

Article

Directional ℓ_0 Sparse Modeling for Image Stripe Noise Removal

Hong-Xia Dou, Ting-Zhu Huang ^{*}, Liang-Jian Deng, Xi-Le Zhao, and Jie Huang

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, P. R. China

^{*} Corresponding author: tingzhuhuang@126.com

Version December 26, 2017 submitted to MDPI

Abstract: Remote sensing images are often polluted by stripe noise, which leads to negative impact on visual performance. Thus, it is necessary to remove stripe noise for the subsequent applications, *e.g.*, classification, recognition, *etc.* This paper commits to remove the stripe noise to enhance the visual quality of images, in the meanwhile preserves image details of stripe-free regions. Instead of solving the underlying image by various algorithms, we first estimate the stripe noise from the degraded images, then computing the final destriping image by the difference of the known stripe image and the estimated stripe noise. In this paper, we propose a non-convex ℓ_0 sparse model for remote sensing image destriping by taking full consideration of the intrinsically directional and structural priors of stripe noise, as well as the locally continuous property of underlying image. Moreover, the proposed non-convex model is solved by a proximal alternating direction method of multipliers (PADMM) based algorithm and we also give the corresponding theoretical analysis of the proposed algorithm. Extensive experimental results on simulated and real data demonstrate that the proposed method outperforms recently state-of-the-art destriping methods, both visually and quantitatively.

Keywords: Non-convex ℓ_0 sparse model; PADMM based algorithm; Mathematical program with equilibrium constraints (MPEC); Stripe noise removal.

1. Introduction

Stripe noise (all denoted as “stripes” in this paper), which is generally caused by the inconsistence of the detecting element scanning or the influence of the detector moving and temperature changes, *etc.*, are an universal phenomenon in remote sensing images. They will result in a bad influence not only on visual quality but also on subsequent applications in remote sensing images. Therefore, it is necessary to remove stripes and simultaneously maintain the healthy pixels from the degraded images. In general, the stripes have strongly directional and structural information, *e.g.*, pixels normally damaged on row by row or column by column.

Recently, many approaches for destriping problems have been proposed, which may be roughly divided into three categories, mainly including filtering-based methods, statistics-based methods and optimization-based methods. Note that the proposed method belongs to the category of optimization-based methods.

The filtering-based methods, which are easy to obtain the results with various filters, have been widely utilized for remote sensing image destriping, see [1–4]. In [1], Chen *et al.* propose an approach for remote sensing image destriping tasks based on a finite-impulse response filter (FIR) in frequency domain, as well as exhibit the results on the experimental CMODIS data. However, the given method unavoidably leads to ringing and ripple artifacts. In [3], the wavelet analysis and adaptive fourier

zero-frequency amplitude normalization are used for hyperspectral image destriping problems, and this wavelet-based method shows promising ability for both stripes and random noise.

The statistics-based methods are mainly to analyze the distribution of stripes. These approaches hold strong directional characters, to formulate excellent priors for the remote sensing image destriping, e.g., [5–11]. In [7], Weinreb *et al.* introduce a method based on matching empirical distribution functions (EDFs) for GOES-7 data, while the limitations and unstable property are caused by assuming the similarity and regularity among the stripes. To conquer the instability when the stripes are irregular or nonlinear, Rakwatin *et al.* [9] introduce a method, using both histogram-matching algorithm and local least squares fitting, to remove the stripes of Aqua MODIS band 6. In [10], spectral moment matching (SpcMM) method, which can remove various frequencies stripes in a specific band automatically, is proposed for Hyperion image destriping. In addition, Shen *et al.* [11] employ a piece-wise destriping method, which uses correction coefficients of each portion by considering neighbouring normal row, for nonlinear and irregular stripes, but it can not automatically select a threshold to divide the image into different parts.

Recently, the optimization-based methods show superiorities for remote sensing image destriping problems, e.g., [12–23]. The image destriping generally results in an ill-posed problem which fails to obtain a meaningful, stable and unique solution. Therefore, a common strategy for ill-posed problem is to construct a regularization model via investigating the underlying image priors. For the optimization-based methods, they focus on searching and discovering the intrinsically prior knowledge to generate reasonable regularization models. In [17], the authors present a unidirectional total variational (UTV) model for MODIS image stripes removal by fully considering the directional information of stripes. The UTV model is motivated by the classical TV model and the analysis of directional stripes. Chang *et al.* [21] propose an optimization model combining the UTV with sparse priors of stripes applying to denoising and destriping simultaneously. In [22], the authors utilize the split Bregman iteration method with an anisotropic spectral-spatial total variation regularization to remove multispectral image stripes.

In summary, although these optimization-based methods can yield excellent results of removing stripes, there still exists much room to improve. Most of them are implemented only from the perspective of noise removal, but without considering the typical properties of stripes, e.g., directional and structural properties. Even though considering these properties, the formulated sparse destriping models fail to accurately depict the typical properties of stripes, see [24], [25]. Moreover, the designed algorithms for non-convex models, e.g., ℓ_0 sparse model, can not obtain the most precise solution. These motivate us to develop a more reasonable model and effectively design the corresponding algorithm, which theoretically guarantees the convergence, to solve the remote sensing destriping problems.

In this paper, to remove the stripes of remote sensing images, we propose a non-convex sparse model which mainly consists of three sparse priors, including an ℓ_0 sparse prior by fully considering the directional property of stripes (y-axis), an ℓ_1 sparse prior by considering the discontinuity of underlying image (x-axis), and the sparsity of stripes by considering the structural property of stripes. Moreover, we design a PADMM based algorithm to solve the proposed non-convex sparse model. In particular, the convergence to the KKT point of the optimization problem is theoretically proven in the work. Results of several simulated and real images show that the proposed method is superior to recently state-of-the-art destriping methods.

The contributions of this work are summarized as follows

1) Fully considering the latent priors of stripes, we formulate an ℓ_0 sparse model which depicts the intrinsically sparse character more accurately than ℓ_1 sparse model.

2) We solve the non-convex model by a designed PADMM based algorithm which we have given the corresponding theoretical analysis of the proposed algorithm by this paper (see Appendix A).

3) The proposed method, which is less sensitive to related parameters, outperforms recently several state-of-the-art image destriping methods.

⁸⁴ The outline of this paper is organized as follows. In Section 2, we will briefly introduce the related
⁸⁵ work. The proposed model and detailed solving algorithm will be shown in Section 3. In section 4, we
⁸⁶ compare the proposed method with some state-of-the-art remote sensing image destriping methods,
⁸⁷ and discuss the results with different stripes. Finally, conclusions are drawn in Section 5.

⁸⁸ 2. Related work

⁸⁹ 2.1. Destriping problem formulation

The striping effects in remote sensing images mainly make up of additive and multiplicative components [15]. However, the multiplicative stripes can be described as additive case by the logarithm [26]. Thus, many researches more focus on the additive stripes model

$$\mathbf{b}(x, y) = \mathbf{u}(x, y) + \mathbf{s}(x, y) \quad (1)$$

where $\mathbf{b}(x, y)$, $\mathbf{u}(x, y)$ and $\mathbf{s}(x, y)$ separately denote the components of the observe image, the underlying image and stripes at the location (x, y) . For convenience, a matrix-vector form can be written as follows

$$\mathbf{b} = \mathbf{u} + \mathbf{s}, \quad (2)$$

⁹⁰ where \mathbf{b} , \mathbf{u} and $\mathbf{s} \in \mathbb{R}^n$ represent the lexicographical order vectors of $\mathbf{b}(x, y)$, $\mathbf{u}(x, y)$ and $\mathbf{s}(x, y)$,
⁹¹ respectively. The purpose of our work is to estimate the stripes \mathbf{s} , then the underlying image will be
⁹² recovered by the formula of $\mathbf{u} = \mathbf{b} - \mathbf{s}$.

⁹³ 2.2. UTV for remote sensing image destriping

The total variation (TV) model, which is first proposed by Rudin, Oshaer and Fatemi (ROF) [27], has shown powerful ability in many image applications, e.g., image unmixing [28], image deblurring [29], image inpainting [30], etc. It has the following form

$$E(\mathbf{u}) = \frac{1}{2} \int_{\Omega} ||\mathbf{u} - \mathbf{b}||^2 + \lambda TV(\mathbf{u}), \quad (3)$$

where λ is a positive regularization parameter, and $TV(\mathbf{u})$ represents the regularization expressed as

$$TV(\mathbf{u}) = \int_{\Omega} |\nabla \mathbf{u}| = \int_{\Omega} \sqrt{\left(\frac{\partial \mathbf{u}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{u}}{\partial y}\right)^2} dx dy. \quad (4)$$

⁹⁴ In many approaches, $\mathbf{s}(x, y)$ is usually regarded as constant in a given line. Although this
⁹⁵ assumption has shown stability in MOS-B, it fails in MODIS. Not only predominant nonlinear effects,
⁹⁶ but also the data quality of random stripes have been obtained in many emissive bands. Thus, more
⁹⁷ realistic assumptions are introduced to design an efficient destriping method.

Without loss the generality, we can assume that the stripes are along the vertical direction (y-axis). Fully considering the directional property of stripes, the authors in [17] consider the following relation

$$\left| \frac{\partial \mathbf{s}(x, y)}{\partial y} \right| \ll \left| \frac{\partial \mathbf{s}(x, y)}{\partial x} \right|, \quad (5)$$

where we denote y-axis is along stripes direction, and x-axis is across stripes direction. By the relation in Eq. (5), we have

$$\int_{\Omega} \left| \frac{\partial \mathbf{s}(x, y)}{\partial y} \right| dx dy \ll \int_{\Omega} \left| \frac{\partial \mathbf{s}(x, y)}{\partial x} \right| dx dy, \quad (6)$$

which means

$$TV_y(\mathbf{s}) \ll TV_x(\mathbf{s}) \quad (7)$$

where TV_x and TV_y are horizontal and vertical variations, respectively. The authors in [17] encourage the robustness of stripes removal by minimizing the unidirectional total variation (UTV) model as follows

$$E(\mathbf{u}) = TV_y(\mathbf{u} - \mathbf{b}) + \lambda TV_x(\mathbf{u}), \quad (8)$$

which can be solved by Euler-Lagrange equation based algorithm.

In [17], the UTV model can effectively deal with remote sensing image destriping problems, which has been demonstrated holding promising ability on Aqua and Terra MODIS data. Although TV model preserves image edges well, it can not accurately depict the specifically directional property of stripes, and leads to undesired results. The UTV model that involves unidirectional constraint can remove stripes excellently in the meanwhile not destroy the underlying image details. Inspired by the UTV model, we fully consider the intrinsically directional and structural priors of stripes and the continuous property of the underlying image. Finally, we form a unidirectional and sparse based optimization model.

3. The proposed method

Combining the stripes model (2), we will give the proposed optimal model with unidirectional prior motivated by the extension of the UTV model. In what follows, the detailed explanations of the proposed model and the corresponding solving algorithm will be exhibited.

3.1. The proposed model

3.1.1. Local smoothness along stripe direction

The stripes of remote sensing images are generally appeared with column-by-column (y-axis) or row-by-row (x-axis), without loss of generality, we view all stripes as column-by-column case to formulate the finally directional model¹. Considering the smoothness within the stripes, the difference between adjacent pixels is quite small, or even close to zero, thus we generally use sparse prior for this character along the stripe direction (y-axis). The first regularization for the difference within the stripes is given as follows

$$R_1 = \|\nabla_y \mathbf{s}\|_0, \quad (9)$$

where ∇_y is a partial difference operator along stripe direction². Comparing with some popular sparse measures, e.g., ℓ_1 -norm and ℓ_p -norm ($0 < p < 1$), the ℓ_0 -norm that stands for the number of non-zero elements of a vector is the most accurate measure to depict sparse property, thus here we employ ℓ_0 -norm to describe the sparsity of $\nabla_y \mathbf{s}$. Although this term will lead to the non-convexity of the proposed model, we utilize the designed PADMM based algorithm to guarantee the solution converging to the KKT point.

3.1.2. Local continuity of the underlying image

In general, the underlying image \mathbf{u} along x-axis is viewed as being continuous. When adding column-by-column stripes \mathbf{s} to the underlying image, the local continuity of \mathbf{u} is broken, which means that we should force $\nabla_x \mathbf{u}$ being small to keep the continuity of \mathbf{u} . By this assumption and the relation $\mathbf{u} = \mathbf{b} - \mathbf{s}$, we utilize the following ℓ_1 -norm regularization to describe the local continuity of the underlying image

$$R_2 = \|\nabla_x(\mathbf{b} - \mathbf{s})\|_1, \quad (10)$$

¹ The row-by-row stripes can be easily rotated to column-by-column stripes to fit in the proposed model.

² $\nabla_y \mathbf{u}$ represents the vector form of $\nabla_y U$ where U is a 2D image. The similar meaning is $\nabla_x \mathbf{u}$.

120 where ∇_x represents the difference operator in the across-stripe direction. Note that this term is
 121 actually the second term of the UTV model (8).

