A New Relaxed Total Generalized Variation (RTGV) Technique for Image Denoising

Mengge Chenga, Huayan Zhangb,*

The School of Computer Science and Technology, Tiangong University, Tianjin, China achengmenge22@163.com, bzhanghy307@163.com
*Corresponding author

Abstract: In this paper, a new relaxed second-order total generalized variational model is proposed for image denoising. In this model, the regularization term is a combination of a gradient operator and weighted divergence operator, while the data fidelity terms we used is 12-norm. The weighted divergence operator in the regularization is used to adjust the higher-order smoothing term, which can reduce the computational complexity and guarantee the discrete accuracy. Moreover, the Augmented Lagrangian algorithm, which produced several closed form solutions is used to solve the proposed RTGV method. Then we applied the denoising model to both gray and color image denoising, and performed numerous experiments. Our algorithm is discussed from several aspects, including influence of parameters, numerical discretization and comparisons with other methods. Numerical experimental results demonstrate that this technique has the significant advantage in preserving image features and can effectively prevent the staircase artifacts when compared to several other classical existing based methods.

Keywords: Relaxed total generalized variation, Augmented Lagrangian method, Image denoising

1. Introduction

Image denoising is a fundamental problem in the field of image processing, which provides the basis for the subsequent segmentation and recognition. Generally, the images acquired in nature are often noisy due to the influence of imaging devices or transmission elements, and image denoising aims to restore clean or original images from corrupted ones, which can be expressed by a mathematical formula as follows:

$$\mathbf{f} = \mathbf{u} + \mathbf{n}$$
 (1)

where f represents the noisy image, which is transformed by the contamination of clean images u, and n represents the random noise like Gaussian noise.

Image denoising is a crucial fundamental research topic. For restoring the clean original image, various methods have been presented such as local based method [1, 2, 3], variational based method [4, 5, 6, 7, 8, 9, 10, 11], learning based method [12] and recently deep learning based method [13, 14, 15, 16]. In the paper, we focus on a new variational based technique.

In 1992, Rudin, Osher and Fatemi first proposed the following classic image denoising variational model (TV), which uses the total variable component to transform the image denoising problem into a extremum problem by solving the energy functional.

$$TV(\mathbf{u}) = \sup\{ \int_{\Omega} \mathbf{u} \operatorname{div}(\mathbf{v}) dx | \mathbf{v} \in \mathbb{C}^{1}_{c}(\Omega), \|\mathbf{v}\|_{2,\infty} \le 1 \} = \int_{\Omega} |\nabla \mathbf{u}| dx \quad (2)$$

The success of the total variation regularization (TV) model lies in its ability to remove noise while maintaining the original unique features of the image. However, this regularization also has a known drawback in that it leads to staircase effect that transforms the smooth regions into piecewise constants. To solve this problem, various weighted total variation regularization has been developed [11, 17], which has the following form:

$$TV_{\phi}(\mathbf{u}) = \int_{\Omega} \phi(k) |\nabla \mathbf{u}| dx$$
 (3)

In addition, Bredies et al. considered a more general total variation method in [18], which is named generalized variation (TGV).

$$\mathsf{TGV}_{\alpha}^k(\mathbf{u}) = \sup\{\int_{\Omega} \mathbf{u} \, \mathrm{div}^k \mathbf{v} dx | \mathbf{v} \in \mathbb{C}_c^k(\Omega, \mathsf{Sym}^k(\mathbb{R}^d)), \|\mathrm{div}^l \mathbf{v}\|_{2,\infty} \leq \alpha_l, l = 0, \cdots, k-1\} \quad (4)$$

where $Sym^k(\mathbb{R}^d)$ represents the k-order symmetric tensor space in \mathbb{R}^d , α_l are fixed positive parameters. Actually, when k=1, Eq. (4) equals to TV. When k=2, Eq. (4) has the following form:

$$TGV_{\alpha}^{2}(\mathbf{u}) = \min_{\mathbf{v}} \alpha_{1} \int_{\Omega} |\nabla \mathbf{u} - \mathbf{v}| dx + \alpha_{0} \int_{\Omega} |\mathcal{E}(\mathbf{v})| dx \quad (5)$$

where $\mathcal{E}(v) = \frac{1}{2}(\nabla v + \nabla v^T)$ denotes the symmetric derivative operator.

