

Article

# Variational Destriping in Remote Sensing Imagery: Total Variation with $L^1$ Fidelity

Igor Yanovsky <sup>1,2,\*</sup> and Konstantin Dragomiretskiy <sup>3</sup>

<sup>1</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

<sup>2</sup> Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA 90095, USA

<sup>3</sup> Department of Mathematics, University of California, Los Angeles, CA 90095, USA; kdragomiretskiy@ucla.edu

\* Correspondence: Igor.Yanovsky@jpl.nasa.gov; Tel.: +1-818-354-2243

Received: 3 January 2018; Accepted: 12 February 2018; Published: 15 February 2018

**Abstract:** This paper introduces a variational method for destriping data acquired by pushbroom-type satellite imaging systems. The model leverages sparsity in signals and is based on current research in sparse optimization and compressed sensing. It is based on the basic principles of regularization and data fidelity with certain constraints using modern methods in variational optimization, namely, total variation (TV),  $L^1$  fidelity, and the alternating direction method of multipliers (ADMM). The proposed algorithm, TV- $L^1$ , uses sparsity-promoting energy functionals to achieve two important imaging effects. The TV term maintains boundary sharpness of the content in the underlying clean image, while the  $L^1$  fidelity allows for the equitable removal of stripes without over- or under-penalization, providing a more accurate model of presumably independent sensors with an unspecified and unrestricted bias distribution. A comparison is made between the TV- $L^2$  model and the proposed TV- $L^1$  model to exemplify the qualitative efficacy of an  $L^1$  striping penalty. The model makes use of novel minimization splittings and proximal mapping operators, successfully yielding more realistic destriped images in very few iterations.

**Keywords:** alternating direction method of multipliers (ADMM); image striping restoration; raster scan; sparse optimization; split Bregman; total variation; variational destriping

## 1. Introduction

Image striping is a well-known phenomenon that arises in multi-detector imaging systems ranging from pushbroom-type instruments, such as the Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI), to atomic force microscopy (AFM). Biases in lateral detection occur as a result of response variation in spatial detectors, such as in satellite imaging systems, or temporal changes, such as in raster scans. Although these systems are optimally precalibrated, post-processing, such as destriping, of data is a prerequisite for accurate and valid analyses. Striping removal has been traditionally performed using either statistically based methods [1,2] or low-pass filtering in the frequency domain [3–7]. These methods, however, do not remove stripes completely and have effects of blurring the image. More recently, wavelet-based filtering methods have been proposed [8–10]. However, such methods also blur the images and produce ringing effects in reconstruction.

In [11], the authors introduced a robust destriping algorithm with a spatially adaptive unidirectional total variation (TV) model. The authors developed a new destriping method that combines the TV–Stokes model and unidirectional TV model in [12]. In [13], the authors introduced a TV and framelet regularization model for destriping. The same authors proposed an anisotropic spectral–spatial TV regularization to enhance the smoothness of the solution along both the spectral

and spatial dimension in [14]. In [13,14], the authors employed the  $L^2$  norm as the fidelity term, and we refer to these models as TV- $L^2$  in this paper.

We follow the pedigree of variational and partial differential equation (PDE)-based methods applied to images [15,16] in order to construct a well-defined, optimizable model yielding fast and high-quality destriping. The focus of our paper is not on creating a sparse wavelet representation of the destriped image, but rather on how to remove the optimal striping mask while preserving high image fidelity. In this paper, we propose using the  $L^1$  penalty for striping size.

Our research is robust to both isotropic and anisotropic versions of TV, whereas [13] argue that the anisotropic case is the only appropriate case. While it is true that the anisotropic case uses decoupled energy for the measure of smoothness and is therefore easier to minimize, isotropic TV is not selective in which direction smoothness is penalized. Image content smoothness (or lack thereof) is not known a priori, and thus no preference should immediately be given to certain directions for evaluating smoothness. Using the  $L^1$  penalty, and depending on the data, the isotropic TV, which theoretically uses more local information, allows for qualitatively better, less invasive and more intelligent destriping. We compare the  $L^1$  and  $L^2$  penalties, ultimately favoring the  $L^1$  as a result of a wider yet tighter distribution of the striping mask. We include detailed derivations and a motivated evolution of the optimization problem with pedagogy in mind so that the proposed TV- $L^1$  method, along with its alternating direction method of multipliers (ADMM) (split Bregman) optimization, can be accessible to all academic disciplines involved with image processing.

We construct a variational model that is well-defined, qualitatively motivated, and easily minimized. The constructed energy uses sparsity-promoting energy functionals, on the basis of TV and  $L^1$  energy, to achieve minimally invasive destriping. Both isotropic and anisotropic TV, along with  $L^1$  energy, are considered in our variational model. The ADMM is used in conjunction with nonlinear proximal operators to efficiently optimize the energy, yielding quick and high-quality results.

## 2. Materials and Methods

In this section, we describe the striping problem, provide the background and motivations, discuss tools from functional analysis to be used in describing the model, develop the proposed TV- $L^1$  model, and provide the algorithm.

### 2.1. Striping Structure

Let  $U(x, y)$  be a stripe-free image of size  $R$  by  $C$ , and let  $G(y)$  be a multiplicative stripe noise of length  $R$ . Then the observed image,  $F$ , can be written as

$$F(x, y) = G(x, y)U(x, y) \quad (1)$$

Taking logarithms of both sides of Equation (1) yields an additive structure more suitable for energy minimization methods. The model can now be written as

$$f(x, y) = g(y) + u(x, y)$$

where  $f(x, y) = \ln(F(x, y))$ ,  $g(x, y) = \ln(G(x, y))$ , and  $u(x, y) = \ln(U(x, y))$ .

Striping in images can be viewed as structured noise, of which variations are mainly concentrated along one axis. This can be mathematically encoded as  $\|\nabla_x G\| \ll \|\nabla_y G\|$ , or with the logarithmic terms, as  $\|\nabla_x g\| \ll \|\nabla_y g\|$ .

## 2.2. Tikhonov Minimization

A classical Tikhonov minimization problem would consist of a smoothness regularizer and a data fidelity term, both easily differentiable, with the striping constraint [17]:

$$\min_u \left\{ \|\nabla u(x, y)\|_2^2 + \frac{\lambda}{2} \|u(x, y) - f(x, y)\|_2^2 \right\}$$

such that  $\|\nabla_x g\| \ll \|\nabla_y g\|$

This constraint can be simplified by the approximation that  $\nabla_x G(x, y) = 0 \quad \forall(x, y)$ , which would make  $G(x, y) = G(y)$ , and  $g(x, y) = g(y)$  functions of only one variable. Using the additive identity between  $f$ ,  $g$ , and  $u$  along with the constraint approximation, the new unconstrained minimization problem is

$$\min_g \left\{ \|\nabla_y(f(x, y) - g(x, y))\|_2^2 + \frac{\lambda}{2} \|g(x, y)\|_2^2 \right\}$$

By taking the first variation of the energy and setting it to zero, the closed-form solution to this minimization problem is

$$\begin{aligned} g(x, y) &= (\nabla_y \cdot \nabla_y + \lambda I)^{-1} (\nabla_y \cdot \nabla_y f(x, y)) \\ &= \left( \frac{\partial^2}{\partial y^2} + \lambda I \right)^{-1} (f_{yy}(x, y)) \end{aligned}$$

However, this solution would cause  $g$  to become bivariate, in contradiction with the constraint. Instead, using the Cartesian regularity of our rectangular domain  $\Omega = I_x \times I_y$  and using  $g(x, y) = g(y)$ , we can come to a solution that is in agreement with the constraint by integrating with respect to  $x$ :

$$\begin{aligned} \int_{\Omega} g(x, y) dx &= \int_{I_x} g(y) dx = g(y) \int_{I_x} dx = g(y) \mu(I_x) = \int_{I_x} \left( \frac{\partial^2}{\partial y^2} + \lambda I \right)^{-1} (f_{yy}(x, y)) dx \Rightarrow \\ g(y) &= \frac{1}{\mu(I_x)} \int_{I_x} \left( \frac{\partial^2}{\partial y^2} + \lambda I \right)^{-1} (f_{yy}(x, y)) dx \end{aligned}$$

## 2.3. Fourier Interpretation

Utilizing the Plancherel Fourier isometry, the solution can be interpreted in spectral form as

$$\hat{g}(\omega_y) = \frac{1}{\mu(I_x)} \left( \frac{\omega_y^2}{\lambda + \omega_y^2} \right) \hat{f} = \left( \frac{\omega_y^2}{\lambda + \omega_y^2} \right) \hat{f}$$

For a specific  $x$ , the striping  $g(y)$  is constant and is of higher frequency, whereas the underlying clean image varies more slowly (has more low-frequency content) and for each  $x$  has somewhat different frequencies. Therefore, the average frequencies of the clean image are low in magnitude and are of lower frequency, while the average frequencies of the striping are high in magnitude and are of higher frequency. Therefore, the average frequencies (averaged over  $\omega_x$ ) of the cleaned image are simply the average frequencies of the original image multiplied by a one-dimensional low-pass filter  $\frac{\lambda}{\lambda + z^2}$ . Likewise, the striping mask on the spectral side,  $\hat{g}$ , is obtained analogously to a one-dimensional high-pass filter  $\frac{z^2}{\lambda + z^2}$ .

