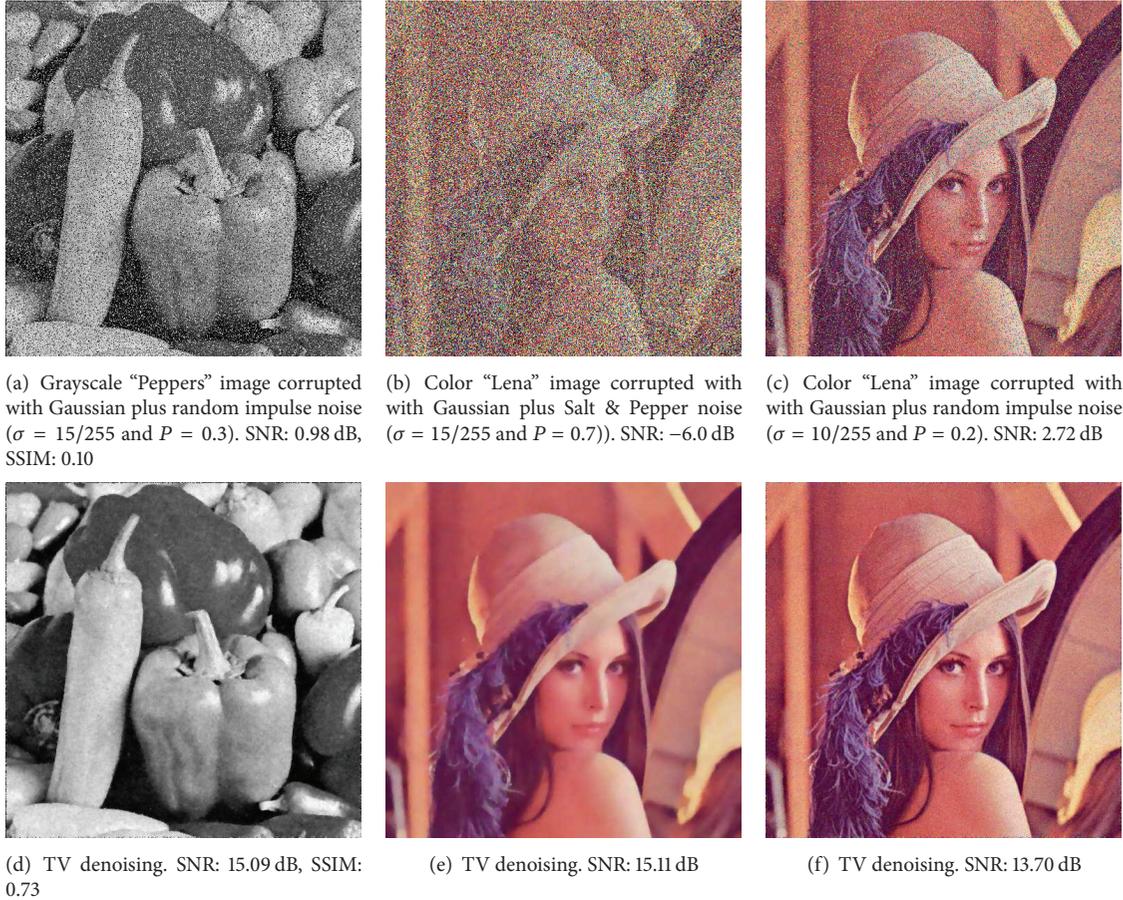


FIGURE 5: TV denoising for the mixed Gaussian-Impulse noise model via [118] (a spatially adaptive algorithm).

TABLE 5: Denoising performance results for the mixed Gaussian Impulse TV method for the grayscale “Peppers” and the color “Lena” images (both 512×512). The Gaussian plus Salt & Pepper noise case is marked as “G + S & P”, whereas the Gaussian plus random-valued impulse noise case is marked as “G + R.”

