

Fast Overlapping Group Sparsity Total Variation Image Denoising Based on Fast Fourier Transform and Split Bregman Iterations

Yingpin Chen^{1,2}, Linna Wu³, Zhenming Peng¹⁺ and Xingguo Liu¹

¹ School of Optoelectronic Information, University of Electronic Science and Technology of China, Chengdu 610054, China

² College of Physics and Information Engineering, Minnan Normal University, Zhangzhou, 363000, China

³ Centers for Biomedical Engineering, University of Science and Technology of China, Hefei, 230027, China

Abstract. The total variation denoising model is considered to be one of the best denoising models. However, the total variation model always introduces stair-case artifacts. To overcome the drawback, we use an overlapping group sparsity total variation instead of total variation denoising model. By introducing fast Fourier transform and split Bregman iteration framework, we propose a fast algorithm to solve the overlapping group sparse model. Experiments are carried out to compare with the traditional TV denoising method and state-of-the-art total generalized variation method. The experiments demonstrate that our algorithm avoids stair-case of traditional TV model.

Keywords: Total variation, image denoising, fast Fourier transform, overlapping group sparsity.

1. Introduction

Rudin, Osher and Fatemi proposed the well-known total variation (TV) denoising model[1]. The TV denoising model achieves promising denoising performance. However, the first order TV assumes images to be piece-wise constant, which introduce stair-case artifacts. To overcome the drawback mentioned above, some extensions of TV have been put forward. For example, total generalized variation (TGV)[2] is proved to be a good way to reduce stair-case artifacts.

Recently, an overlapping group sparsity (OGS) total variation model was firstly put forward for one-dimensional signal denoising[3]. As first-order difference of signal is not only sparse but also contains a form of structured sparsity, the OGS total variation model can fit the prior knowledge and reduce staircase artifacts efficiently. Liu, Huang, et al applied the OGS TV model in image deburring[4] under the framework of alternating direction method of multipliers. Considering that the OGS TV model is a complex problem, we use split Bregman iterations[5] to transform the issue into some sub problems which are easier to solve. Considering that the efficiency of fast Fourier transform (FFT) is high[6], we introduce FFT to improve the sub problem of split Bregman iterations. In this way, the method we proposed can avoid large size matrix multiplication. As a result, the whole computational efficiency is improved prominently.

The rest of this paper is organized as follows. We first review the traditional TV denoising model in section 2. Then, the proposed method is shown in section 3. In section 4, we carry out some experiments to demonstrate the proposed method. At last, we summarize the whole algorithm and provide some possible directions for further studies.

2. Preparation

The TV denoising model is shown as follow [1]. (remark: considering using the FFT in the proposed

⁺ Corresponding author.
E-mail address: zmpeng@uestc.edu.cn.

method, we assume that the image satisfies the periodic boundary condition.)

$$\mathbf{f} = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{f} - \mathbf{g}\|_2^2 + \mu R(\mathbf{f}), \quad (1)$$

where $\mathbf{g} \in \mathbf{R}^{N^2 \times 1}$ is the vectorial form of image with noise (for simplify, we assume the image size is $N \times N$).

$\mathbf{f} \in \mathbf{R}^{N^2 \times 1}$ represents the recovered image by TV denoising model. N is the image size. $R(\mathbf{f})$ represents the regularity term. μ denotes the regularity parameter balancing the fidelity term and the regularity term. The classical anisotropic total variation is defined as

$$R_{ATV}(\mathbf{f}) = \|\nabla_h \mathbf{f}\|_1 + \|\nabla_v \mathbf{f}\|_1, \quad (2)$$

where ∇_h and ∇_v denote horizontal and vertical differential matrix. They can be calculated by (3).

$$\nabla_h = \mathbf{D}_{v0} \otimes \mathbf{I}, \quad \nabla_v = \mathbf{I} \otimes \mathbf{D}_{v0}, \quad (3)$$

where \mathbf{I} is the identity matrix and \otimes denotes the kronecker operator. As we assume the image satisfies the

periodic boundary condition, we define \mathbf{D}_{v0} as $\mathbf{D}_{v0} = \begin{pmatrix} -1 & 1 & & \\ & -1 & 1 & \\ & & \ddots & \ddots \\ & & & -1 & 1 \\ 1 & & & & -1 \end{pmatrix} \in \mathbf{R}^{N \times N}$. The traditional TV model

may lead to stair-case artifacts, which will be shown in section 4. In the next section, we will propose an OGS based model to overcome the drawback of traditional TV model.

