



IMAGE RECONSTRUCTION USING SPARSE BASED MODEL

Reetumala Thakre, Prof. Rajendra Singh, Dr. Soni Changlani
Department of Electronics & Communication Engineering,
Lakshmi Narain College of Technology & Science,
Bhopal, Madhya Pradesh, INDIA.

Abstract: Sparse representation is proven to be a promising approach to image super-resolution, where the low-resolution (LR) image is usually modeled as the down-sampled version of its high-resolution (HR) counterpart after blurring. When the blurring kernel is the Dirac delta function, i.e., the LR image is directly down-sampled from its HR counterpart without blurring, the super-resolution problem becomes an image interpolation problem. In such cases, however, the conventional sparse representation models (SRM) become less effective, because the data fidelity term fails to constrain the image local structures. In natural images, fortunately, many nonlocal similar patches to a given patch could provide nonlocal constraint to the local structure. Sparse representation models code an image patch as a linear combination of a few atoms chosen out from an over-complete dictionary, and they have shown promising results in various image reconstruction applications. However, due to the degradation of the observed image, the sparse representations by conventional models may not be accurate enough for a faithful reconstruction of the original image. To improve the performance of sparse representation-based image restoration, in this the concept of sparse noise is introduced, and the goal of image reconstruction turns to how to suppress the sparse noise.

I. INTRODUCTION

Reconstructing a high quality image from one or several of its degraded (e.g., noisy, blurred, and/or down sampled) versions has many applications, such as medical imaging, remote sensing, surveillance, and entertainment, etc. For an observed image y , the problem of image reconstruction (IR) can be generally formulated by

$$y = Hx + v \quad (1)$$

Where H a degradation matrix, x is the original image vector and v is the additive noise vector. With different settings of matrix H , Eq. (1) can represent different IR problems; for example, image denoising when H is an identity matrix, image deblurring when H is a blurring operator, image super resolution when H is a composite operator of blurring and down-sampling, and compressive sensing when H is a random projection matrix. In the past decades, extensive studies have been conducted on developing various IR approaches. Due to the nature of IR, the regularization-based techniques have been widely used by regularizing the solution spaces. In order for an effective regularization, it is of great importance to find and model the appropriate prior knowledge of natural images, and various image prior models have been developed. The classic regularization models, such as the TV regularization are effective in removing the noise artifacts but tend to over smooth the images due to the piecewise constant assumption. Mathematically, the sparse representation model assumes that a signal $x \in \mathbb{R}^N$ can be represented as $x \approx \Phi\alpha$, where $\Phi \in \mathbb{R}^{n \times M}$ ($N < M$) is an over-complete dictionary, and most entries of the coding vector α are zero or close to zero. The sparse decomposition of x can be obtained by solving an l_0 -

minimization problem, formulated as: $\alpha_x = \arg \min_{\alpha} \|\alpha\|_0, s.t. \|x - \Phi\alpha\|_2 \leq \varepsilon$

It is a pseudo norm that counts the number of non-zero entries in α , and ε is a small constant controlling the approximation error. Since l_0 -minimization is an upward combinatorial optimization problem, it is often relaxed to the convex l_1 -minimization. The l_1 -norm based sparse coding problem can be generally formulated in the following Lagrangian form:

$$\alpha_x = \arg \min_{\alpha} \{\|x - \Phi\alpha\|_2^2 + \lambda \|\alpha\|_1\} \quad (2)$$

where constant λ denotes the regularization parameter. With an appropriate selection of the regularization parameter λ , we can get a good balance between the sparse approximation error of x and the sparsity of α , and the term "sparse coding" refer to this sparse approximation process of x . Many efficient l_1 - minimization techniques have been proposed to solve Eq. (2), such as iterative thresholding algorithm. In addition, compared with the analytically designed dictionaries (e.g., wavelet/curve let dictionary), the dictionaries learned from