Although this minimization problem is readily solvable in closed form and has a motivated physical interpretation, we must abandon the quadratic energy terms so that we may have less penalization for heavier striping and to allow for less smooth solutions. Although the differentiability of terms is easy to work with, enough optimization machinery has been developed that we may tread forward without it. We now investigate and outline some tools from signal processing in order to refine

our model. Stripe and ring artifact removal from this frequency perspective has been accomplished using wavelet and Fourier filtering [18].

#### 2.4. Tools from Functional Analysis

We now introduce the notations and tools from functional analysis that are used in Section 2.5.

##### 2.4.1. Total Variation: Anisotropic vs. Isotropic

The idea of using TV as a regularizer and denoiser that promotes sparsity and piecewise constant smoothness can be traced back to Rudin, Osher, and Fatemi [19,20]. We begin with defining the notion of TV, which is used as a regularizer in the model.

**Definition 1.** The TV of a function  $f \in L^1(\Omega)$  is

$$V(f, \Omega) := \sup \left\{ \int_{\Omega} f(x) \operatorname{div} \phi(x) \, dx : \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\}$$

For a differentiable function  $f \in \Omega$ , with  $\Omega \subseteq \mathbb{R}^n$ , the TV of  $f$  can be written as

$$V(f, \Omega) = \int_{\Omega} |\nabla f(x)| \, dx$$

The choice of vectorial norm inside the integral yields the two different types of TV.

**Definition 2.** Isotropic TV:  $|\cdot|$  denotes the  $l_2$ -norm, in which case

$$V(f, \Omega) = \int_{\Omega} \left( \sum_i^n f_{x_i}^2(x) \right)^{\frac{1}{2}} dx$$

**Definition 3.** Anisotropic TV:  $|\cdot|$  denotes the  $l_1$ -norm, in which case

$$V(f, \Omega) = \int_{\Omega} \sum_i^n |f_{x_i}(x)| dx$$

The isotropic and anisotropic cases differ in terms of the geometries they each preserve. While the decoupled anisotropic TV preserves piecewise constant orthogonal structures, such as rectangular roofs, the coupled isotropic TV preserves piecewise constant radial structures, such as silos. Our model is robust with respect to either choice of TV, and dual derivations of variable updates are shown; however, the experiments and results are based on the anisotropic definition.

##### 2.4.2. Shrinkage Proximal Operator

We introduce a splitting variable and quadratic penalty into the model. The solution to the  $l_1$ -regularized least-squares problem:

$$\arg \min_{\vec{x}} \mu \|\vec{x}\|_1 + \frac{1}{2} \|\vec{x} - \vec{y}\|_2^2$$

is given by the soft threshold proximal mapping operator, shrinkage [21,22]:

**Definition 4.**

$$\operatorname{Shrink}(\vec{x}, \mu) = S_{\mu}(\vec{x}) = \frac{\vec{x}}{|\vec{x}|} \max\{|\vec{x}| - \mu, 0\}$$

If  $\|\vec{x}\|_1 = |x_1| + |x_2|$ , as in the anisotropic case of TV, the shrinkage is decoupled and done component wise. On the other hand, if  $\|\vec{x}\|_1 = \sqrt{|x_1|^2 + |x_2|^2}$ , as in the isotropic case, the terms are coupled and both components are updated simultaneously. Both variants have their merits; while the former is computationally simpler, the latter has the advantage of using more local information and may be more conformant to certain image processing applications.

### 2.5. TV- $L^1$ ADMM Model

We now describe our model using the motivations and tools introduced earlier. The two energy components of the minimization problem are the smoothness regularizer and the data fidelity term. The energy of the data fidelity term,  $\frac{\lambda}{2} \|f(x, y) - u(x, y)\|_2^2 = \frac{\lambda}{2} \|g(y)\|_2^2$ , can be interpreted as the magnitude of the striping mask. The  $L^2$  fidelity overly penalizes stripes of large magnitude and likewise underexaggerates the significance of stripes of small magnitude. In areas of no striping, we intend our (logarithm of the) striping mask to be very close to zero, while in areas of heavy striping, we wish to remove such striping and thus yield a larger magnitude of our striping mask in that region. The  $L^1$  fidelity gives a smaller striping mask in areas of no striping, leaving enough energy to remove the heavier striping in localized areas of the image. Because there is no prior knowledge of the distribution of the stripes and qualitatively we may wish to remove deep striping effects while preserving sharp geometry, we believe it is better to update the model with an  $L^1$  striping penalty,  $\|g\|_1$ .

An  $L^2$  gradient term would cause oversmoothing of the retrieved clean image  $u(x, y)$ . This could cause a loss in boundary sharpness of elements in the image (e.g., lakes, rooftops, etc.), which seems important in the pursuit and usage of destriped images. Implementing a TV-based regularizer would give performance similar to the  $L^2$  gradient but maintains boundary sharpness more natural to the underlying image. Although these terms are not differentiable, impeding a closed-form solution, state-of-the-art nonlinear optimization algorithms are available for fast convergence to qualitatively meaningful minimizers. The unconstrained TV  $L^1$  model (TV- $L^1$ ) is

$$\min_u \left\{ V(u(x, y), I_x \times I_y) + \lambda \|u(x, y) - f(x, y)\|_1 \right\}$$

or equivalently, minimizing with respect to the striping mask  $g$ ,

$$\min_g \left\{ V(f(x, y) - g(y), I_x \times I_y) + \lambda \|g(y)\|_1 \right\}$$

#### 2.5.1. Discretization

For the purpose of application and computation, we now move the problem into a discrete setting. Let  $\Omega = \{x_1, \dots, x_C\} \times \{y_1, \dots, y_R\}$  be an  $R \times C$  image. First variations are approximated via forward differences, so that  $\frac{\partial f}{\partial y}(x_i, y_j) \approx f(x_i, y_{j+1}) - f(x_i, y_j) := \delta_y f_{i,j}$  for  $j = 1, \dots, R - 1$  and analogously  $\frac{\partial f}{\partial x} \approx f(x_{i+1}, y_j) - f(x_i, y_j) := \delta_x f_{i,j}$  for  $i = 1, \dots, C - 1$ . We take Neumann boundary conditions, so that on the forward boundary ( $i = C$  or  $j = R$ ), the derivative is set to zero.

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \in M_{R \times R}(\mathbb{R})$$

$$\text{so that } Du(x_{i_0}, y) = (\delta_y u_{i_0,1}, \dots, \delta_y u_{i_0,R-1}, 0)^T$$

The isotropic TV is

$$|\nabla^I f(x_i, y_j)| \approx \sqrt{(\delta_x f_{i,j})^2 + (\delta_y f_{i,j})^2}$$

The anisotropic TV is

$$|\nabla^{AI} f(x_i, y_j)| \approx |\delta_x f_{i,j}| + |\delta_y f_{i,j}|$$

With these discrete operators defined, the discrete unconstrained TV- $L^1$  minimization problem is

$$\min_g \left\{ \sum_{i,j} \|\langle \delta_x f_{i,j}, \delta_y (f_{i,j} - g_j) \rangle\|_1 + \lambda \|g(y)\|_1 \right\}$$

The two types of the minimization problem are the following:

Anisotropic:

$$\min_g \left\{ \sum_{i,j} |\delta_y (f_{i,j} - g_j)| + \lambda \|g(y)\|_1 \right\}$$

Isotropic:

$$\min_g \left\{ \sum_{i,j} \sqrt{(\delta_x f_{i,j})^2 + (\delta_y (f_{i,j} - g_j))^2} + \lambda \|g(y)\|_1 \right\}$$