Image	Noise			SNR (dB)		SSIM		Time (s)		
	σ	$p_{S\&P}$	p_R	G + S & P	G + R	G + S & P	G + R	G + S & P	G + R	
Peppers (gray)	5/255	0.3	0.1	19.88	18.80	0.87	0.83	10.43	12.05	
		0.5	0.2	18.15	16.43	0.85	0.76	10.85	11.03	
		0.7	0.3	15.99	14.62	0.82	0.70	11.95	10.81	
	10/255	0.3	0.1	18.28	18.13	0.82	0.81	11.28	12.22	
		0.5	0.2	16.97	16.68	0.81	0.78	11.52	12.49	
		0.7	0.3	14.89	15.33	0.78	0.75	12.94	11.89	
	15/255	0.3	0.1	17.02	17.37	0.79	0.78	11.45	12.52	
		0.5	0.2	15.98	16.26	0.78	0.76	11.81	12.52	
		0.7	0.3	13.98	15.09	0.75	0.73	13.17	12.04	
	Lena (color)	5/255	0.3	0.1	21.16	17.38	—	—	38.09	30.62
			0.5	0.2	19.76	15.13	—	—	37.57	29.06
			0.7	0.3	17.51	13.48	—	—	42.24	29.31
10/255		0.3	0.1	18.93	15.04	—	—	38.18	30.50	
		0.5	0.2	17.85	13.70	—	—	40.10	31.10	
		0.7	0.3	16.02	12.68	—	—	44.05	29.46	
15/255		0.3	0.1	17.76	12.78	—	—	39.08	30.19	
		0.5	0.2	16.77	12.15	—	—	40.82	29.18	
		0.7	0.3	15.11	11.41	—	—	45.10	32.65	

Finally, for the mixed model TV (Section 4.5) method, we choose to present simulations for three denoising cases, where in one case the grayscale “Peppers” image (512×512) was corrupted with Gaussian plus random-valued impulse noise generated according to (45) with $\sigma = 15/255$ and $P = 0.3$, and in the other two, color “Lena” image (512×512) was corrupted with Gaussian plus Sal & Pepper noise as well as with Gaussian plus random-valued impulse noise with $\sigma = 15/255$ and $P = 0.7$, and $\sigma = 10/255$ and $P = 0.2$, respectively (according to (45)), see Figure 5. Additionally, in Table 5 we summarize the restoration results for the mixed model TV method for a broader combination of levels of Gaussian and Impulse noises that matches typical experimental setups [118, 120, 121].

6. Conclusions

To the best of our knowledge, we have provided a complete summary of recent TV algorithms for different noise models, that is, Gaussian, Impulse (e.g., Salt & Pepper), Poisson, and Speckle (e.g., Gamma) and the mixed Gaussian and impulse noise models. Although for some noise models, there are particular algorithms that have better reconstruction performance (e.g., patch-based approaches for the Gaussian noise model) we expect that in the coming years TV will improve its current reconstruction performance (and hopefully become competitive with state-of-the-art alternative denoising methods); moreover, we believe that the development of new TV algorithms that spatially adapt the regularization parameter is a possible answer (see the computational results for the adaptive ℓ^1 -TV case) to current reconstruction performance shortcomings when compared to patch-based approaches.

Finally, it is important to highlight the flexibility of the TV method, that allows it to be relevant to applications that come from a variety of fields, ranging from astronomical to SAR based systems, including medical applications such as Ultrasound, PET, and others.

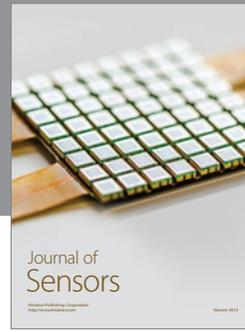
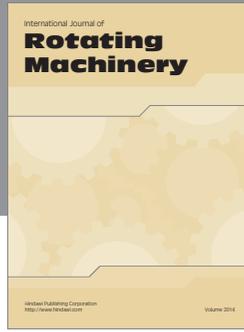
References

- [1] A. C. Bovik, *The Essential Guide To Image Processing*, Academic Press, Boston, Mass, USA, 2009.
