Pakistan Journal of Engineering and Technology
Multidisciplinary | Peer Reviewed | Open Access
Volume: 2, Number: 1, Pages: 51- 58, Year: 2019

Variational Regularization for Multi-Channel Image Denoising

Muhammad Wasim Nawaz[1], Abdesselam Bouzerdoum[2], Senior Member, IEEE, and Son Lam Phung[2], Member, IEEE

[1] Department of Computer Engineering, The University of Lahore, Lahore, 54000, Pakistan
[2] School of Electrical, Computer and Telecommunications Engineering, Northfields Avenue, Wollongong, Australia

Corresponding author: Muhammad Wasim Nawaz (e-mail: Muhammad.wasim@dce.uol.edu.pk).

Abstract- **Image restoration from noisy observations is an inverse problem. Total variation (TV) is widely used to regularize this problem. TV preserves object boundaries better than a quadratic regularizer; however, it performs poor in low-textured image regions because it generates undesirable staircase artefacts. Furthermore, TV can preserve sharp horizontal and vertical edges; however, it causes the unnecessary smoothing of edges at an angle other than 0° or 90°. This problem arises because TV minimizes the gradient magnitude. Therefore, to preserve sharp boundaries, the design of an efficient variational regularizer is crucial. This paper presents a novel regularizer for the denoising of multi-channel vector valued image. The proposed regularizer uses horizontal, vertical as well as diagonal derivatives, and imposes the intensity continuity of partial image derivatives at each pixel of the underlying image. Experiments reveal that the proposed regularizer preserves edges and object boundaries better than TV based regularizers. This regularizer is also able to reduce undesirable staircase artefacts produced by TV in flat image regions.**

Index Terms -- **Image denoising, Regularization, Sparsity, Total variation, Multi-channel images.**

I. INTRODUCTION

Most of image processing tasks are inverse problems where the aim is to find the solution of an unknown signal from noisy observations. Variational methods stabilize the solution of these ill-posed problems by regularizing unknown signals [15]. The regularization is required to obtain physically plausible solutions. A good regularizer ensures a stable estimation of the unknown signal; therefore, the design of an efficient variational regularizer is crucial.

Variational methods impose a smoothness constraint to regularize ill-posed image processing tasks. A quadratic regularizer blurs strong edges and object boundaries by penalizing intensity variations at or across them. To protect sharp edges and boundaries, robust norms are used with the smoothness constraint [4]. The ℓ_1 norm is of particular interest because it makes the variational functional convex. TV regularization has been successfully used in numerous image processing tasks such as image denoising [2, 12, 23, 28], image restoration [18, 19], image deconvolution [8] and image de-blurring [2, 20]. The TV regularizer promotes the sparsity of the computed solution. It preserves object boundaries better than a quadratic regularizer; however, it performs poor in low-textured image regions because it generates undesirable staircase artefacts. Thus, this paper proposes a novel sparsity enhancing regularizer, which aims to overcome shortcomings of the TV regularizer.

The rest of the paper is organized as follows. Section 2 reviews regularization techniques for variational methods. Isotropic, anisotropic and higher order total variation regularizations are focused in this review. Section 3 proposes a novel regularizer that enforces the continuity of partial derivatives of the underlying image. The rotational invariance of the proposed regularizer is proved. Section 4 embeds the proposed regularizer into a variational framework to denoise multi-channel images. It also gives algorithmic details of the proposed method. Section 5 presents experimental results to show the superiority of the proposed regularizer over total variation for noisy image restoration. Section 6 concludes the paper.

II. SPARSITY PROMOTING TV REGULARIZATION

This section presents sparsity promoting isotropic, anisotropic and higher order TV regularizers.

A. ISOTROPIC TOTAL VARIATION

An isotropic quantity does not change its value regardless of its direction of measurement. An isotropic regularizer applies the same amount of regularization in each direction [6]. Let a digital image. *F(i, j)* be defined for the horizontal and vertical co-ordinates *i* and *j*, respectively, over a domain $\mathfrak{D}$. The discrete isotropic TV (iTV) of *F(i, j)* can be defined as the sum of the magnitude of the image gradient at each pixel:

$$\begin{aligned} ||\nabla F(i,j)||_{\text{iTV}} &:= \sum_{i,j\in\mathfrak{D}} |\nabla F(i,j)| \\ &= \sum_{i,j\in\mathfrak{D}} \sqrt{[\nabla_x F(i,j)]^2 + [\nabla_y F(i,j)]^2}. \end{aligned} \quad (1)$$

The sum of the gradient magnitude makes TV a semi ℓ_1 norm. It is well-known that ℓ_1 norm of the gradient promotes the sparsity of the image in the gradient domain [5]. Therefore, a piecewise smooth image is obtained by minimizing (1). An isotropic TV can preserve sharp horizontal and vertical edges; however, it causes the unnecessary smoothing of edges at an angle other than 0^o or 90^o. This problem arises because isotropic TV minimizes the gradient magnitude. This problem can be reduced by using variants of TV, for example, anisotropic TV [13, 25, 26], nonlocal TV [16, 21] or higher order TV [7, 27].

B. ANISOTROPIC TOTAL VARIATION

Anisotropic TV applies a direction dependent regularization to the underlying image. It imposes the smoothing along strong intensity structures but not across them [26]. For a discrete image *F(i, j)*, the anisotropic TV (aTV) can be defined as the sum of the absolute difference of partial image derivatives:

$$||\nabla F(i,j)||_{\text{aTV}} := \sum_{i,j\in\mathfrak{D}} |\nabla_x F(i,j)| + |\nabla_y F(i,j)|. \quad (2)$$

Anisotropic TV regularization performs better than isotropic TV at strong intensity structures such as edges and object boundaries. However, unlike isotropic TV, it is not rotationally invariant to the pixel grid. Thus, it produces suboptimal solutions in the presence of rotations of the camera or the pixel grid [17].