The TGV model has far-reaching implications in field of image restoration [9, 10, 19, 20]. However, the second term $\mathcal{E}(v)$ contains mixed partial derivatives, which involve computational complexity. Moreover, the discretization of these mixed partial derivatives is more complex, and it is difficult to guarantee the accuracy. Learning from the research on 3d mesh [21], it considers using famous divergence operator to approximate the symmetrized derivative operator $\mathcal{E}(v)$. As known, the divergence operator and its variant Laplace operator have been successfully used in various fields [5, 6, 7, 19, 22], due to simple to calculate and can effectively reduce the complexity of the algorithm. Besides, there have been many researches discussing about the discretization of these operators, which guarantee the theoretical support.

Therefore, by introducing a weighted divergence operator, we consider a relaxed total generalized variation (RTGV) model, which can be reformulated as

$$RTGV(\nabla \mathbf{u}) = \min_{\mathbf{v}} \alpha_1 \|\nabla \mathbf{u} - \mathbf{v}\|_1 + \alpha_0 \|\operatorname{div}(w\mathbf{v})\|_1 \quad (6)$$

With the above RTGV model, we then consider an iterative image denoising method. Augmented Lagrangian method has been demonstrated effective in solving non-differential optimization problem. In addition, variable separation method is also useful for such problems. Numerical implementations indicates that the proposed model can effectively remove noise and prevent staircase artifacts while retaining distinctive features and structures. Besides, compared with other existing variational models, our method verifies the superiority in visually and quantitatively.

This paper is organized as follows. In section 2, some notations are given for facilitate understanding. In section 3, we introduce the RTGV image denoising model and describe some details. Section 4 presents the Augmented Lagrangian method for solving the proposed non-differential problem. In section 5, we devoted our proposed method to image denoising experiments and discussed it in various aspects. Finally, we summarize our work and next steps in section 6.

2. Notations

For a given image $f=(f_{ij}):\Omega\to\mathbb{R}$ with η -channel, we represent the Euclidean space $\mathbb{R}^{M\times N\times \eta}$ as U. The discrete gradient operator of U is a mapping $\nabla:U\to V$, where $V=\mathbb{R}^{M\times N\times \eta\times 2}$. Consequently, the divergence operator of V is written as div: $V\to U$.

For $u \in U$, the definition of ∇u as follows:

$$(\nabla \mathbf{u})_{ij} = \begin{bmatrix} \nabla_{x} u_{1}, & \nabla_{y} u_{1} \\ \nabla_{x} u_{2}, & \nabla_{y} u_{2} \\ \dots \\ \nabla_{x} u_{\eta}, & \nabla_{y} u_{\eta} \end{bmatrix}_{ii}$$
(7)

where $i=1,\dots,M,j=1,\dots,N$, the ∇x and ∇y are the partial derivatives in the x directive and the y direction, respectively. For discrete divergence operator div, we denote div(v) and div(wv) by

$$(\operatorname{div}(\mathbf{v}))_{ij} = \left[\frac{\partial v_{11}}{\partial x} + \frac{\partial v_{12}}{\partial y}, \frac{\partial v_{21}}{\partial x} + \frac{\partial v_{22}}{\partial y}, \cdots, \frac{\partial v_{\eta 1}}{\partial x} + \frac{\partial v_{\eta 2}}{\partial y}\right]_{ij},$$

$$(\operatorname{div}(w\mathbf{v}))_{ij} = \left[\frac{\partial w_1 v_{11}}{\partial x} + \frac{\partial w_1 v_{12}}{\partial y}, \frac{\partial w_2 v_{21}}{\partial x} + \frac{\partial w_2 v_{22}}{\partial y}, \dots, \frac{\partial w_3 v_{\eta 1}}{\partial x} + \frac{\partial w_\eta v_{\eta 2}}{\partial y}\right]_{ij}. \quad (8)$$

where w is a weight for capturing the sharp features of images.