### 2.5.2. Augmented Lagrangian

**Anisotropic:** With the discrete forward difference approximation matrix defined above, we can rewrite the minimization problem as follows:

$$\text{Point Form: } \min_{\vec{g}} \left\{ \sum_{i,j} |\delta_y (f_{i,j} - g_j)| + \lambda \|g(y)\|_1 \right\}$$

$$\text{Vector Form: } \min_{\vec{g}} \left\{ \sum_i \|D(f_i - \vec{g})\|_1 + \lambda \|g(y)\|_1 \right\}$$

$$\text{Matrix Form: } \min_{\vec{g}} \left\{ \|D(f - \vec{g} \otimes \vec{1}_C)\|_{1,1} + \lambda \|g(y)\|_1 \right\}$$

To render the constrained minimization problem unconstrained, we introduce auxiliary variables, Lagrangian multipliers (split Bregman), and quadratic penalty terms, so that the augmented Lagrangian is defined as follows:

$$\begin{aligned} \mathcal{L}_{\alpha,\lambda}(b_i, h, g, q_i, r) &= \sum_i \left( \|b_i\|_1 + \frac{\alpha}{2} \|b_i - D(g - f_i)\|_2^2 + \langle q_i, b_i - D(g - f_i) \rangle \right) \\ &\quad + \lambda \left( \|h\|_1 + \frac{\alpha}{2} \|h - g\|_2^2 + \langle r, h - g \rangle \right) \\ &= \sum_i \left( \|b_i\|_1 + \frac{\alpha}{2} \|b_i - D(g - f_i) + \frac{q_i}{\alpha}\|_2^2 \right) \\ &\quad + \lambda \left( \|h\|_1 + \frac{\alpha}{2} \|h - g + \frac{r}{\alpha}\|_2^2 \right) + \mathcal{O}(q_i^2, r^2) \end{aligned}$$

We now solve the unconstrained saddle point problem:

$$\min_{b_i, h, g} \max_{q_i, r} \mathcal{L}_{\alpha,\lambda}(b_i, h, g, q_i, r)$$

The solution to the original constrained minimization problem is now found as the saddle point of the augmented Lagrangian  $\mathcal{L}$  in a sequence of iterative suboptimizations called the ADMM [23–28].

The splitting variables  $b_i$  and  $h$  are updated by the proximal mapping operator:

$$\begin{aligned}
 b_i^{k+1} &= \arg \min_{b_i} \mathcal{L}_{\alpha,\lambda}(b_i, h^k, g^k, q_i^k, r^k) \\
 &= \arg \min_{b_i} \left\{ \|b_i\|_1 + \frac{\alpha}{2} \|b_i - D(g - f_i) + \frac{q_i}{\alpha}\|_2^2 \right\} \\
 &= S_{\frac{1}{\alpha}} \left( D(g - f_i) - \frac{q_i}{\alpha} \right) \\
 h^{k+1} &= \arg \min_h \mathcal{L}_{\alpha,\lambda}(b_i^{k+1}, h, g^k, q_i^k, r^k) \\
 &= \arg \min_{h_i} \left\{ \|h\|_1 + \frac{\alpha}{2} \|h - g\|_2^2 \right\} \\
 &= S_{\frac{1}{\alpha}} \left( g - \frac{r}{\alpha} \right)
 \end{aligned}$$

Because of the introduction of the splitting variables,  $g$  is only contained in quadratic terms and thus is easily solved for

$$\begin{aligned}
 g^{k+1} &= \arg \min_g \mathcal{L}_{\alpha,\lambda}(b_i^{k+1}, h^{k+1}, g, q_i^k, r^k) \\
 &= \arg \min_g \left\{ \frac{\alpha}{2} \sum_i \|b_i - D(g - f_i) + \frac{q_i}{\alpha}\|_2^2 + \frac{\lambda\alpha}{2} \|h - g + \frac{r}{\alpha}\|_2^2 \right\} \Rightarrow \\
 \frac{\delta \mathcal{L}}{\delta g} &= \alpha \sum_i -D^T(b_i - D(g - f_i) + \frac{q_i}{\alpha}) - \lambda\alpha(h - g + \frac{r}{\alpha}) = \\
 \alpha(CD^T D + \lambda I)g - \alpha D^T(\sum_i b_i + Df_i + \frac{q_i}{\alpha}) - \lambda\alpha h + \lambda r &= 0 \Rightarrow \\
 g &= (CD^T D + \lambda I)^{-1} \left( D^T(\sum_i b_i + Df_i + \frac{q_i}{\alpha}) + \lambda(h - \frac{r}{\alpha}) \right)
 \end{aligned}$$

The Langrangian multipliers (split Bregman variables) are updated through gradient ascent:

$$\begin{aligned}
 q_i^{k+1} &= q_i^k + \tau\alpha(b_i^{k+1} - D(g^{k+1} - f_i)) \\
 r^{k+1} &= r^k + \tau\lambda\alpha(h^{k+1} - g^{k+1})
 \end{aligned}$$

**Isotropic:** As a result of the coupling of the terms in this version of the minimization problem, we cannot compactly write the problem with matrices as above; however, the solution is just as readily available. Here the  $\diamond$  denotes the Hadamard matrix power operator, which acts pointwise on the matrix.

Point Form:

$$\min_{\vec{g}} \left\{ \sum_{i,j} \sqrt{(\delta_x f_{i,j})^2 + (\delta_y (f_{i,j} - g_j))^2} + \lambda \|g(y)\|_1 \right\}$$

Matrix Form:

$$\min_{\vec{g}} \left\{ \|[(fD^T)^{\diamond 2} + (D(f - \vec{g} \otimes \vec{1}_C))^{\diamond 2}]^{\diamond \frac{1}{2}}\|_{1,1} + \lambda \|g(y)\|_1 \right\}$$

Just as before, we introduce splitting variables and Lagrangian multipliers to form the augmented Lagrangian:

$$\begin{aligned}
 \mathcal{L}_{\alpha,\lambda}(a_{i,j}, b_{i,j}, h, g, p_{i,j}, q_{i,j}, r) &= \sum_{i,j} \sqrt{|a_{i,j}|^2 + |b_{i,j}|^2} + \frac{\alpha}{2} \|a_{i,j} - \delta_x f_{i,j}\|_2^2 + \langle p_{i,j}, a_{i,j} - \delta_x f_{i,j} \rangle \\
 &\quad + \frac{\alpha}{2} \|b_{i,j} - \delta_y(f_{i,j} - g_j)\|_2^2 + \langle q_{i,j}, b_{i,j} - \delta_y(f_{i,j} - g_j) \rangle \\
 &\quad + \lambda(\|h\|_1 + \frac{\alpha}{2} \|h - g\|_2^2 + \langle r, h - g \rangle) \\
 &= \sum_{i,j} \sqrt{|a_{i,j}|^2 + |b_{i,j}|^2} + \frac{\alpha}{2} \|a_{i,j} - \delta_x f_{i,j} + \frac{p_{i,j}}{\alpha}\|_2^2 \\
 &\quad + \frac{\alpha}{2} \|b_{i,j} - \delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha}\|_2^2 \\
 &\quad + \lambda(\|h\|_1 + \frac{\alpha}{2} \|h - g + \frac{r}{\alpha}\|_2^2) \\
 &\quad + \mathcal{O}(q_{i,j}^2, p_{i,j}^2, r^2)
 \end{aligned}$$

The splitting variables  $a_{i,j}$  and  $b_{i,j}$  are updated by the vectorial proximal mapping operator:

$$\begin{aligned}
 \langle a_{i,j}, b_{i,j} \rangle &= \arg \min_{\langle a_{i,j}, b_{i,j} \rangle} \mathcal{L}_{\alpha,\lambda}(a_{i,j}, b_{i,j}, h^k, g^k, p_i^k, q^k) \\
 &= \arg \min_{\langle a_{i,j}, b_{i,j} \rangle} \left\{ \sqrt{|a_{i,j}|^2 + |b_{i,j}|^2} + \frac{\alpha}{2} \|a_{i,j} - \delta_x f_{i,j} + \frac{p_{i,j}}{\alpha}\|_2^2 \right. \\
 &\quad \left. + \frac{\alpha}{2} \|b_{i,j} - \delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha}\|_2^2 \right\} \\
 &= \arg \min_{\langle a_{i,j}, b_{i,j} \rangle} \left\{ \|\langle a_{i,j}, b_{i,j} \rangle\| + \frac{\alpha}{2} \|\langle a_{i,j}, b_{i,j} \rangle - \langle \delta_x f_{i,j} + \frac{p_{i,j}}{\alpha}, \delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha} \rangle\|_2^2 \right\} \\
 &= S_{\frac{1}{\alpha}}(\langle \delta_x f_{i,j} + \frac{p_{i,j}}{\alpha}, \delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha} \rangle)
 \end{aligned}$$