122 3.1.3. Global sparsity of stripes

In many destriping approaches, *e.g.*, [24,25,31,32], the stripes can be naturally viewed as being sparse when the stripes are not heavy. Inspired by their excellent works, here we take the ℓ_1 -norm to depict the sparsity of stripes, see as follows

$$R_3 = \|\mathbf{s}\|_1. \quad (11)$$

123 Even though the stripes are heavy, this sparse term (11) is still necessary to retain, since it can
 124 effectively avoid the undesired effect and keep the robustness of the proposed method (see more
 125 discussion from the results section).

Combining the above three regularization terms, we finally formulate the ℓ_0 sparse model for remote sensing image destriping,

$$\min_{\mathbf{s}} \|\nabla_y \mathbf{s}\|_0 + \mu \|\mathbf{s}\|_1 + \lambda \|\nabla_x(\mathbf{b} - \mathbf{s})\|_1, \quad (12)$$

126 where μ and λ are two positive parameters.

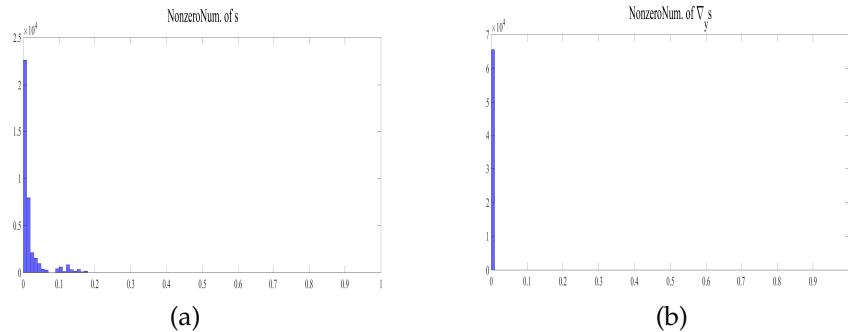


Figure 1. The number of nonzero of \mathbf{s} (a) and $\nabla_y \mathbf{s}$ (b), where \mathbf{s} is estimated from a real image example (see Fig. 4) by the method [24]. It is clear that $\nabla_y \mathbf{s}$ is more sparse than \mathbf{s} .

127 Note that, the proposed model (12) is similar as the model in [24], since they both employ the
 128 directional property of stripes. However, there still exists an important difference that the model in [24]
 129 enforces ℓ_1 norm to $\nabla_y \mathbf{s}$ and ℓ_0 norm to \mathbf{s} whereas our model enforces ℓ_1 norm to \mathbf{s} and ℓ_0 norm to $\nabla_y \mathbf{s}$.
 130 It can be seen that our model is more reasonable than the model in [24], because $\nabla_y \mathbf{s}$ is significantly
 131 more sparse than \mathbf{s} . For instance, Fig. 1 shows the number of non-zeros of \mathbf{s} (Fig. 1(a)) and $\nabla_y \mathbf{s}$ (Fig.
 132 1(b)), where \mathbf{s} is estimated from a real image example by the method [24], it is clear that $\nabla_y \mathbf{s}$ is almost
 133 all around 0, whereas \mathbf{s} is not. The ℓ_0 norm is the best way to depict sparsity, thus our model which
 134 enforces ℓ_0 norm to $\nabla_y \mathbf{s}$.

135 In what follows, we will exhibit how to solve the proposed non-convex sparse model by
 136 introducing the PADMM based algorithm, as well as give the theoretical analysis of the convergence.

137 3.2. The solution

138 Before solving the proposed model (12), we first present an excellent work, *i.e.*, Mathematical
 139 program with equilibrium constraints (MPEC) [31], to transfer the non-convex ℓ_0 regularization term
 140 to the other equivalent one.

141 **Equivalent MPEC reformulation:** For the non-convex ℓ_0 regularization term, there exist many
 approaches to approximate it, *e.g.*, ℓ_1 -norm [33], the logarithm function [34] or the penalty

decomposition algorithm (PDA) [35]. In this work, we are inspired by a recently elegant work, *i.e.*, MPEC, to transfer the ℓ_0 regularization term to an equivalent problem, so that we can design a PADMM based algorithm to efficiently solve the equivalent model, in the meanwhile theoretically guarantee the convergence. **Lemma:** [MPEC equation [31]] For any given $\mathbf{w} \in \mathbb{R}^n$, it holds that

$$\|\mathbf{w}\|_0 = \min_{\mathbf{v} \leq \mathbf{1}} \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle, \text{ s.t. } \mathbf{v} \odot |\mathbf{w}| = 0, \quad (13)$$

¹⁴² and $\mathbf{v}^* = \mathbf{1} - \text{sign}(|\mathbf{w}|)$ is the unique optimal solution of the minimization problem (13).

¹⁴³ **Proof:** See details in [31].

From Lemma 3.2, the ℓ_0 -norm sparse optimization model in Eq. (12) is equivalent to

$$\begin{aligned} & \min_{\mathbf{v} \leq \mathbf{1}, \mathbf{s}} \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle + \mu \|\mathbf{s}\|_1 + \lambda \|\nabla_x(\mathbf{b} - \mathbf{s})\|_1, \\ & \text{s.t. } \mathbf{v} \odot |\nabla_y \mathbf{s}| = 0, \end{aligned} \quad (14)$$

¹⁴⁴ where \odot denotes the elementwise product. According to the analysis of [31], if \mathbf{s}^* is the globally
¹⁴⁵ optimal solution of Eq. (12), then $(\mathbf{s}^*, \mathbf{1} - \text{sign}(|\nabla_y \mathbf{s}^*|))$ is the unique global minimizer of Eq. (14).

¹⁴⁶ Note that the Eq. (14) is still a non-convex problem, and the non-convexity is only caused by the
¹⁴⁷ constraint $\mathbf{v} \odot |\nabla_y \mathbf{s}| = 0$. However, this problem (14) is similar to the main problem in [31], which
¹⁴⁸ is efficiently solved by a PADMM³ based algorithm that theoretically guarantees the convergence.
¹⁴⁹ Therefore, we employ the designed PADMM based algorithm to solve the resulted problem (14), as
¹⁵⁰ well as give the theoretical analysis of the convergence.

¹⁵¹ In the following, we will use the PADMM based algorithm to solve the optimization problem (14).

¹⁵² 3.3. PADMM based Algorithm

Considering the non-smooth ℓ_1 terms in problem (14), we take the following variable substitutions to get the new optimization problem,

$$\begin{aligned} & \min_{\mathbf{v} \leq \mathbf{1}, \mathbf{s}} \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle + \mu \|\mathbf{z}\|_1 + \lambda \|\mathbf{w}\|_1, \\ & \text{s.t. } \mathbf{v} \odot |\mathbf{h}| = 0, \nabla_y \mathbf{s} = \mathbf{h}, \mathbf{s} = \mathbf{z}, \nabla_x(\mathbf{b} - \mathbf{s}) = \mathbf{w}, \end{aligned} \quad (15)$$

with the auxiliary variables $\mathbf{h}, \mathbf{z}, \mathbf{w} \in \mathbb{R}^n$. The augmented Lagrangian function \mathcal{L} of Eq. (15) is as follows

$$\begin{aligned} & \mathcal{L}(\mathbf{h}, \mathbf{z}, \mathbf{w}, \mathbf{v}, \mathbf{s}, \pi_1, \pi_2, \pi_3, \pi_4, \beta_1, \beta_2, \beta_3, \beta_4) \\ & = \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle + \mu \|\mathbf{z}\|_1 + \lambda \|\mathbf{w}\|_1 + \langle \nabla_y \mathbf{s} - \mathbf{h}, \pi_1 \rangle \\ & + \frac{\beta_1}{2} \|\nabla_y \mathbf{s} - \mathbf{h}\|_2^2 + \langle \mathbf{s} - \mathbf{z}, \pi_2 \rangle + \frac{\beta_2}{2} \|\mathbf{s} - \mathbf{z}\|_2^2 \\ & + \langle \nabla_x(\mathbf{b} - \mathbf{s}) - \mathbf{w}, \pi_3 \rangle + \frac{\beta_3}{2} \|\nabla_x(\mathbf{b} - \mathbf{s}) - \mathbf{w}\|_2^2 \\ & + \langle \mathbf{v} \odot |\mathbf{h}|, \pi_4 \rangle + \frac{\beta_4}{2} \|\mathbf{v} \odot |\mathbf{h}|\|_2^2, \end{aligned} \quad (16)$$

¹⁵³ where π_1, π_2, π_3 and π_4 are Lagrange multipliers, and $\beta_1, \beta_2, \beta_3$ and β_4 are positive parameters. The
¹⁵⁴ minimization problem (16) can be solved by the PADMM based algorithm. Next, we discuss the
¹⁵⁵ solution of each subproblem.

³ Actually, PADMM method is an extended version of ADMM method, which has been applied to many image applications, *e.g.*, image deblurring [36], image denoising [37], tensor completion [38], *etc.*

1) The \mathbf{h} -subproblem can be written to the minimized problem as follows

$$\begin{aligned} \min_{\mathbf{h}} & \langle \nabla_y \mathbf{s}^k - \mathbf{h}, \boldsymbol{\pi}_1^k \rangle + \frac{\beta_1}{2} \|\nabla_y \mathbf{s}^k - \mathbf{h}\|_2^2 \\ & + \langle \mathbf{v}^k \odot |\mathbf{h}|, \boldsymbol{\pi}_4^k \rangle + \frac{\beta_4}{2} \|\mathbf{v}^k \odot |\mathbf{h}|\|_2^2. \end{aligned} \quad (17)$$

Now, let h_i is the i-th pixel of \mathbf{h} and we discuss two situations when the element $h_i \neq 0$, if $h_i > 0$,

$$h_i = \frac{(\beta_1(\nabla_y \mathbf{s})_i + (\boldsymbol{\pi}_1^k)_i) - (\boldsymbol{\pi}_4^k)_i \odot (\mathbf{v}^k)_i}{\beta_1 + \beta_4(\mathbf{v}^k)_i \odot (\mathbf{v}^k)_i}, \quad (18)$$

if $h_i < 0$,

$$h_i = (-1) \frac{-(\beta_1(\nabla_y \mathbf{s})_i + (\boldsymbol{\pi}_1^k)_i) - (\boldsymbol{\pi}_4^k)_i \odot (\mathbf{v}^k)_i}{\beta_1 + \beta_4(\mathbf{v}^k)_i \odot (\mathbf{v}^k)_i}. \quad (19)$$

In summary, the \mathbf{h} -subproblem has the closed-form solution as follows

$$\mathbf{h}^{k+1} = \text{sign}(\mathbf{q}^k) * \frac{|\mathbf{q}^k| - \boldsymbol{\pi}_4^k \odot \mathbf{v}^k}{\beta_1 + \beta_4 \mathbf{v}^k \odot \mathbf{v}^k}, \quad (20)$$

where $\mathbf{q}^k = \beta_1 \nabla_y \mathbf{s}^k + \boldsymbol{\pi}_1^k$.

2) The \mathbf{z} -subproblem is given as follows

$$\min_{\mathbf{z}} \mu \|\mathbf{z}\|_1 + \langle \mathbf{s}^k - \mathbf{z}, \boldsymbol{\pi}_2^k \rangle + \frac{\beta_2}{2} \|\mathbf{s}^k - \mathbf{z}\|_2^2, \quad (21)$$

which has the closed-form solution by soft-thresholding strategy [39]

$$\mathbf{z}^{k+1} = \text{Shrink}(\mathbf{s}^k + \frac{\boldsymbol{\pi}_2^k}{\beta_2}, \frac{\mu}{\beta_2}), \quad (22)$$

where $\text{Shrink}(\mathbf{a}, T) = \text{sign}(\mathbf{a}) * \max(|\mathbf{a} - T|, 0)$.

3) Similar to \mathbf{z} -subproblem, \mathbf{w} -subproblem is written as follows

$$\min_{\mathbf{w}} \lambda \|\mathbf{w}\|_1 + \frac{\beta_3}{2} \|\nabla_x(\mathbf{b} - \mathbf{s}^k) - \mathbf{w} + \frac{\boldsymbol{\pi}_3^k}{\beta_3}\|_2^2. \quad (23)$$

The problem (23) has the following closed-form solution by the soft-shrinkage formulation,

$$\mathbf{w}^{k+1} = \text{Shrink}(\mathbf{q}^k, \frac{\lambda}{\beta_3}), \quad (24)$$

where $\mathbf{q}^k = \nabla_x(\mathbf{b} - \mathbf{s}^k) + \frac{\boldsymbol{\pi}_3^k}{\beta_3}$.

4) The \mathbf{v} -subproblem can be written as follows

$$\min_{0 \leq \mathbf{v} \leq \mathbf{1}} \langle \mathbf{v}, \mathbf{c}^k \rangle + \frac{\beta_4}{2} \|\mathbf{v} \odot |\mathbf{h}^{k+1}|\|_2^2, \quad (25)$$

where $\mathbf{c}^k = \mathbf{1} - \boldsymbol{\pi}_4^k \odot |\mathbf{h}^{k+1}|$. Combining with the constraint $\mathbf{0} \leq \mathbf{v} \leq \mathbf{1}$, it has the closed-form solution,

$$\mathbf{v}^{k+1} = \min(\mathbf{1}, \max(\mathbf{0}, \frac{-\mathbf{c}^k}{\beta_4 |\mathbf{h}^{k+1}| \odot |\mathbf{h}^{k+1}|})). \quad (26)$$

Algorithm 1: The algorithm for model (12)

Input: The observed image \mathbf{b} (with stripes), the parameters $\lambda, \mu, \beta_i, i = 1, 2, 3, 4$,
 the constant $\kappa \in (0, \frac{1}{\beta_1 \|\nabla_y^T\|^2 + \beta_2 + \beta_3 \|\nabla_x^T\|^2})$, the maximum number of iterations M_{iter} ,
 and the calculation accuracy tol .