example image patches can improve much the sparse representation performance since they can better characterize the image structures. In the scenario of IR, what we observed is the degraded image signal $\mathbf{y}$ via $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$. To recover $\mathbf{x}$ from $\mathbf{y}$, first $\mathbf{y}$ is sparsely coded with respect to Φ by solving the following minimization problem:

$$\alpha_{\mathbf{y}} = \arg \min_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi\alpha\|_2^2 + \lambda \|\alpha\|_1 \} \quad (3)$$

and then $\mathbf{x}$ is reconstructed by $\hat{\mathbf{x}} = \Phi\alpha_{\mathbf{y}}$. Clearly, it is expected that $\alpha_{\mathbf{y}}$ could be close enough to $\alpha_{\mathbf{x}}$. Due to the degradation of the observed image (e.g., the image is blurry and noisy), however, it is very challenging to recover the true sparse code $\alpha_{\mathbf{x}}$ from $\mathbf{y}$. Using only the local sparsity constraint $\|\alpha\|_1$ in Eq. (3) may not lead to an accurate enough image reconstruction. On the other hand, it is known that image sparse coding coefficients α are not randomly distributed due to the local and nonlocal correlations existing in natural images. One can say that the sparse coding coefficients are correlated, while the strong correlations allow us to develop a much more accurate sparse model by exploiting the local and nonlocal redundancies. Indeed, some recent works, such as are based on such considerations. For example, a group sparse coding scheme was proposed to code similar patches simultaneously, and it achieves impressive denoising results. In this we improve the sparse representation performance by proposing a *non-locally centralized sparse representation* model. To faithfully reconstruct the original image, the sparse code $\alpha_{\mathbf{y}}$ [refer to Eq. (3)] should be as close as possible to the sparse codes $\alpha_{\mathbf{x}}$ [refer to Eq. (2)] of the original image. In other words, the difference $v_{\alpha} = \alpha_{\mathbf{y}} - \alpha_{\mathbf{x}}$, called as sparse coding noise, (SCN) should be reduced and hence the quality of reconstructed image $\hat{\mathbf{x}} = \Phi\alpha_{\mathbf{y}}$ can be improved because $\hat{\mathbf{x}} - \mathbf{x} \approx \Phi\alpha_{\mathbf{y}} - \Phi\alpha_{\mathbf{x}} = \Phi v_{\alpha}$. To reduce the SCN, we centralize the sparse codes to some good estimation of $\alpha_{\mathbf{x}}$. In practice, a good estimation of $\alpha_{\mathbf{x}}$ can be obtained by exploiting the rich amount of nonlocal redundancies in the observed image. The proposed SR model can be solved effectively by conventional iterative shrinkage algorithm, which allows us to adaptively adjust the regularization parameters from a Bayesian viewpoint. The extensive experiments conducted on typical IR problems, including image denoising, deblurring and super-resolution, demonstrate that the proposed SR-based IR method can achieve highly competitive performance to state-of-the-art denoising methods, and outperforms state-of-the-art image deblurring and super-resolution methods.

II. NONLOCALLY CENTRALIZED SPARSE REPRESENTATION (SR)

Using the notation for an image $\mathbf{x} \in \mathbb{R}^N$, let $\mathbf{x}_i = \mathbf{R}_i \mathbf{x}$ denote an image patch of size $\sqrt{n} \times \sqrt{n}$ extracted at location i , where $\mathbf{R}_i$ is the matrix extracting patch $\mathbf{x}_i$ from $\mathbf{x}$ at location i . Given an dictionary $\Phi \in \mathbb{R}^{n \times M}$ ($N < M$) each patch can be sparsely represented as $\mathbf{x}_i \approx \Phi \alpha_{\mathbf{x},i}$ by solving an l_1 -minimization problem $\alpha_{\mathbf{x},i} = \arg \min_{\alpha_i} \{ \|\mathbf{x}_i - \Phi \alpha_i\|_2 + \lambda \|\alpha_i\|_1 \}$. Then the entire image $\mathbf{x}$ can be represented by the set of sparse codes $\{\alpha_{\mathbf{x},i}\}$. The patches can be overlapped to suppress the boundary artifacts, and we obtain a redundant patch-based representation. Reconstructing $\mathbf{x}$ from $\{\alpha_{\mathbf{x},i}\}$ is an over-determined system, and a straightforward least-square solution is:

$$\mathbf{x} \approx \left(\sum_{i=1}^N \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \sum_{i=1}^N \left(\mathbf{R}_i^T \Phi \alpha_{\mathbf{x},i} \right).$$

