International Journal of Modern Communication Technologies & Research (IJMCTR)
ISSN: 2321-0850, Volume-1, Issue-9, November 2013

REMOTE SENSING IMAGE DENOISING USING PDE &AUXILLARY IMAGES AS PRIOR AND COMPARISON WITH OTHER METHODS

Aswani P V

Abstract— **Several classes of denoising algorithms such as total variation (TV) , wavelets and nonlocal means have all achieved much success. These algorithms are based on different theories, and all show good performance in denoising. When denoising an image, the TV method makes use of the geometric features of the image, the wavelet method makes use of the statistical features of the coefficients, and the nonlocal means method makes use of the redundancy in the image texture features. However, the features that have been used by these methods all come from the noisy image itself. In fact, the image features acquired by other sensors from the same scene can also be used as priors in denoising. In many situations involving multicomponent remotesensing images, a single-component image with a higher SNR or higher spatial resolution is often available. In the past, such an auxiliary image was applied for fusion with a multispectral image to improve its spatial resolution. In fact, the auxiliary noise-free image is a more suitable aid to denoising. In an auxiliary image as a prior is used to assist in denoising or deblurring the image, but this is not suitable for a remote-sensing image, and the auxiliary noisy image must come from the same sensor. There has been some research on denoising based on such an auxiliary image in the remotesensing field. The goal of our approach is denoising. In this letter, following the ideas and, the correlation between the different bands of multicomponent images is used. The auxiliary image as the prior is introduced into the TV or partial differential equation (PDE) denoising method. Moreover, the auxiliary image is applied in the form of a "noise-free" single-component image (no image is completely noise-free and by "noise-free" mean "with a high SNR").**

Index Terms— **total variation, wavelet method, pde.**

I. INTRODUCTION

In recent years, several classes of denoising algorithms such as total variation (TV),wavelets and nonlocal means have all achieved much success. These algorithms are based on different theories, and all show good performance in denoising. When denoising an image, the TV method makes use of the geometric features of the image, the wavelet method makes use of the statistical features of the coefficients, and the nonlocal means method makes use of the redundancy in the image texture features. However, the features that have been used by these methods all come from the noisy image itself. In fact, the image features acquired by other sensors from the same scene can also be used as priors in denoising.

Manuscript received November 19, 2013.

ASWANI P V, received Btech in Electronics and Communication Engg from CUSAT University Kerala. Currently pursuing Mtech in Signal Processing at Govt Engg college, Cherthala under Cochin University Of Science And Technology.

In many situations involving multicomponent remote sensing images, a single-component image with a higher SNR or higher spatial resolution is often available. In the past, such an auxiliary image was applied for fusion with a multispectral image to improve its spatial resolution. In fact, the auxiliary noise-free image is a more suitable aid to denoising. Normally an auxiliary image as a prior is used to assist in denoising or deblurring the image, but this is not suitable for a remote-sensing image, and the auxiliary noisy image must come from the same sensor. There has been some research on denoising based on such an auxiliary image in the remotesensing field. In hyperspectral images, the infrared part of the spectrum contains noise near the water-vapor absorption band. To denoise these bands, image bands from other parts of the spectrum can be applied as noise-free images. Recently, a multispectral and hyperspectral image denoising algorithm has been proposed and has achieved good results, where within the Bayesian framework, the extra initial information is included in the form of a noise-free single-band image. The goal of our approach is denoising. In this case,the correlation between the different bands of multicomponent images is used. The auxiliary image as the prior is introduced into the TV or partial differential equation (PDE) denoising method. Moreover, the auxiliary image is applied in the form of a "noise-free" single-component image (no image is completely noise-free and by "noise-free" we mean "with a high SNR"). To illustrate the proposed method, we experiment on the multispectral and hyperspectral remote sensing.

Visual information transmitted in the form of digital images is becoming a major method of communication in the modern age, but the image obtained after transmission is often corrupted with noise. The received image needs processing before it can be used in applications. Image denoising involves the manipulation of the image data to produce a visually high quality image.

II. TOTAL VARIATION DENOISING

Reduces the total-variation of the image. Filters out noise while preserving edges. Textures and fine-scale details are also removed. In this demo the assumption is that a white Gaussian noise is added with a-priori known (or estimated) noise power (variance). The fidelity term to the input image is calculated automatically so that the power of the noise is reduced.

Fig.1: Noisy input   Filtered by TV

A. TEXTURE PRESERVING TV

Reduces selectively the total-variation of the image. Generalization of the TV process to adaptive power constraints. Denoising is strong in smooth regions and weaker in textured regions. Preserves better texture and fine-scale details. This is a two-phase process where the noise and textures are first isolated by scalar TV. The adaptive process then imposes local power constraints based on local variance measures of the first phase.

a) ANALYSIS

We assume that the image acquisition system may be modelled by the following image formation model

$z = h * u + n$, (1)

where u: R2 → R denotes the ideal undistorted image, h: R2 → R is a blurring kernel, z is the observed image which is represented as a function z : R2 → R, and n is an additive Gaussian white noise with zero mean and standard deviation σ.

Let us denote by the interval $(0,N]^2$. As in most of works, in order to simplify this problem, assume that the functions h and u are periodic of period N in each direction. That amounts to neglecting some boundary effects. Therefore, assume that h, u are functions defined in □ and, to fix ideas, we assume that $h, u \in L2$(□). Our problem is to recover as much as possible of u, from our knowledge of the blurring kernel h, the statistics of the noise n, and the observed image z. The problem of recovering u from z is ill-posed due to the ill-conditioning of the operator $Hu = h * u$. Several methods have been proposed to recover u. Most of them can be classified as regularization methods which may take into account statistical properties (Wiener filters), information theoretic properties, a priori geometric models or the functional analytic behavior of the image given in terms of its wavelet coefficients.

The typical strategy to solve this ill-conditioning is regularization. Probably one of the first examples of regularization method consists in choosing between all possible solutions of (1) the one which minimized the Sobolev (semi) norm of u. Usually, the only information we know about the noise is statistical and limited to an estimate of its mean and its variance. In that case, the model equation (1) is incorporated as a set of constraints for (2): a first constraint corresponding to the assumption that the noise has zero mean, and a second one translating the fact that σ is an upper bound of the standard deviation of n.