Output: The stripes \mathbf{s}

Initialize:

1) $k \leftarrow 0, \mathbf{v}^0 \leftarrow \mathbf{1}, \mathbf{s}^0 \leftarrow \mathbf{b}, rho \leftarrow 1$

While $rho > tol$ and $k < M_{iter}$

2) $k \leftarrow k + 1$

3) Solve \mathbf{h}^k by Eq. (20)

4) Solve \mathbf{z}^k by Eq. (22)

5) Solve \mathbf{w}^k by Eq. (24)

6) Solve \mathbf{v}^k by Eq. (26)

7) Solve \mathbf{s}^k by Eq. (28)

8) Update the multipliers $\pi_i, i = 1, 2, 3, 4$, by Eq. (29)

9) Calculate the error

$$rho = \|\nabla_y \mathbf{s}^{k+1} - \mathbf{h}^{k+1}\|_2 + \|\mathbf{s}^{k+1} - \mathbf{z}^{k+1}\|_2 + \|\nabla_x(\mathbf{b} - \mathbf{s}^{k+1}) - \mathbf{w}^{k+1}\|_2 + \|\mathbf{v}^{k+1} \odot |\mathbf{h}^{k+1}| \|_2.$$

Endwhile

5) Here, PADMM based algorithm needs to introduce an extra convex proximal term $\frac{1}{2} \|\mathbf{s} - \mathbf{s}^k\|_{\mathbf{D}}^2$, which is defined as $\|\mathbf{x}\|_{\mathbf{D}}^2 = \mathbf{x}^T \mathbf{D} \mathbf{x}$, and D is a symmetric positive definite matrix. The \mathbf{s} -subproblem becomes a strong convex optimization problem as

$$\begin{aligned} \min_{\mathbf{s}} \quad & \langle \nabla_y \mathbf{s} - \mathbf{h}^{k+1}, \pi_1^k \rangle + \frac{\beta_1}{2} \|\nabla_y \mathbf{s} - \mathbf{h}^{k+1}\|_2^2 \\ & + \langle \mathbf{s} - \mathbf{z}^{k+1}, \pi_2^k \rangle + \frac{\beta_2}{2} \|\mathbf{s} - \mathbf{z}^{k+1}\|_2^2 \\ & + \langle \nabla_x(\mathbf{b} - \mathbf{s}) - \mathbf{w}^{k+1}, \pi_3^k \rangle \\ & + \frac{\beta_3}{2} \|\nabla_x(\mathbf{b} - \mathbf{s}) - \mathbf{w}^{k+1}\|_2^2 + \frac{1}{2} \|\mathbf{s} - \mathbf{s}^k\|_{\mathbf{D}}^2, \end{aligned} \quad (27)$$

where

$$\mathbf{D} = \frac{1}{\kappa} \mathbf{I} - (\beta_1 \nabla_y^T \nabla_y + \beta_2 + \beta_3 \nabla_x^T \nabla_x),$$

$$\kappa \in \left(0, \frac{1}{\beta_1 \|\nabla_y\|_2^2 + \beta_2 + \beta_3 \|\nabla_x\|_2^2} \right).$$

Then, Eq. (27) will be equivalent to:

$$\mathbf{s}^{k+1} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{s} - \mathbf{g}^k\|_2^2, \quad (28)$$

160 where $\mathbf{g}^k = \mathbf{s}^k - \kappa [\beta_1 (\nabla_y \mathbf{s}^k - \mathbf{h}^{k+1}) + \beta_2 (\mathbf{s}^k - \mathbf{z}^{k+1}) - \beta_3 \nabla_x^T (\nabla_x \mathbf{b} - \nabla_x \mathbf{s}^k - \mathbf{w}^{k+1})]$.

6) Finally, we update the Lagrangian multipliers by

$$\begin{aligned} \pi_1^{k+1} &= \pi_1^k + \beta_1 (\nabla_y \mathbf{s}^{k+1} - \mathbf{h}^{k+1}), \\ \pi_2^{k+1} &= \pi_2^k + \beta_2 (\mathbf{s}^{k+1} - \mathbf{z}^{k+1}), \\ \pi_3^{k+1} &= \pi_3^k + \beta_3 (\nabla_x(\mathbf{b} - \mathbf{s}^{k+1}) - \mathbf{w}^{k+1}), \\ \pi_4^{k+1} &= \pi_4^k + \beta_4 (\mathbf{v}^{k+1} \odot |\mathbf{h}^{k+1}|). \end{aligned} \quad (29)$$

Combining steps 1) to 6), we formulate the final algorithm to iteratively solve the proposed ℓ_0 sparse model (12). In particular, the subproblems all have the closed-form solutions to ensure the accuracy of the algorithm. Finally, the solving process has been summarized in Algorithm 1.

In Algorithm 1, λ , μ , β_1 , β_2 , β_3 , β_4 are some pre-defined parameters, tol and M_{iter} represent the positive tolerance value and the maximum iterations, respectively. In this work, we set $tol = 1/255$ and $M_{iter} = 10^3$. In the following, we discuss the convergence of the Algorithm 1.

4. Experiment results

In this section, we compare the proposed method with several state-of-the-art destriping methods, including the wavelet Fourier adaptive filter (WFAF) [3], the statistical linear destriping (SLD) [26], the unidirectional total variation model (UTV) [17], the global sparsity and local variational (GSLV) [24], and the Low-Rank Single-Image Decomposition (LRSID) [25], on both simulated and real remote sensing data. The codes of these methods, except the GSLV method, are available⁴. As suggested in [25], we utilize the same periodic/nonperiodic stripes function adding stripes intensity [0, 255] to the underlying images. By the similar measure as in [25], the degraded images were normalized between [0, 1]. All experiments are conducted in MATLAB (R2016a) on a desktop with 16Gb RAM and Inter(R) Core(TM) CPU i5-4590: @3.30GHz.

To evaluate the effects of different destriping methods, we will compare several qualitative and quantitative assessments. On the qualitative aspect, we show the visual results, the mean cross-track profile and the power spectrum of different methods. We also employ some acknowledged indexes, *i.e.*, peak signal-to-noise ratio (PSNR)[40], structural similarity index (SSIM) [40] and the relative error (ReErr), to evaluate the performance of different approaches. The ReErr formula is as follows,

$$\text{ReErr} = \frac{\|\mathbf{s}_{\text{added}} - \mathbf{s}_{\text{restored}}\|_2}{\|\mathbf{s}_{\text{added}}\|_2},$$

where the $\mathbf{s}_{\text{added}}$ and $\mathbf{s}_{\text{restored}}$ represent the added stripes and restored stripes by different methods, respectively. Then, we will discuss how to select parameters. We note that we test the comparing methods according to the default or suggested parameters in their papers and codes.

4.1. Simulated experiments

In simulated experiments, the stripes with periodic (Per) and nonperiodic (NonPer) noise are mainly determined by “Intensity” and r . Here, the “Intensity” means the added absolute value of the stripe scope, and the r represents the stripes ratio level within the remote sensing images. For convenience to compare, different stripes added to remote sensing images will be denoted as a vector with three elements, *e.g.*, (Per, 10, 0.2) which represents the periodic stripes, the “Intensity” 10 and stripes ratio 0.2.

We take six experimental images, which the first, second, third and sixth examples are available on the website⁵, and the forth and fifth examples are available on the website⁶, to test the performance of different methods. To compare these methods clearly, we zoom in destriping details on the bottom left or bottom right of the image.

1) Periodic Stripes. For the periodic stripes case, we only take one example, *i.e.*, the first column of Fig. 2 with added stripes (Per, 10, 0.2), to compare the performance. Almost of all existing methods performs quite excellent due to the simple structures of periodic stripes. The first column of Fig. 2 also demonstrates the consistent conclusion that all comparing approaches remove the periodic stripes and well preserve the image details of stripe-free regions.

⁴ <http://www.escience.cn/people/changyi/codes.html>.

⁵ DigitalGlobe with <http://www.digitalglobe.com/product-samples>.

⁶ MODIS data with <https://ladsweb.nascom.nasa.gov/>

Table 1. The ReErr results between s_{added} and $s_{restored}$ for different methods

images	(a)	(b)	(c)	(d)	(e)	(f)
WFAF	0.1588	0.2828	0.2519	0.2468	0.2386	0.2574
SLD	0.0874	0.1670	0.1723	0.1664	0.1330	0.1346
UTV	0.0831	0.1542	0.2371	0.1818	0.1314	0.1375
GSLV	0.0867	0.1030	0.2385	0.1926	0.0912	0.1654
LRSID	0.0917	0.1884	0.2731	0.2125	0.1450	0.1897
Ours	0.0193	0.0693	0.0365	0.0892	0.0304	0.0813

2) Nonperiodic Stripes. For the non periodic stripes case, we test five remote sensing images from the second column to the end column in Fig. 2 with added stripes (NonPer, 100, 0.6), (NonPer, 50, 0.2), (NonPer, 60, 0.4), (NonPer, 100, 0.4) and (NonPer, 50, 0.6), respectively. Then, we display the destriping results of WFAF, SLD, UTV, GSLV, LRSID and the proposed method for different simulated remote sensing images starting from third row to the end row. See the visual results of the second column, the WFAF method has a obvious black line and changes the intensity contrast of the underlying image significantly. Although the other comparing methods can remove stripes, some regions change the intensity contrast of the underlying image on the left and the right parts, and the proposed method shows a good performance. Then, from the third to sixth examples, we can clearly observe the residual stripes and blurring effects resulted by the others comparing methods. Moreover, our method not only removes stripes completely but also preserves image details well. From Fig. 3, we display the smaller patches of Fig. 2 for visual quality comparisons, and ours results have a better performance than the others.

Fig. 4 shows the estimated stripes based on Fig. 2. From Fig. 4, we know that the other comparing methods may generate blurring effect and change intensity contrast. Meanwhile, the estimated stripes of the proposed method neither eliminate image structures nor bring in blurring effects for both periodic and nonperiodic stripes cases.

In Fig. 5, we show the difference/residuals between the added stripes and restored ones. Although ours results have some residuals, the proposed method shows a better performance than the others compared methods. Moreover, we utilize the ReErr results to show the differences/residuals of Fig. 5 in quantitative aspect. The ReErr results have shown in Table 1. From Table 1, our results outperform than the other compared methods.

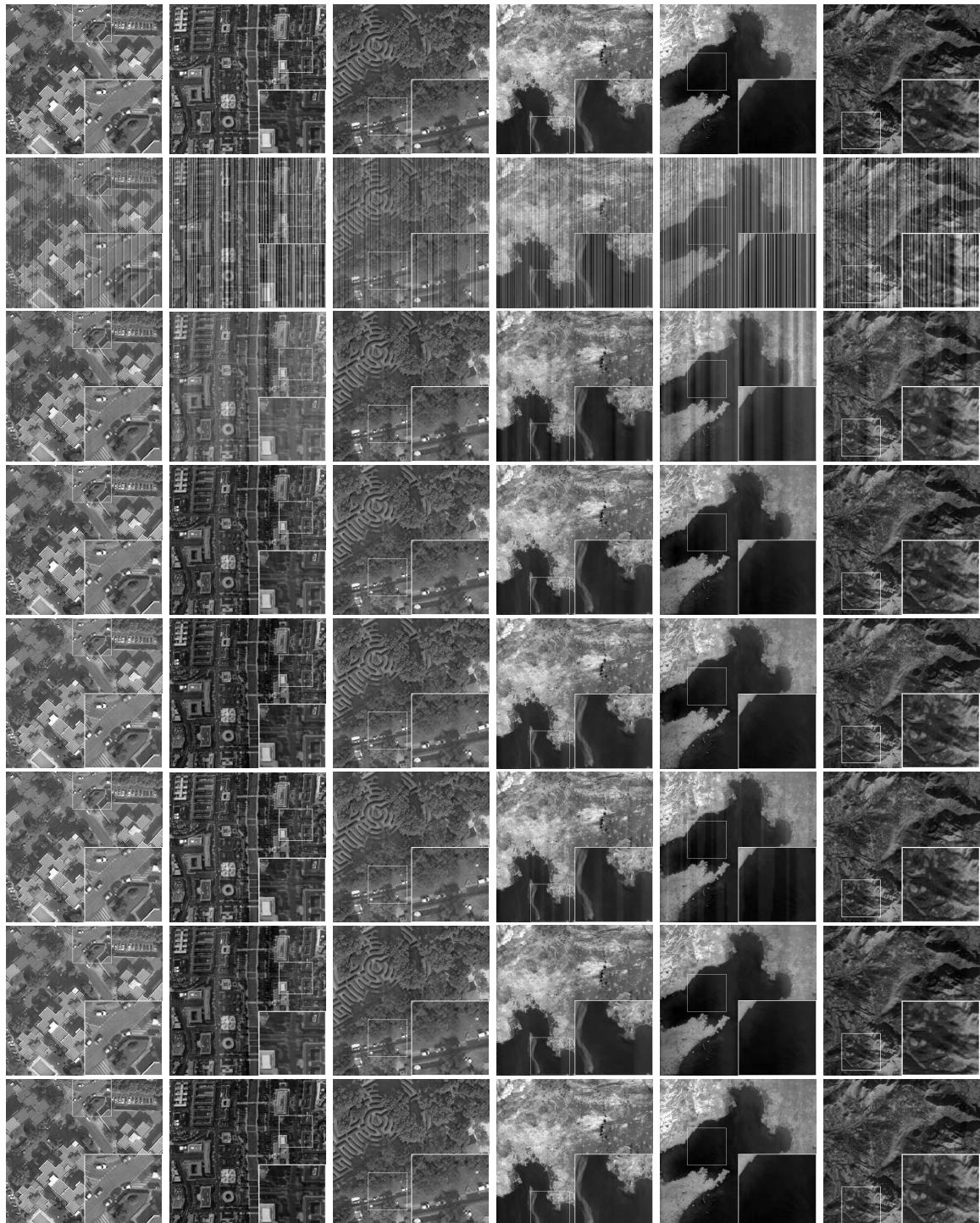


Figure 2. The visual results of different simulated images. From top to bottom: underlying images, degraded images, the destriping results of WFAF, SLD, UTV, GSLV, LRSID and Ours. The degraded images in the second row are respectively added the stripes (from left to right): (Per, 10, 0.2), (NonPer, 100, 0.4), (NonPer, 50, 0.2), (NonPer, 60, 0.4), (NonPer, 100, 0.4) and (NonPer, 50, 0.6). Readers are recommended to zoom in all figures for better visibility.