We also present the inner products for $u^1, u^2 \in U, v^1, v^2 \in V$ and the discrete ∞ -norms on U and V.

$$(\mathbf{u}^{1}, \mathbf{u}^{2}) = \sum_{1 \le k \le \eta} (u_{k}^{1}, u_{k}^{2}),$$
$$(\mathbf{v}^{1}, \mathbf{v}^{2}) = \sum_{1 \le k \le \eta} (v_{k}^{1}, v_{k}^{2}),$$

$$\mathbf{u} \in \mathbf{U} : \|\mathbf{u}\|_{2,\infty} = \max_{ij} \sum_{k} (|u_{k1}| + |u_{k2}| + |u_{k3}|)_{ij},$$

$$\mathbf{v} \in \mathbf{V}: \|\mathbf{v}\|_{2,\infty} = \max_{ij} \sqrt{\sum_{k} (v_{k1}^2 + v_{k2}^2)_{ij}}.$$
 (9)

3. RTGV image denoising method

According to the above notations, the relaxed total generalized variational method RTGV can be expressed as follows

$$RTGV(\nabla \mathbf{u}) = \min_{\mathbf{v}} \alpha_1 \|\nabla \mathbf{u} - \mathbf{v}\|_1 + \alpha_0 \|\operatorname{div}(w\mathbf{v})\|_1$$

$$= \min_{\mathbf{v}} \alpha_1 \sum_{k=1}^{\eta} \sum_{ij} \sqrt{(\nabla_x u_k)_{ij}^2 + (\nabla_y u_k)_{ij}^2} + \alpha_0 \sum_{k=1}^{\eta} \sum_{ij} |\operatorname{div}(w_k \mathbf{v}_k)_{ij}|. \quad (10)$$

where α_1 and α_0 are positive parameters.

Devoted the above RTGV regularization term to image denoising, and incorporate the data fidelity term, we can get the following optimization problem:

$$\min_{\mathbf{u} \in \mathbf{U}} \mathsf{RTGV}(\nabla \mathbf{u}) + \frac{\beta}{2} \|\mathbf{u} - \mathbf{f}\|_2^2$$

$$= \min_{\mathbf{u} \in \mathbf{U}, \mathbf{v} \in \mathbf{V}} \alpha_1 \|\nabla \mathbf{u} - \mathbf{v}\|_1 + \alpha_0 \|\operatorname{div}(w\mathbf{v})\|_1 + \frac{\beta}{2} \|\mathbf{u} - \mathbf{f}\|_2^2 \quad (11)$$

where β is a positive parameter to balance the two terms in the RTGV model.

Since the relaxed model in (11) is non-differentiable, they are difficult to solve by conventional methods. The role of the augmented Lagrange method is to solve optimization problems under the equation constraint. Recently, variable splitting and augmented Lagrangian method (ALM) [23] have been proven to be very efficient for such non-differential problems [8, 24, 25, 26]. In the following, we give the details for solving (11) by the ALM.

4. Algorithm Details

Firstly, we introduce three auxiliary variables p, and q for converting (11) to the following constrained optimization problem:

$$\min_{\mathbf{u}, \mathbf{v}, \mathbf{p}, \mathbf{q}} \alpha_1 \|\mathbf{p}\|_1 + \alpha_0 \|\mathbf{q}\|_1 + \frac{\beta}{2} \|\mathbf{u} - \mathbf{f}\|_2^2,$$

s.t.
$$\mathbf{p} = \nabla \mathbf{u} - \mathbf{v}, \mathbf{q} = \operatorname{div}(w\mathbf{v}).$$
 (12)