This formulation was an important step, but the results were not satisfactory, mainly due to the unability of the previous functional to resolve discontinuities (edges) and oscillatory textured patterns. The smoothness required by the Dirichlet integral is too restrictive and information corresponding to high frequencies of z is attenuated by it. The a priori hypothesis is that functions of bounded variation (the BV model) are a reasonable functional model for many problems in image processing, in particular, for restoration problems. Typically, functions of bounded variation have discontinuities along rectifiable curves, being continuous in some sense (in the measure theoretic sense) away from discontinuities. The discontinuities could be identified with edges. The ability of total variation regularization to recover edges is one of the main features which advocate for the use of this model but its ability to describe textures is less clear, even if some textures can be recovered, up to a certain scale of oscillation.

III. WAVELET DENOISING OF MULTICOMPONENT IMAGES

Here wavelet denoising of multicomponent images using gaussian scale mixture models and a noise-free image as priors is discussed. Bayesian wavelet-based denoising procedure for multicomponent images is proposed. A denoising procedure is constructed that 1) fully accounts for the multicomponent image covariances, 2) makes use of Gaussian scale mixtures as prior models that approximate the marginal distributions of the wavelet coefficients well, and 3) makes use of a noise-free image as extra prior information. It is shown that such prior information is available with specific multicomponent image data of, e.g., remote sensing and biomedical imaging. Experiments are conducted in these two domains, in both simulated and real noisy conditions.

Here, a wavelet-based denoising strategy of multicomponent images is proposed. There, a threshold value was derived using the discrete wavelet transform. The threshold value was *universal*, i.e., independent of the subband, and *soft*, i.e., all wavelet coefficients below the threshold were removed, and all others were shrunk by the threshold value. Later on, threshold values were derived that became adaptive, i.e., dependent of the specific subband. For this, Bayesian approaches were applied, where a prior model for the noise-free signal pdf was assumed. Popular priors are generalized Laplacian models and Gaussian scale mixture (GSM) models. Recently, several wavelet-based procedures for multicomponent images were proposed. Denoising was performed that accounted for the multicomponent image covariances, applying wavelet thresholding, mean squared-error estimation and Bayesian estimation, using different prior models: Gaussian models, GSM models, Laplacian models, and Bernouilli– Gaussian models. From these it is clear that a superior denoising of multicomponent images is obtained when accounting for the full multicomponent image covariance structure, and when a good approximation of the wavelet coefficients marginals is used. A comparative study of these different techniques showed that all heavy tailed prior models outperformed the multinormal model. When compared with each other, they provided similar results and can, thus, be regarded as state of the art denoising techniques for multicomponent images. In this paper, a Bayesian estimator is constructed that makes use of such a heavy tailed multicomponent model for the noise free signal. Moreover, other prior information is applied in the form of a “noise-free” single-component image (of course, no image is completely noise-free, with “noise-free” we mean “of high SNR”). In many practical situations involving multicomponent images, a single-component image of higher spatial resolution and/or SNR is available. In remote sensing, e.g., a higher-resolution sensor might be available. In the past, such auxiliary image was applied for fusion with a multispectral image to improve the latter's spatial resolution while maintaining the spectral information. Several fusion procedures were worked out in the wavelet domain. An example is given by the fusion of Panchromatic data of a SPOT satellite system with Landsat Thematic Mapper multispectral images.In hyperspectral images, the infrared (IR) part of the spectrum contains noisy bands near the water vapor absorption band. To denoise these bands, image bands from other parts of the spectrum could be applied as noise-free images. In medical imaging, different images from the same or different modalities with various resolution and SNR can be acquired, e.g., in functional MR or diffusion tensor MR studies, noisy multicomponent images are acquired using fast acquisition procedures. During the same examination, anatomical information is obtained, that can be applied as an extra diagnostic tool. Such auxiliary image information is regularly applied for visualization or registration purposes. A technique is proposed to enhance the resolution of hyperspectral remote sensing images, with the aid of a high-resolution auxiliary sensor. The goal of our approach is denoising, rather than resolution enhancement. For this, the technique is extended to wavelet domain. In this way, we obtain a multiresolution Bayesian framework that accounts for the correlation between a multicomponent image and an auxiliary noise-free image, in order to improve the SNR of the first. In order to model the wavelet coefficients, multicomponent models should be applied. Since the difference between the three proposed multicomponent heavy-tailed priors in were minor, we conjecture that the choice of the prior in this paper is not too critical as long as it is a heavy tailed multicomponent prior. Since the Laplacian model is not rigorously extendable to the multicomponent case, and the multicomponent Bernouilli–Gaussian model was originally introduced by other authors, in this paper we choose to apply the GSM model. Since the proposed procedure makes use of correlations, it is important that the multicomponent image and the auxiliary image are geometrically coregistered. In some applications, this causes no real problem, since the same sensor is applied for the images. As an example, we will demonstrate the procedure on a hyperspectral remote sensing image to denoise the bands near the water vapor absorbtion band. In some applications the misalignment is small, such as, e.g., in intramodal medical acquisition systems; in others, it should be performed *a priori*. In order to study the misregistration effects on the propoesed procedure, we include a simulation experiment with misaligned images.

A. DENOISING FRAMEWORK

i. Nondecimated Wavelet Transform

The wavelet transform reorganizes image content into a low-resolution approximation and a set of details of different orientations and different resolution scales. A fast algorithm for the discrete wavelet transform is an *iterative filter bank algorithm* of Mallat, where a pair of high-pass and low-pass filters followed by downsampling by two is iterated on the low-pass output. The outputs of the low-pass filter are the so-called *scaling* coefficients and the outputs of the high-pass filter are the wavelet coefficients. At each decomposition level, the filter bank is applied sequentially to the rows and to the columns of the image. Low-pass filtering of both the rows and the columns yields the low-pass LL subband (i.e., approximation subband consisting of the scaling coefficients). Other combinations of low-pass and high-pass filtering yield the wavelet subbands at different orientations: high-pass filtering of rows and low-pass filtering of columns (HL) yields horizontal edges and the opposite combination (LH) yields vertical edges, while high-pass filtering of both the rows and the columns (HH) yields diagonal edges.