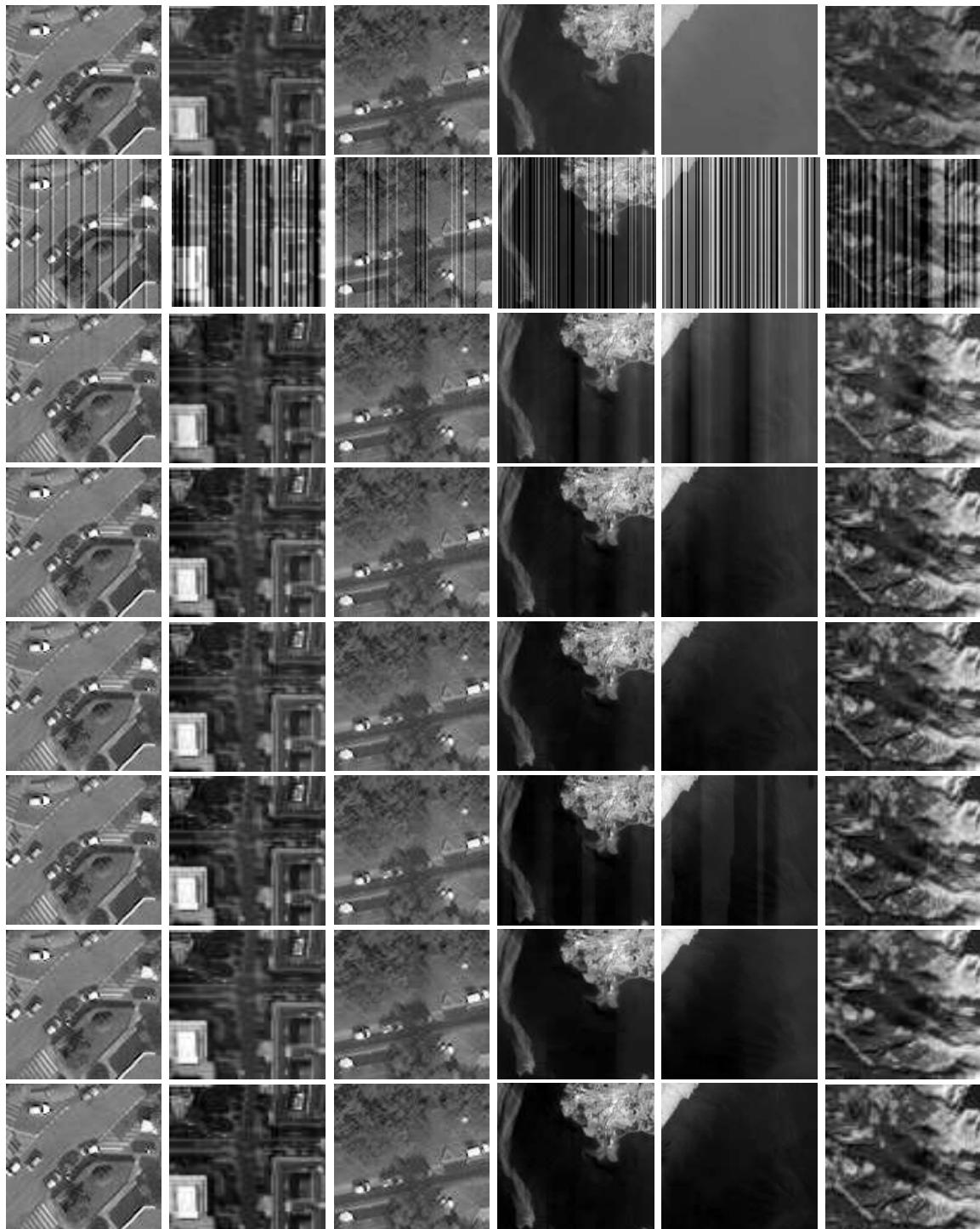


Figure 3. The zoom results of different simulated images in Fig. 2. From top to bottom: zoom of the underlying images, the degrded images, the destriping results of WFAF, SLD, UTV, GSLV, LRSID and Ours. Note that the levels of stripes are same as Fig. 2.

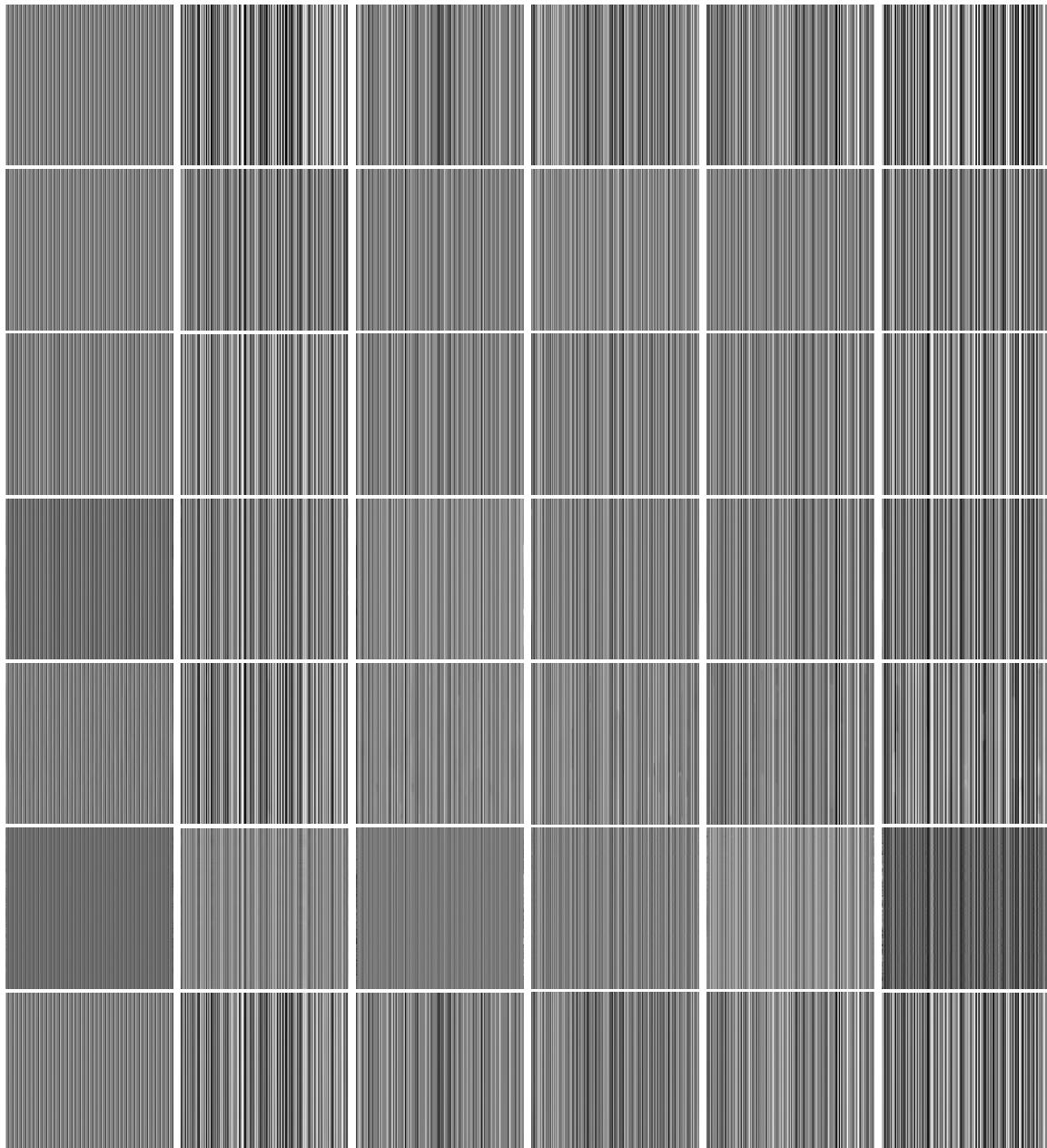


Figure 4. The stripes s of different simulated images in Fig. 2. From top to bottom: the added stripes on the underlying image, the extracted stripe components of WFAF, SLD, UTV, GSLV, LRSID and Ours. Note that the levels of stripes are same as Fig. 2.

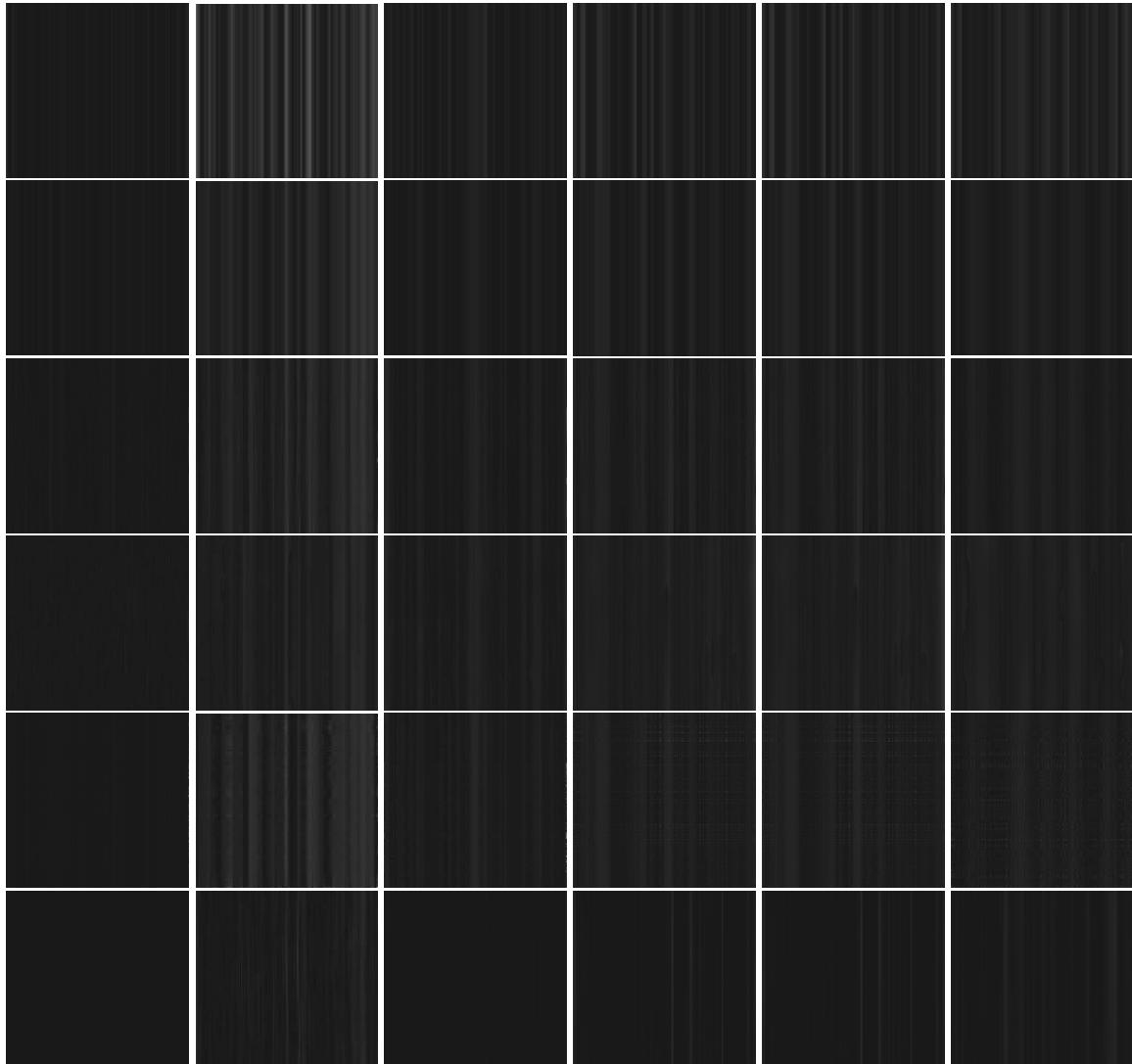


Figure 5. The difference of the added stripes and restored ones. From top to bottom: the difference results of WFAF, SLD, UTV, GSLV, LRSID and Ours. Note that the levels of stripes are same as Fig. 2.