Each component of the vector is updated via shrinkage as follows:

$$\begin{aligned}
 a_{i,j} &= \frac{\delta_x f_{i,j} + \frac{p_{i,j}}{\alpha}}{s} \cdot \max(s - \frac{1}{\alpha}, 0) \\
 b_{i,j} &= \frac{\delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha}}{s} \cdot \max(s - \frac{1}{\alpha}, 0) \\
 s &= \sqrt{(\delta_x f_{i,j} + \frac{p_{i,j}}{\alpha})^2 + (\delta_y(f_{i,j} - g_j) + \frac{q_{i,j}}{\alpha})^2}
 \end{aligned}$$

The splitting variable  $h$ , the striping mask  $g$ , and the Lagrangian multipliers are updated as before due to the common structure between the two models:

$$p_{i,j}^{k+1} = p_{i,j}^k + \tau\alpha(\delta_x f_{i,j} + \frac{p_{i,j}}{\alpha})$$

We now describe the algorithm for the TV- $L^1$  model (see Algorithm 1). The first algorithm (anisotropic) is presented in vector form. The second algorithm (isotropic) is presented in matrix form.

## 2.5.3. Algorithm

**Algorithm 1** Complete ADMM optimization of TV- $L^1$ .

1: Initialize:  $A^0 = (\vec{a}_i), B^0 = (\vec{b}_i), Q^0 = (\vec{q}_i) \leftarrow 0 \in \mathbb{R}^{R \times C}, \vec{g}^0, \vec{h}^0, \vec{r}^0 \leftarrow 0 \in \mathbb{R}^{R \times 1}, n \leftarrow 0 \in \mathbb{R}$   
 2:  $f = (\vec{f}_i) \leftarrow \ln(F), D \leftarrow 0 \in \mathbb{R}^{R \times R}, (D)_{i,i} = -1, (D)_{i,i+1} = 1$  for  $i = 1, \dots, R-1$

3: **repeat**

4:  $n \leftarrow n + 1$

5: **case** Anisotropic:

6: **for**  $i = 1 : C$  **do**

7: Update splitting variable for smoothness regularizer term via shrinkage:

$$\vec{b}_i^{n+1} \leftarrow S_{\frac{1}{\alpha}} \left( D(\vec{g}^n - \vec{f}_i) - \frac{\vec{q}_i^n}{\alpha} \right)$$

8: Update Lagrangian multiplier for regularizer term via dual ascent:

$$\vec{q}_i^{n+1} \leftarrow \vec{q}_i^n + \tau\alpha(\vec{b}_i^{n+1} - D(\vec{g}^n - \vec{f}_i))$$

9: **end for**

10: **case** Isotropic:

11: **for**  $i = 1 : C, j = 1 : R$  **do**

12: Update splitting variables for smoothness regularizer term via vectorial shrinkage:

$$\langle a_{i,j}^{n+1}, b_{i,j}^{n+1} \rangle \leftarrow S_{\frac{1}{\alpha}} \left( \langle \delta_x f_{i,j} + \frac{p_{i,j}^n}{\alpha}, \delta_y(f_{i,j} - g_j^n) + \frac{q_{i,j}^n}{\alpha} \rangle \right)$$

13: Update Lagrangian multipliers for regularizer term via dual ascent:

$$p_{i,j}^{n+1} \leftarrow p_{i,j}^n + \tau\alpha(\delta_x f_{i,j} + \frac{p_{i,j}^n}{\alpha})$$

$$q_{i,j}^{n+1} \leftarrow q_{i,j}^n + \tau\alpha(b_{i,j}^{n+1} - \delta_y(g_j^n - f_{i,j}))$$

14: **end for**

15: Update splitting variable for data fidelity term via shrinkage:

$$\vec{h}^{n+1} \leftarrow S_{\frac{1}{\alpha}} \left( \vec{g}^n - \frac{\vec{r}^n}{\alpha} \right)$$

16: Update striping mask:

$$\vec{g}^{n+1} \leftarrow (CD^T D + \lambda I)^{-1} \left( D^T \left( \sum_i \vec{b}_i^{n+1} + D\vec{f}_i + \frac{\vec{q}_i^{n+1}}{\alpha} \right) + \lambda \vec{h}^{n+1} - \frac{\lambda}{\alpha} \vec{r}^n \right)$$

17: Update Lagrangian multiplier for data fidelity term via dual ascent:

$$\vec{r}^{n+1} \leftarrow \vec{r}^n + \tau\lambda\alpha(\vec{h}^{n+1} - \vec{g}^{n+1})$$

18: Update energy terms:

$$E_1^{n+1} \leftarrow \sum_i \|D(\vec{f}_i - \vec{g}^{n+1})\|_1, \quad E_2^{n+1} \leftarrow \lambda \|\vec{g}^{n+1}\|_1$$

$$E^{n+1} \leftarrow E_1^{n+1} + E_2^{n+1}$$

19: **until** convergence:

$$\|\vec{g}^{n+1} - \vec{g}^n\|_2^2 / \|\vec{g}^n\|_2^2 < \epsilon_g \quad \text{and} \quad |E^{n+1} - E^n|^2 / |E^n|^2 < \epsilon_E$$

20: Retrieve clean image:

$$u \leftarrow f - \vec{g}^{n+1} \otimes [1_1, 1_2, \dots, 1_C]$$

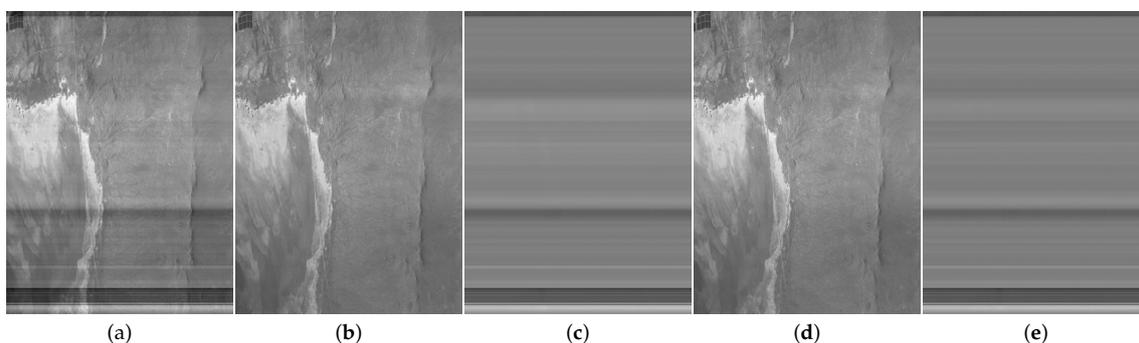
$$U \leftarrow \exp(u)$$

### 3. Results

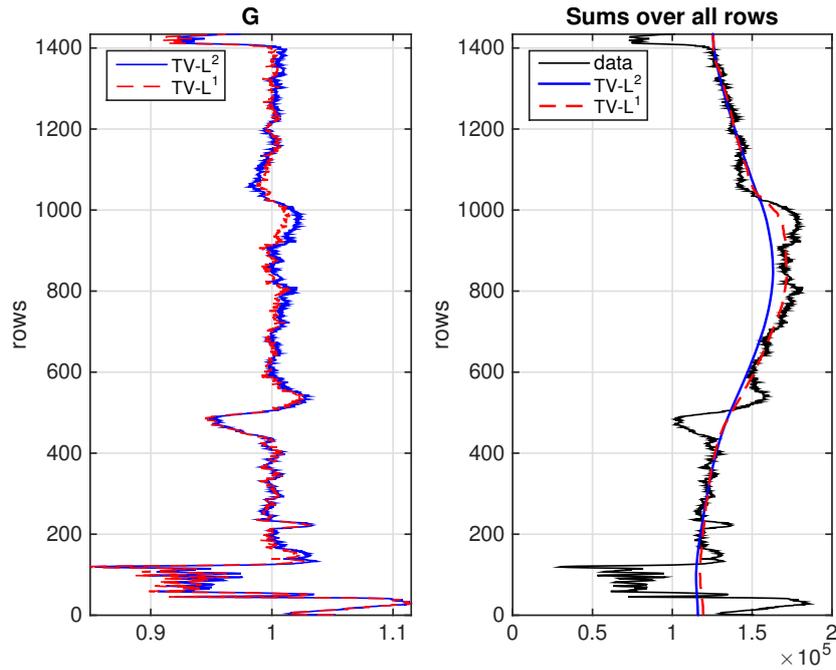
In our experiments, we used data acquired by the AirMSPI. The AirMSPI is an airborne prototype instrument similar to that of the future satellite-borne MSPI instrument for obtaining multi-angle polarization imagery [29]. The instrument was built for NASA by the Jet Propulsion Laboratory in Pasadena, California and has been flying aboard the NASA ER-2 high altitude aircraft since October 2010.