3. Method

In this section, we introduce the overlapping group sparsity TV model to improve the denoising performance, that is

$$\mathbf{f} = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{f} - \mathbf{g}\|_2^2 + \mu [\varphi(\nabla_h \mathbf{f}) + \varphi(\nabla_v \mathbf{f})], \quad (4)$$

where $\varphi(\mathbf{v}) = \sum_{i=1}^N \sum_{j=1}^N \|\tilde{\mathbf{V}}_{i,j,K,K}\|_2$. K is the group size. When $K=1$, the OGS total variation model becomes TV

model. $\tilde{\mathbf{V}}_{i,j,K,K}$ is defined as

$$\tilde{\mathbf{V}}_{i,j,K,K} = \begin{bmatrix} \mathbf{V}_{i-K_l, j-K_r} & \mathbf{V}_{i-K_l, j-K_r+1} & \cdots & \mathbf{V}_{i-K_l, j-K_r} \\ \mathbf{V}_{i-K_l+1, j-K_r} & \mathbf{V}_{i-K_l+1, j-K_r+1} & \cdots & \mathbf{V}_{i-K_l+1, j-K_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{V}_{i+K_r, j-K_l} & \mathbf{V}_{i+K_r, j-K_l+1} & \cdots & \mathbf{V}_{i+K_r, j+K_r} \end{bmatrix} \in \mathbf{R}^{K \times K}, \quad (5)$$

where $\mathbf{v} \in \mathbf{R}^{N^2 \times 1}$ is stacked in column-wisely, that is, the (i, j) th entry of a matrix $\mathbf{V} \in \mathbf{R}^{N \times N}$ is assigned to

be the $(j-1)N + i$ th entry of the vector \mathbf{v} . $K_l = \lfloor \frac{K-1}{2} \rfloor$, $K_r = \lfloor \frac{K}{2} \rfloor$.

As is seen in formula (4), the structured sparsity information of neighbouring pixels is considered. As a result, the stair-case artifacts can be eliminated. To solve the problem defined in (4), we can introduce some split variables, which are $\mathbf{z}_1 = \nabla_h \mathbf{f}$, $\mathbf{z}_2 = \nabla_v \mathbf{f}$, Then we get

$$\mathbf{f} = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{f} - \mathbf{g}\|_2^2 + \mu [\varphi(\mathbf{z}_1) + \varphi(\mathbf{z}_2)] \quad s.t \quad \mathbf{z}_1 = \nabla_h \mathbf{f}, \mathbf{z}_2 = \nabla_v \mathbf{f}. \quad (6)$$

The optimization problem in (6) can be solved by split Bregman iterations as follows

$$(\mathbf{f}^{(k+1)}, \mathbf{z}_1^{(k+1)}, \mathbf{z}_2^{(k+1)}) = \min_{\mathbf{f}, \mathbf{z}_1, \mathbf{z}_2} \frac{1}{2} \|\mathbf{f} - \mathbf{g}\|_2^2 + \mu[\varphi(\mathbf{z}_1) + \varphi(\mathbf{z}_2)] + \frac{\beta}{2} \left(\left\| \mathbf{z}_1 - \nabla_h \mathbf{f} - \tilde{\mathbf{z}}_1 \right\|_2^2 + \left\| \mathbf{z}_2 - \nabla_v \mathbf{f} - \tilde{\mathbf{z}}_2 \right\|_2^2 \right), \quad (7)$$

$$\tilde{\mathbf{z}}_1^{(k+1)} = \tilde{\mathbf{z}}_1^{(k)} + (\nabla_h \mathbf{f} - \mathbf{z}_1^{(k+1)}), \quad \tilde{\mathbf{z}}_2^{(k+1)} = \tilde{\mathbf{z}}_2^{(k)} + (\nabla_v \mathbf{f} - \mathbf{z}_2^{(k+1)}),$$

where $\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2$ are the dual variables of $\mathbf{z}_1, \mathbf{z}_2$.