For the convenience of expression let

$$\mathbf{x} \approx \Phi \circ \alpha_{\mathbf{x}} = \left(\sum_{i=1}^N \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \sum_{i=1}^N \left(\mathbf{R}_i^T \Phi \alpha_{\mathbf{x},i} \right) \quad (4)$$

where $\alpha_{\mathbf{x}}$ denotes the concatenation of all $\alpha_{\mathbf{x},i}$. The above equation is nothing but the overall image is reconstructed by averaging each reconstructed patch of $\mathbf{x}_i$. In the scenario of image reconstruction (IR), the observed image is modeled as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$. The sparsity-based IR method recovers $\mathbf{x}$ from $\mathbf{y}$ by solving the following minimization problem:

$$\alpha_{\mathbf{y}} = \arg \min_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 \}. \quad (5)$$

The image $\mathbf{x}$ is then reconstructed as $\hat{\mathbf{x}} = \Phi \alpha_{\mathbf{y}}$.

A. SPARSE CODING NOISE

In order for an effective IR, the sparse codes $\alpha_{\mathbf{y}}$ obtained by solving the objective function in Eq. (5) are expected to be as close as possible to the true sparse codes $\alpha_{\mathbf{x}}$ of the original image $\mathbf{x}$. However, due to the degradation of the observed image $\mathbf{y}$ (e.g., noisy and blurred), the sparse code $\alpha_{\mathbf{y}}$ will deviate from $\alpha_{\mathbf{x}}$ and the IR quality depends on the level of the sparse coding noise (SCN), which is defined as the difference between $\alpha_{\mathbf{y}}$ and $\alpha_{\mathbf{x}}$.

$$v_{\mathbf{x}} = \alpha_{\mathbf{y}} - \alpha_{\mathbf{x}} \quad (6)$$

To investigate the statistical property of SCN v_α , we perform some experiments on typical IR problems. From that we can see that the empirical distributions of SCN v_α can be well characterized by Laplacian distributions, while the Gaussian distributions have much larger fitting errors.

B. MODELING OF NCSR

The definition of SCN v_α indicates that by suppressing the SCN v_α we could improve the IR output $\hat{x}$. However, the difficulty lies in that the sparse coding vector α_x is unknown so that v_α cannot be directly measured. Nonetheless, if we could have some reasonably good estimation of α_x denoted by β , available, then $\alpha_y - \beta$ can be a good estimation of the SCN v_α . To suppress v_α and improve the accuracy of α_y and thus further improve the objective function of Eq. (5), we can propose the following centralized sparse representation (CSR) model:

$$\alpha_y = \arg \min_{\alpha} \left\{ \|y - H\Phi \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i\|_1 + \gamma \sum_i \|\alpha_i - \beta_i\|_p \right\} \quad (7)$$

where β_i is some good estimation of α_i , γ is the regularization parameter and p can be 1 or 2. In the above CSR model, while enforcing the sparsity of coding coefficients α_i , the sparse codes are also centralized to some estimate of α_x (i.e., β) so that SCN v_α can be suppressed. One important issue of sparsity-based IR is the selection of dictionary. Conventional analytically designed dictionaries, such as DCT, wavelet and curvelet dictionaries, are insufficient to characterize the so many complex structures of natural images. In general the learned dictionaries are required to be very redundant such that they can represent various image local structures. However, it has been shown that sparse coding with an over complete dictionary is unstable, especially in the scenario of image restoration. We cluster the training patches extracted from a set of example images into K clusters, and learn a PCA sub-dictionary for each cluster. Then for a given patch, one compact PCA sub-dictionary is adaptively selected to code it, leading to a more stable and sparser representation, and consequently better image restoration results.