The jth decomposition level yields the coefficients at the *resolution scale* 2^j. A full signal representation consists of the scaling coefficients at the resolution level J and of all the wavelet coefficients at the resolution levels 1 to J. The total number of these coefficients, due to downsampling by 2 at each stage, equals the number of the input image samples. This is a *critically sampled* transform with precisely enough coefficients for the perfect reconstruction. Despite their mathematical elegance and a remarkable signal/ image compression ability, critically sampled representations are less attractive for denoising. These representations lack shift invariance, meaning that the wavelet coefficients of a shifted signal differ from the shifted wavelet coefficients of the unshifted signal. In image denoising, better results are offered by *redundant* wavelet representations. In this paper, we use a nondecimated wavelet transform implemented with the algorithm *àtrous*. The algorithm inserts 2^j-1 zeroes between the filter coefficients at the resolution level j. Downsampling of the filter outputs is excluded, so the size of each wavelet subband equals the size of the input image.

ii. Wavelet Processing of Multicomponent Data

A natural way of exploiting the multicomponent correlations is by vector-based processing, operating on all the components simultaneously. Let $s_l^{(j,o,b)}$denote the noise-free wavelet coefficient at *spatial* position l , *resolution* level j, *orientation* subband o , and *image component b* . $x_l^{(j,o,b)}$, and $n_l^{(j,o,b)}$ are the corresponding wavelet coefficients of the observed noisy image and noise, respectively. A vector processing approach groups the $x_l^{(j,o,b)}$ wavelet coefficients of all the components at a given spatial position, within a subband of a given orientation and resolution level into a -dimensional vector

$$\mathbf{x}_l^{(j,o)} = \left[x_l^{(j,o,1)}, \ldots, x_l^{(j,o,B)}\right]^T .$$

Equivalent processing is typically applied to all the wavelet subbands, and, hence, we shall omit the indices that denote the resolution level j and orientation. In each wavelet subband, a multicomponent pixel obeys the additive noise model

$$x=s+n$$

where the probability density function (i.e., *density*) of **n** is a multivariate Gaussian of zero mean and covariance matrix $\mathbf{C_n}: p(\mathbf{n}) = \phi(\mathbf{n}; \mathbf{C_n}). \phi(\mathbf{x}; \mathbf{C})$ is defined as the zero mean multivariate Gaussian distribution with covariance C . Cx denotes the covariance matrix of the noisy vector **x**, and the $\hat{C}$s estimate of the covariance matrix of the unknown noise-free vector **s**. The noise covariance in each wavelet subband is, in general, a scaled version of the input image noise covariance, where the scaling factors depend on the wavelet filter coefficients. With the orthogonal wavelet families that we use in this paper, the noise covariance in all the wavelet subbands is equal to the input image noise covariance. Let Cn denote the noise covariance in the input image. When conducting experiments with real noisy data, the noise covariance will be estimated separately. The signal covariance matrix is estimated as

$$\hat{\mathbf{C}}_{\mathbf{s}} = \mathbf{C}_{\mathbf{x}} - \mathbf{C}_{\mathbf{n}}.$$

Since $\hat{\mathbf{C}}\mathbf{s}$ is a covariance matrix, it needs to be positive semi-positive definite. This is assured by performing an eigen value decomposition and clipping possible negative eigen values to zero. In the rare event of negative eigen values occurring, they were observed to be negligible.

B. ESTIMATION APPROACH AND OPTIMIZATION CRITERION

Various linear and nonlinear (adaptive) methods can be applied in data denoising. We focus on the *Bayesian* approach, where *a priori* knowledge about the distribution of the noisefree data is assumed. In particular, we impose a multicomponent prior distribution (to be called hereafter *prior*) on the noise-free wavelet coefficients in a given subband. As an optimization criterion, we adopt *minimization of the mean squared error*, i.e., the Bayesian risk is a quadratic loss function. The minimum mean-squared error (MMSE) estimate is the posterior conditional mean

$$\begin{aligned}\mathbb{E}\{\mathbf{s}\,|\,\mathbf{x}\} &= \int_{-\infty}^{\infty} \mathbf{s}p(\mathbf{s}|\mathbf{x})d\mathbf{x} \\ &= \frac{\int_{-\infty}^{\infty} \mathbf{s}p(\mathbf{x}|\mathbf{s})p(\mathbf{s})d\mathbf{s}}{\int_{-\infty}^{\infty} p(\mathbf{x}|\mathbf{s})p(\mathbf{s})d\mathbf{s}} \\ &= \frac{\int_{-\infty}^{\infty} \mathbf{s}\phi(\mathbf{x}-\mathbf{s};\mathbf{C}_{\mathbf{n}})p(\mathbf{s})d\mathbf{s}}{\int_{-\infty}^{\infty} \phi(\mathbf{x}-\mathbf{s};\mathbf{C}_{\mathbf{n}})p(\mathbf{s})d\mathbf{s}.}\end{aligned}$$

Assuming, e.g., a Gaussian prior for the noise-free signal as

$$p(\mathbf{s}) = \phi(\mathbf{s};\mathbf{C}_{\mathbf{s}})$$

the above MMSE estimate is the Wiener filter

$$\hat{\mathbf{s}} = \mathbb{E}\{\mathbf{s}|\mathbf{x}\} = \hat{\mathbf{C}}_{\mathbf{s}}(\hat{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_{\mathbf{n}})^{-1}\mathbf{x}.$$

The proposed technique has some advantages over standard denoising of each image component. It accounts for the full covariance structure of the multicomponent image, it makes use of a prior that approximates well the marginal distributions of the wavelet coefficients, and it makes use of a noise-free image as extra prior information. In order to demonstrate the impact of these three advantages on the improvement of the performance, the following denoising strategies are applied to compare.

• Single component denoising, where each band image is treated independently. A Gaussian prior model is assumed. This is the Wiener filter in wavelet domain.

• Multicomponent denoising, where the correlations between the different components are exploited. A Gaussian prior is assumed.

• Multicomponent denoising, assuming the GSM model as a prior to better model the marginal distributions of the wavelet coefficients.