The AirMSPI is an eight-band (355, 380, 445, 470, 555, 660, 865, and 935 nm) pushbroom camera, measuring polarization in the 470, 660, and 865 nm bands, mounted on a gimbal to acquire multi-angular observations over a  $\pm 67^\circ$  along-track range. Two principal observing modes are employed: step-and-stare, in which  $11 \text{ km} \times 11 \text{ km}$  targets are observed at a discrete set of view angles with a spatial resolution of  $\sim 10 \text{ m}$ ; and continuous sweep, in which the camera slews back and forth along the flight track within  $\pm 67^\circ$  to acquire wide-area coverage (11 km swath at nadir; target length of 108 km) with a  $\sim 25 \text{ m}$  spatial resolution. Step-and-stare provides more angles, but continuous sweep gives greater coverage. Multiple observing modes can be programmed into the instrument and activated under cockpit control. Multi-angle radiance and polarization imagery from the AirMSPI will provide 3D scene context where clouds and aerosol plumes are present. It will also enable the retrieval of aerosol and cloud macrophysical properties (distribution and height), microphysical properties (size distribution, single scattering albedo, and shape), and optical depth.

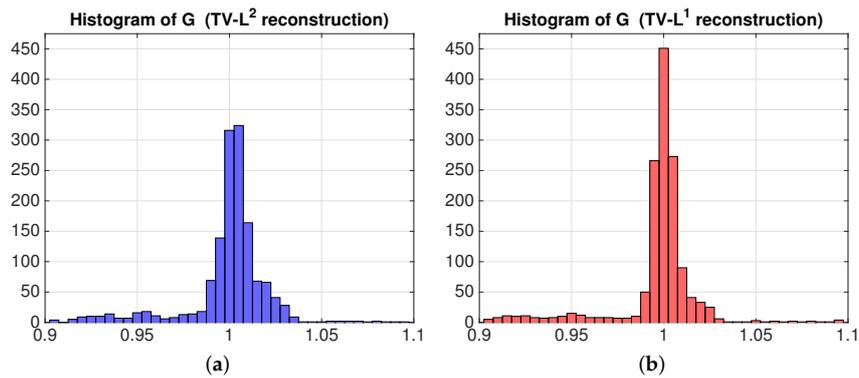
We first compare destriping results generated using  $\text{TV-L}^2$  and the proposed  $\text{TV-L}^1$  models. Figure 1 shows the 355 nm ultraviolet (UV) channel image with stripes captured by the AirMSPI instrument from the nadir angle at Mojave, California. The image has been destriped using the  $\text{TV-L}^2$  and  $\text{TV-L}^1$  models. As we can see from the destriped images and corresponding differences between the captured and destriped images, the  $\text{TV-L}^2$  model does not preserve radiometric intensities in regions where no stripes are present. Figure 2 (left) shows plots of the recovered function  $G$  for the  $\text{TV-L}^2$  and  $\text{TV-L}^1$  destriped images. The  $\text{TV-L}^1$  recovered function  $G$  is closer to the identity, particularly for the rows with no stripes, suggesting the  $\text{TV-L}^1$  reconstruction is more accurate than the  $\text{TV-L}^2$  reconstruction. Figure 2 (right) shows plots of sums over all rows of the original image with stripes (from Figure 1), as well as sums of rows for the  $\text{TV-L}^2$  and  $\text{TV-L}^1$  destriped images. These plots indicate that the  $\text{TV-L}^1$  model preserves radiometric intensities of the original images better, and the  $\text{TV-L}^2$  model produces more artificial smoothing throughout the image. Figure 3 shows histograms of the function  $G$  for the  $\text{TV-L}^2$  and  $\text{TV-L}^1$  reconstructions. The  $\text{TV-L}^1$  reconstruction is pointier than the  $\text{TV-L}^2$  reconstruction. It is also centered at 1, as opposed to the  $\text{TV-L}^2$  reconstruction, which further indicates better accuracy of the  $\text{TV-L}^1$  model. We note that the true stripes, at around  $G \approx 0.95$ , are represented in the histograms by small bumps.



**Figure 1.** (a) The 355 nm channel image with stripes captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument from nadir angle at Mojave, California. (b) Destriped image using total variation  $\text{TV-L}^2$  model. (c) Difference between captured image from (a) and  $\text{TV-L}^2$  destriped image from (b). (d) Destriped image using  $\text{TV-L}^1$  model. (e) Difference between captured image from (a) and  $\text{TV-L}^1$  destriped image from (d).

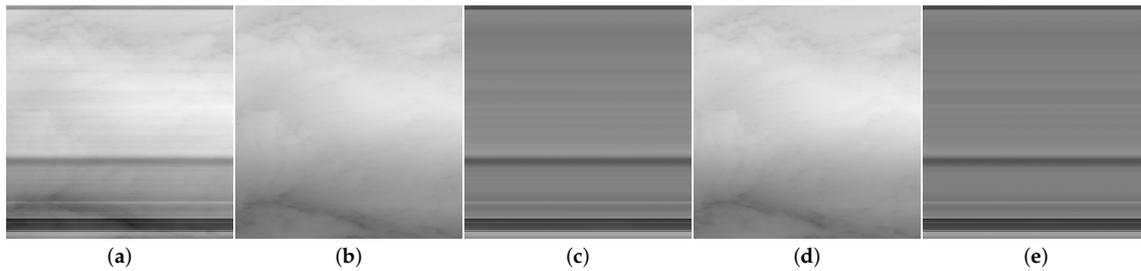


**Figure 2.** Comparisons of total variation  $TV-L^2$  and  $TV-L^1$  destriping for the results shown in Figure 1. On the left, plots of recovered function  $G$  for  $TV-L^2$  destriped image (blue) and  $TV-L^1$  destriped image (red) are shown.  $TV-L^1$  recovered function  $G$  is closer to the identity, particularly for the rows with no stripes, suggesting  $TV-L^1$  reconstruction is more accurate than  $TV-L^2$  reconstruction. On the right, plots of sums over all rows of original image with stripes (black),  $TV-L^2$  destriped image (blue), and  $TV-L^1$  destriped image (red) are shown. These plots indicate that  $TV-L^1$  model preserves radiometric intensities of the original images better, and  $TV-L^2$  model produces more artificial smoothing throughout the image.

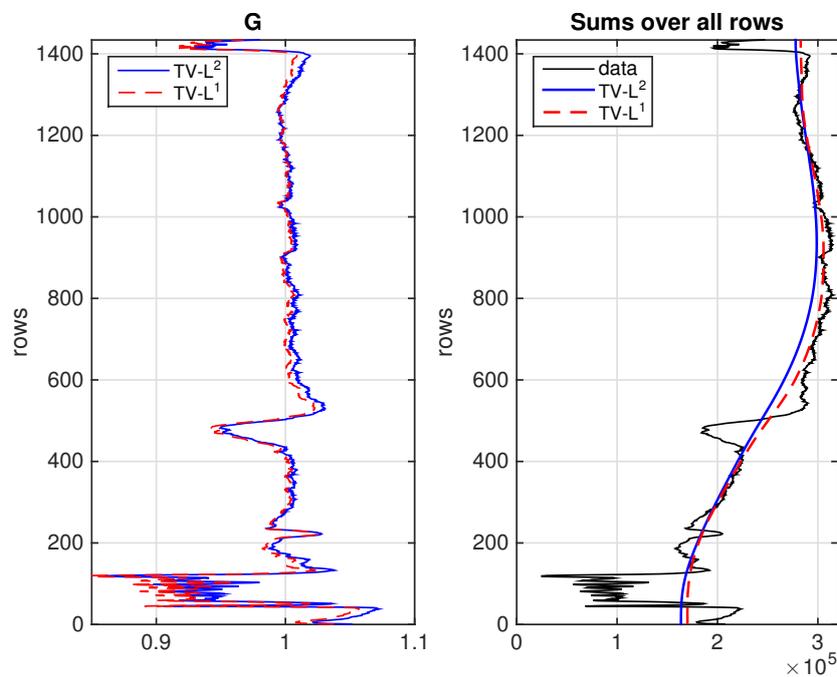


**Figure 3.** Comparisons of total variation  $TV-L^2$  and  $TV-L^1$  destriping for the results shown in Figure 1. (a) Histogram of function  $G$  for  $TV-L^2$  reconstruction. (b) Histogram of function  $G$  for  $TV-L^1$  reconstruction.  $TV-L^1$  reconstruction is pointier than  $TV-L^2$  reconstruction. It is also centered at 1, as opposed to  $TV-L^2$  reconstruction, which further indicates better accuracy of  $TV-L^1$  model.