According to (7), the \mathbf{f} sub problem can be solved by the sub objective function shown as follow

$$\mathbf{f}^{(k+1)} = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{f} - \mathbf{g}\|_2^2 + \frac{\beta}{2} \left(\left\| \mathbf{z}_1^{(k)} - \nabla_h \mathbf{f} - \tilde{\mathbf{z}}_1^{(k)} \right\|_2^2 + \left\| \mathbf{z}_2^{(k)} - \nabla_v \mathbf{f} - \tilde{\mathbf{z}}_2^{(k)} \right\|_2^2 \right). \quad (8)$$

Setting the derivative of \mathbf{f} in formula (8) to zero, then we have

$$\mathbf{0} = (\mathbf{f} - \mathbf{g}) + \beta[\nabla_h^T (\nabla_h \mathbf{f} - \mathbf{z}_1^{(k)} + \tilde{\mathbf{z}}_1^{(k)}) + \nabla_v^T (\nabla_v \mathbf{f} - \mathbf{z}_2^{(k)} + \tilde{\mathbf{z}}_2^{(k)})]. \quad (9)$$

Formula (9) can be solved by using conjugate gradient method (CGM). However, In (9), there are two large size matrix (for example, $\nabla_h \in \mathbf{R}^{N^2 \times N^2}$) multiplications. Because the computational complexity of multiplication is $\mathcal{O}((N^2)^3)$, the computing efficiency is not very high by using CGM.

To avoid large size matrix multiplication in (9), we rewrite formula (8) as matrix convolution form. In this way, the computational complexity becomes $\mathcal{O}(N^2 \log_2(N))$.

$$\mathbf{F}^{(k+1)} = \underset{\mathbf{F}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{F} - \mathbf{G}\|_2^2 + \frac{\beta}{2} \left(\left\| \mathbf{Z}_1^{(k)} - \mathbf{K}_h * \mathbf{F} - \tilde{\mathbf{Z}}_1^{(k)} \right\|_2^2 + \left\| \mathbf{Z}_2^{(k)} - \mathbf{K}_v * \mathbf{F} - \tilde{\mathbf{Z}}_2^{(k)} \right\|_2^2 \right), \quad (10)$$

where $\mathbf{K}_h = [1, -1]$ is the horizontal convolution kernel and $\mathbf{K}_v = [1; -1]$ is the vertical convolution kernel. Therefore, \mathbf{K}_h and \mathbf{K}_v are consistent with equation (3). The symbol $*$ is the two-dimensional convolution operator. $\mathbf{Z}_i^{(k)} (i=1,2) \in \mathbf{R}^{N \times N}$ are the matrix form of $\mathbf{z}_i (i=1,2)$. $\tilde{\mathbf{Z}}_i^{(k)} (i=1,2) \in \mathbf{R}^{N \times N}$ are the matrix form of $\tilde{\mathbf{z}}_i (i=1,2)$. $\mathbf{G} \in \mathbf{R}^{N \times N}$ is the matrix form of the observed image. $\|\bullet\|_2$ represents the entrywise L_2 norm of matrix.