Discontinuities in an image occur along object boundaries where the image gradient is high. Therefore, making the regularization adaptive to the image structure can preserve sharp boundaries better than a non-adaptive regularization. To this end, anisotropic TV regularization is sometimes weighted by an image-driven weight function $w(|\nabla I|)$ as

$$E_{\text{reg}}(F(i,j)) = \sum_{i,j\in\mathfrak{D}} w(|\nabla I|)\big(|\nabla_x F(i,j)| + |\nabla_y F(i,j)|\big). \quad (3)$$

For small positive numbers α and β, $w(|\nabla I|)$ can be chosen as $w(|\nabla I|) = \exp(-\alpha|\nabla I|\beta)$. Anisotropic TV is easier to minimize than its isotropic counterpart. Therefore, numerous convex minimization methods can be used to minimize anisotropic TV. These include gradient methods [3, 14], primal dual methods [6], iterative shrinkage or thresholding-based methods [2, 4], and graph cuts based methods [11, 17].

C. HIGHER ORDER TOTAL VARIATION

Mathematically, higher order total variation (HOTV) can be given as

$$||\nabla F(i,j)||_{\text{HOTV}} := \sum_{i,j\in\mathfrak{D}} |\nabla F(i,j)| + \alpha|\nabla^k F(i,j)|, \quad (4)$$

where k represents the order of the regularization, and a positive constant α balances the effect of gradient and higher order derivatives. The idea of using higher order derivatives has been modified to include Laplacian Δ with the gradient for the regularization of unknown images as

$$||\nabla F(i,j)||_{\text{Lap-TV}} := \sum_{i,j\in\mathfrak{D}} |\nabla F(i,j)| + \alpha|\Delta F(i,j)|. \quad (5)$$

Isotropic and anisotropic TV produce staircase artefacts in low-textured and flat image regions. To reduce this undesirable effect, higher order total variation regularization has been proposed [7].

The use of higher order derivatives may result in the blurring of sharp image boundaries. Thus, higher order TV regularization uses an adaptive functional which makes the regularizer act as ordinary TV at sharp boundaries, whereas it uses higher order derivatives in textured and flat image regions.

Total generalized variation (TGV) has been proposed as a generalization of higher order TV regularization [22, 27]. By changing the order of the regularizer, TGV allows to reconstruct piecewise smooth, affine and quadratic images. The TGV regularizer can be used to obtain a globally optimal solution because, similar to the TV regularizer, it is also convex. TGV and HOTV are computationally more expensive than the ordinary TV because of the calculation of higher order derivatives.

III. PROPOSED REGULARIZER

This section proposes a novel regularizer capable of avoiding the shortcomings of TV based regularizers. The proposed regularizer is based on the variational measure introduced in [24], which imposes the intensity continuity of partial image derivatives at each pixel in a small neighborhood. The variational measure is defined for scalar images only. However, the proposed regularizer is designed to handle multi-channel vector valued images. This regularizer, in contrast to TV based regularizers, can preserve edges and object boundaries which are not either horizontal or vertical. It is also able to reduce undesirable staircase artefacts produced by TV in flat regions where there is a little intensity change. First, the regularizer is formulated for multi-channel images. Second, the rotational invariance of the proposed regularizer is proved for multi-channel images.

A. THE FORMULATION

Let $\mathbf{F} = [F_1\ F_2 \cdots F_c]^T$ be a multi-channel image with c number of channels. Discrete partial derivatives of this image $\mathbf{F}$ at pixel location (i, j) can be given as forward differences:

$$\nabla_{x(i,j)}\mathbf{F}(i,j) = \mathbf{F}(i+1,j) - \mathbf{F}(i,j),$$

and

$$\nabla_{y(i,j)}\mathbf{F}(i,j) = \mathbf{F}(i,j+1) - \mathbf{F}(i,j).$$

Now, for each channel of $\mathbf{F}$, let us consider the continuity of its partial derivatives in a 2 × 2 neighborhood. Partial

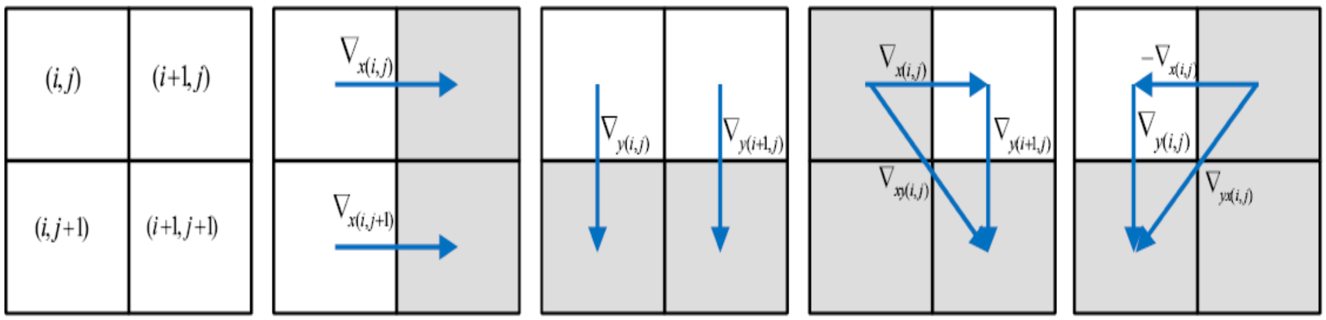

Figure 1: The gradient continuity: (a) a 2 × 2 neighbourhood showing pixel positions, (b)-(e) the vertical, horizontal, diagonal 45° and diagonal 135° boundaries, respectively. Required derivatives for the gradient continuity are shown in blue colour in (b)-(e), for each direction of the boundary.

derivatives $\nabla_{x(i,j)}$ and $\nabla_{y(i,j)}$ can be continuous along all directions except their own directions because they are desired to be discontinuous to preserve sharp edges and boundaries. For example, $\nabla_{x(i,j)}$ enforces the continuity along all directions except the horizontal direction. The continuity of partial derivatives depends upon the direction associated with the boundary. Continuity constraints for different boundary directions in a 2 × 2 neighborhood are as follows:

- vertical, $\nabla_{xx(i,j)} = \nabla_{x(i,j+1)} - \nabla_{x(i,j)} = 0;$
- horizontal, $\nabla_{yy(i,j)} = \nabla_{y(i+1,j)} - \nabla_{y(i,j)} = 0;$
- diagonal 45°, $\nabla_{xy(i,j)} = \nabla_{x(i,j)} + \nabla_{y(i+1,j)} = 0;$
- diagonal 135°, $\nabla_{yx(i,j)} = \nabla_{y(i,j)} - \nabla_{x(i,j)} = 0.$

Figure 1 shows these four boundaries along with associated image derivatives in a 2 × 2 neighborhood. For vertical, horizontal, diagonal 45° and 135° boundaries, the directional continuity of partial derivatives ∇_x and ∇_y is enforced by minimizing $\nabla_{xx(i,j)}$, $\nabla_{yy(i,j)}$, $\nabla_{xy(i,j)}$ and $\nabla_{yx(i,j)}$, respectively. To regularize the multi-channel image **F**, we minimize the ℓ1 norm of aforementioned partial derivatives as

$$E_{\text{reg}}(\mathbf{F}) = ||(\nabla_x \mathbf{F})||_1^2 + ||(\nabla_y \mathbf{F})||_1^2 + ||(\nabla_{xy} \mathbf{F})||_1^2 + ||(\nabla_{yx} \mathbf{F})||_1^2 + ||(\nabla_{xx} \mathbf{F})||_1^2 + ||(\nabla_{yy} \mathbf{F})||_1^2. \quad (6)$$

A careful inspection of continuity constraints reveals that

$$\nabla_{xx}\mathbf{F} = \mathbf{F}(i+1, j+1) + \mathbf{F}(i,j) - \mathbf{F}(i+1,j) - \mathbf{F}(i,j+1) = \nabla_{yy}\mathbf{F}. \quad (7)$$

Given the goal to recover sparsest partial derivatives, $\nabla_{xx(i,j)} = 0$ or $\nabla_{yy(i,j)}, = 0$ implies zero partial derivative along the horizontal or vertical direction, respectively. This is equivalent to $\nabla_{x(i,j)} = 0$ or $\nabla_{x(i,j)} = 0$. In this case, the minimization of $||\nabla_{xx}\mathbf{F}||_1$ and $||\nabla_{yy(i,j)}\mathbf{F}||_1$ is redundant under the minimization of either $||\nabla_x \mathbf{F}||_1$ or $||\nabla_y \mathbf{F}||1$. Therefore, these two terms can be omitted in (6). Since $\mathbf{F} = [F_1\, F_2 \cdots F_c]^{\mathrm{T}}$, we penalize the magnitude of

Horizontal, vertical and diagonal derivatives of each channel, and denote this regularizer as HVD (**F**):

$$\begin{aligned}\mathrm{HVD}(\mathbf{F}) := E_{\text{reg}}(\mathbf{F}) &= ||\sqrt{(\nabla_x F_1)^2 + (\nabla_x F_2)^2 + \cdots + (\nabla_x F_c)^2}||_1^2 \\ &+ ||\sqrt{(\nabla_y F_1)^2 + (\nabla_y F_2)^2 + \cdots + (\nabla_y F_c)^2}||_1^2 \\ &+ ||\sqrt{(\nabla_{xy} F_1)^2 + (\nabla_{xy} F_2)^2 + \cdots + (\nabla_{xy} F_c)^2}||_1^2 \\ &+ ||\sqrt{(\nabla_{yx} F_1)^2 + (\nabla_{yx} F_2)^2 + \cdots + (\nabla_{yx} F_c)^2}||_1^2.\end{aligned} \quad (8)$$

The regularizer HVD(**F**) enforces the continuity of partial image derivatives at each pixel, and minimizes their ℓ1 norm separately.

Therefore, regularizing an image using (8) is expected to preserve sharp horizontal, vertical as well as diagonal edges. Furthermore, the inclusion of diagonal derivatives along with the horizontal and vertical derivatives in a neighborhood around each pixel imposes more constraints on the image to be restored. Consequently, the HVD regularizer reduces staircase artefacts in flat regions. In addition, it is more robust against outliers than traditional TV regularizer. The separate minimization of partial derivatives favors a solution which is sparser than the solution obtained by using TV. As an implication, HVD(**F**) requires fewer number of measurements than TV for the estimation of unknown signals.

B. ROTATIONAL INVARIANCE

We prove the rotational invariance of the proposed regularizer given in (8) for multi-channel images. We use 2D rotations to give proof for 2-channel images; nevertheless, by using higher dimensional rotations, it is easy to show that the regularizer is invariant to rotations for multi-channel images. Let **R** be a 2D rotation matrix for a 2- channel image $\mathbf{F} = [F_1\, F_2]^{\mathrm{T}}$. When the camera is rotated by an angle θ, the rotated image **RF** is given as

$$\mathbf{RF} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \quad (9)$$

or

$$\mathbf{RF} = \mathbf{R}_1 F_1 + \mathbf{R}_2 F_2, \quad (10)$$

where $\mathbf{R} = (\mathbf{R}_1 \;\; \mathbf{R}_2)$. For the proposed regularizer, we will prove that HVD(**F**) = HVD(**RF**). Considering the first term in the square root of (8), i.e., $(\nabla_x F_1)^2 + (\nabla_x F_2)^2$, and substituting rotated image **RF** into this term, we get

$$\begin{aligned}(\nabla_x \mathbf{R}_1 F_1)^2 + (\nabla_x \mathbf{R}_2 F_2)^2 &= (\nabla_x F_1 \cos\theta - \nabla_x F_2 \sin\theta)^2 \\ &+ (\nabla_x F_1 \sin\theta + \nabla_x F_2 \cos\theta)^2, \\ &= (\nabla_x F_1 \cos\theta)^2 + (\nabla_x F_2 \sin\theta)^2 \\ &- 2(\nabla_x F_1 \cos\theta)(\nabla_x F_2 \sin\theta) \\ &+ (\nabla_x F_1 \sin\theta)^2 + (\nabla_x F_2 \cos\theta)^2 \\ &+ 2(\nabla_x F_1 \sin\theta)(\nabla_x F_2 \cos\theta).\end{aligned}$$

