A Shearlet-TV Regularization Approach for Image Denoising in the ADMM Framework

Xingguo Liu [1, 2], Yingpin Chen [1], Zhenming Peng [1 +] and Juan Wu [3]

[1] School of Optoelectronic Information, University of Electronic Science and Technology of China, Chengdu, 610054, P. R. China

[2] Chongqing College of Electronic Engineering, Chongqing, 401331, P. R. China

[3] College of Communication Engineering, Chongqing University, Chongqing, 400044, P. R. China

Abstract. The Shearlet transform is widely used in images denoising. However, the Shearlet based denoising method suffers from Gibbs artifacts because of the cut off in Fourier domain. In order to suppress the Gibbs artifacts, we propose a multi-constrained denoising model by using Shearlet transform and total variation. Considering that the proposed model is a multi-constrained problem, alternating direction method of multipliers is used to decouple the complex issue into some sub problems, which are easier to solve. Experiments are carried out to show the performance of proposed method. The experimental results indicate that the proposed method effectively relive the Gibbs artifacts of Shearlet transform.

Keywords: shearlet, TV, ADMM, image denoising

1. Introduction

Image restoration is a typical ill-conditioned problem in image processing. Regularization is an important tool to solve this issue. The basic idea of regularization is to find a mathematical model fitting with the prior knowledge with respect to the processed image. As a result, the restoration becomes the result balancing with observed image with prior regularization. There are many regularization items, including total variation regularization[1], wavelet regularization and so on. Recently, sparse regularization theory has been widely used in image denoising[2], image super-resolution reconstruction[3], compressed sensing[4] and so on. Sparse regularization assumes that the image has a sparse distribution in a certain transformation domain. However, sparse regularization only considers the local similarity structure of image, which may cause the loss of image texture and structure information. Typically, the Shearlet transform (ST) has the best sparse representation for two-dimensional signals. Nevertheless, it introduces the Gibbs artifacts. Considering that the anisotropic total variation (ATV) is good at keeping the edge of processed image, we use ATV method to reduce the Gibbs artifacts in Shearlet denoising.

In the following, we first review related work in section 2. Then, the proposed method is shown in detail in section 3. We carry out some experiments to demonstrate the proposed method in section 4, and draw the conclusions in section 5.

2. Related work

The shear matrix $\boldsymbol{B}_s$ and the anisotropic scale matrix $\boldsymbol{A}_a$ are defined as[5]

$$\boldsymbol{B}_s = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}, s \in R, \tag{1}$$

[+] Corresponding author.

E-mail address: zmpeng@uestc.edu.cn.

$$A_a = \begin{bmatrix} a & 0 \\ a & \sqrt{a} \end{bmatrix}, a \in R , \tag{2}$$

where s represents the shear parameter and a represents the dilation coefficient which gives the parabolic scale in the transform.

The shearlet basis function $\Psi_{ast}(\varphi)$ is described as[6]

$$\Psi_{ast}(\varphi) = \{\varphi_{ast}(\boldsymbol{x}) = a^{-\frac{3}{4}}\varphi(M_{as}^{-1}(\boldsymbol{x} - \boldsymbol{t}))\}, \tag{3}$$

where $\boldsymbol{x}$ and $\boldsymbol{t}$ are the two-dimensional signal and time shift parameter, respectively. M_{as} provides the scale transformation and geometrical transformations, such as shear transformation, which is defined as

$$M_{as} = B_s \cdot A_a = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} = \begin{pmatrix} a & \sqrt{a}s \\ 0 & \sqrt{a} \end{pmatrix}. \tag{4}$$

For all $\boldsymbol{f} \in L^2(R^2)$, the shearlet coeffcients can be calculated by equation as follows

$$\mathrm{SH}_{\Psi}\boldsymbol{f}(a,s,\boldsymbol{t}) = <\boldsymbol{f}, \Psi_{ast}>, a \in R, s \in R, \boldsymbol{t} \in R^2 , \tag{5}$$

Different from the traditional wavelet transform, the shearlet transform has an anisotropic dilation which means that, in case of two dimensions, the dilations in the x direction and y direction can be different.