Setting the first-order derivative with regard to \mathbf{F} to zero, then we have

$$\mathbf{F}^{(k+1)} = F_{2D}^{-1} \frac{F_{2D}(\mathbf{G}) + \beta(F_{2D}(\mathbf{K}_h))^* \circ F_{2D}(\mathbf{Z}_1^{(k)} - \tilde{\mathbf{Z}}_1^{(k)}) + (F_{2D}(\mathbf{K}_v))^* \circ F_{2D}(\mathbf{Z}_2^{(k)} - \tilde{\mathbf{Z}}_2^{(k)})}{\mathbf{1} + \beta[(F_{2D}(\mathbf{K}_h))^* \circ F_{2D}(\mathbf{K}_h) + (F_{2D}(\mathbf{K}_v))^* \circ F_{2D}(\mathbf{K}_v)]}, \quad (11)$$

where F_{2D} and F_{2D}^{-1} represent the 2D FFT and IFFT operator respectively. The symbol \circ denotes componentwise multiplication. The division in (11) is componentwise as well. $\mathbf{f}^{(k+1)}$ can be obtained by vectorize $\mathbf{F}^{(k+1)}$. The symbol $\mathbf{1}$ represents the matrix with every element is 1.

The $\mathbf{z}_i (i=1,2)$ sub problem can be easily solved by the sub objective function shown in (12).

$$\mathbf{z}_1^{(k+1)} = \arg \min_{\mathbf{z}_1} \mu\varphi(\mathbf{z}_1) + \frac{\beta}{2} \left\| \mathbf{z}_1 - \nabla_h \mathbf{f}^{(k+1)} - \tilde{\mathbf{z}}_1^{(k)} \right\|_2^2, \quad (12)$$

$$\mathbf{z}_2^{(k+1)} = \arg \min_{\mathbf{z}_2} \mu\varphi(\mathbf{z}_2) + \frac{\beta}{2} \left\| \mathbf{z}_2 - \nabla_v \mathbf{f}^{(k+1)} - \tilde{\mathbf{z}}_2^{(k)} \right\|_2^2.$$

According to the majorization minimization (MM) method[7], the $\mathbf{z}_i (i=1,2)$ sub problem can be easily solved by the recursive iterations shown as follows

$$\mathbf{z}_{1(n+1)}^{(k+1)} = [\mathbf{I} + \frac{\mu}{\beta} \Lambda^2(\mathbf{z}_{1(n)}^{(k+1)})]^{-1} \mathbf{z}_{1(0)}^{(k+1)},$$

$$\mathbf{z}_{2(n+1)}^{(k+1)} = [\mathbf{I} + \frac{\mu}{\beta} \Lambda^2(\mathbf{z}_{2(n)}^{(k+1)})]^{-1} \mathbf{z}_{2(0)}^{(k+1)} \quad (n=1,2,\dots), \quad (13)$$

where $\mathbf{I} \in \mathbf{R}^{N^2 \times N^2}$ denotes the unit matrix. $\mathbf{z}_{i(n)}^{(k+1)}$ ($i=1,2$) denotes the n -th inner iteration of MM method in the $(k+1)$ -th external circulation. $\mathbf{z}_{1(0)}^{(k+1)} = \nabla_h \mathbf{f}^{(k+1)} + \tilde{\mathbf{z}}_1^{(k)}$, $\mathbf{z}_{2(0)}^{(k+1)} = \nabla_v \mathbf{f}^{(k+1)} + \tilde{\mathbf{z}}_2^{(k)}$. $\mathbf{\Lambda} \in \mathbf{R}^{N^2 \times N^2}$ is a diagonal matrix with each diagonal component is defined as

$$[\mathbf{\Lambda}(\mathbf{z})]_{m,m} = \sqrt{\sum_{i=-K_l}^{K_r} \sum_{j=-K_l}^{K_r} \left\{ \sum_{k_1=-K_l}^{K_r} \sum_{k_2=-K_l}^{K_r} |\mathbf{z}_{m-i+k_1, m-j+k_2}|^2 \right\}^{-\frac{1}{2}}}, \quad (14)$$

where \mathbf{Z} is the matrix form of \mathbf{z} .