Then, the augmented Lagrangian dual functional for the (13) can be written as:

$$\mathcal{L}(\mathbf{u}, \mathbf{v}, \mathbf{p}, \mathbf{q}; \lambda_{\mathbf{p}}, \lambda_{\mathbf{q}}) = \alpha_1 \|\mathbf{p}\|_1 + \alpha_0 \|\mathbf{q}\|_1 + \frac{\beta}{2} \|\mathbf{u} - \mathbf{f}\|_2^2$$

$$+ \langle \lambda_{\mathbf{p}}, \mathbf{p} - (\nabla \mathbf{u} - \mathbf{v}) \rangle + \frac{r_{\mathbf{p}}}{2} \|\mathbf{p} - (\nabla \mathbf{u} - \mathbf{v})\|_2^2$$

$$+ \langle \lambda_{\mathbf{q}}, \mathbf{q} - \operatorname{div}(w\mathbf{v}) \rangle + \frac{r_{\mathbf{q}}}{2} \|\mathbf{q} - \operatorname{div}(w\mathbf{v})\|_2^2 \qquad (13)$$

It has been proven that the solution of (14) can be translated into the following saddle-point problem:

$$\max_{\lambda_{\textbf{p}}, \lambda_{\textbf{q}}} \min_{\textbf{u}, \textbf{v}, \textbf{p}, \textbf{q}} \mathcal{L}(\textbf{u}, \textbf{v}, \textbf{p}, \textbf{q}; \lambda_{\textbf{p}}, \lambda_{\textbf{q}}) \quad (14)$$

To facilitate solving the above saddle-point problem, we divide (14) into five sub-problems.

u sub-problem:

$$\min_{\mathbf{u}} \langle \lambda_{\mathbf{p}}, -\nabla \mathbf{u} \rangle + \frac{r_{\mathbf{p}}}{2} \|\mathbf{p} - (\nabla \mathbf{u} - \mathbf{v})\|_{2}^{2} + \frac{\beta}{2} \|\mathbf{u} - \mathbf{f}\|_{2}^{2} \quad (15)$$

v sub-problem:

$$\min_{\mathbf{v}} \langle \lambda_{\mathbf{p}}, \mathbf{v} \rangle + \frac{r_{\mathbf{p}}}{2} \|\mathbf{p} - (\nabla \mathbf{u} - \mathbf{v})\|_{2}^{2} + \langle \lambda_{\mathbf{q}}, -\operatorname{div}(w\mathbf{v}) \rangle + \frac{r_{\mathbf{q}}}{2} \|\mathbf{q} - \operatorname{div}(w\mathbf{v})\|_{2}^{2}$$
 (16)

p sub-problem:

$$\min_{\mathbf{p}} \alpha_1 \|\mathbf{p}\|_1 + \langle \lambda_{\mathbf{p}}, \mathbf{p} \rangle + \frac{r_{\mathbf{p}}}{2} \|\mathbf{p} - (\nabla \mathbf{u} - \mathbf{v})\|_2^2 \quad (17)$$

it has the following solution:

$$\mathbf{p}_{ij} = \begin{cases} (1 - \frac{\alpha_1}{r_p | c_{ij}|}) \mathbf{c}_{ij}, & |\mathbf{c}_{ij} > \frac{\alpha_1}{r_p}|\\ 0, & |\mathbf{c}_{ij} \le \frac{\alpha_1}{r_p}| \end{cases}$$
(18)

where
$$c_{ij} = (\nabla u - v - \frac{\lambda_p}{r_n})$$
.

q sub-problem:

$$\min_{\mathbf{q}} \alpha_0 \|\mathbf{q}\|_1 + \langle \lambda_{\mathbf{q}}, \mathbf{q} \rangle + \frac{r_{\mathbf{q}}}{2} \|\mathbf{q} - \operatorname{div}(w\mathbf{v})\|_2^2 \quad (19)$$

which has a unique closed form solution:

$$\mathbf{q}_{ij} = \begin{cases} (1 - \frac{\alpha_0}{r_{\mathbf{q}} |\mathbf{w}_{ij}|}) \mathbf{w}_{ij}, & |\mathbf{w}_{ij} > \frac{\alpha_0}{r_{\mathbf{q}}}|\\ 0, & |\mathbf{w}_{ij} \le \frac{\alpha_0}{r_{\mathbf{q}}}| \end{cases}$$
(20)

where
$$w_{ij} = (div(wv) - \frac{\lambda_q}{r_q})$$
.