• Multicomponent denoising, using a noise-free single component image, using a Gaussian prior.

• Multicomponent denoising, using a noise-free single component image, using a GSM prior.

If the noise is simulated, the performance of the different denoising techniques can be quantitatively described by the PSNR (in dB), defined for 8-bit images as

$$\mathrm{PSNR}(\hat{S}, S) = 10\log_{10}\left(\frac{255^2}{\mathrm{mse}(\hat{S}, S)}\right)$$

When adding the noise-free image, the difference between Gaussian and GSM model becomes clear; with the Gaussian prior, the noise is not completely removed, while when using the GSM model, the noise is effectively removed and sharp edges are clearly retained. Bayesian wavelet-based denoising procedure for multicomponent images was proposed. The procedure makes use of a noise-free single component image as prior information. The prior model for the wavelet coefficient marginals is a GSM model. Experiments were performed with simulated noise on MRI imagery, multispectral and hyperspectral remote sensing images. The results show a gradual improvement, when applying a technique that 1) fully accounts for the multicomponent image covariances, 2) makes use of Gaussian Scale Mixtures as prior models that approximate well the marginal distributions of the wavelet coefficients, and 3) makes use of a noise-free image as extra prior information.

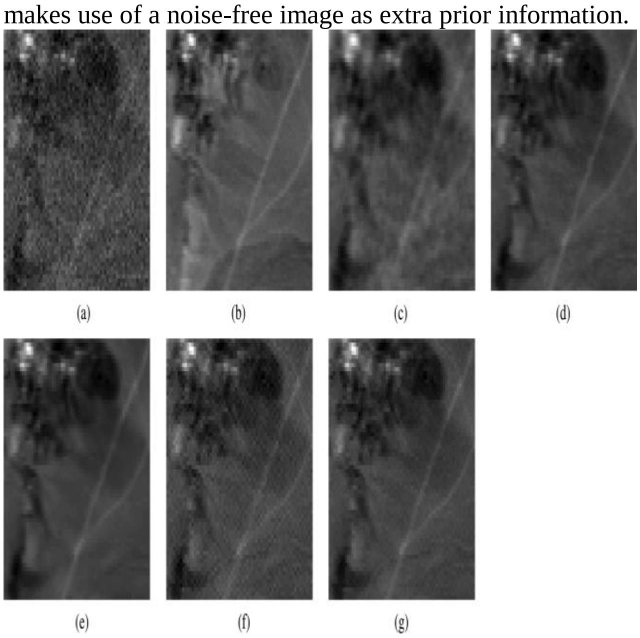


Fig.2.f(a) Detail of original 2.479-_mband of the AVIRIS cuprite image; (b) 1.991_mband of the cuprite image, used as the noise-free image; (c)–(g) results forthe 2.479-_m band after denoising; (c) single-component denoising with Gaussian prior; (d) multicomponent denoising with Gaussian prior; (e) multicomponent denoising with GSM prior; (f) multicomponent denoising with Gaussian prior and noise-free image; (g) multicomponent denoising with GSM prior and noise-free image

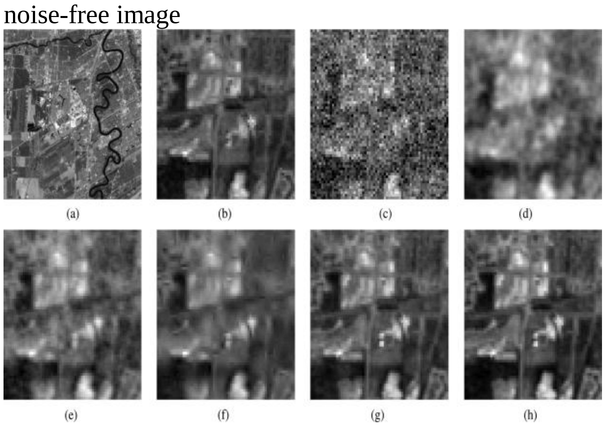


Fig.3.(a) One band of Landsat multispectral image; (b) detail of image; (c) detail image with simulated Gaussian noise (σ = 15); (d) single-component denoising with Gaussian prior; (e) multicomponent denoising with Gaussian prior; (f) multicomponent denoising with GSM prior; (g) multicomponent denoising with Gaussian prior and noise-free image; (h) multicomponent denoising with GSM prior and noise-free image.

IV. REMOTE-SENSING IMAGE DENOISING USING PARTIAL DIFFERENTIAL EQUATIONS AND AUXILIARY IMAGES AS PRIORS

The goal of our approach is denoising. In this approach, the correlation between the different bands of multicomponent images is used. The auxiliary image as the prior is introduced into the TV or partial differential equation (PDE) denoising method. Moreover, the auxiliary image is applied in the form of a "noise-free" single-component image (no image is completely noise-free and by "noise-free" we mean "with a high SNR"). To illustrate the proposed method, we experiment on the multispectral and hyperspectral remotesensing images.

A. IMAGE DENOISING BASED ON PDEs

When an image is corrupted by noise, the following is used:

$$Io = I + n. \qquad (1)$$

In (1), Io is the observed image, I is the original image, and n is the additive noise in the observed image. Usually, n is assumed to follow a Gaussian distribution with a zero mean and a variance of $\sigma 2$. The TV denoising model,

$$\mathrm{TV}(I) = \int_S |\nabla I| dxdy + \frac{\lambda}{2}\int_S (I - I_o)^2 dxdy. \qquad (2)$$

Here, S is the support area of the image, $\int_S |\nabla I| dxdy$ is the regularization term, and λ is the regularization parameter. ∇I denotes the gradient of I, and $|\nabla I|$ is the modulus of ∇I. Minimizing the object function of (2) with respect to I, we obtain (3) for I the following:

$$\frac{\partial \mathrm{TV}(I)}{\partial I} = \mathrm{div}\left(\frac{\nabla I}{|\nabla I|}\right) + \lambda(I - I_o) = 0. \qquad (3)$$

In (3), $div(\cdot)$ is the divergence operator. To solve (3), the time marching method can be employed, and for the nth iteration, we obtain

$$I^{n+1} = I^n + dt\left(\mathrm{div}\left(\frac{\nabla I^n}{|\nabla I^n|}\right) + \lambda(I^n - I_o)\right). \qquad (4)$$