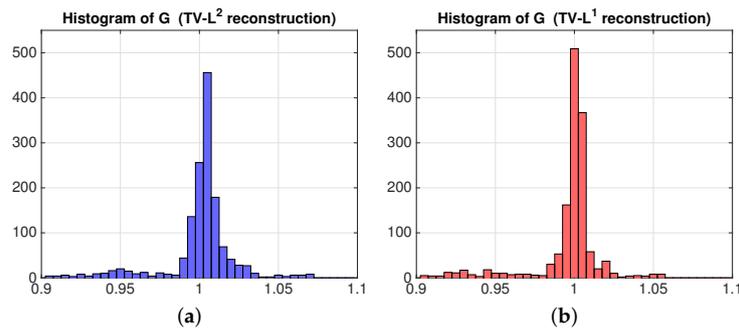
Figures 4–6 display similar results to those shown in Figures 1–3 for the 355 nm channel image with stripes depicting clouds over the Pacific Ocean captured by the AirMSPI instrument from the 66.0°F angle.



**Figure 4.** (a) The 355 nm channel image with stripes depicting clouds over the Pacific Ocean captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument from  $66.0^\circ$ F angle. (b) Destriped image using total variation  $TV-L^2$  model. (c) Difference between captured image from (a) and  $TV-L^2$  destriped image from (b). (d) Destriped image using  $TV-L^1$  model. (e) Difference between captured image from (a) and  $TV-L^1$  destriped image from (d).

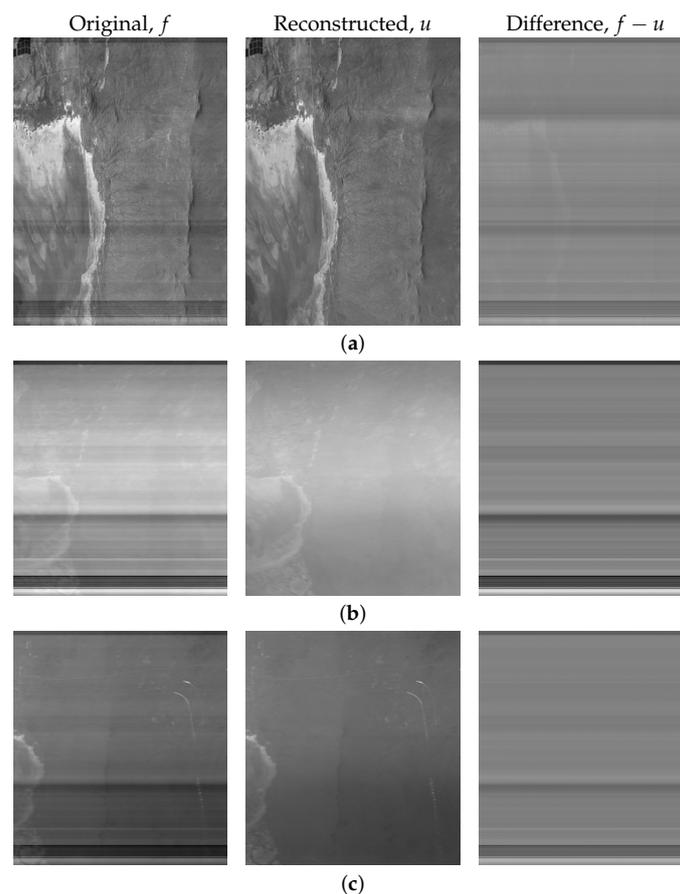


**Figure 5.** Comparisons of total variation  $TV-L^2$  and  $TV-L^1$  destriping for the results shown in Figure 4. On the left, plots of recovered function  $G$  for  $TV-L^2$  destriped image (blue) and  $TV-L^1$  destriped image (red) are shown.  $TV-L^1$  recovered function  $G$  is closer to the identity, particularly for the rows with no stripes, suggesting  $TV-L^1$  reconstruction is more accurate than  $TV-L^2$  reconstruction. On the right, plots of sums over all rows of original image with stripes (black),  $TV-L^2$  destriped image (blue), and  $TV-L^1$  destriped image (red) are shown. These plots indicate that  $TV-L^1$  model preserves radiometric intensities of the original images better, and  $TV-L^2$  model produces more artificial smoothing throughout the image.

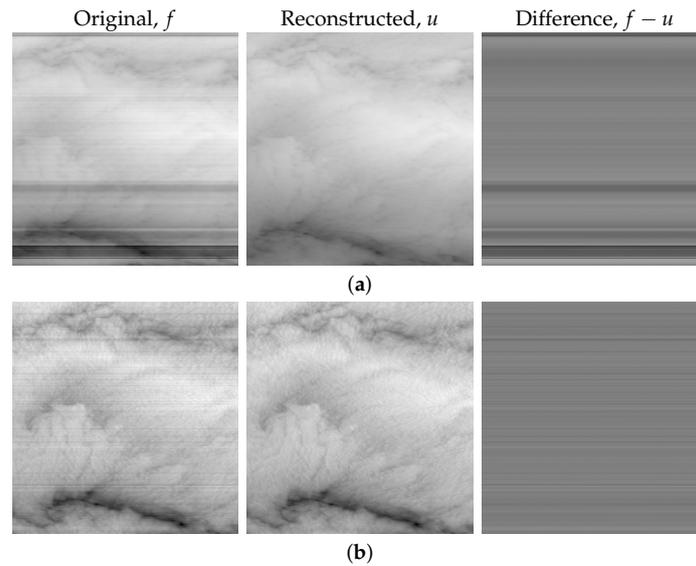


**Figure 6.** Comparisons of total variation  $TV-L^2$  and  $TV-L^1$  destriping for the results shown in Figure 4. (a) Histogram of function  $G$  for  $TV-L^2$  reconstruction. (b) Histogram of function  $G$  for  $TV-L^1$  reconstruction.  $TV-L^1$  reconstruction is pointier than  $TV-L^2$  reconstruction. It is also centered at 1, as opposed to  $TV-L^2$  reconstruction, which further indicates better accuracy of  $TV-L^1$  model.

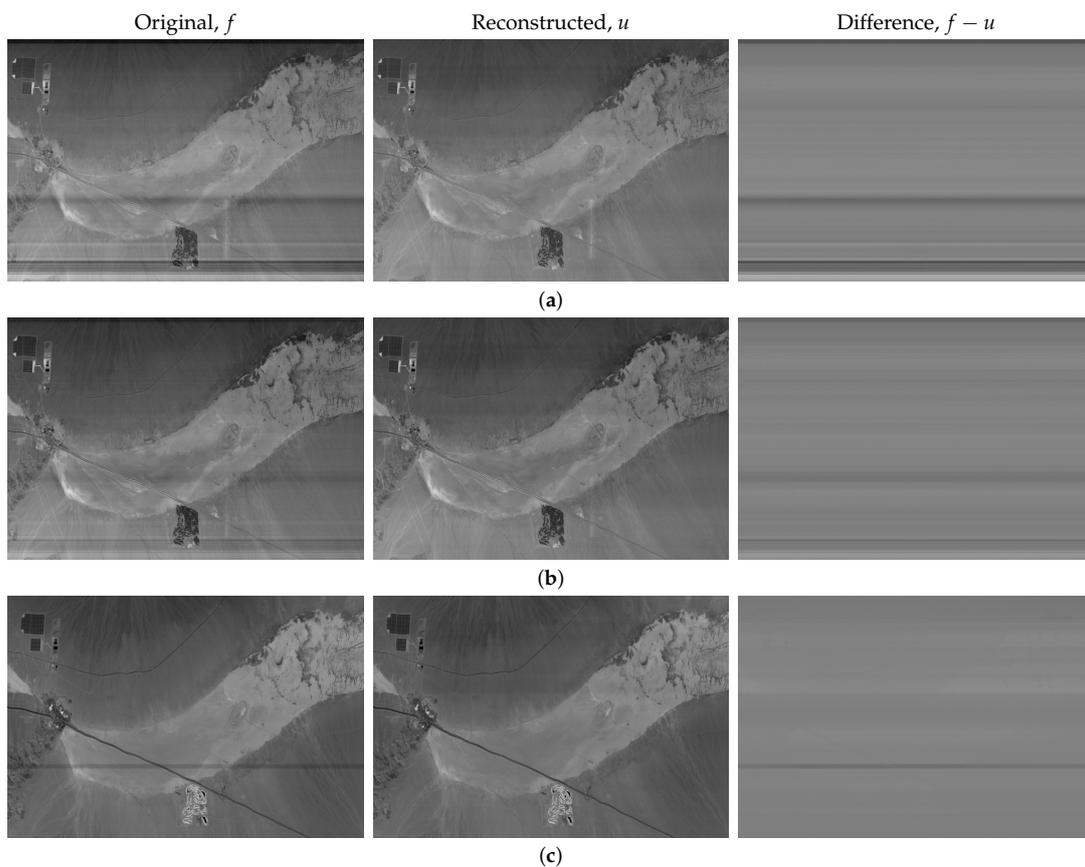
Because the  $TV-L^1$  model generates results with better accuracy than the  $TV-L^2$  model, we focus on the  $TV-L^1$  model in the following examples. Figures 7–9 show more examples of the  $TV-L^1$  reconstruction of images captured using the continuous-sweep observing mode. Figures 10 and 11 display images captured using the step-and-stare observing mode as well as destriped results using the  $TV-L^1$  model.