We extract image patches from image x and cluster the patches into K clusters (typically $K = 70$) by using the K -means clustering method. Since the patches in a cluster are similar to each other, there is no need to learn an over-complete dictionary for each cluster. Therefore, for each cluster we learn a dictionary of PCA bases and use this compact PCA dictionary to code the patches in this cluster. In this, for each patch to be coded, we adaptively select one sub-dictionary from the trained K PCA sub-dictionaries to code it. This actually enforces the coding coefficients of this patch over the other sub-dictionaries to be 0, leading to a very sparse representation of the given patch. In other words, our algorithm will naturally ensure the sparsity of the coding coefficients, and thus the local sparsity regularization term $\|\alpha_i\|_1$ can be removed. Hence we propose the following sparse coding model:

$$\alpha_y = \arg \min_{\alpha} \left\{ \|y - H\Phi \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i - \beta_i\|_p \right\} \quad (8)$$

There is only one regularization term $\|\alpha_i - \beta_i\|_p$ in the above model. In the case that $p = 1$, and the estimate β_i is obtained by using the nonlocal redundancy of natural images, this regularization term will become a non-locally centralized sparsity term, and we call this model *non-locally centralized sparse representation*. Let's discuss how to obtain a good estimation β_i of the unknown sparse coding vectors α_i .

C. NONLOCAL ESTIMATE OF UNKNOWN SPARSE CODE

Generally, there can be various ways to make an estimate of α_x depending on how much the prior Knowledge of α_x we have. If we have many training images that are similar to the original image x , we could learn the estimate β of α_x from the training set. However, in many practical situations the training images are simply not available. On the other hand, the strong nonlocal correlation between the sparse coding coefficients, allows us to learn the estimate β from the input data. Based on the fact that natural images often contain repetitive structures, i.e., the rich amount of nonlocal redundancies, we search the nonlocal similar patches to the given patch i in a large window centered at pixel i . For higher performance, the search of similar patches can also be carried out across different scales at the expense of higher computational complexity, as shown in. Then a good estimation of α_i , i.e., β_i , can be computed as the weighted average of those sparse codes associated with the nonlocal similar patches (including patch i) to patch i . For each patch x_i , we have a set of its similar patches, denoted by i . Finally β_i can be computed from the sparse codes of the patches within i . Denote by $\alpha_{i,q}$ the sparse codes of patch $x_{i,q}$ within set i . Then β_i can be computed as the weighted average of $\alpha_{i,q}$

$$\beta_i = \sum_{q \in \Omega_i} \omega_{i,q} \alpha_{i,q} \quad (9)$$

where $\omega_{i,q}$ is the weight. Similar to the nonlocal means approach we set the weights to be inversely proportional to the distance between patches x_i and $x_{i,q}$

$$\omega_{i,q} = \frac{1}{W} \exp(-\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_{i,q}\|_2^2/h) \quad (10)$$

Such a procedure is iterated until convergence. In the l th iteration, the sparse vector is obtained by solving the following

$$\alpha_y^{(l)} = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i - \beta_i^{(l)}\|_p \right\} \quad (11)$$

$$\hat{\mathbf{x}}^{(l)} = \Phi \circ \alpha_y^{(l)}.$$

The restored image is then updated as

In the above iterative process, the accuracy of sparse coding coefficient $\alpha_y^{(l)}$ is gradually improved, which in turn improves the accuracy of β_i . The improved β_i are then used to improve the accuracy of α_y , and so on. Finally, the desired sparse code vector α_y is obtained when the alternative optimization process falls into a local minimum.