Thus, the problem of TV denoising is converted to computing the PDE given by (4). In fact, we can consider the corresponding form of $div(\nabla In/|\nabla In|)$ from the perspective of differential geometry as follows:

$$\mathrm{div}\left(\frac{\nabla I^n}{|\nabla I^n|}\right) = \frac{1}{|\nabla I^n|}\xi_{I^n}^T H_{I^n}\xi_{I^n}. \qquad (5)$$

In (5), ξI^n is the unit vector perpendicular to the direction of the image gradient, and $\xi T \,.\, I^n$ is the transposition of ξI^n. HI^n represents the Hessian matrix of the image I^n. Substituting (5) into (4) gives

$$I^{n+1} = I^n + dt\left(\frac{1}{|\nabla I^n|}\xi_{I^n}^T H_{I^n}\xi_{I^n} + \lambda(I^n - I_o)\right). \qquad (6)$$

With regard to the effects of the regularization term $(1/|\nabla In|)\xi^T{}_{In}H_{In}\xi_{In}$, we can see that, on one hand, due to the anisotropic nature of $\xi^T{}_{In}H_{In}\xi_{In}$,the smoothing is always in the direction of ξI^n, which is tangential to the edge, whereas on the other hand, due to the uneven property of $(1/|\nabla In|)$, the smoothing is always weakened where the gradient of the image is large, and this further preserves the edge of the image. In the next section, another smoothing term that relates to an auxiliary image of the same scene is constructed and introduced into (6). The direction of the edges and the strength of the gradients of the auxiliary image will be used in the PDE-based remotesensing image denoising.

B. INTRODUCING AN AUXILIARY IMAGE AS A PRIOR INTO PDE DENOISING

In many situations in the remote-sensing area, multicomponent images are often acquired. Although an image comprised of several bands is corrupted by noise, a single-component image with a higher SNR is often available. For multispectral and hyperspectral images, there are often noise-free image bands that can be used as priors in the denoising process. In this letter, the auxiliary image from another sensor is denoted as the reference image u. Image u is both similar to and different from the noisy image Io, i.e., the image intensity distributions are different, but the edge directions and texture information are similar. If we smooth the noise-free image u using (4) and (5), there is a similar smoothing term as given by

$$\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{1}{|\nabla u|}\xi_u^T H_u \xi_u. \qquad (7)$$

Similar to (5), (7) has anisotropic and uneven properties. However, we do not need to smooth the noise-free u but use the priors of the edges of u. To make use of the priors of u in denoising I^n, we need to construct the new smoothing term referring to the specific smoothing direction and specific smoothing intensity of image u and (7). Since ξ_u determines the direction of the smoothing and $|\nabla u|$ determines the strength of the smoothing, we can still refer to the local feature ξ_u and $|\nabla u|$ of the image u, and there is a proposed hybrid smoothing term, as shown in Fig. 4.

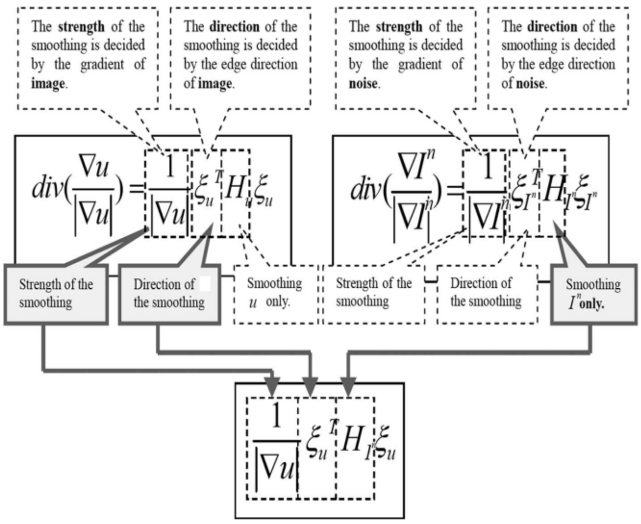


Fig. 4. Introducing the prior of the auxiliary image into the smoothing term

In Fig. 4, the new term $(1/|\nabla u|)\xi^T{}_u H_{In}\xi_u$ that contains both the information of the noise image *In* and the reference image *u* is constructed.For the arbitrary iteration *n*, the edges and gradients in both images *In* and *u* should be taken into consideration in denoising; therefore, we add $(1/|\nabla u|)\xi_u^T H_{I^n}\xi_u$ into (6) to give

$$I^{n+1}=I^n+dt\left(\frac{1}{|\nabla I^n|}\xi_{I^n}^T H_{I^n}\xi_{I^n}+\frac{1}{|\nabla u|}\xi_u^T H_{I^n}\xi_u+\lambda(I^n-I_o)\right). \tag{8}$$

Now, when we denoise an image using (8), the information of the auxiliary image will be taken into consideration. In denoising, the direction of the smoothing refers not only to ξI^n but also to ξ_u, and the strength of the smoothing refers not only to $|\nabla I^n|$ but also to $|\nabla u|$. Therefore, when the auxiliary image *u* is noise free or not as noisy as *Io*, $|\nabla u|$ and ξ_u also contain less noise, and PDE-based denoising, as given by (8), should give better results. When referring to the form of the time marching method [1], we use (9) and (10) to define the discrete forms of $(1/|\nabla I^n|)\xi_{I^n}^T H_{I^n}\xi_{I^n}$ and $(1/|\nabla u|)\xi_u^T H_{I^n}\xi_u$ in (8)

$$\frac{1}{|\nabla I^n|}\xi_{I^n}^T H_{I^n}\xi_{I^n}=\frac{I_{xx}^n\left(I_y^n\right)^2-2I_{xy}^n I_x^n I_y^n+I_{yy}^n\left(I_x^n\right)^2}{\left(\left(I_y^n\right)^2+\left(I_x^n\right)^2+\beta\right)^{\frac{3}{2}}} \tag{9}$$

$$\frac{1}{|\nabla u|}\xi_u^T H_{I^n}\xi_u=\frac{I_{xx}^n(u_y)^2-2I_{xy}^n u_x u_y+I_{yy}^n(u_x)^2}{((u_y)^2+(u_x)^2+\beta)^{\frac{3}{2}}}. \tag{10}$$

In (9) and (10), β is a small constant to prevent a zero denominator. u_x and u_y are the derivatives of *u*. $I^n{}_x$, $I^n{}_{xx}$, $I^n{}_{xy}$,$I^n{}_{xy}$, and $I^n{}_y$ are the derivatives of I^n, the upwind scheme is used in computing u_x, u_y, $I^n{}_x$, $I^n{}_{xx}$, $I^n{}_{xy}$,$I^n{}_{xy}$, and $I^n{}_y$. In the next section, we will do some denoising experiments using (8) and compare its performance with other algorithms.

