2020 16th Specialist Meeting on Microwave Radiometry and Remote Sensing for the Environment (MicroRad) | 978-1-7281-7093-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/MicroRad49612.2020.9342611

An Effective Super-Resolution Reconstruction Method for Geometrically Deformed Image Sequences

Jing Qin* and Igor Yanovsky†‡
*Department of Mathematics, University of Kentucky, Lexington, KY, USA
†Jet Propulsion Laboratory, California Institute of Technology, Los Angeles, CA, USA
‡Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA

***Abstract*—Despite of the technology advancements, remote sensing images usually suffer from a poor spatial resolution. To resolve this issue, a lot of research efforts have been devoted to developing resolution enhancement methods which retrieve a high-resolution image out of its low-resolution degraded versions. In this paper, we consider a nonlocal total variation (NLTV) based super-resolution method which handles low-resolution images with geometric deformations. In particular, we apply the framework of alternating direction method of multipliers (ADMM) to deduce an effective algorithm, which involves soft thresholding and gradient descent. Effectiveness and robustness to noise of the proposed method are verified by various numerical experiments.**

***Index Terms*—Remote sensing images, super-resolution image reconstruction, nonlocal total variation, alternating direction method of multipliers (ADMM)**

I. Introduction

Remote sensing images have been playing an important role in many areas including geology, oceanography, weather forecasting, astrophysics, and radio astronomy. During the acquisition and transition, image quality is usually deteriorated due to limited imaging capabilities which results in a poor *spatial resolution*, i.e., the side length of an image pixel. In the meanwhile, other image degradations including noise, blurring artifacts and geometric deformations are inevitably present. Therefore, it is demanding to develop super-resolution (SR) methods, which create a high-resolution (HR) image using its low-resolution (LR) degraded versions. In this paper, by assuming that LR images are aligned via registration in advance, we study the multi-frame SR image enhancement and recovery given blurring and sampling operators.

A lot of research efforts have been devoted to applying various regularizations to solve the SR recovery problem. For example, the total variation (TV) has been widely used to preserve edge-like discontinuities during the SR reconstruction [1], [2], [3], [4]. However, by assuming that the ground truth is piecewise constant, TV based methods usually cause spurious staircasing artifacts in practice. As one of the remedies, bilateral TV (BTV) was integrated into the SR recovery framework by considering image smoothness in large neighborhoods together with the L_1-norm based fidelity [5]. Furthermore, high-order TV regularized SR methods have been developed, such as the fractional-order TV [6].

Nonlocal total variation (NLTV), which fully exploits patch-wise similarity in the gradient domain of an image, has shown great potential in image processing [7]. By taking more nearby pixels into consideration, NLTV can capture more neighborhood similarity than BTV. Thus NLTV and its variants have been applied to SR recovery with promising performance [8], [9]. Moreover, to deal with sophisticated noise in real applications, we adopt the L_1-norm fidelity because of its outstanding performance in image processing over the traditional L_2-norm fidelity [10], [11]. On the other hand, implementation of NLTV involves intensive computation of patchwise similarity which becomes a bottleneck of iterative least-square solvers. To circumvent this problem, we recently proposed a SR model based on NLTV and the L_1-norm fidelity (NLTV-L1) [12], which is solved by a primal-dual type of algorithm. Duality brings the flexibility of converting nonsmooth subproblems into smooth ones but introduces redundant dual variables and corresponding subproblems. To further improve efficiency and reduce the number of intermediate variables, we resort to the operator splitting technique. In this work, we convert the NLTV-L1 model into a linear constrained optimization problem, and then apply the alternating direction method of multipliers (ADMM) [13]. ADMM has recently gained popularity in image processing and solving inverse problems [14], [15], [16] since there are closed-form solutions to the resultant subproblems that can be either implemented by the Fast Fourier Transform (FFT) or expressed by simple proximal operators. Furthermore, due to the presence of highly ill-conditioned operators – blurring, geometric deformation, and downsampling – in our problem, we solve the least squares subproblem using the gradient descent in the Gauss-Seidel manner. Numerical experiments have shown that this proposed algorithm based on the inexact ADMM yields smaller reconstruction error than the primal-dual type of algorithm.