III. ALGORITHM

In Eq. (8) or Eq. (11) the parameter λ that balances the fidelity term and the centralized sparsity term should be adaptively determined for better IR performance. In this subsection we provide a Bayesian interpretation of the NCSR model, which also provides us an explicit way to set the regularization parameter λ . In the literature of wavelet denoising, the connection between *Maximum a Posterior* (MAP) estimator and sparse representation has been established and here we extend the connection from the local sparsity to nonlocally centralized sparsity. For the convenience of expression, let's define

$$\theta = \alpha - \beta.$$

For a given β , the MAP estimation of θ can be formulated as

$$\theta_y = \arg \max_{\theta} \log P(\theta|\mathbf{y}) = \arg \max_{\theta} \{\log P(\mathbf{y}|\theta) + \log P(\theta)\}. \quad (12)$$

The likelihood term is characterized by the Gaussian distribution

$$P(\mathbf{y}|\theta) = P(\mathbf{y}|\alpha, \beta) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2\right) \quad (13)$$

Where θ and β are assumed to be independent. In the prior probability $P(\theta)$, θ reflects the variation of α from its estimation β . If we take β as a very good estimation of the sparse coding coefficient of unknown true signal, then $\theta_y = \alpha_x - \beta$ is basically the SCN associated with α_y , and we have seen in Fig. 2 that the SCN signal can be well characterized by the Laplacian distribution. Thus, we can assume that θ follows i.i.d. Laplacian distribution, and the joint prior distribution $P(\theta)$ can be modeled as

$$P(\theta) = \prod_i \prod_j \left\{ \frac{1}{\sqrt{2}\sigma_{i,j}} \exp\left(-\frac{|\theta_i(j)|}{\sigma_{i,j}}\right) \right\} \quad (14)$$

where $\theta_i(j)$ are the j th elements of θ_i and $\sigma_{i,j}$ is the standard deviations of $\theta_i(j)$

Substituting Eqs. (13) and (14) into Eq. (12), we obtain

$$\theta_y = \arg \min_{\theta} \left\{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + 2\sqrt{2}\sigma_n^2 \times \sum_i \sum_j \frac{1}{\sigma_{i,j}} |\theta_i(j)| \right\}. \quad (15)$$

Hence, for a given β the sparse codes α can then be obtained by minimizing the following objective function

$$\alpha_y = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + 2\sqrt{2}\sigma_n^2 \times \sum_i \sum_j \frac{1}{\sigma_{i,j}} |\alpha_i(j) - \beta_i(j)| \right\}. \quad (16)$$

Compared with Eq. (8), we can see that the l_1 -norm (i.e., $p=1$) should be chosen to characterize the SCN term $\alpha_i - \beta_i$. Comparing Eq. (16) with Eq. (8), we have

$$\lambda_{i,j} = \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i,j}}. \quad (17)$$

In order to have robust estimations of $\alpha_{i,j}$, the image nonlocal redundancies can be exploited. In practice, we estimate $\alpha_{i,j}$ using the set of θ_i computed from the nonlocal similar patches $\lambda_{i,j}$ is then updated with the updated

θ in each iteration or in several iterations to save computational cost. Next we present the detailed algorithm of the proposed NCSR scheme.

Algorithm: NCSR-Based Image Restoration

1. Initialization:

(a) Set the initial estimate as $\hat{x} = y$ for image denoising and deblurring, or initialize $\hat{x}$ by bicubic interpolator for image super-resolution;

(b) Set initial regularization parameter λ and δ ;

2. Outer loop (dictionary learning and clustering): iterate on $l = 1, 2, \dots, L$

(a) Update the dictionaries $\{\Phi_k\}$ via k-means and PCA;

(b) Inner loop (clustering): iterate on $j = 1, 2, \dots, J$

(I) $\hat{x}^{(j+1/2)} = \hat{x}^{(j)} + \delta \mathbf{H}^T (y - \mathbf{H} \hat{x}^{(j)})$ where δ is the pre-determined constant;