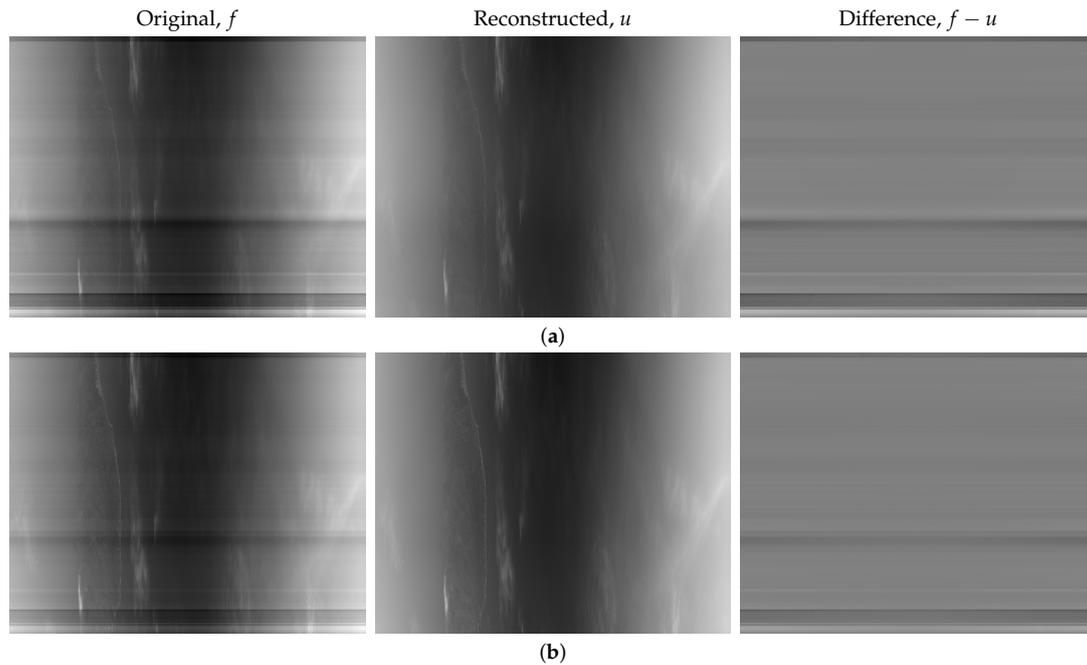
**Figure 7.** Images with stripes captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument at Mojave, California (left); destriped images using total variation  $TV-L^1$  model (center); and differences between captured and destriped images (right) are shown. The bands and viewing angles are (a) 380 nm band at nadir angle, (b) 355 nm band at  $66.1^\circ$  angle, and (c) 355 nm band at nadir angle; 355 nm band at nadir angle is shown in Figure 1.



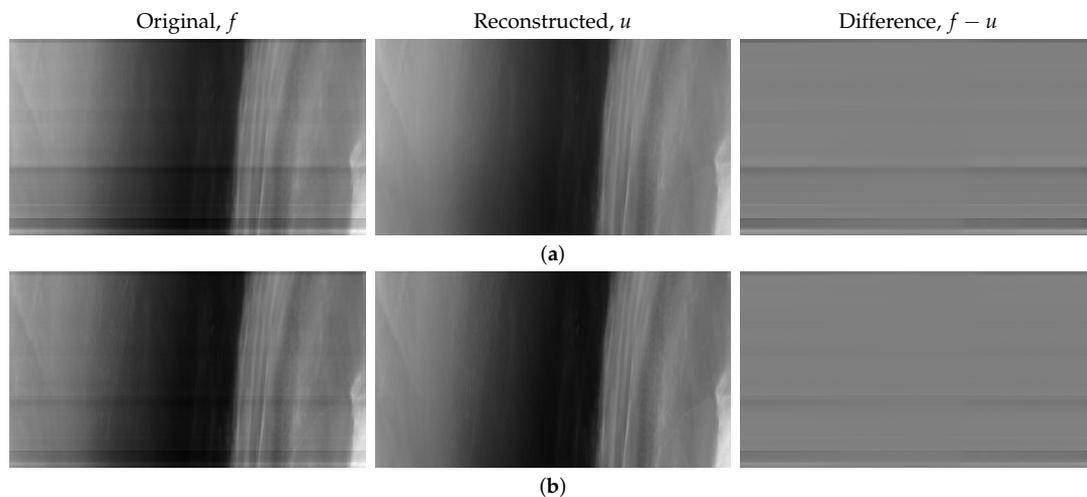
**Figure 8.** Images with stripes depicting clouds over the Pacific Ocean captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument (**left**), destriped images using total variation  $TV-L^1$  model (**center**), and differences between captured and destriped images (**right**) are shown. The bands are (a) 380 nm, and (b) 660 nm, all at  $66.0^\circ F$ ; 355 nm band at  $66.0^\circ F$  angle is shown in Figure 4.



**Figure 9.** Images with stripes of dry Ivanpah Lake, California, captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument (**left**); destriped images using total variation  $TV-L^1$  model (**center**); and differences between captured and destriped images (**right**) are shown. The bands are (a) 355 nm, (b) 380 nm, and (c) 865 nm, all at nadir angle.



**Figure 10.** Images captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument at Avalon, California (**left**); destriped images using total variation  $TV-L^1$  model (**center**); and differences between captured and destriped images (**right**) are shown. The bands are (a) 355 nm, and (b) 380 nm, both captured using the continuous-sweep observing mode.



**Figure 11.** Images with stripes captured by Airborne Multi-angle Spectro Polarimetric Imager (AirMSPI) instrument at Fallbrook, California (**left**); destriped images using total variation  $TV-L^1$  model (**center**); and differences between captured and destriped images (**right**) are shown. The bands are (a) 355 nm, and (b) 380 nm, both captured using the continuous-sweep observing mode.

#### 4. Discussion

We have provided detailed derivations and a motivated evolution of the optimization problem so that the proposed  $TV-L^1$  method can be accessible to all academic disciplines involved with image processing. We have presented a variational model that is well-defined, qualitatively motivated, and easily minimized. The constructed energy uses sparsity-promoting energy functionals, on the basis of TV and  $L^1$  energy, to achieve minimally invasive destriping. Both isotropic and anisotropic TV,

along with  $L^1$  energy, are considered in our variational model. The ADMM is used in conjunction with nonlinear proximal operators to efficiently optimize the energy, yielding quick and high-quality results.

The  $L^2$  fidelity overly penalizes stripes of large magnitude and likewise underexaggerates the significance of stripes of small magnitude. In areas of no striping, we intend our striping mask to be very close to the identity, while in areas of heavy striping, we wish to remove such striping and thus yield a larger magnitude of our striping mask in that region. The  $L^1$  fidelity gives a smaller striping mask in areas of no striping, leaving enough energy to remove the heavier striping in localized areas of the image. Because there is no prior knowledge of the distribution of the stripes and qualitatively we may wish to remove deep striping effects while preserving sharp geometry, we conclude that the model with an  $L^1$  striping penalty is more effective.