The remainder of the paper is structured as follows. In Section II, we present the NLTV-regularized super-resolution

This 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 research of Qin is supported by the NSF grant DMS-1941197.

reconstruction model and propose an efficient ADMM-based algorithm. A variety of numerical simulations are conducted to justify effectiveness of our algorithm on producing a HR image out of a set of geometrically deformed LR remote sensing images with various noise settings in Section III.

II. Proposed Super-Resolution Method

A. Nonlocal Operators

Given an image in its vector form $\{u(i)\}_{i=1}^{M}$, the discrete nonlocal gradient is defined as $D_\omega u(i,j) = \sqrt{\omega(i,j)}(u(i) - u(j))$. Here we adopt the following symmetric weight function $\omega(i,j) = \exp(-\sum_t k_\alpha(t)|u(i+t) - u(j+t)|^2/(2h^2))$ where k_α is the Gaussian convolution kernel with the standard deviation α, and h acts as a filtering scale. According to the nonlocal gradient, the NLTV is characterized as $\|D_\omega u\|_1 = \sum_i \sqrt{\sum_j D_\omega u(i,j)^2}$. By extending the divergence theorem to the nonlocal setting, the nonlocal divergence $\mathrm{div}_\omega u \in \mathbb{R}^M$ is given by $(\mathrm{div}_\omega u)(i) = \sum_j (\sqrt{\omega(i,j)}u(i) - \sqrt{\omega(j,i)}u(j))$. Notice that the summands in the nonlocal operators can be empirically narrowed down to a search window of small size instead of the whole image space. Straightforward computations lead to $\langle D_\omega u, v\rangle = \langle u, \mathrm{div}_\omega v\rangle$ which implies that div_ω is the adjoint operator of D_ω.

B. Proposed SR Reconstruction Method

Let $B_k \in \mathbb{R}^M$ with $k = 1,\dots,N$ be the degraded LR images of size $\sqrt{M}\times\sqrt{M}$ in lexicographic order, and $X \in \mathbb{R}^{r^2 M}$ the desired HR image where r denotes the sampling factor. Assume that the relationship between the HR image X and the observed LR image sequence $\{B_k\}_{k=1}^N$ is expressed by the linear equation $B_k = P_k H_k F_k X + \varepsilon_k$ for $k = 1,\dots,N$. Here F_k stands for the geometric deformation between the HR image and its k-th LR observation, which can be estimated by image registration. The blur operator $H_k \in \mathbb{R}^{r^2 M \times r^2 M}$ consists of atmospheric blur and sensor blur, and the downsampling operator $P_k \in \mathbb{R}^{M\times r^2 M}$ involves the factor $r > 1$. Moreover, ε_k is typically additive noise, e.g., Gaussian noise. According to the properties of circulant matrices, we have $F_k H_k = H_k F_k$. To simplify the discussion, P_k, H_k and F_k are provided in advance and H_k is restricted to the Gaussian blurring operator. Motivated by the robustness of the L_1-norm fidelity to noise and its advantage in preserving image contrast [5], [10], [11], we propose the following NLTV regularized SR image reconstruction model with a L_1-norm fidelity

$$\min_X \sum_{k=1}^{N} \|P_k H_k F_k X - B_k\|_1 + \lambda \|D_w X\|_1 . \tag{1}$$

To solve the above model, we first split variables by substituting $P_k H_k F_k X - B_k$ and $D_w X$ with other variables, and then directly apply ADMM. Specifically, we convert (1) into its equivalent form by the change of variables $U_k = A_k X - B_k$ and $V = D_w X$:

$$\begin{aligned} \min_{X,U_k,V} &\sum_{k=1}^{N} \|U_k\|_1 + \|V\|_1 , \\ &\text{s.t.} \quad A_k X - B_k = U_k, \quad D_w X = V. \end{aligned} \tag{2}$$