- [2] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [3] T. Chan, S. Esedoglu, F. Park, and A. Yip, “Recent developments in total variation image restoration,” in *Handbook of Mathematical Models in Computer Vision*, Springer, New York, NY, USA, 2005.
- [4] D. Strong and T. Chan, “Edge-preserving and scale-dependent properties of total variation regularization,” *Inverse Problems*, vol. 19, no. 6, pp. S165–S187, 2003.
- [5] P. Kvam and B. Vidakovic, *Nonparametric Statistics with Applications to Science and Engineering*, Wiley-Interscience, New York, NY, USA, 2007.
- [6] N. P. Galatsanos, A. K. Katsaggelos, R. T. Chin, and A. D. Hillery, “Least squares restoration of multichannel images,” *IEEE Transactions on Signal Processing*, vol. 39, no. 10, pp. 2222–2236, 1991.
- [7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 3rd edition, 1991.
- [8] M. Lindenbaum, M. Fischer, and A. Bruckstein, “On Gabor’s contribution to image enhancement,” *Pattern Recognition*, vol. 27, no. 1, pp. 1–8, 1994.
- [9] M. R. Banham and A. K. Katsaggelos, “Digital image restoration,” *IEEE Signal Processing Magazine*, vol. 14, no. 2, pp. 24–41, 1997.
- [10] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proceedings of the IEEE 6th International Conference on Computer Vision*, pp. 839–846, January 1998.
- [11] S. M. Smith and J. M. Brady, “SUSAN—a new approach to low level image processing,” *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, 1997.
- [12] D. L. Donoho, “De-noising by soft-thresholding,” *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [13] A. Buades, B. Coll, and J. M. Morel, “A review of image denoising algorithms, with a new one,” *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [14] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [15] R. J. Marks, G. L. Wise, D. G. Haldeman, and J. L. Whited, “Detection in laplace noise,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 14, no. 6, pp. 866–872, 1978.
- [16] S. Alliney, “Digital filters as absolute norm regularizers,” *IEEE Transactions on Signal Processing*, vol. 40, no. 6, pp. 1548–1562, 1992.
- [17] S. Alliney, “A property of the minimum vectors of a regularizing functional defined by means of the absolute norm,” *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 913–917, 1997.
- [18] S. Alliney and S. A. Ruzinsky, “Algorithm for the minimization of mixed l_1 and l_2 norms with application to Bayesian estimation,” *IEEE Transactions on Signal Processing*, vol. 42, no. 3, pp. 618–627, 1994.
- [19] M. Nikolova, “Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers,” *SIAM Journal on Numerical Analysis*, vol. 40, no. 3, pp. 965–994, 2002.
- [20] T. F. Chan and S. Esedoglu, “Aspects of total variation regularized L^1 function approximation,” *SIAM Journal on Applied Mathematics*, vol. 65, no. 5, pp. 1817–1837, 2005.
- [21] J. Tukey, “Nonlinear (nonsuperimposable) methods for smoothing data,” in *Proceedings of the IEEE Electronics and Aerospace Conference (EASCON ’74)*, p. 673, Conference Records, 1974.
- [22] H. Hwang and R. A. Haddad, “Adaptive median filters: new algorithms and results,” *IEEE Transactions on Image Processing*, vol. 4, no. 4, pp. 499–502, 1995.
- [23] S.-J. Ko and Y. H. Lee, “Center weighted median filters and their applications to image enhancement,” *IEEE Transactions on Circuits and Systems*, vol. 38, no. 9, pp. 984–993, 1991.
- [24] Z. Wang and D. Zhang, “Progressive switching median filter for the removal of impulse noise from highly corrupted images,” *IEEE Transactions on Circuits and Systems II*, vol. 46, no. 1, pp. 78–80, 1999.
- [25] R. Garnett, T. Huegerich, C. Chui, and W. He, “A universal noise removal algorithm with an impulse detector,” *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1747–1754, 2005.

- [26] S. Schulte, M. Nachtgaele, V. de Witte, D. van der Weken, and E. E. Kerre, "A fuzzy impulse noise detection and reduction method," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1153–1162, 2006.