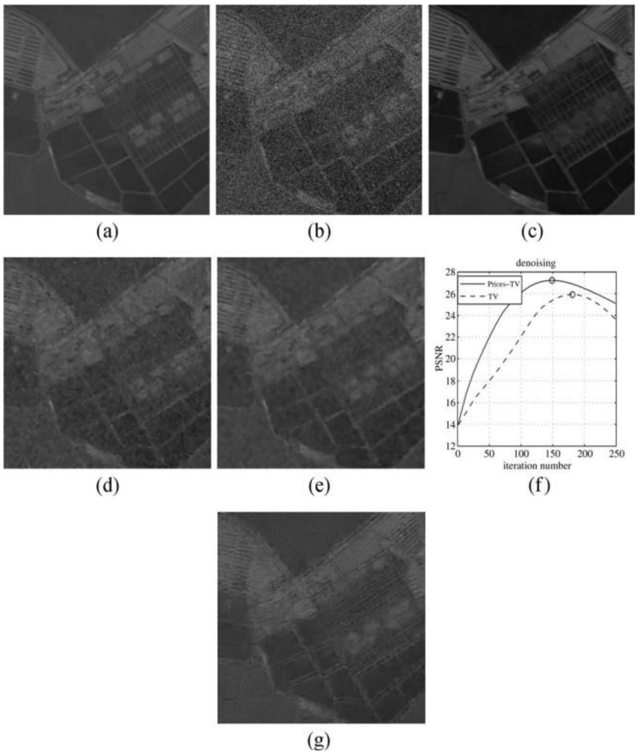


Fig. 5. Denoising the simulated noisy Image of CEBERS. σ = 20. (a) Original image (b) Noisy image. (c) Reference image. (d)TV, 25.92 dB. (e) priors–TV, 27.05 dB. (f) The PSNR in the iteration. (g) Priors–wavelet, 25.61 dB.

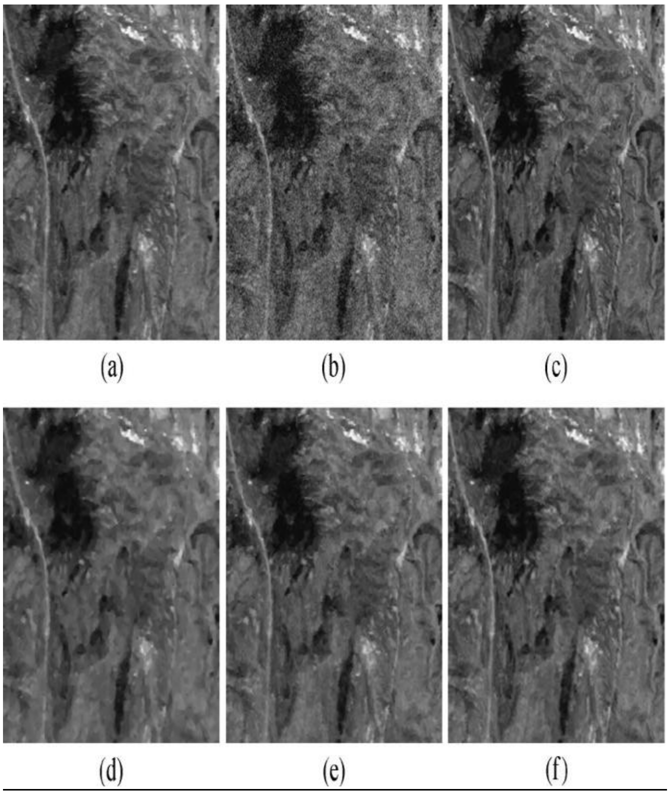


Fig. 6. Denoising the simulated noisy image of AVIRIS. σ = 15 (a) Original image. (b) Noisy image. (c) Reference image. (d) TV, 30.08 dB. (e) Priors–TV, 31.31 dB. (f) Priors–wavelet, 31.24 dB.

V. COMPARISON

To validate and compare the proposed method, we perform the simulation experiments and real-data experiments on different data sets. Multispectral and hyperspectral sensors acquire multicomponent images. These data sets contain both the noisy and noise-free images. A higher quality image can be obtained from one of the sensors or from another part of the reflectance spectrum with a higher SNR. We will apply such an image as the noise-free image in the proposed method.

A. EXPERIMENTS ON SIMULATED NOISY IMAGES

The multispectral images come from the CEBERS satellite. There are five bands in the multispectral image from CEBERS. The noisy image is simulated by contaminating the original image with additive Gaussian noise with the standard deviation σ, and we denoise the simulated noisy third band and use the fourth band as the reference image for the prior in Fig. 5. For the hyperspectral images, the AVIRIS images over Cuprite, Nevada, were taken. The noisy image is simulated by contaminating the original image with additive Gaussian noise with standard deviation σ, and we denoise the simulated noisy hyperspectral band and use the 12th band as the reference image for the prior in Fig. 6. A state-of-the-art wavelet-based method that has also introduced a noise-free image as a prior was compared with the proposed method, and this method was denoted as the priors–wavelet. The TV denoising method without image priors is denoted as TV in the experiments. The proposed PDE method with the priors of the noise-free image is denoted by priors–TV in the experiments. When the noise is simulated, the performance of the different denoising techniques can be quantitatively described by the PSNR (in decibels) .

The PSNR is defined as

$$\mathrm{PSNR}(I,\hat{I}) = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}(I,\hat{I})}\right) \qquad (11)$$

where MSE(I,I^) is the MSE between the original noiseless image I and the denoised image I^.