By introducing the dual variables $\hat{U}_k$ and $\hat{V}$, we define the corresponding augmented Lagrangian function

$$\begin{aligned} \mathcal{L} = &\sum_{k=1}^{N} \|U_k\|_1 + \|V\|_1 + \frac{\rho_1}{2}\left\|A_k X - B_k - U_k + \hat{U}_k\right\|_2^2 \\ &+ \frac{\rho_2}{2}\left\|D_w X - V + \hat{V}\right\|_2^2 . \end{aligned}$$

Next we separate the primal variables into two groups: $(U_1,\dots,U_N,V)$ and X. After applying the two-block ADMM, we obtain the subproblems

$$\begin{aligned} X^{n+1} &= \operatorname*{argmin}_X \mathcal{L}(X, U_1^n, \dots, U_N^n, V^n) \\ (U_1^{n+1},\dots,U_N^{n+1},V_k^{n+1}) &= \operatorname*{argmin}_{U_1,\dots,U_N,V} \mathcal{L}(X^{n+1}, U_1,\dots,U_N,V), \end{aligned}$$

together with the update of the dual variables $\hat{U}_1,\dots,\hat{U}_N,\hat{V}$. Due to the separability of U_k's and V in the $(U_1,\dots,U_N,V)$-subproblem, we can further get three groups of subproblems

$$\begin{aligned} X^{n+1} &= \operatorname*{argmin}_X \frac{\rho_1}{2}\sum_{k=1}^{N}\left\|A_k X - B_k - U_k^n + \hat{U}_k^n\right\|_2^2 \\ &\qquad + \frac{\rho_2}{2}\left\|D_w X - V^n + \hat{V}^n\right\|_2^2 , \\ U_k^{n+1} &= \operatorname*{argmin}_{U_k} \|U_k\|_1 + \frac{\rho_1}{2}\left\|A_k X - B_k - U_k^n + \hat{U}_k^n\right\|_2^2 , \\ V^{n+1} &= \operatorname*{argmin}_V \|V\|_1 + \frac{\rho_2}{2}\left\|D_w X - V^n + \hat{V}^n\right\|_2^2 . \end{aligned}$$

The X-subproblem has no closed-form solution expressed by FFT since ∇_w cannot be diagonalized under the Fourier transform. Moreover, matrix factorization based least-squares solvers will easily get stuck due to the ill-conditioned A_k's. Thus we apply the gradient descent to inexactly solve the X-subproblem but in the Gauss-Seidel manner, i.e., updating X or its copy whenever it is available. The detailed algorithm is described in Algorithm 1, where the soft thresholding operator reads as $\mathcal{S}(I,\rho)_i = \mathrm{sign}(I_i)\times\max(|I_i| - \rho, 0)$.

III. Numerical Results

To demonstrate the performance of the proposed algorithm, we test an image which simulates the 150-GHz microwave channel of the hurricane Rita in 2005. The data is generated to match the realistic configuartion of the Advanced Microwave Sounding Unit (AMSU-B) instrument, which is designed to measure microwave radiation from multiple layers of the atmosphere and record the sounding information. AMSU-B data has played an important role in the weather analysis and prediction. We use the cloud resolving numerical weather prediction model (WRF) [17] to generate a test image of size

Algorithm 1 Robust Super-Resolution Reconstruction Algorithm Based on ADMM

INPUT: LR images $\{B_k\}_{k=1}^N$, sampling factor $r \in \mathbb{N}$, geometric deformation F_k, blurring operator H_k, parameters $\rho_1, \rho_2 > 0$, maximum iteration number N_0 and tolerance tol.

OUTPUT: HR image X.

For $k = 1, \ldots, N$, set the downsampling operator P_k and $A_k = P_k H_k F_k$.

Set $X^0 = \text{MedianAndShift}(B_1, \ldots, B_N)$.