According to the principle of split Bregman iteration, the dual variable can be update as follows

$$\tilde{\mathbf{z}}_1^{(k+1)} = \tilde{\mathbf{z}}_1^{(k)} + \nabla_h \mathbf{f}^{(k+1)} - \mathbf{z}_1^{(k)}, \quad \tilde{\mathbf{z}}_2^{(k+1)} = \tilde{\mathbf{z}}_2^{(k)} + \nabla_v \mathbf{f}^{(k+1)} - \mathbf{z}_2^{(k)}. \quad (15)$$

The whole algorithm is summarized in Table 1.

Table 1. Pseudo-code of proposed method

Algorithm 1 Pseudo-code of proposed method

Initialize: $K, \tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2 = \mathbf{0}, \mathbf{z}_1, \mathbf{z}_2 = \mathbf{0}, \mathbf{f}^{(k)} = \mathbf{0}, \mu, \beta, tol$

While $\|\mathbf{f}^{(k+1)} - \mathbf{f}^{(k)}\|_2 / \|\mathbf{f}^{(k)}\|_2 > tol$ **do**

Update $\mathbf{f}^{(k+1)}$ using equation (11).

While $\|\mathbf{z}_{i(n+1)}^{(k+1)} - \mathbf{z}_{i(n)}^{(k+1)}\|_2 / \|\mathbf{z}_{i(n+1)}^{(k+1)}\|_2 > tol$ **do**

Update $\mathbf{z}_{i(n+1)}^{(k+1)}$ ($i=1,2$) using equation (13).

$n=n+1$.

End while

$\mathbf{z}_i^{(k+1)} = \mathbf{z}_{i(n+1)}^{(k+1)}$ ($i=1,2$).

Update $\tilde{\mathbf{z}}_1^{(k+1)}, \tilde{\mathbf{z}}_2^{(k+1)}$ using equation (15).

$k = k + 1$.

End while

4. Experiment

In this section, we show the denoising result of our method (we name it OGSTV-FFT for short) comparing with the traditional TV model and total generalized variation model which is state-of-the-art TV model. Fig 1 shows the denoising results by different methods. In our experiment, we set $K=3$.

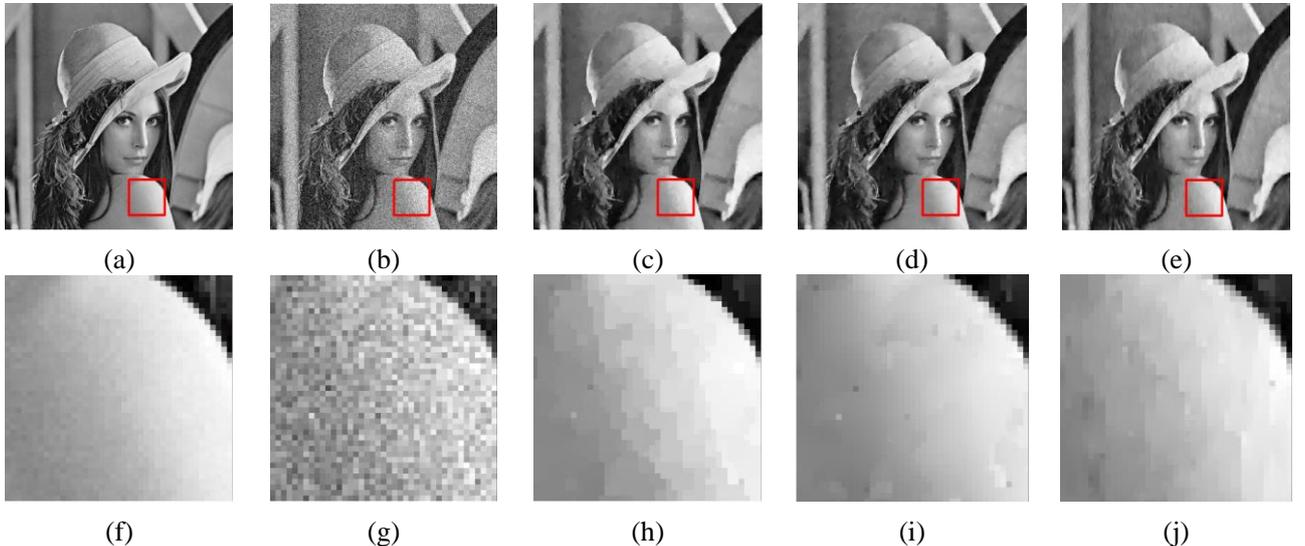


Fig 1: Denoising results of different methods.(a) Ground truth.(b) Image with noise (the noise is the white Gaussian