By canceling common terms and after some rearrangements, we obtain

$$\begin{aligned}(\nabla_x \mathbf{R}_1 F_1)^2 + (\nabla_x \mathbf{R}_2 F_2)^2 &= (\nabla_x F_1)^2(\cos^2\theta + \sin^2\theta) \\ &+ (\nabla_x F_2)^2(\cos^2\theta + \sin^2\theta), \\ &= (\nabla_x F_1)^2 + (\nabla_x F_2)^2,\end{aligned}$$

which is identical to the image without rotation. The similar proof can be provided for terms involving ∇_y, ∇_{xy} and ∇_{yx}. Hence the proposed regularizer is invariant to camera rotations.

IV. IMAGE DENOISING USING PROPOSED REGULARIZER

In this section, we apply the proposed regularizer to the problem of image denoising. Let $\mathbf{f} = [\mathbf{f}_1 \; \mathbf{f}_2 \; \cdots \; \mathbf{f}_c]^T$ be the lexicographically vectorized multi-channel image **F**, **g** be the degraded version of **f** and S be a linear matrix operator that represents the degradation process. The image restoration model can now be given as

$$\mathbf{g} = S\mathbf{f} + \eta, \qquad (11)$$

where η denotes multichannel noise. One popular example of the restoration process is image denoising. When the matrix S is assumed to be an identity matrix I_n of size n × n, we get **g** to be a noisy version of the original image **f**. We denoise (11) using the proposed regularizer. The variational energy E(**f**) incorporating (11) and the proposed regularizer can now be given as

$$E(\mathbf{f}) = \min_{\mathbf{f}} ||\mathbf{f} - \mathbf{g}||_1^2 + \lambda[\mathrm{HVD}(\mathbf{f})], \qquad (12)$$

where λ is a regularization parameter. The first term in (12) represents the data fidelity term that enforces the restored image to be close to the noisy image. We have used the robust ℓ_1 norm with the data fidelity term to handle outliers in the restoration process. Note that both data and regularization terms in (12) are convex; thus, convex optimization methods can be used to solve for f from the resulting energy E(**f**). Here, we demonstrate how a fast algorithm, NESTA, presented in [3] can be modified to solve Equation (12). NESTA has been used to solve large-scale variational problems [9]. We give the algorithmic details for the image denoising. NESTA uses a differentiable Huber norm approximation to the ℓ_1 norm; therefore, it can handle smooth as well as non-smooth convex functionals. We modify NESTA to solve image restoration problem.

Table 1: The proposed algorithm for image restoration

Initialization: $\mathbf{f}^0 = 0$.
Set iteration index $k = 1$,
while(not converged & $k \le$ max_iter)
1. compute $\partial_{\mathbf{f}^k} E(\mathbf{f}^k)$ from (14),
2. compute $\gamma^k = \frac{1}{2}(k+1)$ and $\tau^k = \frac{2}{k+3}$,
3. compute $\mathbf{p}^k = \mathbf{f}^k - \frac{1}{L}\partial_{\mathbf{f}^k} E(\mathbf{f}^k)$,
4. compute $\mathbf{q}^k = \mathbf{f}^0 - \frac{1}{L}\sum_i^k \gamma^i, \partial_{\mathbf{f}^i} E(\mathbf{f}^i)$,
5. update $\mathbf{f}^k = \tau^k \mathbf{p}^k + (1-\tau^k)\mathbf{q}^k$,
$k = k + 1$,
end while

The Huber norm is given as

$$||x||_\epsilon = \begin{cases} \frac{x^2}{2\epsilon}, & \text{if } |x| \le \epsilon, \\ |x| - \frac{\epsilon}{2}, & \text{otherwise.} \end{cases}$$

The derivative of the Huber norm is given by

$$\frac{\partial}{\partial x}||x||_\epsilon = \frac{x}{\max(\epsilon, |x|)}.$$

We use differentiable Huber norm in place of the ℓ_1 norm in (12). The combined data and the regularization energy E(**f**) is now given as

$$\begin{aligned}E(\mathbf{f}) = \min_{\mathbf{f}} &||\mathbf{f} - \mathbf{g}||_\epsilon^2 \\ &+ \lambda\Big(||\sqrt{(\boldsymbol{\nabla}_x \mathbf{f}_1)^2 + (\boldsymbol{\nabla}_x \mathbf{f}_2)^2 + \cdots + (\boldsymbol{\nabla}_x \mathbf{f}_c)^2}||_\epsilon^2 \\ &+ ||\sqrt{(\boldsymbol{\nabla}_y \mathbf{f}_1)^2 + (\boldsymbol{\nabla}_y \mathbf{f}_2)^2 + \cdots + (\boldsymbol{\nabla}_y \mathbf{f}_c)^2}||_\epsilon^2 \\ &+ ||\sqrt{(\boldsymbol{\nabla}_{xy} \mathbf{f}_1)^2 + (\boldsymbol{\nabla}_{xy} \mathbf{f}_2)^2 + \cdots + (\boldsymbol{\nabla}_{xy} \mathbf{f}_c)^2}||_\epsilon^2 \\ &+ ||\sqrt{(\boldsymbol{\nabla}_{yx} \mathbf{f}_1)^2 + (\boldsymbol{\nabla}_{yx} \mathbf{f}_2)^2 + \cdots + (\boldsymbol{\nabla}_{yx} \mathbf{f}_c)^2}||_\epsilon^2\Big),\end{aligned} \qquad (13)$$

where $\boldsymbol{\nabla}_x$, $\boldsymbol{\nabla}_y$, $\boldsymbol{\nabla}_{xy}$ and $\boldsymbol{\nabla}_{yx}$ are sparse difference matrices to calculate derivatives of vectorized images. An iterative scheme is used to find the minimum of (13) at iteration k as