218 **2) Averagely quantitative performance on 32 test images.** To quantitatively test robustness
 219 and effectiveness of the proposed method, Table 2 and Table 3 report the averagely quantitative
 220 comparisons on 32 remote sensing images, which are randomly selected from three websites⁷. In the
 221 tables, the best PSNR and SSIM results have been identified in bold. Especially, we compare these
 222 methods on 32 remote sensing images with fixed parameters for each method.

223 Table 2 shows the PSNR and SSIM results on periodic stripes with different stripe levels. Although
 224 variance of PSNR is not the smallest, the SSIM of the proposed method holds the best performance,
 225 and SSIM is an important index to indicate stability on structural similarity of one method. Moreover,
 226 our method has the best mean value results of PSNR and SSIM which show the significant advantages
 227 than the other comparing methods.

⁷ 1) “DigitalGlobe” with <http://www.digitalglobe.com/product-samples>. 2) some subimages of “hyperspectral image of Washington DC Mall” with <https://engineering.purdue.edu/~biehl/MultiSpec/>. 3) “MODIS” data with <https://ladsweb.nascom.nasa.gov/>

Table 2. The mean value of PSNR and SSIM of 32 images with periodic noise

Intensity		Intensity=10		Intensity=50		Intensity=100	
	Ratio	r=0.2	r=0.6	r=0.2	r=0.6	r=0.2	r=0.6
PSNR	WFAF	41.400±3.601	41.702±3.870	37.160±1.975	37.553±1.975	32.196±1.457	32.501±1.732
	SLD	42.037±2.927	41.048±2.909	41.710±2.930	41.957±2.928	40.614±2.549	41.644±2.836
	UTV	42.030±3.229	41.032±2.886	40.920±2.773	43.086±2.298	41.470±3.385	41.058±3.299
	GSLV	42.552±2.955	42.630±2.886	42.202±3.058	43.533±2.856	43.431±3.091	43.801±2.705
	LRSID	43.948±2.104	42.775±2.010	42.308±2.169	44.548±1.976	43.779±2.500	44.035±2.014
SSIM	Ours	52.918±4.074	49.497±3.956	52.853±4.910	49.212±4.390	52.854±4.902	49.182±4.368
	WFAF	0.9934±0.0058	0.9936±0.0062	0.9887±0.0084	0.9905±0.0078	0.9818±0.0103	0.9847±0.0085
	SLD	0.9966±0.0029	0.9966±0.0029	0.9965±0.0031	0.9965±0.0032	0.9962±0.0033	0.9964±0.0037
	UTV	0.9959±0.0027	0.9959±0.0027	0.9911±0.0025	0.9928±0.0023	0.9954±0.0024	0.9937±0.0076
	GSLV	0.9991±0.0077	0.9968±0.0076	0.9916±0.0079	0.9903±0.0082	0.9966±0.0085	0.9969±0.0053
	LRSID	0.9990±0.0107	0.9945±0.0056	0.9932±0.0044	0.9947±0.0032	0.9936±0.0047	0.9957±0.0031
Ours	Ours	0.9994±0.0007	0.9987±0.0011	0.9994±0.0013	0.9986±0.0016	0.9994±0.0062	0.9986±0.0019

Table 3. The mean value of PSNR and SSIM of 32 images with nonperiodic noise

Intensity		Intensity=10		Intensity=50		Intensity=100	
	Ratio	r=0.2	r=0.6	r=0.2	r=0.6	r=0.2	r=0.6
PSNR	WFAF	40.971±2.523	39.372±2.249	30.536±1.508	37.609±2.263	24.849±1.573	22.594±1.541
	SLD	41.476±2.592	40.935±2.201	35.964±1.510	42.007±3.020	30.963±1.414	28.403±1.729
	UTV	41.153±2.880	38.615±2.041	35.648±1.527	42.505±3.010	31.055±4.687	31.599±2.578
	GSLV	42.282±2.359	39.018±1.654	41.985±1.239	39.838±2.903	36.184±1.399	35.408±2.472
	LRSID	42.672±1.418	39.034±1.302	42.814±1.349	40.497±2.024	37.779±1.212	33.559±1.132
SSIM	Ours	48.801±3.985	44.700±3.784	49.057±4.791	49.057±4.492	44.365±5.106	39.452±4.494
	WFAF	0.9925±0.0056	0.9903±0.0069	0.9744±0.0104	0.9905±0.0081	0.9364±0.0207	0.9029±0.0565
	SLD	0.9965±0.0031	0.9952±0.0031	0.9950±0.0041	0.9964±0.0032	0.9907±0.0060	0.9823±0.0142
	UTV	0.9958±0.0029	0.9934±0.0052	0.9937±0.0042	0.9914±0.0056	0.9886±0.0193	0.9851±0.0122
	GSLV	0.9982±0.0016	0.9917±0.0042	0.9962±0.0101	0.9967±0.0088	0.9956±0.0091	0.9933±0.0152
	LRSID	0.9983±0.0032	0.9934±0.0113	0.9891±0.0070	0.9962±0.0042	0.9975±0.0091	0.9924±0.0402
Ours	Ours	0.9991±0.0006	0.9956±0.0035	0.9990±0.0010	0.9986±0.0016	0.9979±0.0012	0.9942±0.0042

228 For the nonperiodic stripes, we show the mean value results in Table 3. The WFAF method shows
 229 the instability, and the PSNR and SSIM of LRSID method are consistent with the results in [25]. From
 230 the two tables, our method always shows a good performance significantly.

231 In Fig. 6, we take two examples of Table 2 to show the PSNR and SSIM performance of all
 232 comparing methods on each image. The y-axis stands for the value of PSNR or SSIM and the x-axis
 233 represents the i -th image of 32 examples. Fig. 6 (I) and Fig. 6 (II) are the PSNR and SSIM performance
 234 of stripes (Per, 100, 0.6), and Fig. 6 (III) and Fig. 6 (IV) are that of stripes (NonPer, 50, 0.2), respectively.
 235 Although the PSNR results fluctuate with respect to different images, our method holds the best
 236 PSNR results on almost of all images. Moreover, the SSIM results show the best performance with the
 237 smallest variance, which is consistent with the results of Table 2 and Table 3. From Fig. 6, our method
 238 is superior to the other comparing methods.

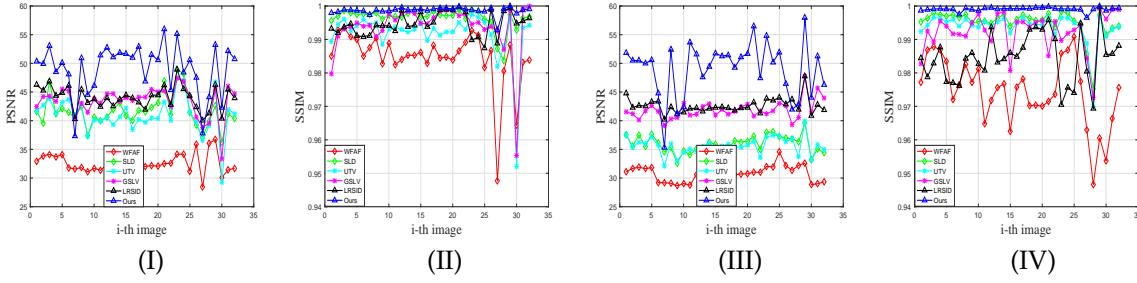


Figure 6. The PSNR and SSIM performance on 32 images for the stripes (Per, 100, 0.6) and (NonPer, 50, 0.2). The x-axis represents each image and the quantitative results are shown in y-axis. (I) and (II) are the PSNR and SSIM results for the stripes (Per, 100, 0.6), respectively. (III) and (IV) respectively represent the PSNR and SSIM performance of the stripes (NonPer, 50, 0.2).

239 4.2. Real experiments

240 We also display the destriping results of six methods for six real remote sensing images, which
 241 are also available on the website⁸, see Fig. 7. Similar to Fig. 2, the six real images with different stripes
 242 are shown in the first row, and the destriping results of all comparing methods are presented from the
 243 second row to the end row.

244 In Fig. 7, for the first, fifth and last real images, the proposed method not only removes the stripes
 245 completely, but also preserves image details on stripe-free regions well. Note that the methods GSLV
 246 and LRSID fail to obtain excellent results for the first image as the mentioned in their papers. For the
 247 forth column, there are also several stripe residuals with WFAF and SLD, and the wide black shadow
 248 areas appear by the UTV, GSLV and LRSID methods. Moreover, the destriping results of the WFAF
 249 and SLD leave significant stripes for the second image, and still exist the wispy stripes for the third
 250 example. According to several real experiments, the results demonstrate the universal effectiveness
 251 and stability of the proposed method.

252 4.3. More discussion

253 **1) Qualitative Analysis.** For the further comparisons of different destriping methods for
 254 simulated and real remote sensing images, we show the following two assessments. One is the
 255 mean cross-track profile that the x-axis stands for the column number of an image and the y-axis
 256 represents the mean value of each column, see Fig. 8 and Fig. 10. The other is the power spectrum that
 257 the x-axis is the normalized frequencies of an image, and the y-axis shows the spectral magnitude with
 258 a logarithmic scale, see Fig. 9 and Fig. 11.

259 In simulated experiments, the mean cross-track profile of the first image of Fig. 2 has been shown
 260 in Fig. 8. Note that Fig. 8 (a) shows the mean cross-track profile of the underlying image, and Fig. 8 (b)
 261 is the result of the degraded image. Moreover, Fig. 8 (c)-(f) are the mean cross-track profile results of
 262 the six destriping methods, respectively. From the overall perspective, Fig. 8 (d) and Fig. 8 (e) have
 263 obvious change of the intensity contrast. Seeing the details, Fig. 8 (c)-(g) have some mild fluctuations
 264 which are different with the underlying image in Fig. 8 (a). The proposed method shows the best
 265 performance, since it is almost same as the original one.

266 In addition, the power spectrum results of the second image of Fig. 2 has been shown in Fig. 9.
 267 We denote the power spectrum results as Fig. 9 (a)-(h) which represent the power spectrum results of
 268 the underlying image, the degraded image and the destriping results of six methods, respectively. Fig.
 269 9 (c)-(g) have more fluctuations which indicate these methods may have the stripe residuals or bring a

8 <https://ladsweb.nascom.nasa.gov/>



Figure 7. The visual results of different real images. From top to bottom: the real images, the destriping results of WFAF, SLD, UTV, GSLV, LRSID and Ours. Readers are recommended to zoom in all figures for better visibility.

270 little new noise in their destriping processes. For our method, *i.e.*, Fig. 9 (f), it not only removes all
271 stripes, but also preserves almost the essential details such as edges.

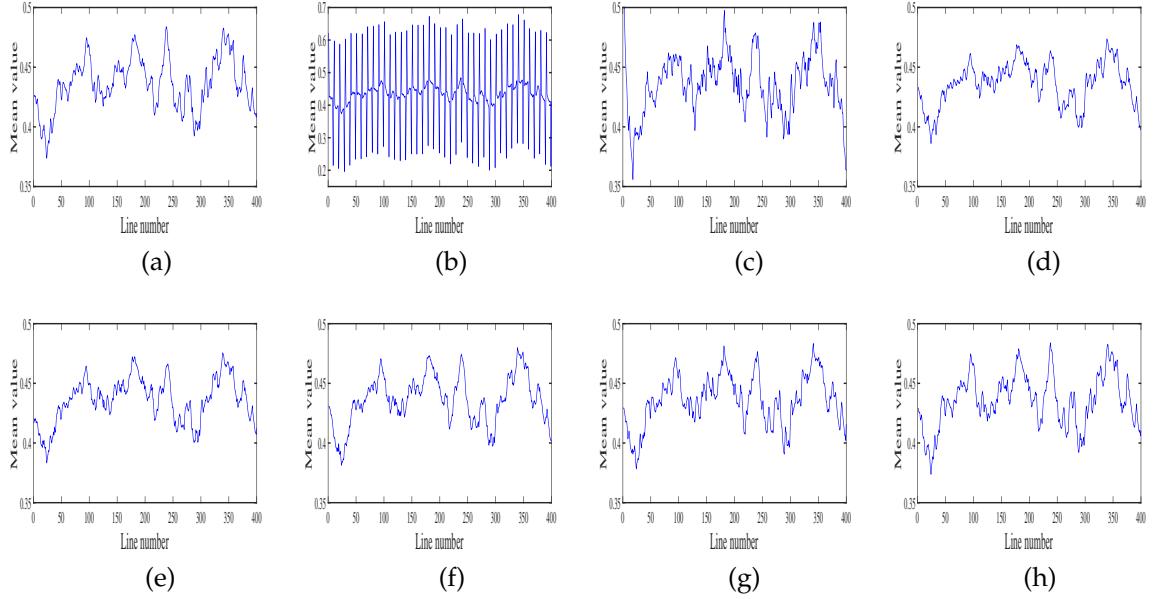


Figure 8. Spatial mean cross-track profiles for simulated image of the first simulated example of Fig. 2.
(a) Underlying image. (b) Degraded image. Destriping results by (c) WFAF, (d) SLD, (e) UTV, (f) GSLV,
(g) LRSID, (h) Ours.