Set $U_k^0, V^0, \hat{U}_k^0, \hat{V}^0 = \mathbf{0}$.

for $n = 0, 1, 2, \ldots, N_0$ **do**

  Set $Z^0 = X^n$

  for $k = 0, 1, \ldots, N-1$ **do**

    $Z^{k+1} = Z^k - \rho_1 A_k^*(A_k Z^k - B_k - U_k^n + \hat{U}_k^n)$

  end for

  Update $X^{n+1} = Z^N - \rho_2 \,\text{div}_w(D_w Z^N - V^n + \hat{V}^n)$

  for $k = 0, 1, \ldots, N-1$ **do**

    $U_k^{n+1} = \mathcal{S}(A_k X^{n+1} - B_k + \hat{U}_k^n, 1/\rho_1)$

  end for

  $V^{n+1} = \mathcal{S}(D_w X^{n+1} + \hat{V}^{n+1}, \lambda/\rho_2)$

  $\hat{U}_k^{n+1} = \hat{U}_k^n + (A_k X^{n+1} - B_k - U_k^{n+1})$

  $\hat{V}^{n+1} = \hat{V} + (D_w X^{n+1} - V^{n+1})$

  if $\left\|X^{n+1} - X^n\right\|_2 / \left\|X^{n+1}\right\|_2 < tol$ **then**

    the algorithm terminates.

  end if

end for

402×402 with the spatial resolution 1.3 km per pixel. Methods to be compared include: Algorithm 1, SR method based on BTV and L_1-norm fidelity (BTV-L1) [5] and our previous work based on the primal-dual algorithm (NLTV-PD) [12]. Throughout the experiments, we use the same pre-processing step and initial guess for all methods: F_k's in (1) are created by the registration process, and the initial guess is generated by applying a median filter to the upsampled and translated LR images. Finally, we crop all the reconstructed images to remove dark boundaries. To quantify the performance, we use the peak signal-to-noise ratio (PSNR) defined as $PSNR = 20\log(I_{\max}/\|X - \widehat{X}\|_2)$, where $\widehat{X}$ is the estimation of the noise-free image X and $I_{\max}$ is the maximum possible image intensity. All the experiments are implemented using MATLAB R2019a for Windows 10 on a desktop PC with 64GB RAM and a 3.10GHz Intel Core i9-9960X CPU.

Following [12], we create a set of 64 noise-free LR images of size 100×100 based on the relationship between B_k and X in Section II. More specifically, the ground truth is first rotated clockwise/counterclockwise by i degrees, or translated diagonally/antidiagonally by i pixels for $i = 1, \ldots, 16$. Note that dark boundaries incurred from the geometric deformation are removed. Next a 3×3 Gaussian blur with standard deviation one is applied to all deformed images. Finally, the blurry images are downsampled with the factor 4. As an illustration, Fig. 1 displays two such degraded LR test images.

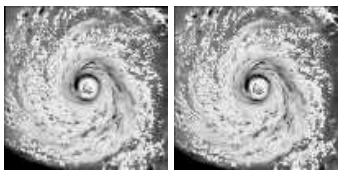
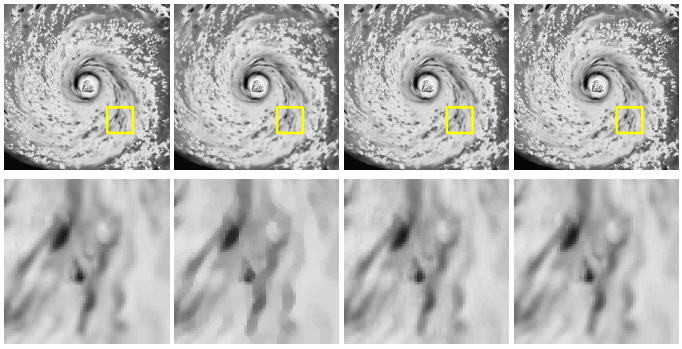


Fig. 1. LR test images. From left to right: clockwise rotated by 5 degrees and diagonal translated by 5 pixels.



Ground truth BTV-L1 NLTV-PD Algorithm 1

Fig. 2. SR results without noise. PSNR(dB): 27.05 (BTV-L1), 28.28 (NLTV-PD), 33.88 (Algorithm 1).