$$\begin{aligned}\partial_{\mathbf{f}^k} E(\mathbf{f}^k) = \Bigg[&2(\mathbf{f}^k - \mathbf{g}) \\ &+ \lambda\Bigg(\boldsymbol{\nabla}_x^T \frac{(\boldsymbol{\nabla}_x \mathbf{f}^k)}{\max(\epsilon, |\boldsymbol{\nabla}_x \mathbf{f}^k|)} + \boldsymbol{\nabla}_y^T \frac{(\boldsymbol{\nabla}_y \mathbf{f}^k)}{\max(\epsilon, |\boldsymbol{\nabla}_y \mathbf{f}^k|)} \\ &+ \boldsymbol{\nabla}_{xy}^T \frac{(\boldsymbol{\nabla}_{xy} \mathbf{f}^k)}{\max(\epsilon, |\boldsymbol{\nabla}_{xy} \mathbf{f}^k|)} + \boldsymbol{\nabla}_{yx}^T \frac{(\boldsymbol{\nabla}_{yx} \mathbf{f}^k)}{\max(\epsilon, |\boldsymbol{\nabla}_{yx} \mathbf{f}^k|)}\Bigg)\Bigg],\end{aligned} \qquad (14)$$

where $\partial_{\mathbf{f}^k} E(\mathbf{f}^k) = \partial E(\mathbf{f}^k) / \partial \mathbf{f}^k$. The algorithm computes two auxiliary variables p_k and q_k at each iteration from $\partial_{\mathbf{f}^k} E(\mathbf{f}^k)$. It then combines both auxiliary variables to get next estimate $\mathbf{f}^k$. The choice of ε plays an important role in the algorithm. The speed of the convergence is shown to have direct relationship with this approximation constant [3]. A small value of ε gives good accuracy at the cost of slow convergence and vice versa. The proposed algorithm uses Lipschitz continuity; therefore, a Lipschitz constant L is required for the computation of auxiliary variables from (14). Lipschitz constant L depends on λ, Huber norm parameter ε, and the norms of sparse difference matrices.

To compute the Lipschitz constant L, we need to find the upper bound for the norm of HVD. It has been shown in [10] that difference matrices used to calculate TV are bounded above by 8. A similar analysis can be made for difference matrices used in HVD. The ℓ_1 norm of any matrix is maximum absolute column sum of that matrix.

HVD consists of four sparse difference matrices: ∇_x, ∇_y, ∇_{xy} and ∇_{yx}, which are used to compute discrete differences. These matrices have exactly two nonzero entries +1 or −1. Therefore, they satisfy $||\nabla_x||1 = 2$, $||\nabla_y||1 = 2$, $||\nabla_{xy}||1 = 2$ and $||\nabla_{yx}||1 = 2$. Since we minimize the magnitude of each partial flow derivative in HVD,

$$||\nabla_x^{\mathrm{T}}\nabla_x||_1 + ||\nabla_y^{\mathrm{T}}\nabla_y||_1 + ||\nabla_{xy}^{\mathrm{T}}\nabla_{xy}||_1 + ||\nabla_{yx}^{\mathrm{T}}\nabla_{yx}||_1 = 4+4+4+4 = 16.$$

Hence, difference matrices used in HVD are bounded above by 16. Lipschitz constant L is then given as $L = 16\lambda/\varepsilon$. The algorithm runs for a fixed number of iterations or until it reaches the convergence. The summary of the algorithm is shown in Table 1.

V. EXPERIMENTAL RESULTS

The proposed regularizer has been tested on several images to evaluate its performance under noise. For the validation of the proposed regularizer, a comparative analysis is conducted with isotropic [6], anisotropic [26] and higher order TV [27] regularizers by assessing the quality of the denoised images. First, the experimental setup, describing images used in experiments, is presented. Second, the performance is analyzed on images corrupted by a controlled amount of noise.

A. EXPERIMENTAL SETUP

All experiments have been performed on publically available real world images which are used as benchmarks for various vision and image processing tasks. The image dataset used in these experiments comprises of greyscale images *Cameraman*, *Barbara*, *Boat* and *Man*, and colour images *Baboon*, *House*, *Monarch* and *Pepper*. These images are corrupted by a controlled amount of noise. The quality of denoised images have been assessed by calculating the peak signal to noise ratio (PSNR), which is given as

$$\mathrm{PSNR} = 20 * \log(\frac{f_{\max}^2}{\mathrm{MSE}}),$$

where $f_{\max}$ = 255 for an 8 bit image and MSE is the mean squared error

$$\mathrm{MSE} = \frac{1}{n}\sum_n(\mathbf{f}-\mathbf{g})^2.$$

To conduct a fair comparison, the energies of proposed regularizer and the three TV regularizers are minimized using NESTA ([3]). It should be mentioned that the use of different regularizers alter the variational energy to be minimized. Consequently, the values of optimum regularization parameters for these regularizers also change. In these experiments, we have manually tuned regularization parameters of these regularizers to get best denoising results for all of these regularizers. These experiments are conducted using λ_{iTV} = 0.05, λ_{aTV} = 0.05, λ_{HVD} = 0.01 and λ_{HOTV} = 0.006.

B. IMAGE DENOISING

These experiments have been conducted to test the capability of the proposed regularizer to denoise images. A controlled amount of Gaussian noise is added to images described above. Since clean images are available, performances of the proposed and TV based regularizers have been measured quantitatively by calculating the PSNR for denoised images. The influence of the controlled noise is also analyzed qualitatively on denoised images.

Table 2 shows quantitative results for all four regularizers on eight images. These results have been taken for a standard deviation of noise σ = 25. Optimum values of regularization parameters are used for all regularizers. It can be observed that the HOTV regularizer performs better than both isotropic and anisotropic TV regularizers. However, the proposed HVD regularizer outperforms the HOTV for most of images.