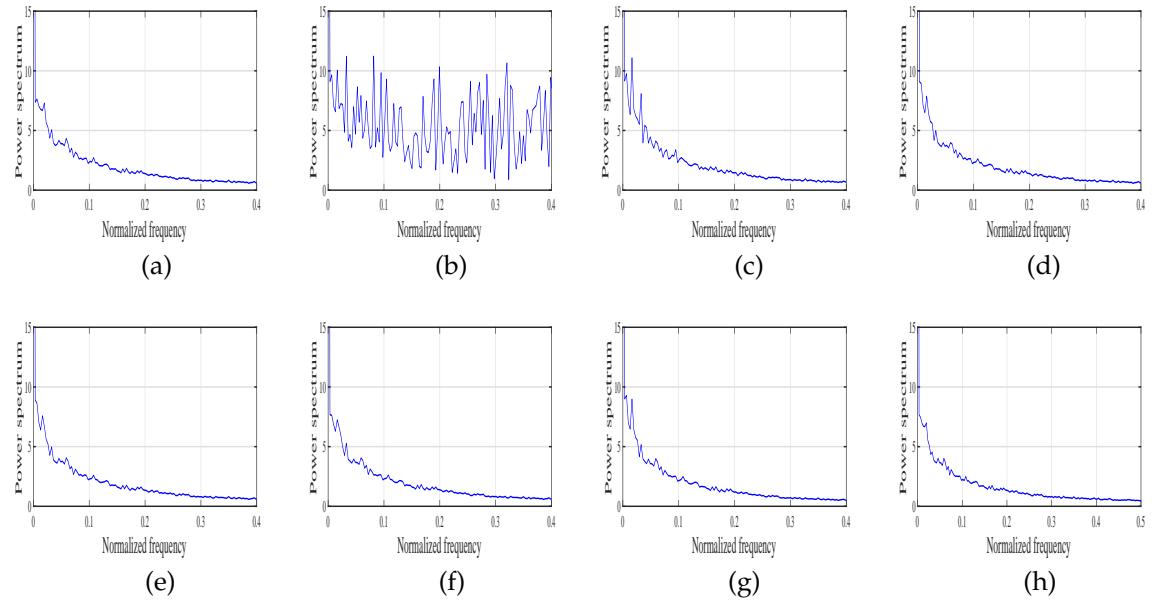


Figure 9. Power spectrum for simulated image of the second example of Fig. 7. (a) Underlying image.
(b) Degraded image. Destriping results by (c) WFAF, (d) SLD, (e) UTV, (f) GSLV, (g) LRSID, (h) Ours.

272 In real experiments, we also show the mean cross-track profile and the power spectrum in Fig.
273 10 and Fig. 11, respectively. Fig. 10 shows the mean cross-track profile results of the first column of
274 Fig. 7. Note that Fig. 10 (a) is the mean cross-track profile result of the first real remote sensing image,
275 and Fig. 10 (b)-(g) show the profile results of six destriping methods, respectively. In general, the

276 profiles of the destriping method should smoothen huge fluctuates and maintain primary structure
 277 information. However, the profiles of WFAF and LRSID have obvious fluctuations where the stripes
 278 still exist, and that of SLD is over-smooth missing a lot of underlying image details. In Fig. 10 (d) and
 279 Fig. 10 (e), although stripes are mostly removed, the destriping profiles have some mild burrs and
 280 too much smoothness because of the unidirectional property of UTV and the global sparsity of GSLV,
 281 respectively. In addition, the profile of the proposed method, *i.e.*, Fig. 10 (g), can realize the desired
 282 result both on removing stripes and keeping underlying image details.

283 In Fig. 11, the power spectrum results of the forth example of Fig. 7 are plotted. Fig. 11 (a)-(h)
 284 represent the power spectrum results of the forth real remote sensing image and six destriping methods,
 285 respectively. We observe that the real remote sensing image in Fig. 11 (a) has much fluctuates where
 286 stand for stripes. According to the power spectrum results of the six methods in Fig. 11 (b)-(f), although
 287 the stripes are almost removed well, there are still some slight blurring regions, while the proposed
 288 method shows the best performance in Fig. 11 (g).

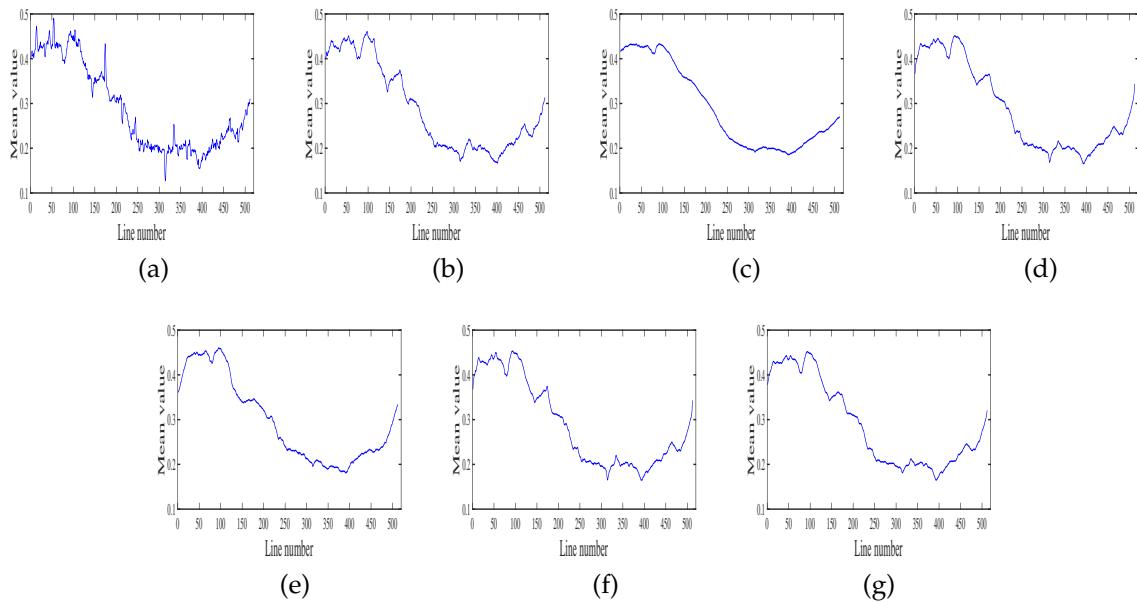


Figure 10. Spatial mean cross-track profiles for the first real example of Fig. 7. (a) Real image.
 Destriping results by (b) WFAF, (c) SLD, (d) UTV, (e) GSLV, (f) LRSID, (g) Ours.

289 **2) The influence of different regularization terms in the proposed model.** Fully considering
 290 the destriping problem (2) and the optimization model (12), we assume that R_2 is a necessary term,
 291 since R_2 is the only term to describe the property of the underlying image \mathbf{u} . To confirm whether both
 292 R_1 and R_2 are necessary priors as well as have significant contribution for destriping performance,
 293 in Fig. 12, we give the mean value of PSNR and SSIM for 32 images as before. Here, R_{12} represents
 294 $R_1 + R_2$, R_{23} stands for $R_2 + R_3$ and R_{123} represents $R_1 + R_2 + R_3$ (*i.e.*, the proposed model). Please
 295 find the definitions of R_1 , R_2 , R_3 from Eq. (9), Eq. (10) and Eq. (11), respectively.

296 Fig. 12 (I) and Fig. 12 (II) show the mean value of PSNR and the mean value of SSIM on 32 images
 297 same as before for periodic stripes. The periodic stripe levels (a)-(f) are (Per, 10, 0.2), (Per, 10, 0.6), (Per,
 298 50, 0.2), (Per, 50, 0.6), (Per, 100, 0.2) and (Per, 100, 0.6), respectively. Moreover, Fig. 12 (III) and Fig. 12
 299 (IV) display the mean value of PSNR and the mean value of SSIM on 32 images for nonperiodic stripes.
 300 The nonperiodic stripe levels (a)-(f) stand for (NonPer, 10, 0.2), (NonPer, 10, 0.6), (NonPer, 50, 0.2),
 301 (NonPer, 50, 0.6), (NonPer, 100, 0.2) and (NonPer, 100, 0.6), respectively.

302 From the results in Fig. 12, we can conclude three points. 1) The results both PSNR and SSIM of
 303 the proposed model (*i.e.*, R_{123}) perform the best than those of the other two models. 2) For R_{12} and R_{23} ,
 304 R_{23} shows more stability than R_{12} as the green bars do not significantly change with different stripes.

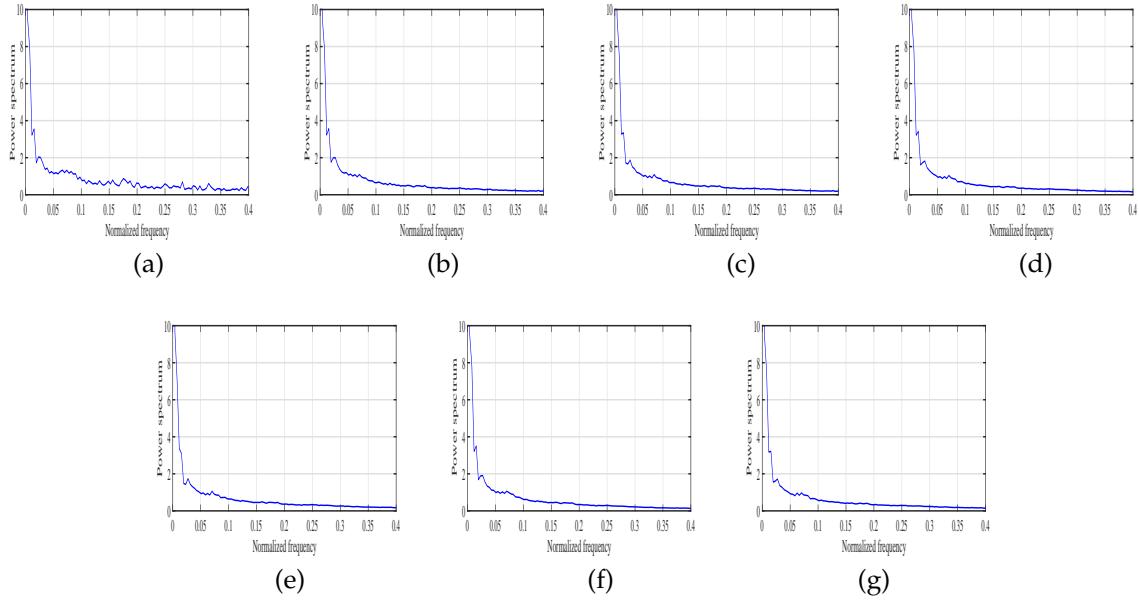


Figure 11. Power spectral for the forth real example of Fig. 7. (a) Real image. Despeckling results by (b) WFAF, (c)SLD, (d) UTV, (e) GSLV, (f) LRSID, (g) Ours.

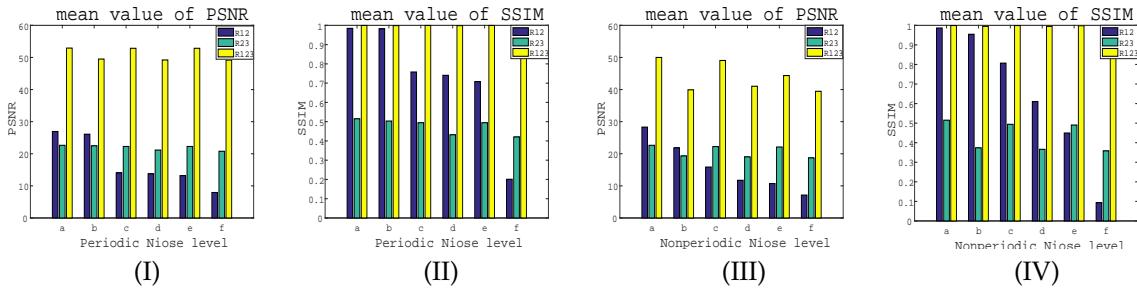


Figure 12. The influence of different terms in the proposed model. R_{12} represents $R_1 + R_2$, R_{23} stands for $R_2 + R_3$ and R_{123} represents $R_1 + R_2 + R_3$ (*i.e.*, the proposed model). (I) The mean PSNR performance on 32 images for periodic stripes with different stripe levels; (II) The mean SSIM performance on 32 images for periodic stripes with different stripe levels; The stripe levels (a)-(f) stand for (Per, 10, 0.2), (Per, 10, 0.6), (Per, 50, 0.2), (Per, 50, 0.6), (Per, 100, 0.2) and (Per, 100, 0.6), respectively. (III) The mean PSNR performance on 32 images for nonperiodic stripes with different stripe levels; (IV) The mean SSIM performance on 32 images for nonperiodic stripes with different stripe levels. The stripe levels (a)-(f) stand for (NonPer, 10, 0.2), (NonPer, 10, 0.6), (NonPer, 50, 0.2), (NonPer, 50, 0.6), (NonPer, 100, 0.2) and (NonPer, 100, 0.6), respectively.