A. Experiment 1: Test Noise-free Data

First, we test the noise-free data. In the NLTV-PD, we choose the parameters $\lambda = 0.04$, $\mu = 0.03$, $\tau = 0.8$, and $\theta = 1$ for which the performance is optimized. In addition, the sizes of similarity and search window are set to 3×3 and 5×5, respectively, and the filtering scale $h = 0.25$ in the weight function ω. In Algorithm 1, we set $\rho_1 = 1$ and $\rho_2 = 10^{-2}$. All the algorithms terminate when 100 iterations are reached. Numerical experiments have shown that Algorithm 1 can achieve good performance within very few iterations. Recovered HR images are shown in Fig. 2, where close-up comparisons of a small region are outlined by yellow lines. Visual quality comparison illustrates that BTV-L1 tends to produce staircase artifacts especially in the smooth regions. By contrast, the proposed method preserves local geometries of an image. This phenomenon can be explained by the formulation of NLTV which involves more neighboring pixels for calculating the similarity. In addition, there is a trade-off between search/similarity window size and computational time/complexity. Furthermore, the runtimes in seconds (s) per iteration for all the methods are listed as follows: 0.0643s for BTV-L1, 0.5016s for NLTV-PD, and 0.7004s for Algorithm 1. Despite being more computationally expensive per iteration than NLTV-PD, the proposed algorithm yields a high-resolution image with higher accuracy and better visual quality.

B. Experiment 2: Test Gaussian Noise

Next we test the robustness of the algorithms to the Gaussian noise. Specifically, we add Gaussian noise with zero mean and standard deviation $\sigma \in \{1, 2, 3, 4, 5\}$ to the noise-free LR test images. The maximum iteration number for both BTV-L1 and NLTV-PD is set as 300 while 100 for Algorithm 1.

Method\σ	1	2	3	4	5
BTV-L1	26.82	26.62	26.39	26.30	25.92
NLTV-PD	27.44	27.26	26.79	26.60	26.10
Proposed	31.18	29.68	28.62	28.25	26.99

TABLE I
SR RESULT COMPARISON WITH GAUSSIAN NOISE IN PSNR (DB).

Method\θ	5	10	15	20	25
BTV-L1	27.05	27.05	27.07	27.03	26.62
NLTV-PD	28.27	28.24	28.17	27.90	26.93
Proposed	33.83	33.57	32.93	31.53	28.96

TABLE II
SR RESULT COMPARISON WITH SALT-AND-PEPPER NOISE IN PSNR (DB).

For NLTV-PD, we choose $\tau = 0.8, 0.2, 0.2, 0.1, 0.1$ and use the same parameters as those in the previous experiment for various Gaussian noise settings. Similarly, we use the same parameters as those in Experiment 1 for our algorithm. Empirical results suggest that τ has to be small when the noise level is high. Table I exhibits the quantitative performance comparison of all three algorithms in terms of PSNR under various settings of Gaussian noise. Note that the proposed method is not sensitive to the choice of parameters, in the sense that a set of appropriate parameters work for the same category of test images with various levels of noise. Surprisingly, the proposed algorithm performs much better than NLTV-PD.

C. Experiment 3: Test Salt-and-Pepper Noise

To further verify the robustness to the noise, we test the Salt-and-Pepper noise setting. Specifically, a sequence of noisy LR images are generated by redefining the image intensities of the noise-free data as follows:

$$\hat{B}_k(i) = \begin{cases} B_{\max}, & \text{if } B_k(i) > B_{\max} - \theta; \\ B_k(i), & \text{if } B_{\min} + \theta \leq B_k(i) \leq B_{\max} - \theta; \\ B_{\min}, & \text{if } B_k(i) \leq B_{\min} + \theta. \end{cases}$$

Here $B_{\max}$ and $B_{\min}$ are the maximum and minimum intensities of the LR images B_k's, respectively. The value of θ varies in $\{5, 10, 15, 20, 25\}$ corresponding to the percentages of wrong pixels: 0.28%, 0.6%, 1.22%, 2.53% and 5.63%. The parameters for BTV-L1 and NLTV-PD are the same as those in Experiment 2. We set $\rho_1 = 1$ and $\rho_2 = 10^{-1}$ in Algorithm 1. The performance comparisons in terms of PSNR under various θ's are shown in Table II. The results show that the proposed algorithm performs best for most cases in terms of PSNR.