In Fig. 5, it compares the proposed priors–TV method with the TV method and thepriors–wavelet method by using the multispectral image. Fig. 5(a) is the original image, Fig. 5(b) is the simulated noisy image, Fig. 5(c) is the reference image, and σ is the standard deviation of the initial noise. For the priors–TV method and the TV method, the number of iterations will affect their performance; therefore, the PSNR in the iteration is given in Fig. 5(f), and the optimal location is shown by the circle. Fig. 5(d) and (e) are the optimal results for TV and priors–TV in terms of the iterations. We always use the optimal denoising result of TV and priors–TV in all the experiments when comparing them with priors–wavelet.

Priors–wavelet is not an iteration method; therefore, we do not need to select its optimal result. We can see that priors–TV in Fig. 5(e) is better than TV in Fig. 5(d). There is less noise left in priors–TV than in TV. The similarity of the edges between the noisy image and the reference image helps priors–TV denoising to conserve more details in the smoothing. In Fig. 5(g), there is also less noise in priors–wavelet, but it is over smoothed and artificially overdone. We can see that for multispectral images, although the edges between the noisy image and the reference image are similar, their image intensity distribution is obviously different for the different spectral response functions of the different sensors. In priors–wavelet denoising, the noise-free reference image also enhances the edges of resulting image in Fig. 5(g).

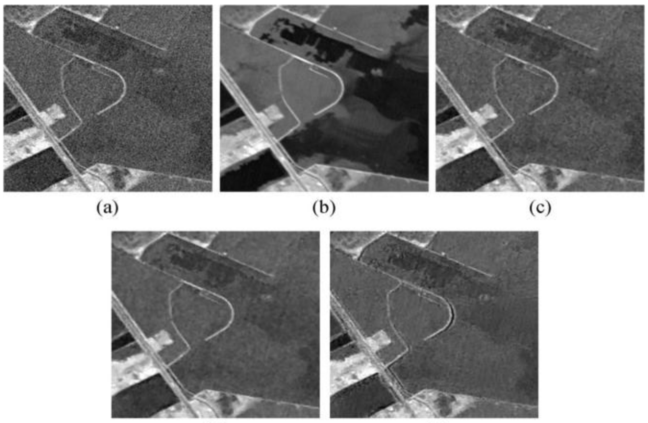
However, because of the noticeable difference in the distribution of the intensity between the different bands of the multispectral images, many fake edges are introduced into the resulting image in Fig. 5(g). In Fig. 6, we compare the proposed method with the TV method and the priors wavelet method using the hyperspectral image. Fig. 6(a) shows the original image, Fig. 6(b) is the simulated noisy image, and Fig. 6(c) is the reference image. σ is the standard deviation of the noise. We can see that, for the hyperspectral image in Fig. 6, the reference image is more similar to the original image than in Fig. 5.

Because the spectrum of the original image in Fig. 6(a) is near 394.9 nm, and the reference image Fig. 6(c) is the 12th band whose spectrum is near 472.7 nm. They are close to each other, and the distributions of the image intensity are similar. However, the situation is different for the multispectral image in Fig. 5. Therefore, in Fig. 6, steadier and stronger priors are introduced into the denoising by the reference image. In this situation, the performances of priors–TV in Fig. 6(e) and priors–wavelet in Fig. 6(f) are very similar, and there are no artificial or fake edges in Fig. 6(f). However, the results of TV in Fig. 6(d) are oversmoothed (relative to priors–TV and priors–wavelet). To further validate the proposed priors–TV method, we compared these methods with different levels of noise and for different images in Table I. In Table I, Images 1 and 2 are multispectral images, and Images 3 and 4 are the hyperspectral images. σ is the standard deviation of the simulated noise. We can see that when the variance of noise is small, priors–TV and priors–wavelet are similar, although each has its own strong point. However, the performance of priors–TV is always better than TV whether the noise is large or small. When the noise is large, the performance of priors–TV is better than that for priors–wavelet.

For the multispectral images (i.e., Images 1 and 2), the proposed priors–TV shows more advantages than priors–wavelet. For the hyperspectral images (i.e., Images 3 and 4), Image 3 is near 394.9 nm, and the reference for Image 3 is the 12th band whose spectrum is near 472.7 nm; they are close to each other, and the distributions of the image intensity are similar. Image 4 is the 189th band of the hyperspectral images whose spectrum is near 2238.4 nm, but the reference image for Image 4 is the 175th band whose spectrum is near 2008.3 nm; therefore, the difference in their image intensity distributions is large. When the reference image is very similar to the original image such as Image 3 in Table I or Fig. 3, the performances of priors–wavelet and priors–TV are close. However, when the spectrum of the reference image is far from the original image such as for Image 4 in Table I, the performance of priors–TV is obviously better than that of priors–wavelet.

TABLE I
COMPARISON OF THE DIFFERENT METHODS WHEN THE SIMULATED NOISY IMAGES HAVE DIFFERENT LEVELS OF NOISE

σ	Method	Image1	Image2	Image3	Image4
5	TV	33.05 dB	30.69 dB	36.01 dB	31.37 dB
	Priors-TV	**33.22** dB	30.92 dB	36.25 dB	31.40 dB
	Priors-Wavelet	32.21 dB	**30.96** dB	**36.38** dB	**31.49 dB**
10	TV	30.26 dB	28.41 dB	32.44 dB	27.90 dB
	Priors-TV	30.63 dB	**28.58** dB	32.81 dB	**28.05** dB
	Priors-Wavelet	**31.19** dB	28.21 dB	**32.92** dB	27.15 dB
15	TV	28.31 dB	27.10 dB	30.08 dB	25.18 dB
	Priors-TV	**29.10** dB	**27.35** dB	**31.31** dB	**25.92** dB
	Priors-Wavelet	28.05 dB	26.33 dB	31.24 dB	25.33 dB
20	TV	25.92 dB	26.51 dB	29.91 dB	24.31 dB
	Priors-TV	**27.05** dB	**26.70** dB	**30.18** dB	**24.93** dB
	Priors-Wavelet	25.61 dB	24.99 dB	30.10 dB	24.24 dB
25	TV	24.78 dB	25.81 dB	28.60 dB	23.91 dB
	Priors-TV	**25.21** dB	**26.16** dB	**29.48** dB	**24.55** dB
	Priors-Wavelet	23.70 dB	24.06 dB	29.27 dB	23.47 dB



(a) (b) (c)
(d) (e)

Fig. 7. Denoising the image of CEBERS. (a) Noisy second band of CEBERS image. (b) Reference image from the fourth band of CEBERS image. (c) Result of TV method. (d) Result of priors–TV method. (e) Result of priors–wavelet method. (a) Noisy image. (b) Reference image. (c) TV. (d) priors–TV.
(e) priors–wavelet.