The u subproblem (15) and v subproblem (16) are quadratic problems, which can be solved directly with the fast Fourier transform or various numerical packages.

The ALM algorithm is described in Algorithm 1.

Algorithm 1 ALM for solving (15)

1. Initialization

1.1
$$\lambda_{\mathbf{p}}^{-k} = 0, \lambda_{\mathbf{q}}^{-k} = 0, \mathbf{v}^{-k} = 0, l = 0, \varepsilon = 10^{-5};$$

2. Repeat

- 2.1 For fixed \mathbf{v}^{l-1} , \mathbf{p}^{l-1} , compute \mathbf{u}^{l} by (15);
- 2.2 For fixed $\mathbf{u}, \mathbf{p}^{l-1}, \mathbf{q}^{l-1}$, compute \mathbf{v}^l by (16);
- 2.3 For fixed $\mathbf{u}^l, \mathbf{v}^l$, compute \mathbf{p}^l by (18);
- 2.4 For fixed \mathbf{v}^l , compute \mathbf{q}^l , by (20);
- 2.5 Update Lagrange multiplier $\lambda_{\mathbf{p}}^{l}$, $\lambda_{\mathbf{q}}^{l}$:

$$\lambda_{\mathbf{p}}^l = \lambda_{\mathbf{p}}^{l-1} + r_{\mathbf{p}} \big(\mathbf{p}^l - (\nabla \mathbf{u}^l - \mathbf{v}^l) \big), \lambda_{\mathbf{q}}^l = \lambda_{\mathbf{q}}^{l-1} + r_{\mathbf{q}} \big(\mathbf{q}^l - \operatorname{div}(w\mathbf{v}^l) \big)$$

Until ($\|\mathbf{u}^l - \mathbf{u}^{l-1}\|^2 < \varepsilon$ or steps > 500).

5. Numerical Experiments

In this section, we devoted our proposed RTGV method to image denoising experiments. We present in Fig.1 some of the images employed in this paper, which contains both gray and color images. Our algorithm is performed in Window 10 and VS2022 on a computer with Intel(R) Core (TM) i5-7200 CPU @2.70GHz and 8GB memory (the other results are generated by MATLAB 2017)



Figure 1: Some of the test image.

In the following, we will discuss our algorithm from multiple angles include discrete of differential operators, influences of parameters, and comparing with other typical existing variational methods. And then we evaluate these algorithms according to structure noise ratio (SNR) and structure similarity index measure (SSIM), the definitions (see [27] for details) are given in the following

$$SNR_k = 10 \log_{10} \frac{\|u - \overline{u}\|^2}{\|u - \widehat{u}_{k+1}\|^2}$$
 (21)

where u denotes the original image, \overline{u} denotes the mean value of u and \hat{u}_{k+1} denotes the denoised image.

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$
 (22)

where x denotes the clean image, y denotes the denoised image; μx , μy being the mean values of x and y; c1, c2 are two constants for avoiding zero denominators during calculations; σx , σy and σxy are the variances of x, y and the covariance of x, y.