3. Proposed method

In this section, a novel model of Shearlet-TV regularization suppressing the Gibbs artifacts in Shearlet denoising is presented. In the proposed model, Shearlet transform is firstly adopted to efficiently and sparsely represent anisotropic information which could make a series of sub-bands with different scales and directions. Then, TV is used to suppress noise of the sub-bands and the Gibbs artifacts. At last, ADMM[7] iteration method is used to solve the model.

The model described in this paper can be written as follows

$$\min_{f} \frac{1}{2}\|\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \mu_0 \sum_{i=1}^{N} \|\mathrm{SH}_i(\boldsymbol{f})\|_1 + \mu_1 R_{ATV}(\boldsymbol{f}) , \tag{6}$$

where μ_0 and μ_1 are positive regularization parameters representing the degree of sparsity in the Shearlet and TV respectively. $\boldsymbol{f} \in \boldsymbol{R}^{M^2 \times 1}$ represents the recovered image, M is the image size. $\boldsymbol{g} \in \boldsymbol{R}^{M^2 \times 1}$ is the image with noise. N is used as the number of sub-bands from the Shearlet transform. $\|\mathrm{SH}_i(\boldsymbol{f})\|_1$ is Shearlet sparse regularization term. $\mathrm{SH}_i(\boldsymbol{f})$ is the ith sub-band. $R_{ATV}(\boldsymbol{f})$ is the anisotropic total variation regularization term, whichis defined as[8]

$$R_{ATV}(\boldsymbol{f}) = \|\nabla_h \boldsymbol{f}\|_1 + \|\nabla_v \boldsymbol{f}\|_1 , \tag{7}$$

where $\|\nabla_h \boldsymbol{f}\|_1$ is the L_1 norm of horizontal difference, and $\|\nabla_v \boldsymbol{f}\|_1$ is the L_1 norm of vertical difference. ∇_h and ∇_v represent horizontal difference matrix and vertical difference matrix respectively.

Three auxiliary variables $\boldsymbol{Z}_1$, $\boldsymbol{Z}_2$ and $\boldsymbol{Z}_3$ are taken to decouple the last two regularization terms. Then, the equation (6) can be reformulated as

$$\min_{f} \frac{1}{2}\|\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \mu_0 \|\boldsymbol{Z}_1\|_1 + \mu_1(\|\boldsymbol{Z}_2\|_1 + \|\boldsymbol{Z}_3\|_1)$$
$$s.t. \ \boldsymbol{Z}_{1i} = \mathrm{SH}_i(\boldsymbol{f}), \boldsymbol{Z}_2 = \nabla_h \boldsymbol{f}, \boldsymbol{Z}_3 = \nabla_v \boldsymbol{f} \tag{8}$$

Equation (8) could be converted to an unconstrained optimization problem by the Lagrangian method. The augmented Lagrangian is following form of:

$$L_\rho(\boldsymbol{f}, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \boldsymbol{Z}_3, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3) = \frac{1}{2}\|\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \mu_0 \|\boldsymbol{Z}_1\|_1 + \mu_1(\|\boldsymbol{Z}_2\|_1 + \|\boldsymbol{Z}_3\|_1)$$
$$+ \langle \boldsymbol{Y}_1, (\boldsymbol{Z}_1 - \mathrm{SH}(\boldsymbol{f})) \rangle + \langle \boldsymbol{Y}_2, (\boldsymbol{Z}_2 - \nabla_h \boldsymbol{f}) \rangle + \langle \boldsymbol{Y}_3, (\boldsymbol{Z}_3 - \nabla_v \boldsymbol{f}) \rangle. \tag{9}$$
$$+ \frac{\rho}{2}(\|\boldsymbol{Z}_1 - \mathrm{SH}(\boldsymbol{f})\|_2^2 + \|\boldsymbol{Z}_2 - \nabla_h \boldsymbol{f}\|_2^2 + \|\boldsymbol{Z}_3 - \nabla_v \boldsymbol{f}\|_2^2)$$