- [27] Y. Dong and S. Xu, "A new directional weighted median filter for removal of random-valued impulse noise," *IEEE Signal Processing Letters*, vol. 14, no. 3, pp. 193–196, 2007.
- [28] W. Richardson, "Bayesian-based iterative method of image restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55–59, 1972.
- [29] L. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, p. 745, 1974.
- [30] J. Starck and F. Murtagh, *Astronomical Image and Data Analysis (Astronomy and Astrophysics Library)*, Springer, New York, NY, USA, 2006.
- [31] B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for poisson noise removal," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1093–1108, 2008.
- [32] F.-X. Dupe, J. M. Fadili, and J.-L. Starck, "A proximal iteration for deconvolving Poisson noisy images using sparse representations," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 310–321, 2009.
- [33] M. Mäkitalo and A. Foi, "Optimal inversion of the anscombe transformation in low-count poisson image denoising," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 99–109, 2011.
- [34] F. Anscombe, "The transformation of poisson, binomial and negativebinomial data," *Biometrika*, vol. 35, no. 3-4, pp. 246–254, 1948.
- [35] M. Freeman and J. Tukey, "Transformations related to the angular and the square root," *The Annals of Mathematical Statistics*, vol. 21, no. 4, pp. 607–611, 1950.
- [36] J. W. Goodman, "Some fundamental properties of speckle," *Journal of the Optical Society of America*, vol. 66, no. 11, pp. 1145–1150, 1976.
- [37] C. Loizou and C. Pattichis, *Despeckle Filtering Algorithms and Software For Ultrasound Imaging*, Synthesis Lectures on Algorithms and Software in Engineering, Morgan & Claypool Publishers, 2008.
- [38] G. Aubertt and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM Journal on Applied Mathematics*, vol. 68, no. 4, pp. 925–946, 2008.
- [39] C. Sheng, Y. Xin, Y. Liping, and S. Kun, "Total variation-based speckle reduction using multi-grid algorithm for ultrasound images," in *Image Analysis and Processing ICIAP 2005*, vol. 3617, pp. 245–252, Springer, Berlin, Germany, 2005.
- [40] B. Li, Q. Liu, J. Xu, and X. Luo, "A new method for removing mixed noises," *Science China Information Sciences*, vol. 54, no. 1, pp. 51–59, 2011.
- [41] R. Rojas, *Automatic regularization parameter selection for the total variation mixed noise image restoration framework [M.S. thesis]*, Pontifical Catholic University of Peru (PUCP), 2012.
- [42] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 2nd edition, 2006.
- [43] D. G. Luenberge and Y. Ye, *Linear and Nonlinear Programming*, Springer, New York, NY, USA, 3rd edition, 2008.
- [44] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational Pde, Wavelet, and Stochastic Methods*, SIAM, 2005.
- [45] A. Bonnet, "On the regularity of edges in image segmentation," *Annales de L'Institut Henri Poincaré C*, vol. 13, no. 4, pp. 485–528, 1996.
- [46] P. Blomgren and T. F. Chan, "Color TV: total variation methods for restoration of vector-valued images," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 304–309, 1998.
- [47] P. Huber, "Robust regression: asymptotics, conjectures and monte carlo," *The Annals of Statistics*, vol. 1, no. 5, pp. 799–821, 1973.
- [48] Y. Li and F. Santosa, "A computational algorithm for minimizing total variation in image restoration," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 987–995, 1996.
- [49] C. Vogel, *Computational Methods For Inverse Problems*, SIAM, 2002.
- [50] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM Journal on Scientific Computing*, vol. 17, no. 1, pp. 227–238, 1996.
- [51] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 813–824, 1998.
- [52] T. F. Chan, G. H. Golub, and P. Mulet, "Nonlinear primal-dual method for total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 20, no. 6, pp. 1964–1977, 1999.
- [53] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.
- [54] D. Krishnan, P. Lin, and A. M. Yip, "A primal-dual active-set method for non-negativity constrained total variation deblurring problems," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2766–2777, 2007.