5.1. Numerical Discretization

For given image $u = (u_{ij})$ with size M \times N. We define the periodic partial derivatives as follows:

$$\nabla_x^+ u_{ij} = u_{i+1j} - u_{ij}, \nabla_y^+ u_{ij} = u_{ij+1} - u_{ij},$$

$$\nabla_x^- u_{ij} = u_{ij} - u_{i-1j}, \nabla_y^- u_{ij} = u_{ij} - u_{ij-1}.$$
 (23)

For $p = (p^1, p^2)$, the divergence is defined by

$$\nabla \cdot p = \operatorname{div}(p) = \nabla_{r}^{-} p^{1} + \nabla_{r}^{-} p^{2}. \quad (24)$$

In addition, as the weight w in the second term of RTGV model plays important role in capturing the sharp features, it needs to be well defined. In our method, we give the definition of $w = (w_{ij})$ as follows, and plenty of experiments show that this definition is effective in filtering noise and restoring more sharp features and structures.

$$\begin{aligned} w_{ij} &= e^{-\frac{d_{ij}^2}{2d}}, d_{ij}^2 = d_{xij}^2 + d_{yij}^2, \\ d_{xij}^2 &= (\nabla_x^+ u_{ij})^2 + ((u_{ij} + u_{ij+1} + u_{i-1j} + u_{i-1j+1})/4)^2, \\ d_{yij}^2 &= (\nabla_y^+ u_{ij})^2 + ((u_{ij} + u_{ij-1} + u_{i+1j} + u_{i+1j-1})/4)^2, \\ \bar{d} &= average(d_{ij}^2). \end{aligned}$$

where average(\cdot) denotes the average value of the sum over all d_{ii}^2 .

5.2. Parameters

Our image denoising algorithms has algorithm parameter r_p , r_q , which have impacts on the convergence speed of the algorithm. In addition, there are three model parameters α_0 , α_1 , β , which affect the quality of restoration.

Firstly, according to lots of experimental tests, α_1 , r_p , r_q , r_z are fixed as $\alpha_1 = 0.8$, $r_p = 5$, $r_q = 5$. Fig. 2 shows examples of different $\alpha_1 = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2$ with other parameters held constant. Furthermore, α_0 impacts the sharp features. The larger the α_1 is, the more sharp features are restored. α_0 can be fixed by 1000 for images containing many sharp features, and by 0.5 for images with smooth regions.



Figure 2: Denoising results of different α_1 with other parameters fixed.

Secondly, parameter β affects the effect of denoising. The smaller the β , the smoother the denoising result will be. However, it will be failed to restore image if the β is too large. Besides, for this parameter, we can not get it automatically, and need to tune it manually.

5.3. Comparisons with other methods

For proving the effectiveness of our method, we compare the proposed method with other variational methods, including Wang18 [3], TGV [18] and TC (Total Curvature)[11]. The results of Wang2018 are provided by the authors, and the other results we adjusted the parameters according to SNR and SSIM to get the best results.

Figure 3 displays the denoising results of images with mixed features. As shown, both of these methods are effective in removing noise. However, the TC method produces staircase effects. By comparison, our method can effectively mitigate the staircase effects and recover more features. Furthermore, our method produces slightly larger SNRs and SSIMs, which further prove the priority of our method.

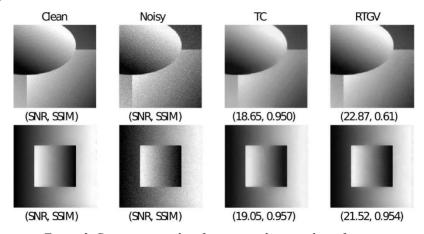


Figure 3: Denoising results of images with many sharp features.

Figure 4 shows the denoising results of some grayscale images. As observed, our method produces better results than the TC method both visually and in terms of the corresponding SNRs and SSIMs for each result. Furthermore, we also test the effect of different levels of noise on denoising, and the results are shown in Figure 5. In theory, the higher the noise level is, the more difficult it is to recover. We can see that the SNRs of the restored images keep decreasing with the noise level increasing. However, the results demonstrate that our algorithm can produce satisfying results for different levels of noise.

To further demonstrate the robustness of our algorithm, we apply the algorithm to color images. The denoising results are shown in Figure 6. The results show that the SNRs and SSIMs of our algorithm are significantly larger than those of the TC method, and the visually restored images are closer to the original clean image.