The TV- $L^1$  model preserves radiometric intensities of the original images better, and the TV- $L^2$  model produces more artificial smoothing throughout the image. TV- $L^1$  recovers the function  $G$  that is closer to the identity, particularly for the rows with no stripes, suggesting that the TV- $L^1$  reconstruction is more accurate than the TV- $L^2$  reconstruction. The histograms of the function  $G$  for the TV- $L^1$  reconstruction are more compact around 1 than those for the TV- $L^2$  reconstruction. They are also centered at 1, as opposed to the TV- $L^2$  reconstruction, which further indicates the better accuracy of the TV- $L^1$  model.

## 5. Conclusions

In this paper, we have presented a novel variational method for image destriping through fast minimization techniques of appropriately modeled energy functionals, namely TV and the  $L^1$  data fidelity term. Our destriping model solves the inverse problem as follows: it minimally removes a univariate multiplicative striping mask from the data, such that the clean image is somewhat smooth and the removed stripe has low energy. We assess the smoothness of the clean image using the TV, which maintains sharp image features and preserves definition and contrast. We address both isotropic and anisotropic TV in this paper, each having their respective strengths and weaknesses. We use  $L^1$  and, for comparison,  $L^2$  energy to measure the removed striping, ensuring minimal data removal and thus a clean image of high fidelity.

The variational problem is solved efficiently using an ADMM approach: it introduces splitting variables and quadratic penalties for deviations from the splitting variables to allow for efficient optimization via proximal shrinkage operators, explicit quadratic solutions, and simple gradient ascent for the Lagrangian multipliers. In our experiments, we have shown that the proposed TV- $L^1$  method yields qualitatively good results and removes minimal masking, and it does so quickly in terms of both iterations and time. From the histogram distributions of  $G$ , we observe a narrower spread around 1, yet a wider, more equidistributed support, suggesting that most of the time, there is minimal masking removal (multiplier close to 1); however, in areas of heavy striping, the destriping effect is more prevalent and is of greater magnitude.

Applications of this algorithm are not limited to satellite imagery, and they may be analogized to other fields, such as raster scans in microscopy. Any scientific measurements (of images) made mostly along a curve—parameterizable by a single dimension—may be susceptible to such striping biases and may be a candidate for similar destriping. Future work will expand this model to multi-modal images and color images, and it may incorporate other specific priors on the data.

**Acknowledgments:** This work was supported by the National Science Foundation under Grant DMS 1217239 and by the W. M. Keck Foundation. The research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The authors would like to thank Andrea Bertozzi, Michael Bull, David Diner, Michael Garay, Veljko Jovanovic, Stanley Osher, Brian Rheingans, Luminita Vese, and Dominique Zosso for their helpful comments and discussions. The authors also thank the AirMSPI team for providing data for the experiments.

**Author Contributions:** Igor Yanovsky and Konstantin Dragomiretskiy designed the algorithm, conceived and designed the experiments, analyzed the data, and wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ADMM	Alternating direction method of multipliers
AFM	Atomic force microscopy
AirMSPI	Airborne Multi-angle Spectro Polarimetric Imager
PDE	Partial differential equation
TV	Total variation
UV	Ultraviolet

## References

1. Carfantan, H.; Idier, J. Statistical Linear Destriping of Satellite-Based Pushbroom-Type Images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1860–1871.
2. Shen, H.; Zhang, L. A MAP-based algorithm for destriping and inpainting of remotely sensed images. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1492–1502.
3. Srinivasan, R.; Cannon, M.; White, J. Landsat data destriping using power spectral filtering. *Opt. Eng.* **1988**, *27*, 939–943.
4. Crippen, R.E. A simple spatial filtering routine for the cosmetic removal of scan-line noise from LANDSAT TM P-tape imagery. *Photogramm. Eng. Remote Sens.* **1989**, *55*, 327–331.
5. Helder, D.L.; Quirk, B.K.; Hood, J.J. A technique for the reduction of banding in LANDSAT Thematic Mapper images. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 1425–1431.
6. Simpson, J.J.; Stitt, J.R.; Leath, D.M. Improved finite impulse response filters for enhanced destriping of geostationary satellite data. *Opt. Eng.* **1998**, *66*, 235–249.
7. Chen, J.; Shao, Y.; Guo, H.; Wang, W.; Zhu, B. Destriping CMODIS data by power filtering. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2119–2124.
8. Torres, J.; Infante, S.O. Wavelet analysis for the elimination of striping noise in satellite images. *Opt. Eng.* **2001**, *40*, 1309–1314.
9. Chen, J.; Lin, H.; Shao, Y.; Yang, L. Oblique striping removal in remote sensing imagery based on wavelet transform. *Int. J. Remote Sens.* **2006**, *27*, 1717–1723.
10. Rakwatin, P.; Takeuchi, W.; Yasuoka, Y. Stripe noise reduction in MODIS data by combining histogram matching with facet filter. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1844–1856.
11. Zhou, G.; Fang, H.; Yan, L.; Zhang, T.; Hu, J. Removal of stripe noise with spatially adaptive unidirectional total variation. *Opt. Int. J. Light Electron Opt.* **2014**, *125*, 2756–2762.
12. Zhang, Y.; Zhou, G.; Yan, L.; Zhang, T. A destriping algorithm based on TV-Stokes and unidirectional total variation model. *Opt. Int. J. Light Electron Opt.* **2016**, *127*, 428–439.
13. Chang, Y.; Fang, H.; Yan, L.; Liu, H. Robust destriping method for unidirectional total variation and framelet regularization. *Opt. Express* **2013**, *21*, 23307–23323.
14. Chang, Y.; Yan, L.; Fang, H.; Luo, C. Anisotropic Spectral-Spatial Total Variation Model for Multispectral Remote Sensing Image Destriping. *IEEE Trans. Image Process.* **2015**, *24*, 1852–1866.
15. Dobrosotskaya, J.A.; Bertozzi, A.L. A Wavelet-Laplace Variational Technique for Image Deconvolution and Inpainting. *IEEE Trans. Image Process.* **2008**, *17*, 657–663.
16. Chan, T.F.; Vese, L.A. Active Contours Without Edges. *IEEE Trans. Image Process.* **2001**, *10*, 266–277.
17. Tikhonov, A.N.; Arsenin, V.Y. *Solutions of Ill-Posed Problem*; Winston and Sons: Washington, DC, USA, 1977.
18. Muench, B.; Trtik, P.; Marone, F.; Stampanoni, M. Stripe and ring artifact removal with combined wavelet–Fourier filtering. *Opt. Express* **2009**, *17*, 8567–8591.
19. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D.* **1992**, *60*, 256–268.
20. Rudin, L.; Osher, S. Total variation based image restoration with free local constraints. In Proceedings of the IEEE International Conference Image Processing, Austin, TX, USA, 13–16 November 1994, Volume 1, pp. 31–35.
21. Donoho, D.L.; Johnstone, I.M. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Stat. Assoc.* **1995**, *90*, 1200–1224.

22. Wang, Y.; Yin, W.; Zhang, Y. *A Fast Algorithm for Image Deblurring With Total Variation Regularization*; CAAM Tech. Rep. TR07-10; Rice University: Houston, TX, USA, 2007.
23. Rockafeller, R.T. A dual approach to solving nonlinear programming problems by unconstrained optimization. *Math. Program.* **1973**, *5*, 354–373.
24. Glowinski, R.; Marrocco, A. Sur l'approximation, par elements finis d'ordre un, et la resolution, par penalisation-dualité, d'une classe de problems de Dirichlet non lineares. *Rev. Fr. Autom. Inform. Rech. Opérationnelle* **1975**, *9*, 41–76.
25. Gaba, D.; Mercier, B. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Comput. Math. Appl.* **1976**, *2*, 17–40.
26. Hestenes, M.R. Multiplier and Gradient Methods. *J. Optim. Theory Appl.* **1969**, *4*, 303–320.
27. Wang, Y.; Yang, J.; Yin, W.; Zhang, Y. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.* **2008**, *1*, 248–272.
28. Goldstein, T.; Osher, S. The split bregman method for L1 regularized problems. *SIAM J. Imaging Sci.* **2009**, *2*, 323–343.
29. Diner, D.J.; Xu, F.; Garay, M.J.; Martonchik, J.V.; Rheingans, B.E.; Geier, S.; Davis, A.; Hancock, B.R.; Jovanovic, V.M.; Bull, M.A.; et al. The Airborne Multiangle SpectroPolarimetric Imager (AirMSPI): A new tool for aerosol and cloud remote sensing. *Atmos. Meas. Tech.* **2013**, *6*, 2007–2025.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).