- [55] A. E. Beaton and J. W. Tukey, "The fitting of power series, meaning polynomials illustrated on band-spectroscopic data," *Technometrics*, vol. 16, no. 2, pp. 147–185, 1974.
- [56] R. Byrd and D. Payne, "Convergence of the iteratively reweighted least squares algorithm for robust regression," Tech. Rep. 313, Johns Hopkins University, 1979.
- [57] S. A. Ruzinsky and E. T. Olsen, " L_1 and L_∞ minimization via a variant of Karmarkar's algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 2, pp. 245–253, 1989.
- [58] I. Daubechies, R. DeVore, M. Fornasier, and C. G. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38, 2010.
- [59] P. Rodríguez and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 322–332, 2009.
- [60] M. A. T. Figueiredo, J. B. Dias, J. P. Oliveira, and R. D. Nowak, "On total variation denoising: a new majorization-minimization algorithm and an experimental comparison with wavelet denoising," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 2633–2636, Atlanta, GA, USA, October 2006.
- [61] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.
- [62] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.
- [63] T. Goldstein and S. Osher, "The Split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

- [64] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.
- [65] R. Glowinski and A. Marrocco, "Sur l'approximation par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires," *Revue Française d'Automatique, Informatique, Recherche Opérationnelle*, vol. 9, no. 2, pp. 41–72, 1975.
- [66] M. A. T. Figueiredo and J. M. Bioucas-Dias, "Restoration of poissonian images using alternating direction optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.
- [67] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Multiplicative noise removal using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1720–1730, 2010.
- [68] P. Rodriguez and B. Wohlberg, "A comparison of the computational performance of iteratively reweighted least squares and alternating minimization algorithms for ℓ_1 inverse problems," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '12)*, pp. 3069–3072, 2012.
- [69] F. Sha, Y. Lin, L. K. Saul, and D. D. Lee, "Multiplicative updates for nonnegative quadratic programming," *Neural Computation*, vol. 19, no. 8, pp. 2004–2031, 2007.
- [70] P. O'Grady and S. Rickard, "Recovery of non-negative signals from compressively sampled observations via non-negative quadratic programming," in *Proceedings of the Signal Processing with Adaptive Sparse Structured Representations (SPARS '09)*, Saint Malo, France, April 2009.
- [71] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition-modeling, algorithms, and parameter selection," *International Journal of Computer Vision*, vol. 67, no. 1, pp. 111–136, 2006.
- [72] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, "Total variation-based image deconvolution: a majorization-minimization approach," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, pp. II861–II864, Toulouse, France, May 2006.
- [73] M. A. T. Figueiredo, J. B. Dias, J. P. Oliveira, and R. D. Nowak, "On total variation denoising: a new majorization-minimization algorithm and an experimental comparison with wavelet denoising," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 2633–2636, Atlanta, GA, USA, October 2006.
- [74] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society B*, vol. 39, no. 1, pp. 1–38, 1977.
- [75] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.
- [76] Y. Wang, W. Yin, and Y. Zhang, "A fast algorithm for image deblurring with total variation regularization," Tech. Rep., Rice University, 2007.
- [77] H. Fu, M. Ng, M. Nikolova, and J. Barlow, "Efficient minimization methods of mixed ℓ_2 - ℓ_1 and ℓ_1 - ℓ_1 norms for image restoration," *SIAM Journal on Scientific Computing*, vol. 27, no. 6, pp. 1881–1902, 2006.
- [78] T. Chan, S. Kang, and J. Shen, "Total variation denoising and enhancement of color images based on the CB and HSV color models," *The Journal of Visual Communication and Image Representation*, vol. 12, pp. 422–435, 2001.
- [79] X. Bresson and T. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems and Imaging*, vol. 2, no. 4, pp. 455–484, 2008.
- [80] Y.-W. Wen, M. K. Ng, and Y.-M. Huang, "Efficient total variation minimization methods for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2081–2088, 2008.
- [81] P. Rodriguez and B. Wohlberg, "A generalized vector-valued total variation algorithm," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '09)*, pp. 1309–1312, Cairo, Egypt, November 2009.