Figure 2 demonstrates the denoising of the greyscale image *Cameraman* when it is contaminated with a noise of σ = 25. Highlighted parts of images in Figure 2 (c) and (d) show staircase artefacts, whereas highlighted parts in Figure 2 (e) and (f) show a significant reduction in these artefacts. However, the image denoised by the HVD has a higher value of PSNR = 30.21 as compared to the image denoised by the HOTV regularizer with a PSNR = 29.38, as given in Table 2. Highlighted parts of Figure 2 are also shown in Figure 3 as 3D plots for a better visualization of staircase artefacts.

Denoising results on the colour image *Monarch* are shown in Figure 4. A qualitative comparison of images in Figure 4 (b) and (c) with (d) and (e) reveals that HVD and HOTV regularizers outperform anisotropic and isotropic TV especially in highlighted textured regions of denoised images. The image denoised by the HVD regularizer in Figure 4 (d) has a higher PSNR = 30.92 than anisotropic, isotropic and higher order TV regularizers with PSNR = 28.58, 28.07 and 30.15, respectively. Moreover, staircase artefacts can be observed in highlighted parts of images in Figure 4 (b) and (c). HVD and HOTV regularizers do not show noticeable staircase artefacts.

Figure 5 presents PSNR as a function of standard deviation of noise σ for all four regularizers. An average PSNR has been calculated over all eight images, and the results are reported in Figure 5 (a). These results clearly indicate that the proposed regularizer outperforms TV based regularizers for increasing values of standard deviation of noise. Similar kind of results can be observed for *Cameraman* and *Pepper* in Figure 5 (b) and (c), respectively

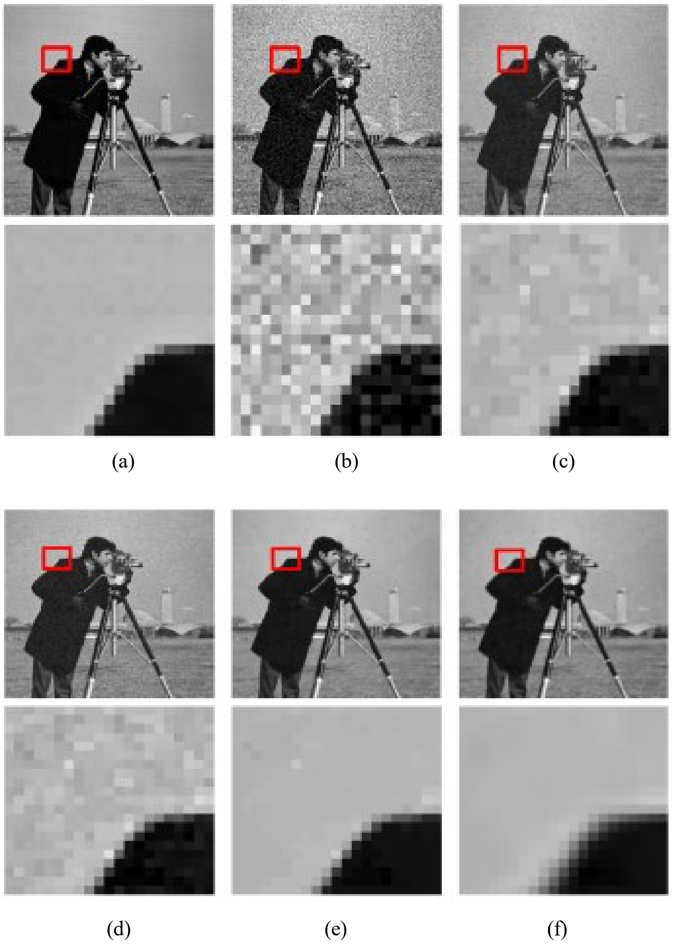

Figure 2: The denoising of the greyscale image *Cameraman*. (a) Original image, (b) image corrupted by a Gaussian noise of σ = 25, image denoised by (c) anisotropic TV, (d) isotropic TV, (e) the proposed HVD and (f) higher order TV regularizers. Highlighted parts of these images are also given in the bottom.

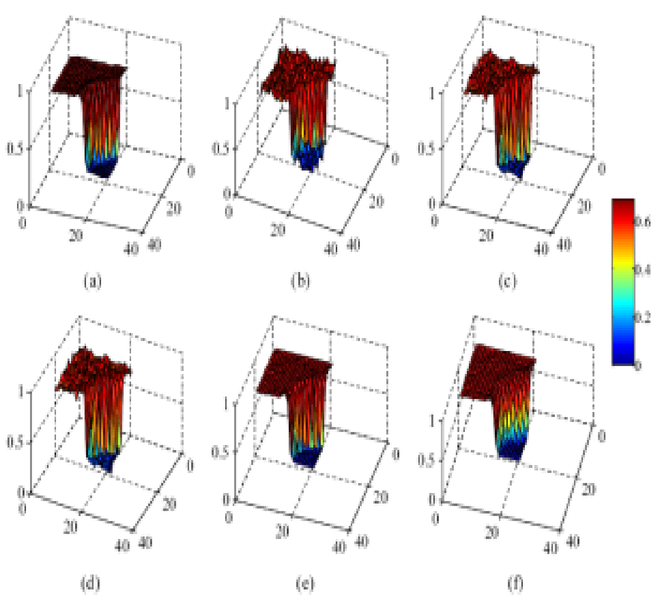

Figure 3: Highlighted parts of Figure3 shown as 3D plots for the visualization of staircase artefacts.