IV. CONCLUSIONS

In remote sensing, reconstruction of a HR image from multiple LR degraded images with complex noise is highly ill-posed as a consequence of image quality degradation caused by blur, noise, sampling and deformation. In this paper, we propose an effective SR method which combines NLTV and L_1-norm data fidelity in the pursuit of reconstruction accuracy and robustness improvement. Specifically, a novel algorithm is derived by following the framework of ADMM, whose subproblems can be solved by Gauss-Seidel type of gradient descent and shrinkage. A large amount of numerical results show that the proposed algorithm is capable of retrieving HR images with higher accuracy and visual quality than the BTV and L_1 based method and our previous work based on the primal-dual algorithm. The proposed approach is also robust to various noise types, including Gaussian and Salt-and-Pepper noise. As for the future work, we will integrate tensor structures into the proposed work to reduce the computational cost for the patchwise similarity in NLTV.

REFERENCES

[1] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang, "A total variation regularization based super-resolution reconstruction algorithm for digital video," *EURASIP Journal on Advances in Signal Processing*, 2007.

[2] J. Qin, I. Yanovsky, and W. Yin, "Efficient simultaneous image deconvolution and upsampling algorithm for low resolution microwave sounder data," *Journal of Applied Remote Sensing*, vol. 9, no. 1, p. 095035, 2015.

[3] I. Yanovsky, B. H. Lambrigtsen, A. B. Tanner, and L. A. Vese, "Efficient deconvolution and super-resolution methods in microwave imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 9, pp. 4273–4283, 2015.

[4] I. Yanovsky and B. Lambrigtsen, "Multispectral super-resolution of tropical cyclone imagery using sparsity-based approaches," *International Journal of Remote Sensing*, vol. 37, no. 11, pp. 2494–2509, 2016.

[5] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *Image processing, IEEE Transactions on*, vol. 13, no. 10, pp. 1327–1344, 2004.

[6] Z. Ren, C. He, and Q. Zhang, "Fractional order total variation regularization for image super-resolution," *Signal Processing*, vol. 93, no. 9, pp. 2408–2421, 2013.

[7] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, "Image recovery via nonlocal operators," *Journal of Scientific Computing*, vol. 42, no. 2, pp. 185–197, 2010.

[8] W. Zeng and X. Lu, "A robust variational approach to super-resolution with nonlocal tv regularisation term," *The Imaging Science Journal*, vol. 61, no. 2, pp. 268–278, 2013.

[9] A. Laghrib, A. Ghazdali, A. Hakim, and S. Raghay, "A multi-frame super-resolution using diffusion registration and a nonlocal variational image restoration," *Computers & Mathematics with Applications*, vol. 72, no. 9, pp. 2535–2548, 2016.

[10] T. Chan and S. Esedoglu, "Aspects of total variation regularized L1 function approximation," *SIAM Journal on Applied Mathematics*, vol. 65, no. 5, pp. 1817–1837, 2005.

[11] F. Li, S. Osher, J. Qin, and M. Yan, "A Multiphase Image Segmentation Based on Fuzzy Membership Functions and L1-norm Fidelity," *Journal of Scientific Computing*, vol. 69, no. 1, pp. 82–106, 2016.

[12] J. Qin and I. Yanovsky, "Robust super-resolution image reconstruction method for geometrically deformed remote sensing images," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 8050–8053.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[14] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[15] W. Guo, J. Qin, and W. Yin, "A new detail-preserving regularization scheme," *SIAM journal on imaging sciences*, vol. 7, no. 2, pp. 1309–1334, 2014.

[16] Y. Li, J. Qin, Y.-L. Hsin, S. Osher, and W. Liu, "s-SMOOTH: Sparsity and Smoothness Enhanced EEG Brain Tomography," *Frontiers in neuroscience*, vol. 10, 2016.

[17] J. Michalakes, "Design of a next-generation regional weather research and forecast model." Argonne National Lab., IL (US), Tech. Rep., 1999.