- [82] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.
- [83] L. Bar, A. Brook, N. Sochen, and N. Kiryati, "Deblurring of color images corrupted by impulsive noise," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1101–1111, 2007.
- [84] P. Rodriguez, "Multiplicative updates algorithm to minimize the generalized total variation functional with a non-negativity constraint," in *Proceedings of the 17th IEEE International Conference on Image Processing, ICIP 2010*, pp. 2509–2512, Hong Kong, China, September 2010.
- [85] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.
- [86] R. Chartrand and B. Wohlberg, "Total-variation regularization with bound constraints," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '10)*, pp. 766–769, Dallas, TX, USA, March 2010.
- [87] P. Rodriguez, "A non-negative quadratic programming approach to minimize the generalized vector-valued total variation functional," in *Proceedings of the European Signal Processing Conference (EUSIPCO '10)*, pp. 314–318, Aalborg, Denmark, 2010.
- [88] D. Strong and T. Chan, "Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing," Division in Image Processing, UCLA Math Department CAM Report, 1996.
- [89] S. Ramani, T. Blu, and M. Unser, "Monte-Carlo sure: a black-box optimization of regularization parameters for general denoising algorithms," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1540–1554, 2008.
- [90] Y. Lin, B. Wohlberg, and H. Guo, "UPRE method for total variation parameter selection," *Signal Processing*, vol. 90, no. 8, pp. 2546–2551, 2010.
- [91] M. Nikolova, "A variational approach to remove outliers and impulse noise," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 99–120, 2004.
- [92] A. Chambolle, "Total variation minimization and a class of binary MRF models," in *Proceedings of the 5th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, vol. 3757, pp. 136–152, 2005.
- [93] J. Darbon and M. Sigelle, "Image restoration with discrete constrained total variation part I: fast and exact optimization," *Journal of Mathematical Imaging and Vision*, vol. 26, no. 3, pp. 261–276, 2006.

- [94] J. Darbon and M. Sigelle, "Image restoration with discrete constrained total variation part II: levelable functions, convex priors and non-convex cases," *Journal of Mathematical Imaging and Vision*, vol. 26, no. 3, pp. 277–291, 2006.
- [95] D. Goldfarb and W. Yin, "Parametric maximum flow algorithms for fast total variation minimization," Tech. Rep. CAAM TR07-09, Department of Computational and Applied Mathematics, Rice University, 2007.
- [96] D. S. Hochbaum, "An efficient algorithm for image segmentation, Markov random fields and related problems," *Journal of the ACM*, vol. 48, no. 4, pp. 686–701, 2001.
- [97] D. Goldfarb and W. Yin, "Second-order cone programming methods for total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 27, no. 2, pp. 622–645, 2006.
- [98] R. H. Chan, C.-W. Ho, and M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1479–1485, 2005.
- [99] Y. Dong, M. Hintermüller, and M. Rincon-Camacho, "Automated regularization parameter selection in Multi-Scale total variation models for image restoration," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 82–104, 2010.
- [100] R. Rojas and P. Rodríguez, "Spatially adaptive total variation image denoising under salt and pepper noise," in *Proceedings of the European Signal Processing Conference (EUSIPCO '11)*, pp. 314–318, Barcelona, Spain, 2011.
- [101] E. Jonsson, S. Huang, and T. Chan, "Total-variation regularization in positron emission tomography," UCLA CAM Report 98-48, 1998.
- [102] N. Dey, L. Blanc-Feraud, C. Zimmer et al., "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," *Microscopy Research and Technique*, vol. 69, no. 4, pp. 260–266, 2006.
- [103] R. H. Chan and K. Chen, "Multilevel algorithm for a Poisson noise removal model with total-variation regularization," *International Journal of Computer Mathematics*, vol. 84, no. 8, pp. 1183–1198, 2007.
- [104] J. M. Bardsley and A. Luttmann, "Total variation-penalized Poisson likelihood estimation for ill-posed problems," *Advances in Computational Mathematics*, vol. 31, no. 1–3, pp. 25–59, 2009.