In addition, we also compare our method with the classical total generalized variation (TGV) method [18] and Wang2018 [3] in Fig. 7. As observed, all the four methods can filter the noise effectively. However, the TGV method blurs some features. There are some noise left in the result of Wang2018 [3]. The TC method smoothes out some details. By comparison, our method performs well, and the SNRs

and SSIMs further indicate the priority of our method.

5.4. Computational costs

In the experiments, we found that both the noise level and the choice of parameter β affect the convergence speed of the algorithm. Understandably, the higher the noise, the slower the convergence. Furthermore, the convergence speed will be slow if β is too small. According to our experiments, for a gray image with size 183×180 , the CPU costs of our algorithm is about 4s. Generally, at a reasonable CPU cost, our method is able to recover more details than other methods.

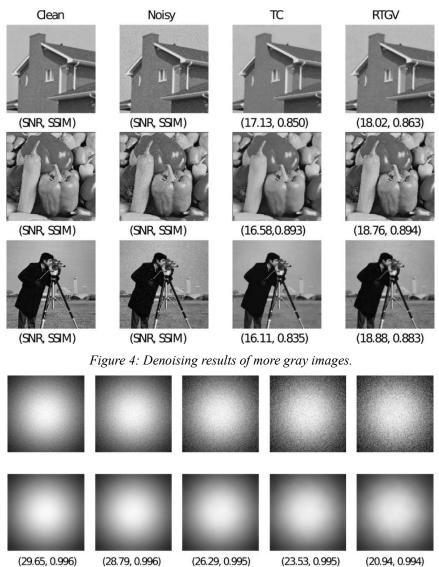


Figure 5: Denoising results of images with different noise levels. The first row is the images with different noise levels, and the second rows are results produced by RTGV.



Figure 7: Denoising results for MRI images.

5.5. Limitation

Our approach has been proven to be effective. However, it still has some limitations. Such as our method cannot recover some details of the image when the noise level is too high. In Fig. 8, we display denoising results for images with higher level noise by TGV, TC and our method. As shown, all the three

methods cannot get satisfying results. In contrast, our results are a little better than that of the other two methods. In addition, the parameters β and α_0 cannot be computed automatically.

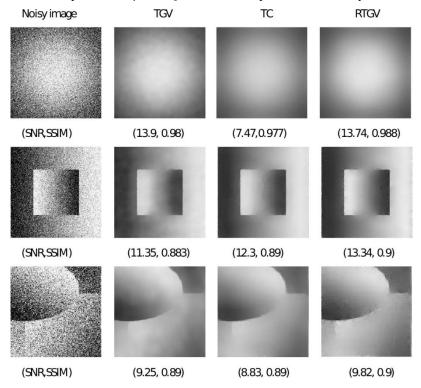


Figure 8: Denoising results for higher level noise.

6. Conclusion

In this paper, we propose a new relaxed total generalized variation technique for image denoising. The proposed technique tries to simplify the classical TGV model by the weighted divergence operator, which can reduce the computational complexity and guaranteed the discrete accuracy. In addition, we have iteratively solved our relaxed total generalized variation image denoising method by augmented Lagrangian method. Our algorithms are discussed from influence of parameters, numerical discretization and comparisons with other classical variation methods. The experimental results show that compared with other variational models, the proposed method is more advantageous in maintaining structures information and alleviate most of staircase effect. In the next, we intend to extend this method to image segmentation and image inpainting.

References

- [1] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. Multiscale Modeling and Simulation, 4(2):460–489, 2005.
- [2] Y. Romano and M. Elad. Boosting of image denoising algorithms. SIAM Journal on Imaging Sciences, 8(2):1187–1219, 2015.
- [3] W. Wang, S. Wen, C. Wu, and J. Deng. An iterative method based on selective averaging and outlier removal. Numerical Mathematics: Theory, Methods and Applications, accepted, 2018.
- [4] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 60(1):259–268, 1992.
- [5] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. SIAM Journal on Scientific Computing, 22(2):503–516, 2000.
- [6] J. C. De los Reyes, C.-B. Schönlieb, and T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. Journal of Mathematical Imaging and Vision, 57(1):1–25, 2017.