305 3) R_3 actually plays a more important role than R_1 with respect to PSNR (see Fig. 12 (I) and Fig. 12
306 (III)). On the contrary, R_1 plays a more important role than R_3 with respect to SSIM (see Fig. 12 (II) and
307 Fig. 12 (IV)). Fig. 12 demonstrates the effectiveness of the proposed model and the importance of the
308 three terms.

309 **3) Parameters selection.** In this paper, the proposed method mainly involves six parameters λ , μ ,
310 β_1 , β_2 , β_3 and β_4 . The stripes of different types can be removed by setting different parameters. For
311 example, if the stripes are heavy, the μ should be small and the λ should be large.

312 For the simulated experiments, the parameters have the following setting: 1) For the periodic
313 stripes, we empirically set the parameters as $\lambda = 1$, $\mu = 0.1$, $\beta_1 = \beta_2 = \beta_3 = 100$, $\beta_4 = 1000$.
314 Under this parameter setting, it can generate a good performance for most of all examples. 2) For
315 the non-periodic stripes, the parameters are empirically set as $\lambda = 1$, $\mu = 0.1$, $\beta_1 = 100$, $\beta_2 = 10$,

³¹⁶ $\beta_3 = 10$ and $\beta_4 = 1000$. Similarly, the proposed method can obtain an excellent performance under
³¹⁷ this parameter setting for most of images. For the real experiments, most of all examples show the
³¹⁸ superior results with $\lambda = 10$, $\mu = 1$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$. Note that, if fine tuning parameters
³¹⁹ for each images would get better results. To simplify the process of the parameter adjustment, we
³²⁰ unify parameters to demonstrate the stability of the proposed method.

³²¹ 5. Conclusion

³²² In this paper, we proposed a directionally non-convex ℓ_0 sparse model for remote sensing image
³²³ destriping. This model was efficiently solved by the designed PADMM algorithm based on the MPEC
³²⁴ reformulation. Furthermore, we also theoretically gave the corresponding proof of the convergence
³²⁵ to the KKT point by this work. Experimental results on simulated and real data demonstrated
³²⁶ the effectiveness of the proposed method, both quantitatively and visually. Moreover, the mean
³²⁷ performance of Table 2 and Table 3 also exhibited the stability of our method to parameters and
³²⁸ different stripes.

³²⁹ In the future, we will extend the proposed model to the oblique stripes removal by fully
³³⁰ considering the latent properties of oblique stripes. Furthermore, the proposed method was only
³³¹ applied to single-band image stripe removal. We may extend our framework to multispectral or
³³² hyperspectral image stripe removal by some intrinsic properties, e.g., low-rank and non-local priors.

³³³ Acknowledgment

³³⁴ This research is supported by NSFC (61772003, 61702083) and the Fundamental Research Funds
³³⁵ for the Central Universities (ZYGX2016KYQD142, ZYGX2016J132, ZYGX2016J129).

³³⁶ Appendix A Convergence of the proposed method

³³⁷ In fact, the global convergence of the ADMM algorithm has been proved under some conditions
³³⁸ [41], and that of the generalized ADMM is also verified in [42]. Wen *et al.* [43] show that the sequence
³³⁹ formed by ADMM can converge to a KKT point. Moreover, some researches give the convergence
³⁴⁰ property of proximal ADMM (PADMM), see [31] and [44]. Considering our non-convex optimization
³⁴¹ model, convergence to a stationary point (local minimum) is the best convergence property. Similarly,
³⁴² in this paper, we design a PADMM based algorithm to solve the remote sensing image destriping
³⁴³ problem, as well as prove the convergence of the proposed algorithm which can converge to the KKT
³⁴⁴ point. Here, we denote that the limitation of vector is defined as pointwise convergence. For instance,
³⁴⁵ for $\mathbf{x}^k = (x_1^k, \dots, x_n^k)^T$, $\lim_{k \rightarrow \infty} \mathbf{x}^{k+1} - \mathbf{x}^k = 0$ represents that $\lim_{k \rightarrow \infty} x_i^{k+1} - x_i^k = 0$, $i = 1, \dots, n$.

³⁴⁶ **theorem**[Convergence of Algorithm 1] Let $\mathbf{P} \triangleq (\mathbf{h}, \mathbf{z}, \mathbf{w}, \mathbf{v}, \mathbf{s})$, $\mathbf{Q} \triangleq (\pi_1, \pi_2, \pi_3, \pi_4)$. $\{\mathbf{P}^k, \mathbf{Q}^k\}_{k=1}^\infty$
³⁴⁷ is a sequence of the solution of Algorithm 1 after k -th iteration. Assume that $\lim_{k \rightarrow \infty}$
³⁴⁸ $(\mathbf{Q}^{k+1} - \mathbf{Q}^k) = 0$ and $\lim_{k \rightarrow \infty} \mathbf{s}^{k+1} - \mathbf{s}^k = 0$, then the accumulation point of the subsequence $\{\mathbf{P}^k, \mathbf{Q}^k\}$
³⁴⁹ is the KKT point which satisfies the KKT conditions.

Proof: For convenience, we define

$$\Delta \triangleq \{\mathbf{z} \mid \mathbf{0} \leq \mathbf{z} \leq \mathbf{1}\}.$$

Recall our optimization model

$$\begin{aligned} & \min_{\mathbf{v} \in \Delta, \mathbf{s}} \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle + \mu \|\mathbf{z}\|_1 + \lambda \|\mathbf{w}\|_1 \\ & \text{s.t. } \mathbf{v} \odot |\mathbf{h}| = 0, \nabla_y \mathbf{s} = \mathbf{h}, \mathbf{s} = \mathbf{z}, \nabla_x (\mathbf{b} - \mathbf{s}) = \mathbf{w}. \end{aligned} \quad (\text{A1})$$

The Lagrange function L is

$$\begin{aligned} L(\mathbf{h}, \mathbf{z}, \mathbf{w}, \mathbf{v}, \mathbf{s}, \pi_1, \pi_2, \pi_3, \pi_4) &= \langle \mathbf{1}, \mathbf{1} - \mathbf{v} \rangle + \mu \|\mathbf{z}\|_1 + \lambda \|\mathbf{w}\|_1 + \langle \nabla_y \mathbf{s} - \mathbf{h}, \pi_1 \rangle \\ &+ \langle \mathbf{s} - \mathbf{z}, \pi_2 \rangle + \langle \nabla_x (\mathbf{b} - \mathbf{s}) - \mathbf{w}, \pi_3 \rangle + \langle \mathbf{v} \odot |\mathbf{h}|, \pi_4 \rangle, \end{aligned} \quad (\text{A2})$$

where π_1, π_2, π_3 and π_4 are Lagrange multipliers. Now, we give the first-order optimal conditions of the proposed problem for $L(\mathbf{h}^*, \mathbf{z}^*, \mathbf{w}^*, \mathbf{v}^*, \mathbf{s}^*, \pi_1^*, \pi_2^*, \pi_3^*, \pi_4^*)$.

$$\begin{aligned}
0 &= \nabla_y^T \pi_1^* + \pi_2^* - \nabla_x^T \pi_3^*, \\
0 &\leq \langle \pi_4^* \odot |\mathbf{h}^*| - \mathbf{1}, \mathbf{v} - \mathbf{v}^* \rangle, \quad \forall \mathbf{v} \in \Delta \\
0 &\in -\pi_1^* + \pi_4^* \odot \mathbf{v}^* \odot \partial ||\mathbf{h}^*||_1 \\
0 &\in -\pi_2^* + \mu \partial ||\mathbf{z}||_1 \\
0 &\in -\pi_3^* + \lambda \partial ||\mathbf{w}||_1 \\
0 &= \nabla_y \mathbf{s}^* - \mathbf{h}^* \\
0 &= \mathbf{s}^* - \mathbf{z}^* \\
0 &= \nabla_x (\mathbf{b} - \mathbf{s}^*) - \mathbf{w}^* \\
0 &= \mathbf{v}^* \odot |\mathbf{h}^*|
\end{aligned} \tag{A3}$$

The Robinson's constraint qualification can guarantee the existence of the optimization solution.
350 Next, we will confirm the convergence property of the designed PADMM based algorithm with a
351 convergence sequence under the similar assumption condition in [43]. The augmented Lagrangian
352 function $\mathcal{L}(\mathbf{h}, \mathbf{z}, \mathbf{w}, \mathbf{v}, \mathbf{s}, \pi_1, \pi_2, \pi_3, \pi_4, \beta_1, \beta_2, \beta_3, \beta_4)$, which is in Eq. (16), is denoted as \mathcal{L}_β . Note that,
353 the Lagrangian function L is used to get the KKT conditions. Then we prove that the solution of the
354 augmented Lagrangian function \mathcal{L} , which is solved by Algorithm 1, can satisfy the KKT conditions.
355

(i) According to the limit of \mathbf{Q}^k and the update formula of the multipliers \mathbf{Q}^{k+1} , we can get

$$\lim_{k \rightarrow \infty} \nabla_y \mathbf{s}^{k+1} - \mathbf{h}^{k+1} = \mathbf{0}, \tag{A4}$$

$$\lim_{k \rightarrow \infty} \mathbf{s}^{k+1} - \mathbf{z}^{k+1} = \mathbf{0}, \tag{A5}$$

$$\lim_{k \rightarrow \infty} \nabla_x (\mathbf{b} - \mathbf{s}^{k+1}) - \mathbf{w}^{k+1} = \mathbf{0}, \tag{A6}$$

$$\lim_{k \rightarrow \infty} \mathbf{v}^{k+1} \odot |\mathbf{h}^{k+1}| = \mathbf{0}. \tag{A7}$$

(ii) According to the limit of π_1^k, π_4^k , and the \mathbf{h}^{k+1} subproblem of \mathcal{L}_β in Eq. (17), we can get

$$\lim_{k \rightarrow \infty} \mathbf{h}^{k+1} \in \arg \min_{\mathbf{h}} \langle \nabla_y \mathbf{s}^{k+1} - \mathbf{h}, \pi_1^{k+1} \rangle + \frac{\beta_1}{2} ||\nabla_y \mathbf{s}^{k+1} - \mathbf{h}||^2 + \langle \mathbf{v}^{k+1} \odot |\mathbf{h}|, \pi_4^{k+1} \rangle + \frac{\beta_4}{2} ||\mathbf{v}^{k+1} \odot |\mathbf{h}|||^2,$$

By the first optimality condition of \mathbf{h} , we have

$$\lim_{k \rightarrow \infty} -\pi_1^{k+1} + \pi_4^{k+1} \odot \mathbf{v}^{k+1} \odot \partial ||\mathbf{h}^{k+1}||_1 \ni \mathbf{0}. \tag{A8}$$

(iii) According to the limit of π_2^k , and the \mathbf{z}^{k+1} subproblem of \mathcal{L}_β in Eq. (21), we can get

$$\lim_{k \rightarrow \infty} \mathbf{z}^{k+1} \in \arg \min_{\mathbf{z}} \mu ||\mathbf{z}||_1 + \langle \mathbf{s}^{k+1} - \mathbf{z}, \pi_2^{k+1} \rangle + \frac{\beta_2}{2} ||\mathbf{s}^{k+1} - \mathbf{z}||^2,$$

By the first optimality condition of \mathbf{z} , we have

$$\lim_{k \rightarrow \infty} -\pi_2^{k+1} + \mu \partial ||\mathbf{z}^{k+1}||_1 \ni \mathbf{0}. \tag{A9}$$

(iv) According to the limit of π_3^k , and the \mathbf{w}^{k+1} subproblem of \mathcal{L}_β in Eq. (23), we can get

$$\lim_{k \rightarrow \infty} \mathbf{w}^{k+1} \in \arg \min_{\mathbf{w}} \lambda ||\mathbf{w}||_1 + \frac{\beta_3}{2} ||\nabla_x (\mathbf{b} - \mathbf{s}^{k+1}) - \mathbf{w} + \frac{\pi_3^{k+1}}{\beta_3}||^2,$$

By the first optimality condition of \mathbf{w} , we have

$$\lim_{k \rightarrow \infty} -\boldsymbol{\pi}_3^{k+1} + \lambda \partial ||\mathbf{w}^{k+1}||_1 \ni \mathbf{0}. \quad (\text{A10})$$

(v) According to the limit of $\boldsymbol{\pi}_4^k$, \mathbf{h}^k and the \mathbf{v}^{k+1} subproblem of \mathcal{L}_β in Eq. (25), we can get

$$\lim_{k \rightarrow \infty} \mathbf{v}^{k+1} \in \arg \min_{0 \leq \mathbf{v} \leq \mathbf{1}} \langle \mathbf{v}, \boldsymbol{\pi}_4^{k+1} \odot |\mathbf{h}^{k+1}| - \mathbf{1} \rangle + \frac{\beta_4}{2} ||\mathbf{v} \odot |\mathbf{h}^{k+1}|||^2,$$