- [105] A. Sawatzky, C. Brune, J. Müller, and M. Burger, "Total variation processing of images with Poisson statistics," in *Proceedings of the 13th International Conference on Computer Analysis of Images and Patterns (CAIP '09)*, vol. 5702, pp. 533–540, 2009.
- [106] R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Poisson image reconstruction with total variation regularization," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 4177–4180, Hong Kong, China, September 2010.
- [107] I. C. Rodrigues and J. M. R. Sanches, "Convex total variation denoising of poisson fluorescence confocal images with anisotropic filtering," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 146–160, 2011.
- [108] P. Rodriguez, "Total variation regularization for Poisson vector-valued image restoration with a spatially adaptive regularization parameter selection," in *Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis (ISPA '11)*, pp. 402–407, Dubrovnik, Croatia, September 2011.
- [109] R. Chartrand and V. Staneva, "A quasi-newton method for total variation regularization of images corrupted by non-gaussian noise," *IET Image Processing*, vol. 2, pp. 295–303, 2008.
- [110] S. Osher, N. Paragios, L. Rudin, and P. Lions, "Multiplicative denoising and deblurring: theory and algorithms," in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, pp. 103–119, Springer, New York, NY, USA, 2003.
- [111] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Total variation restoration of speckled images using a split-Bregman algorithm," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '09)*, pp. 3717–3720, Cairo, Egypt, November 2009.
- [112] J. Darbon, M. Sigelle, and F. Tupin, "A note on nice-levelable MRFs for SAR image denoising with contrast preservation," Rapport Interne ENST 2006D006, Ecole Nat. Supérieure des Télécommunications, Paris, France, 2006.
- [113] S. Durand, J. Fadili, and M. Nikolova, "Multiplicative noise removal using l_1 fidelity on frame coefficients," *Journal of Mathematical Imaging and Vision*, vol. 36, no. 3, pp. 201–226, 2010.
- [114] J. Shi and S. Osher, "A nonlinear inverse scale space method for a convex multiplicative noise model," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 294–321, 2008.
- [115] L. Xiao, L.-L. Huang, and Z.-H. Wei, "Multiplicative noise removal via a novel variational model," *Eurasip Journal on Image and Video Processing*, vol. 2010, Article ID 250768, 2010.
- [116] J. Shen, "On the foundations of vision modeling: I. Weber's law and Weberized TV restoration," *Physica D*, vol. 175, no. 3–4, pp. 241–251, 2003.
- [117] P. Rodríguez, "Nonconvex total variation speckled image restoration via nonnegative quadratic programming algorithm," in *Proceedings of the European Signal Processing Conference (EUSIPCO '11)*, pp. 288–292, Barcelona, Spain, August 2011.
- [118] P. Rodríguez, R. Rojas, and B. Wohlberg, "Mixed gaussian-impulse noise image restoration via total variation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '12)*, pp. 1077–1080, Kyoto, Japan, March 2012.
- [119] J. Cai, R. Chan, and M. Nikolova, "Two-phase methods for deblurring corrupted by impulse plus Gaussian noise," *Inverse Problems and Imaging*, vol. 2, no. 2, pp. 187–204, 2008.
- [120] J. Cai, R. Chan, and M. Nikolova, "Fast two-phase image deblurring under impulse noise," *Journal of Mathematical Imaging and Vision*, vol. 36, no. 1, pp. 46–53, 2010.
- [121] Y. Xiao, T. Zeng, J. Yu, and M. K. Ng, "Restoration of images corrupted by mixed Gaussian-impulse noise via l_1 - l_0 minimization," *Pattern Recognition*, vol. 44, no. 8, pp. 1708–1720, 2011.
- [122] Signal and I. P. I. of the University of Southern California, "The USC-SIPI image database," <http://sipi.usc.edu/database/>.
- [123] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [124] P. Rodriguez, "IRLS based TV algorithms for different noise models, Matlab central file," <http://www.mathworks.com/matlabcentral/>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