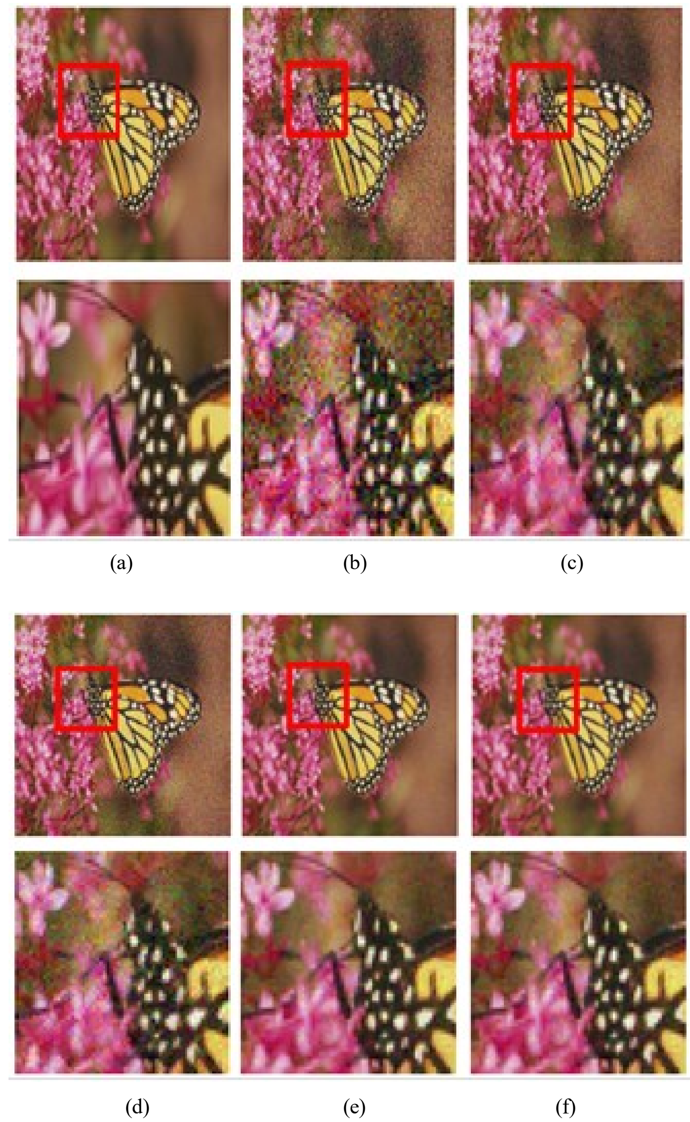

Figure 4: The denoising of the colour image *Monarch*. (a) Original image, (b) image corrupted by a Gaussian noise of σ = 25, image denoised by (c) anisotropic TV, (d) isotropic TV, (e) the proposed HVD and (f) higher order TV regularizers. Highlighted parts of these images are also given in the bottom.

VI. CONCLUSION

This paper presented a novel sparsity enhancing variational regularizer for multi-channel image denoising. It investigated sparsity enhancing regularizers in the context of variational methods. The proposed HVD regularizer was proven to be rotationally invariant to camera motions. TV regularizer is known to generate staircase artefacts in the computed solution. However, the HVD regularizer was shown to reduce these artefacts significantly. The proposed regularizer was applied to the problem of image denoising. Experiments were conducted to show that the proposed regularizer can produce results better than TV based regularizers in terms of higher PSNR and better visual quality.

Table 2: PSNR results for the denoising of all eight images using optimum values of regularization parameters.

Image	Anisotropic TV	Isotropic TV	Proposed HVD	Higher order TV
Cameraman	27.85	27.18	**30.21**	29.38
Barbara	28.65	28.27	30.28	**30.43**
Boat	26.6	26.94	**30.46**	30.07
Man	27.36	27.02	**31.09**	30.42
Baboon	26.95	26.15	29.84	**30.67**
House	27.61	27.09	**30.48**	30.12
Monarch	28.58	28.07	**30.92**	30.15
Pepper	28.64	28.56	**30.67**	29.81

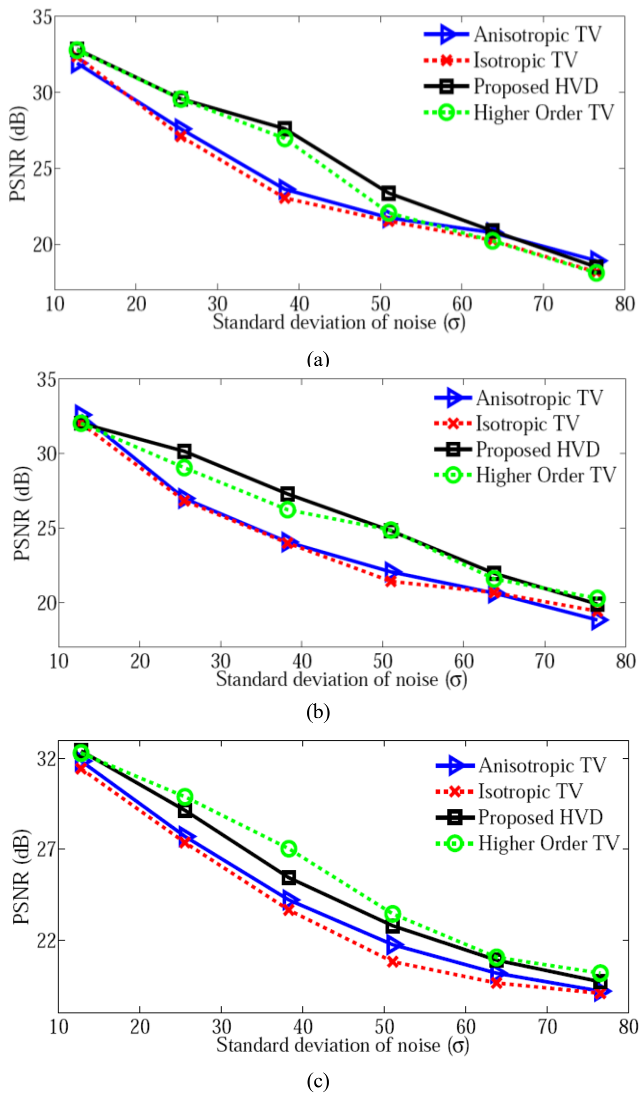

Figure 5: The PSNR as a function of the standard deviation σ of noise for anisotropic TV, isotropic TV, proposed HVD and higher order TV regularizers. Results for (a) the whole dataset of 8 images (b) *Cameraman* and (c) *Pepper*. Regularization parameters of all four regularizers have been tuned to get best PSNR performances.