The ADMM can be used to solve l1-regularized optimization problems, which is intended to blend the decomposability of dual ascent with the superior convergence properties of the method of multipliers. When it comes to the model shown in equation (9), the f sub problem can be solved by the sub objective function shown as follow

$$f^{k+1} = \arg\min_f \{L_\rho(f, Z_1^k, Z_2^k, Z_3^k, Y_1^k, Y_2^k, Y_3^k)\}$$

$$= \arg\min_f \{\frac{1}{2}\|f - g\|_2^2 + \mu_0\|Z_1^k\|_1 + \mu_1(\|Z_2^k\|_1 + \|Z_3^k\|_1)$$

$$+ \langle Y_1^k, (Z_1^k - \mathrm{SH}(f)) \rangle + \langle Y_2^k, (Z_2^k - \nabla_h f) \rangle + \langle Y_3^k, (Z_3^k - \nabla_v f) \rangle \qquad (10)$$

$$+ \frac{\rho}{2}(\|Z_1^k - \mathrm{SH}(f)\|_2^2 + \|Z_2^k - \nabla_h f\|_2^2 + \|Z_3^k - \nabla_v f\|_2^2)\}$$

Equation (10) is a typical quadratic programming problem, which can be solved by conjugate gradient method or Gauss-Seidel method.

The Z_i (i=1,2,3) sub problem can be easily solved by the sub objective function shown in (11).

$$Z_1^{k+1} = \arg\min_{Z_{1i}} \{\mu_0\|Z_1\|_1 + \langle Y_1^k, (Z_1 - \mathrm{SH}(f^{k+1})) \rangle + \frac{\rho}{2}\|Z_1 - \mathrm{SH}(f^{k+1})\|_2^2\}$$

$$\Rightarrow Z_1^{k+1} = S_{\mu_0/\rho}(\mathrm{SH}(f^{k+1}) + Y_1^k/\rho)$$

$$Z_2^{k+1} = \arg\min_{Z_2} \{\mu_1\|Z_2\|_1 + \langle Y_2^k, (Z_2 - \nabla_h f^{k+1}) \rangle + \frac{\rho}{2}\|Z_2 - \nabla_h f^{k+1}\|_2^2\}$$

$$\Rightarrow Z_2^{k+1} = S_{\mu_1/\rho}(\nabla_h f^{k+1} + Y_2^k/\rho) \qquad (11)$$

$$Z_3^{k+1} = \arg\min_{Z_3} \{\mu_1\|Z_3\|_1 + \langle Y_3^k, (Z_3 - \nabla_v f^{k+1}) \rangle + \frac{\rho}{2}\|Z_3 - \nabla_v f^{k+1}\|_2^2\}$$

$$\Rightarrow Z_3^{k+1} = S_{\mu_1/\rho}(\nabla_v f^{k+1} + Y_3^k/\rho)$$

where $S_\lambda(x) = \mathrm{sgn}(x)\max\{|x| - \lambda, 0\}$ is soft thresholding operator.

The Y_i (i=1,2,3) is dual variable, which can be obtained using the coordinate ascent method as follows

$$Y_1^{k+1} = Y_1^k + \rho(Z_1^{k+1} - \mathrm{SH}(f^{k+1}))$$

$$Y_2^{k+1} = Y_2^k + \rho(Z_2^{k+1} - \nabla_h f^{k+1}) \qquad (12)$$

$$Y_3^{k+1} = Y_3^k + \rho(Z_3^{k+1} - \nabla_v f^{k+1})$$

Algorithm 1 provides a pseudo-code implementation of our proposed method. The algorithm gives all the steps in the algorithm including the calculation and updating order in iteration.

Algorithm 1： Shearlet and ATV denoising method by ADMM iteration

1: **Input:** f, ρ, μ_0, μ_1, tol
2: **Initialize:** Y_1, Y_2, Y_3, Z_1, Z_2, Z_3,
3: **While** $\|f^{k+1} - f^k\|_2 / \|f^k\|_2 \geq tol$ **do**
4:　Update f^{k+1};
5:　Update Z_1, Z_2, Z_3;
6:　Update Y_1, Y_2, Y_3;
7:　$k=k+1$;
8: **End while**
9: **Output:** f^{k+1}

---

4. Experiment results

In this section, image "peppers" is used to validate the proposed method, which has the size of 256*256 . There are also three methods for comparison: median filtering, Shearlet shrinkage and TV. We synthesize

mixed noise by the summation of white Gaussian noise, denoting zero mean and standard deviation (σ). Fig 1 shows the original image and the image corrupted by mixed noise ($\sigma = 20$).