 [7] M. Lysaker, A. Lundervold, and X-C. Tai. Noise removal using fourth-order partial differential.
- [7] M. Lysaker, A. Lundervold, and X.-C. Tai. Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. IEEE Transactions on Image Processing, 12(12):1579–1590, 2003.
- [8] C. Wu, J. Zhang, and X.-C. Tai. Augmented lagrangian method for total variation restoration with

- non-quadratic fidelity. Inverse Problems & Imaging, 5(1):237–261, 2011.
- [9] F. Knoll, K. Bredies, T. Pock, and R. Stollberger. Second order total generalized variation (tgv) for mri. Magnetic Resonance in Medicine, 65(2):480–491, 2011.
- [10] M. Kang, M. Kang, and M. Jung. Total generalized variation based denoising models for ultrasound images. Journal of Scientific Computing, 72(1):172–197, 2017.
- [11] Q. Zhong, Y. Li, Y. Yang, and Y. Duan. Minimizing discrete total curvature for image processing. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [12] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In International Conference on Computer Vision, pages 479–486. IEEE, 2011.
- [13] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Transactions on Image Processing, 2017.
- [14] C. Tian, Y. Xu, Z. Li, W. Zuo, and H. Liu. Attention-guided cnn for image denoising. Neural Networks, 124:117–129, 2020.
- [15] S. Laine, T. Karras, J. Lehtinen, and T. Aila. High-quality self-supervised deep image denoising. 2019.
- [16] A. Xl, A. Jx, A. Yz, A. Yy, A. Nl, A. Xw, A. Sw, and B. Sbg. Detail retaining convolutional neural network for image denoising. Journal of Visual Communication and Image Representation, 71, 2020.
- [17] M. Yashtini and S. H. Kang. A fast relaxed normal two split method and an effective weighted tv approach for euler's elastica image inpainting. Siam Journal on Imaging Sciences, 9(4):1552–1581, 2016.
- [18] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. SIAM Journal on Imaging Sciences, 3(3):492–526, 2010.
- [19] K. Papafitsoros and C.-B. Schönlieb. A combined first and second order variational approach for image reconstruction. Journal of Mathematical Imaging and Vision, 48(2):308–338, 2014.
- [20] K. Bredies and H. P. Sun. Preconditioned douglas—rachford algorithms for tv-and tgv-regularized variational imaging problems. Journal of Mathematical Imaging and Vision, 52(3):317–344, 2015.
- [21] H. Zhang, Z. He, and X. Wang. A novel mesh denoising method based on relaxed second-order total generalized variation. SIAM Journal on Imaging Sciences, 15(1):1–22, 2022.
- [22] X.-C. Tai, J. Hahn, and G. J. Chung. A fast algorithm for euler's elastica model using augmented lagrangian method. SIAM Journal on Imaging Sciences, 4(1):313–344, 2011
- [23] R. Glowinski and P. Le Tallec. Augmented Lagrangian and operator-splitting methods in nonlinear mechanics, volume 9. SIAM, 1989.
- [24] C. Wu and X.-C. Tai. Augmented lagrangian method, dual methods, and split bregman iteration for ROF, vectorial TV, and high order models, SIAM Journal on Imaging Sciences, 3(3):300–339, 2010.
- [25] C. Wu, J. Zhang, Y. Duan, and X.-C. Tai. Augmented lagrangian method for total variation based image restoration and segmentation over triangulated surfaces. Journal of Scientific Computing, 50(1):145–166, 2012.
- [26] J. Zhang and K. Chen. A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution. SIAM Journal on Imaging Sciences, 8(4):2487–2518, 2015.
- [27] 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, 13(4):600–612, 2004.