noise with zero mean and standard deviation $\sigma = 20$).(c) The ATV denoising result. (d) The TGV denoising result. (e) The proposed method. (f)-(j) are the partial enlarged detail of (a)-(e).

As is seen in Fig 1(h), the ATV model has stair-case artifacts. The denoising result obtained by our method reduce the stair-case artifacts obviously (shown in Fig 1(j)). To evaluate the proposed algorithm performance objectively, we use peak signal to noise ratio (PSNR) and structural similarity (SSIM) shown in Table 2, comparing with the ATV model and TGV model. The best records are marked in bold.

Table 2. Comparison of experimental results

σ	Methods	PSNR	SSIM	Time
10	ATV	32.237	0.883	1.27
	TGV	32.430	0.890	1.39
	OGSTV-FFT	32.650	0.892	1.20
20	ATV	28.632	0.800	1.44
	TGV	28.974	0.817	2.06
	OGSTV-FFT	28.990	0.819	1.23
50	ATV	24.581	0.687	2.45
	TGV	24.708	0.682	2.19
	OGSTV-FFT	24.939	0.698	2.31
100	ATV	21.221	0.561	2.50
	TGV	21.289	0.552	2.33
	OGSTV-FFT	21.323	0.569	7.31

As shown in Table 2, the PSNR and SSIM obtained by our method are higher than the ones obtained by the other two models. What's more, for the application of FFT, our method works very efficiently although the denoising model becomes more complex than ATV model.

5. Conclusion

In this paper, we propose an efficient method solving the OGS total variation denoising model based on FFT and split Bregman iteration. The proposed method considers structured sparsity of total variation instead of the sparsity feature of total variation, alleviating the stair-case artifacts appeared in total variation. The experimental results indicate that the proposed method performs better than ATV and TGV method. By using FFT, it is very efficient to solve the sub problem of whole algorithm. Since the proposed TV model is a general regular term, it can be applied in the other kinds of noise in further studies.

6. Acknowledgements

This work is supported by National Natural Science Foundation of China [grant numbers 61571096, 41274127], the Natural Science Foundation of Fujian Province [grant number 2015J01270]

7. References

- [1] L. I. Rudin, S. Osher, E. Fatemi. Nonlinear total variation based noise removal algorithms[J]. *Physica D: Nonlinear Phenomena*, 1992, **60**(1-4): 259-268
- [2] K. Bredies, K. Kunisch, T. Pock. Total generalized variation[J]. *SIAM Journal on Imaging Sciences*, 2010, **3**(3): 492-526
- [3] I. W. Selesnick, P.-Y. Chen. Total variation denoising with overlapping group sparsity[C]. *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, 5696-5700
- [4] J. Liu, T.-Z. Huang, I. W. Selesnick, et al. Image restoration using total variation with overlapping group sparsity[J]. *Information Sciences*, 2015, **295**: 232-246
- [5] T. Goldstein, S. Osher. The split Bregman method for L1-regularized problems[J]. *SIAM journal on imaging sciences*, 2009, **2**(2): 323-343
- [6] Y. Wang, J. Yang, W. Yin, et al. A new alternating minimization algorithm for total variation image reconstruction[J]. *SIAM Journal on Imaging Sciences*, 2008, **1**(3): 248-272
- [7] G. Liu, T.-Z. Huang, J. Liu, et al. Total variation with overlapping group sparsity for image deblurring under impulse noise[J]. *PLoS one*, 2015, **10**(4): e0122562