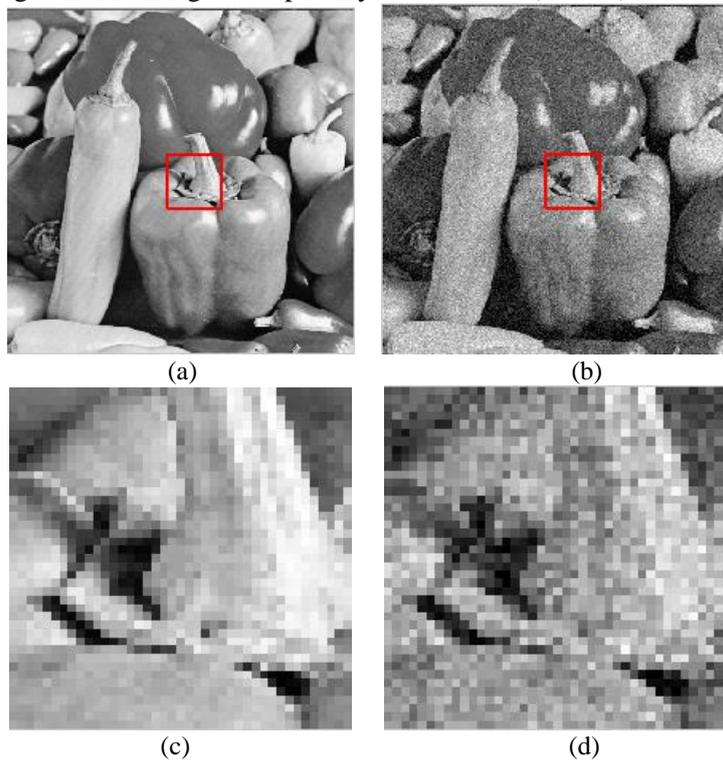(a)   (b)

(c)   (d)

Fig. 1: The original image and the image corrupted by mixed noise,

(a) the original image, (b) the image corrupted by mixed noise ($\sigma = 20$),

(c) and (d) are the partial enlarged detail of (a) and (b)

The results of image denoising for "peppers" are shown in Fig.2, using methods of median filtering, Shearlet shrinkage, TV and the proposed approach.

(a)   (b)   (c)   (d)
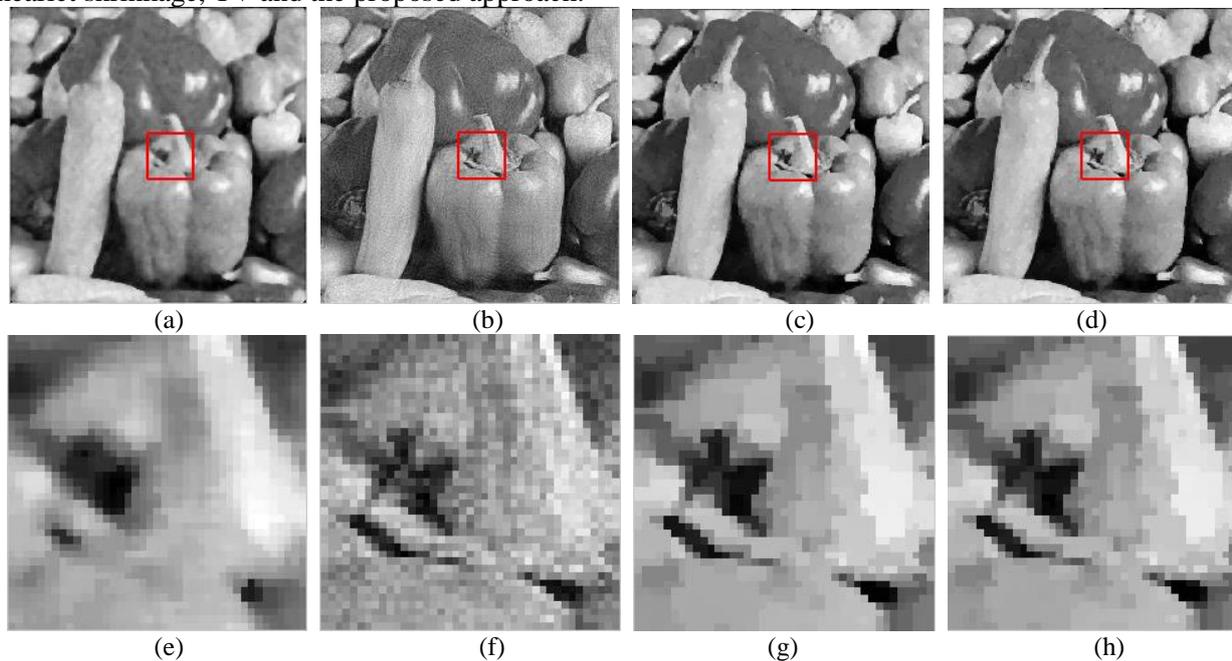
(e)   (f)   (g)   (h)