By the first optimality condition of \mathbf{v} , we have

$$\lim_{k \rightarrow \infty} \langle \boldsymbol{\pi}_4^{k+1} \odot |\mathbf{h}^{k+1}| - \mathbf{1}, \mathbf{v} - \mathbf{v}^{k+1} \rangle \geq 0, \quad \forall \mathbf{v} \in \Delta. \quad (\text{A11})$$

(vi) According to the limit of \mathbf{s}^{k+1} and the update formula of \mathbf{s}^{k+1} subproblem of \mathcal{L}_β in Eq. (27), we have the first optimality condition of \mathbf{s} is

$$\begin{aligned} \nabla_y^T \boldsymbol{\pi}_1^{k+1} + \beta_1 \nabla_y^T (\nabla_y \mathbf{s}^{k+1} - \mathbf{h}^{k+1}) + \boldsymbol{\pi}_2^{k+1} + \beta_2 (\mathbf{s}^{k+1} - \mathbf{z}^{k+1}) - \nabla_x^T \boldsymbol{\pi}_3^{k+1} \\ - \beta_3 \nabla_x^T (\nabla_x^T (\mathbf{b} - \mathbf{s}^{k+1}) - \mathbf{w}^{k+1}) + \mathbf{D}(\mathbf{s}^{k+1} - \mathbf{s}^k) = \mathbf{0}. \end{aligned}$$

Combining it with (A4), (A5), (A6) and (A7), then we have

$$\lim_{k \rightarrow \infty} \nabla_x^T \boldsymbol{\pi}_1^{k+1} + \boldsymbol{\pi}_2^{k+1} - \nabla_y^T \boldsymbol{\pi}_3^{k+1} + \mathbf{D}(\mathbf{s}^{k+1} - \mathbf{s}^k) = \mathbf{0}. \quad (\text{A12})$$

Since the formula $\lim_{k \rightarrow \infty} \mathbf{s}^{k+1} - \mathbf{s}^k = \mathbf{0}$ and the matrix \mathbf{D} is a positive definite, so we have $\lim_{k \rightarrow \infty} \mathbf{D}(\mathbf{s}^{k+1} - \mathbf{s}^k) = \mathbf{0}$. Thus, we have

$$\lim_{k \rightarrow \infty} \nabla_x^T \boldsymbol{\pi}_1^{k+1} + \boldsymbol{\pi}_2^{k+1} - \nabla_y^T \boldsymbol{\pi}_3^{k+1} = \mathbf{0}. \quad (\text{A13})$$

Combining (A4), (A5), (A6), (A7), (A8), (A9), (A10), (A11) and (A13), we conclude that the $\{\mathbf{P}^k, \mathbf{Q}^k\}$ is the sequence generated by the Algorithm 1, and as $k \rightarrow \infty$, there exists a subsequence $\{\mathbf{P}^k, \mathbf{Q}^k\}$, whose accumulation point satisfies the KKT conditions in Eq. (A3).

References

1. Chen, J.S.; Shao, Y.; Guo, H.D.; Wang, W.M.; Zhu, B.Q. Destriping CMODIS data by power filtering. *IEEE Transactions on Geoscience and Remote Sensing* **2003**, *41*, 2119–2124.
2. Münch, B.; Trtik, P.; Marone, F.; Stampaconi, M. Stripe and ring artifact removal with combined wavelet Fourier filtering. *Optics Express* **2009**, *17*, 8567–8591.
3. Pande-Chhetri, R.; Abd-Elrahman, A. De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering. *ISPRS Journal of Photogrammetry and Remote Sensing* **2011**, *66*, 620–636.
4. Pal, M.K.; Porwal, A. Destriping of Hyperion images using low-pass-filter and local-brightness-normalization. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. *IEEE 2015*, pp. 3509–3512.
5. Gadallah, F.L.; Csillag, F.; Smith, E.J.M. Destriping multisensor imagery with moment matching. *International Journal of Remote Sensing* **2000**, *21*, 2505–2511.
6. Horn, B.K.P.; Woodham, R.J. Destriping Landsat MSS images by histogram modification. *Computer Graphics and Image Processing* **1979**, *10*, 69–83.
7. Weinreb, M.P.; Xie, R.; Lienesch, J.H.; Crosby, D.S. Destriping GOES images by matching empirical distribution functions. *Remote Sensing of Environment* **1989**, *29*, 185–195.
8. Wegener, M. Destriping multiple sensor imagery by improved histogram matching. *International Journal of Remote Sensing* **1990**, *11*, 859–875.

- 377 9. Rakwatin, P.; Takeuchi, W.; Yasuoka, Y. Restoration of Aqua MODIS band 6 using histogram matching and
378 local least squares fitting. *IEEE Transactions on Geoscience and Remote Sensing* **2009**, *47*, 613–627.
- 379 10. Sun, L.X.; Neville, R.; Staenz, K.; White, H.P. Automatic destriping of Hyperion imagery based on spectral
380 moment matching. *Canadian Journal of Remote Sensing* **2008**, *34*, S68–S81.
- 381 11. Shen, H.F.; Jiang, W.; Zhang, H.Y.; Zhang, L.P. A piece-wise approach to removing the nonlinear and
382 irregular stripes in MODIS data. *International Journal of Remote Sensing* **2014**, *35*, 44–53.
- 383 12. Fehrenbach, J.; Weiss, P.; Lorenzo, C. Variational algorithms to remove stationary noise: applications to
384 microscopy imaging. *IEEE Transactions on Image Processing* **2012**, *21*, 4420–4430.
- 385 13. Fehrenbach, J.; Weiss, P. Processing stationary noise: model and parameter selection in variational methods.
386 *SIAM Journal on Imaging Sciences* **2014**, *7*, 613–640.
- 387 14. Escande, P.; Weiss, P.; Zhang, W.X. A variational model for multiplicative structured noise removal. *Journal*
388 *of Mathematical Imaging and Vision* **2017**, *57*, 43–55.
- 389 15. Shen, H.F.; Zhang, L.P. A MAP-based algorithm for destriping and inpainting of remotely sensed images.
390 *IEEE Transactions on Geoscience and Remote Sensing* **2009**, *47*, 1492–1502.
- 391 16. Chen, Y.; Huang, T.Z.; Zhao, X.L.; Deng, L.J.; Huang, J. Stripe noise removal of remote sensing images by
392 total variation regularization and group sparsity constraint. *Remote Sensing* **2017**, *9*, 559.
- 393 17. Bouali, M.; Ladjal, S. Toward optimal destriping of MODIS data using a unidirectional variational model.
394 *IEEE Transactions on Geoscience and Remote Sensing* **2011**, *49*, 2924–2935.
- 395 18. Zhang, H.Y.; He, W.; Zhang, L.P.; Shen, H.F.; Yuan, Q.Q. Hyperspectral image restoration using low-rank
396 matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing* **2014**, *52*, 4729–4743.
- 397 19. Zhou, G.; Fang, H.Z.; Yan, L.X.; Zhang, T.X.; Hu, J. Removal of stripe noise with spatially adaptive
398 unidirectional total variation. *Optik—International Journal for Light and Electron Optics* **2014**, *125*, 2756–2762.
- 399 20. Liu, H.; Zhang, Z.L.; Liu, S.Y.; Liu, T.T.; Chang, Y. Destriping algorithm with L0 sparsity prior for remote
400 sensing images. *IEEE International Conference on Image Processing (ICIP)* **2015**, pp. 2295–2299.
- 401 21. Chang, Y.; Yan, L.X.; Fang, H.Z.; Liu, H. Simultaneous destriping and denoising for remote sensing images
402 with unidirectional total variation and sparse representation. *IEEE Geoscience and Remote Sensing Letters*
403 **2014**, *11*, 1051–1055.
- 404 22. Chang, Y.; Yan, L.X.; Fang, H.Z.; Luo, C.N. Anisotropic Spectral-Spatial Total Variation Model for
405 Multispectral Remote Sensing Image Destriping. *IEEE Transactions on Image Processing* **2015**, *24*, 1852–1866.
- 406 23. He, W.; Zhang, H.Y.; Zhang, L.P.; Shen, H.F. Total-variation-regularized low-rank matrix factorization for
407 hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing* **2016**, *54*, 178–188.
- 408 24. Liu, X.X.; Lu, X.L.; Shen, H.F.; Yuan, Q.Q.; Jiao, Y.L.; Zhang, L.P. Stripe Noise Separation and Removal in
409 Remote Sensing Images by Consideration of the Global Sparsity and Local Variational Properties. *IEEE*
410 *Transactions on Geoscience and Remote Sensing* **2016**, *54*, 3049–3060.
- 411 25. Chang, Y.; X., Y.L.; Wu, T.; Zhong, S. Remote Sensing Image Stripe Noise Removal: From Image
412 Decomposition Perspective. *IEEE Transactions on Geoscience and Remote Sensing* **2016**, *54*, 7018–7031.
- 413 26. Carfantan, H.; Idier, J. Statistical linear destriping of satellite-based pushbroom-type images. *IEEE*
414 *Transactions on Geoscience and Remote Sensing* **2010**, *48*, 1860–1871.
- 415 27. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D:*
416 *Nonlinear Phenomena* **1992**, *60*, 259–268.
- 417 28. Zhao, X.L.; Wang, F.; Huang, T.Z.; Ng, M.K.; Plemmons, R.J. Deblurring and sparse unmixing for
418 hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing* **2013**, *51*, 4045–4058.
- 419 29. Ma, T.H.; Huang, T.Z.; Zhao, X.L.; Lou, Y.F. Image Deblurring With an Inaccurate Blur Kernel Using a
420 Group-Based Low-Rank Image Prior. *Information Sciences* **2017**, *408*, 213–233.
- 421 30. Getreuer, P. Total variation inpainting using split Bregman. *Image Processing On Line* **2012**, *2*, 147–157.
- 422 31. Yuan, G.Z.; Ghanem, B. l0TV: A new method for image restoration in the presence of impulse noise. *IEEE*
423 *Conference on Computer Vision and Pattern Recognition* **2015**, pp. 5369–5377.
- 424 32. Shen, H.F.; Li, X.H.; Cheng, Q.; Zeng, C.; Yang, G.; Li, H.F. Missing information reconstruction of remote
425 sensing data: A technical review. *IEEE Geoscience and Remote Sensing Magazine* **2015**, *3*, 61–85.
- 426 33. Dong, B.; Zhang, Y. An efficient algorithm for ℓ_0 minimization in wavelet frame based image restoration.
427 *Journal of Scientific Computing* **2013**, *54*, 350–368.
- 428 34. Fan, Y.R.; Huang, T.Z.; Ma, T.H.; Zhao, X.L. Cartoon-texture image decomposition via non-convex low-rank
429 texture regularization. *Journal of the Franklin Institute* **2017**, *354*, 3170–3187.

- 430 35. Lu, Z.S.; Zhang, Y. Sparse approximation via penalty decomposition methods. *SIAM Journal on Optimization*
431 **2013**, *23*, 2448–2478.
- 432 36. Liu, J.; Huang, T.Z.; Selesnick, I.W.; Lv, X.G.; Chen, P.Y. Image restoration using total variation with
433 overlapping group sparsity. *Information Sciences* **2015**, *295*, 232–246.
- 434 37. Mei, J.J.; Dong, Y.Q.; Huang, T.Z.; Yin, W.T. Cauchy noise removal by nonconvex ADMM with convergence
435 guarantees. *Journal of Scientific Computing* **First Online: 30 May 2017**, DOI: [10.1007/s10915-017-0460-5](https://doi.org/10.1007/s10915-017-0460-5).
- 436 38. Ji, T.Y.; Huang, T.Z.; Zhao, X.L.; Ma, T.H.; Deng, L.J. A non-convex tensor rank approximation for tensor
437 completion. *Applied Mathematical Modelling* **2017**, *48*, 410–422.
- 438 39. Donoho, D.L. De-noising by soft-thresholding. *IEEE transactions on information theory* **1995**, *41*, 613–627.
- 439 40. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to
440 Structural Similarity. *IEEE Transactions on Image Processing* **2004**, *13*, 600–612.
- 441 41. He, B.S.; Yuan, X.M. On the O(1/n) Convergence Rate of the Douglas Rachford Alternating Direction
442 Method. *SIAM Journal on Numerical Analysis* **2012**, *50*, 700–709.
- 443 42. Deng, W.; Yin, W.T. On the global and linear convergence of the generalized alternating direction method
444 of multipliers. *Journal of Scientific Computing* **2016**, *663*, 889–916.
- 445 43. Wen, Z.W.; Yang, C.; Liu, X.; Marchesini, S. Alternating direction methods for classical and ptychographic
446 phase retrieval. *Inverse Problems* **2012**, *28*, 115010.
- 447 44. Fazel, M.; Pong, T.K.; Sun, D.F.; Tseng, P. Hankel matrix rank minimization with applications to system
448 identification and realization. *SIAM Journal on Matrix Analysis and Applications* **2013**, *34*, 946–977.

449 © 2017 by the authors. Submitted to MDPI for possible open access publication under the terms and conditions
450 of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).