REFERENCES

1. D. Bappy, I. Jeon, 'Combination of hybrid median filter and total variation minimization for medical x-ray image restoration', *IET Image Processing*, 2016, 57, (11), pp 1413-1457.
2. A. Beck, M. Teboulle, 'Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems', *IEEE Transactions on Image Processing*, 2009, 18, (11), pp 2419-2434.
3. S. Becker, J. Bobin and E. J. Candes, 'NESTA: a fast and accurate first-order method for sparse recovery', *SIAM Journal on Imaging Sciences*, 2011, 4, (1), pp 1-39.
4. J. Bioucas-Dias, M. Figueiredo, 'A New TwIST: Two-Step Iterative Shrinkage/ Thresholding Algorithms for Image Restoration', *IEEE Transactions on Image Processing*, 2007, 16, (12), pp 2992-3004.
5. E. J. Candes, M. B. Wakin and S. Boyd, 'Enhancing Sparsity by Reweighted L1 Minimization', *Journal of Fourier Analysis and Applications*, 2008, 14, (5), pp 877-905.
6. A. Chambolle, T. Pock, 'A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging', *Journal of Mathematical Imaging and Vision*, 2011, 40, (1), pp 120-145.
7. T. F. Chan, A. Mulet, 'High-order total variation-based image restoration', *SIAM Journal on Scientific Computing*, 2000, 22, (2), pp 503-516.
8. T. F. Chan, A. M. Yip and F. E. Park, 'Simultaneous total variation image inpainting and blind deconvolution', *International Journal of Imaging Systems and Technology*, 2005, 15, (1), pp 92-102.
9. K. Choi, J. Wang, L. Zhu, T-S. Suh, S. Boyd and L. Xing, 'Compressed sensing based cone-beam computed tomography reconstruction with a first-order method', *Medical Physics*, 2010, 37, (9), pp 5113-5125.
10. J. Dahl, P. C. Hansen, S. H. Jenson and T. L. Jensen, 'Algorithms and software for total variation image reconstruction via first-order methods', *Numerical Algorithms*, 2009, 53, (1), pp 67-92.
11. J. Darbon and M. Sigelle, 'Image restoration with discrete constrained total variation, Part I: fast and exact optimization', *Journal of Mathematical Imaging and Vision*, 2006, 26, (3), pp 261-276.
12. M. Elad M. and M. Aharon,'Image denoising via sparse and redundant representations over learned dictionaries', *IEEE Transactions on Image Processing*, 2006, 15, (12), pp 3736-3745.
13. S. Esedoglu and S. Osher, 'Decomposition of Images by the Anisotropic Rudin-Osher-Fatemi Model', *Communications on Pure and Applied Mathematics*, 2004, 57, (1), pp 1609-1626.
14. M. A. T. Figueiredo, R. D. Nowak and S. J. Wright, 'Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems', *IEEE Journal of Selected Topics in Signal Processing*, 2006, 1, (4), pp 586-597.
15. I. M. Gelfand and S. V. Fomin, *'Calculus of Variations*', Courier Dover Publications, 2005.

16. G. Gilboa and S. Osher, 'Nonlocal operators with applications to image processing', *Multiscale Modeling and Simulation*, 2008, 7, (3), pp 1005-1028.
17. D. Goldfarb and W. Yin, 'Parametric Maximum Flow Algorithms for Fast Total Variation Minimization', *SIAM Journal on Scientific Computing*, 2009, 31, (5), pp 3712-3743.
18. M. A. Kitchener, A. Bouzerdoum and S. L. Phung, 'Adaptive regularization for image restoration using a variational inequality approach', *In Proceedings of the IEEE International Conference on Image Processing*, Sep. 2010, pp 2513-2516.
19. A. Matakos, S. Ramani, J. A. Fessler, 'Accelerated edge-preserving image restoration without boundary artifacts', *IEEE transactions on image processing*, 2013, 22, (5), pp 2019-2029.
20. J. P. Oliveira, J. Bioucas-Dias and M. Figueiredo, 'Adaptive total variation image deblurring: A majorization-minimization approach', *Signal Processing*, 2009, 89, (9), pp 1683-1693.
21. G. Peyre, S. Bougleux and L. Cohen, 'Non-local Regularization of Inverse Problems', *In Proceedings of the European Conference on Computer Vision*, 2008, pp 57-68. Springer Berlin Heidelberg.
22. R. Ranftl, K. Bredies and T. Pock, 'Non-local total generalized variation for optical flow estimation', *In Proceedings of the European Conference on Computer Vision*, 2014, pp 439-454.
23. Z. F. Pang, Y. M. Zhou, T. Wu and D. J. Li, 'Image denoising via a new anisotropic total-variation-based model', *Signal Processing: Image Communication*, 2019, 74, 140 – 152.
24. X. Shu and N. Ahuja, 'Hybrid compressive sampling via a new total variation TVL1', *In Proceedings of the European Conference on Computer Vision*, Sep. 2010, pp 393-404.
25. Y. Wang, J. Yang, W. Yin and Y. Zhang, 'A New Alternating Minimization Algorithm for Total Variation Image Reconstruction', *SIAM Journal on Imaging Sciences*, 2008, 1, (3), pp 248-272.
26. L. Wanga , L. Xiaoa , J. Zhang and Z. Weia, 'New image restoration method associated with tetrolets shrinkage and weighted anisotropic total variation', *Signal Processing*, 2013, 93, (4), pp 661-670.
27. L. Xinwu, 'Weighted total generalised variation scheme for image restoration', *IET Image Processing*, 2016, 10, (1), pp 80 - 88.
28. A. B. Said,; R. Hadjidj.and S. Foufou, 'Total Variation for Image Denoising Based on a Novel Smart Edge Detector: An Application to Medical Images', *Journal of Mathematical Imaging and Vision*, 2019, 61, 106-121.