(II) Compute

$$\mathbf{v}^{(j)} = [\Phi_{k_1}^T \mathbf{R}_1 \hat{x}^{(j+1/2)}, \dots, \Phi_{k_N}^T \mathbf{R}_N \hat{x}^{(j+1/2)}]$$

where Φ_{k_i} is the dictionary assigned to patch $\hat{x}_i = \mathbf{R}_i \hat{x}^{(j+1/2)}$;

(III) Compute α_i^{j+1} using the shrinkage operator

(IV) If $\text{mod}(j, J) = 0$ update the parameters $\lambda_{i,j}$ and $\{\beta_i\}$ using Eqs. (17) and (9), respectively;

(V) Image estimate update: $\hat{x}^{(j+1)} = \Phi \circ \alpha_y^{j+1}$ using Eq. (4).

In Algorithm, for fixed parameters $\lambda_{i,j}$ and $\{\beta_i\}$ the objective function is convex and can be efficiently solved by the iterative shrinkage algorithm in the inner loop, and its convergence has been well established in

IV. EXPERIMENTAL RESULTS

The basic parameter setting of SR is as follows: the patch size is 7×7 and $K = 70$. For image denoising, $\delta = 0.02$, $L = 3$, and $J = 3$; for image deblurring and super-resolution, $\delta = 2 \times 10^{-4}$, $L = 5$, and $J = 160$. To evaluate the quality of the restored images, the PSNR and the recently proposed powerful perceptual quality metric FSIM are calculated.



Fig. 1. Deblurring performance comparison on the Cameraman image.

From left to right and top to bottom: noisy and blurred image (9×9 uniform blur, $\sigma_n = 1.414$), the deblurred images by IDD-BM3D (PSNR = 28.56 dB; FSIM = 0.9007), ASDS-Reg (PSNR = 28.08 dB; FSIM = 0.8950), and the proposed NCSR (PSNR = 28.62 dB; FSIM = 0.9026).

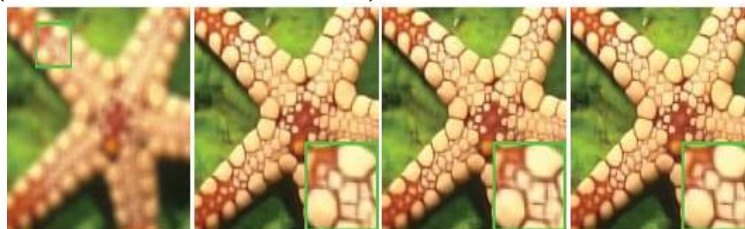


Fig. 2. Deblurring performance comparison on the Starfish image.

From left to right and top to bottom: noisy and blurred image (9×9 uniform blur, $\sigma_n = \sqrt{2}$), the deblurred images by IDD-BM3D [41] (PSNR = 29.48 dB; FSIM = 0.9167), ASDS-Reg [21] (PSNR = 29.72 dB; FSIM = 0.9208), and the proposed NCSR (PSNR = 30.28 dB; FSIM = 0.9293).

V. CONCLUSION

In this paper we presented a novel non-locally centralized sparse representation (SR) model for image restoration. The sparse coding noise (SCN), which is defined as the difference between the sparse code of the degraded image and the sparse code of the unknown original image, should be minimized to improve the performance of sparsity-based image restoration. To this end, we proposed a centralized sparse constraint, which exploits the image nonlocal redundancy, to reduce the SCN. The Bayesian interpretation of the NCSR model was provided and this endows the NCSR model an iteratively reweighted implementation. An efficient iterative shrinkage function was presented for solving the l_1 -regularized NCSR minimization problem. Experimental results on image denoising, deblurring and super-resolution demonstrated that the NCSR approach can achieve

highly competitive performance to other leading denoising methods, and outperform much other leading image deblurring and super-resolution methods.

REFERENCES

- [1] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" IEEE Trans. Inf. Theory, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [2] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [3] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [4] M. Bertero and P. Boccacci, Introduction to Inverse Problems Imaging. Bristol, U.K.: IOP Publishing, 1998.
- [5] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D, Nonlinear Phenomena, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.