B. EXPERIMENTS ON REAL NOISY IMAGES

Figs. 5 and 6, and Table I are all simulation experiments. To further validate the proposed priors–TV method, we compared the three methods by using real noisy images in Figs. 7 and 8. In Fig. 7, we denoise the multispectral image of CEBERS and in Fig. 8, we denoise the hyperspectral image of AVIRIS. For multispectral images with relative large noise in Fig. 7, there is less noise left in priors–TV of Fig. 7(d) than in TV of Fig. 7(c). Furthermore, unlike priors–wavelet in Fig. 7(e), there are no fake edges in priors–TV of Fig. 7(d). For hyperspectral images in Fig. 8, the proposed priors–TV in Fig. 8(e) and (g) is also better than TV in Fig. 8(d). When Reference 1 in Fig. 8(b) is used (i.e., the spectrum of the reference image is very similar to the noisy image), the performance of priors–wavelet in Fig. 8(f) is similar to priors–TV in Fig. 8(e), or each has its own strong point; however, if Reference 2 in Fig. 8(c) is used (i.e., the spectrum of the reference image is far from the noisy image), the performance of priors–TV in Fig. 8(g) is better than that of priors–wavelet in Fig. 8(h). Priors–TV mainly makes use of the direction and strength of the edges of the reference image, but the intensity distribution of the reference image has less effect on priors–TV. In general, the experiment results on the real noisy images are similar to the simulation experiments.

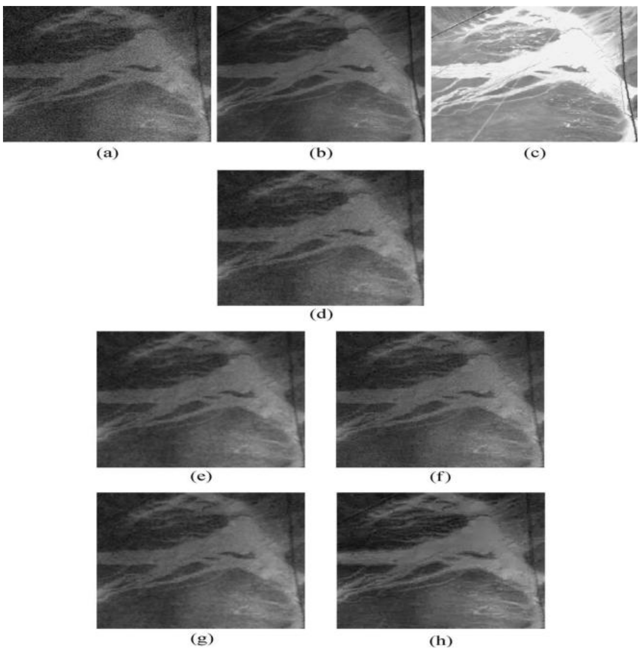


Fig. 8. Denoising the image of AVIRIS. (a) Noisy image is the159th band whose spectrum is near 1871.2 nm. (b) Reference 1 is the 155th band whose spectrum is near 1831.4 nm. (c) Reference 2 is the 130th band whose spectrum is near 1582.3 nm. (d) Result of TV. (e) Result of priors–TV by using Reference 1. (f) Result of priors–wavelet by using Reference 1. (g) Result of priors–TV by using Reference 2. (h) Result of priors–wavelet by using Reference 2. (a) Noisy image. (b) Reference 1 (c) Reference 2. (d) TV.
(e) Priors–TV by Reference 1. (f) Priors–wavelet by Reference 1. (g) Priors–TV
by Reference 2 (h) Priors–wavelet by Reference2.

VI. CONCLUSION

The auxiliary noise-free image has been used as a prior when we denoise one of the noisy images in the multicomponent remote-sensing image. The edge information of the reference image is fully considered, and a new smoothing term reference to the edges is constructed in the proposed method. Comprehensive experiments using different multispectral and hyperspectral images with different levels of noise were carried out. The goal of our approach is denoising. In this case,the correlation between the different bands of multicomponent images is used. The auxiliary image as the prior is introduced into the TV or partial differential equation (PDE) denoising method. Moreover, the auxiliary image is applied in the form of a “noise-free” single-component imageIn particular, when the variance of the noise in the multispectral image is large, the advantage of the proposed method is moreobvious.

REFERENCES

[1]. Peng Liu, Fang Huang, Guoqing Li, and Zhiwen Liu Remote-Sensing Image Denoising Using Partial Differential Equations and Auxiliary Images as Priors Ieee Geoscience And Remote Sensing Letters, Vol. 9, No. 3, May 2012

[2] P. Liu, D. Liu, and Z. Liu, “Total variation restoration of the defocus image based on spectral priors,” in *Proc. SPIE*, 2010, vol. 7830, p. 783 018.

[3] F. Abramovitch, T. Sapatinas, and B.W. Silverman, “Wavelet thresholding via a Bayesian approach,” *J. R. Statist Soc. Ser. B*, vol. 60, no. 4, pp. 725– 749, 1998.

[4]P. Scheunders and S. De Backer, “Wavelet denoising of multicomponent images using gaussian scale mixture models and a noise-free image as priors, image processing,” *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1865–1872, Jul. 2007

ASWANI P V, received Btech in Electronics and Communication Engg from CUSAT University Kerala. Currently pursuing Mtech in Signal Processing at Govt Engg college, Cherthala under Cochin University Of Science And Technology. Areas of interest includes Signal Processing, Image Processing, Artificial Neural Networks.