Fig. 2: Image denoising results,

(a) median filtering, (b) Shearlet shirnkage, (c) TV, (d) the proposed method,

(e)-(f) are the partial enlarged detail of (a)-(d)

Fig. 2 shows that the the denoising result of Shearlet shrinkage has obvious Gibbs artifacts, while that of the proposed method has a significant improvement on the Gibbs artifacts.

Table 1 shows the PSNR, SSIM and Run-time of various denoising methods. The experiment is performed in MATLAB 2014b environment using an Intel Core 2 Duo, 2.1 GHz processor with 8GB of memory, and under Microsoft Windows 7 operating system.

According to the processing results of each estimate, the proposed method attains higher PSNR and SSIM than the other three shemes in all tested situations. Because the proposed method combines the two regularized terms of Shearlet and TV, the processing time is longer than the other three denoising methods.

Table 1.The results comparison of various denoising algorithm.

Noise Variance	Algorithm	PSNR(dB)	SSIM	Run-time(s)
10^2	Median filtering	27.518959	0.604381	0.09
	Shearlet shrinkage	29.932794	0.686005	18.21
	TV	32.720723	0.739078	3.07
	Proposed method	32.787883	0.745161	27.21
20^2	Median filtering	26.280330	0.534854	0.09
	Shearlet shrinkage	25.456839	0.513489	17.08
	TV	29.211581	0.622295	3.48
	Proposed method	29.217259	0.636586	26.55
50^2	Median filtering	22.617847	0.368643	0.09
	Shearlet shrinkage	19.025389	0.252651	17.39
	TV	24.377701	0.442330	4.35
	Proposed method	24.415797	0.456875	27.91
100^2	Median filtering	18.653741	0.238899	0.14
	Shearlet shrinkage	14.891835	0.095202	17.27
	TV	20.374769	0.288740	5.35
	Proposed method	20.428788	0.307070	27.81

5. Conclusions

In this paper, a hybrid regularization model based on ST and TV regularization is proposed. As the total variation is good at holding the edge of image while the ST is good at maintaining the detail of image, the proposed method absorb the merits of the ST and TV, relieving the Gibbs artifacts of ST. The experimental results show that the proposed method can effectively preserve the detail of image. The image restored by the proposed method has better visual effect than ST, MF and TV method, and has higher PSNR and SSIM values.

6. Acknowledgements

7. References

[1]. Li, W., Q. Li, et al. Total variation blind deconvolution employing split Bregman iteration. *Journal of Visual Communication and Image Representation*, 2012. **23**(3): 409-417.

[2]. Elad, M., M.A. Figueiredo, et al. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, 2010. **98**(6): 972-982.

[3]. DENG, C., W. TIAN, et al. Super-resolution reconstruction of approximate sparsity regularized infrared images. *Opt. Precision Eng*, 2014. **22**(6): 1648-1654.

[4]. Donoho, D.L., Compressed sensing. *IEEE Transactions on Information Theory*, 2006. **52**(4): 1289-1306.

[5]. Kutyniok, G. and T. Sauer, Adaptive directional subdivision schemes and shearlet multiresolution analysis. *SIAM Journal on Mathematical Analysis*, 2009. **41**(4): 1436-1471.

[6]. Guo, K. and D. Labate, Optimally sparse multidimensional representation using shearlets. *Siam Journal on Mathematical Analysis*, 2007. **39**(1): 298-318.

[7]. Boyd, S., N. Parikh, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2011. **3**(1): 1-122.

[8]. Beck, A. and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2009. **2**(1): 183-202.