IMAGE FUSION VIA DYNAMIC GRADIENT SPARSITY AND ANISOTROPIC SPECTRAL-SPATIAL TOTAL VARIATION

Chao-Chao Zheng, Ting-Zhu Huang, Liang-Jian Deng, Xi-Le Zhao, Hong-Xia Dou

School of Mathematical Sciences, University of Electronic Science and Technology of China, China

ABSTRACT

In this paper, we develop a sparsity based model for the fusion of a high spatial-resolution image and a multispectral image. The given model is based on the combination of a dynamic gradient sparsity (DGS) and an anisotropic spectral-spatial total variation (ASSTV). We design an alternating direction method of multipliers (ADMM) based algorithm to solve the proposed model. In contrast to existing approaches, the proposed method can generate more spatial details as well as preserve favorable spectral information. Experimental results demonstrate that the proposed approach outperforms several state-of-the-art image fusion methods both quantitatively and visually, in terms of both pansharpening application of remote sensing images and fusion application of natural color images.

1. INTRODUCTION

Image fusion aims at fusing a high spatial-resolution image and a low spatial-resolution multispectral image to generate a multispectral image with high spatial-resolution. It is quite important in many applications, e.g., pansharpening of remote sensing images which fuses a high spatial-resolution panchromatic (PAN) image and a multispectral image, and the fusion of a high spatial-resolution gray image and a RGB color natural image. In this paper, we mainly address these two applications.

There exist many approaches on image fusion application from different methodologies [1, 2], which can be roughly divided into three categories, *i.e.*, component substitution (CS) methods, multiresolution analysis (MRA) methods and variational methods. The CS methods depend on the substitution of a component after a spectral transformation of the multispectral image, with a high spatial-resolution image, *e.g.*, intensity-hue-saturation (IHS) method [3], principal component analysis (PCA) based method [4] and Gram-Schmidt (GS) spectral sharpening [5]. The MRA approaches are based on the injection into the multispectral (MS) image of spatial details which are obtained from the high spatial-resolution image, *e.g.*, additive wavelet luminance proportional (AWLP) [6]. The variational methods are a very powerful tool for image fusion [7, 8]. In [7], Yokoya *et al*. proposed a promissing method for hyperspectral and multispectral image fusion from the perspective of coupled nonnegative matrix factorization (CNMF) unmixing. This method can also be applied to the related applications of this paper. In [8], authors proposed a convex model by considering local spectral consistency and DGS, and solved the proposed model by the fast iterative shrinkage thresholding algorithm (FISTA) framework based algorithm.

In this paper, we propose a sparsity based variational model by considering the intrinsic properties of high spatial-resolution image and multispectral image. The given model has a spectral-data fidelity term by the formulation of generating low spatial-resolution multispectral image, a DGS regularizer used in [8], and an ASSTV term with different weights which takes into account the correlation among spectral images. In addition, we design an ADMM-based algorithm within an iterative manner to solve the proposed convex model. The proposed method is positively compared with several state-of-the-art image fusion methods on a real pansharpening dataset acquired by the WorldView-2 sensor and a color image fusion example from CAVE dataset, and the results demonstrate the effectiveness of our method.

This paper is organized as follows. In Sect. 2, we will present the proposed model and the corresponding solving algorithm. The experimental results are detailedly exhibited in Sect. 3. Finally, we draw conclusions in Sect. 4.

2. THE PROPOSED METHOD

2.1. The related work

Chen *et al.* [8] proposed a convex model to minimize a linear combination of a least-squares fidelity term and a DGS regularizer for the fusion of a high spatial-resolution panchromatic image and a low spatial-resolution multispectral image. The model is shown as follows

$$\min_X J(X) = E_1 + \lambda E_2, \tag{1}$$

where λ is a positive regularization parameter. The data fidelity term E_1 is defined as follows

$$E_1 = \frac{1}{2}\|\psi X - Y_{MS}\|_F^2, \tag{2}$$

where the ψ represents the blurring and downsampling operator, $X \in R^{m_1 \times n_1 \times z}$ is the underlying multispectral image, $Y_{MS} \in R^{m_2 \times n_2 \times z}$ is the low-resolution multispectral image, $m_1 = s \cdot m_2$, $n_1 = s \cdot n_2$ where s is the scaling factor. In addition, the authors employed a joint total variation (JTV) to depict the DGS E_2,

$$\begin{aligned} E_2 &= \|\nabla X - \nabla D(P)\|_{2,1} \\ &= \sum_i \sum_j \sqrt{\sum_d \sum_q (\nabla_q X_{i,j,d} - \nabla_q P_{i,j})^2}, \end{aligned} \tag{3}$$

where $q = 1, 2$, $P \in R^{m_1 \times n_1}$ is the panchromatic image and $D(P)$ means duplicating P to all bands. ∇_1 and ∇_2 denote the forward finite difference operators on the first and second coordinates, respectively.

He *et al.* [9] introduced spatial and spectral sparsity prior for pansharpening, which promotes spatially and spectrally aligned discontinuities across MS bands, as expected in natural images. To the best of our knowledge, in this paper we first utilize another excellently spatial and spectral sparsity prior, *i.e.*, ASSTV (see [10]), to solve the image fusion problem. The ASSTV can improve the smoothness of solutions and well describing the different correlations along both the spatial and spectral dimensions. The ASSTV is defined as follows

$$\|X\|_{ASSTV} = \sum_{i=1}^{3} \omega_i \|\nabla_i X\|_1, \tag{4}$$

where ω_i $(i = 1, 2, 3)$ are the regularization parameters.

In this paper, we will combine ASSTV regularizaer and DGS regularizer to obtain a satisfactory balance between spatial details preserving and spectral correlation preserving.

2.2. The proposed model and its solution

Combining the ASSTV and the DGS regularizers with the least-squares fidelity term, the proposed image fusion model is given as follows

$$\begin{aligned} \min_X \frac{1}{2}\|\psi X - Y_{MS}\|_F^2 + \lambda \|\nabla X - \nabla D(P)\|_{2,1} \\ + \sum_{i=1}^{3} \omega_i \|\nabla_i X\|_1, \end{aligned} \tag{5}$$

where λ is the positive parameter, and ω_i $(i = 1, 2, 3)$ are the weighted parameters of ASSTV term.

It is obvious that the model (5) is convex and has a globally optimal solution if solved by FISTA framework which has been proven holding a worst-case complexity of $O(1/k^2)$ (refer to [11, 12] for more details).

In the FISTA framework, we need to solve the following minimization problem,

$$\min_X \frac{L}{2}\|X - Y\|_F^2 + \lambda \|\nabla X - \nabla D(P)\|_{2,1} + \sum_{i=1}^{3} \omega_i \|\nabla_i X\|_1, \tag{6}$$

where Y is obtained by $\psi^T(\psi X - Y_{MS})$ which has the Lipschitz constant L, and ψ^T denotes the inverse operator of ψ. Due to the non-differentiable convex problem (6), we could utilize primal-dual method [13] or ADMM method [14, 15, 16, 17] to solve it. Here, we address it by ADMM method. Making variable substitutions $U = \nabla X - \nabla D(P)$ and $V_i = \nabla_i X$ to get the following augmented Lagrangian problem,

$$\begin{aligned} &\mathcal{L}(X, U, V_i, A, B_i) = \\ &\frac{L}{2}\|X - Y\|_F^2 + \lambda \|U\|_{2,1} + \frac{\beta_1}{2}\|\nabla X - \nabla D(P) - U - A\|_F^2 \\ &+ \sum_{i=1}^{3} \omega_i \|V_i\|_1 + \frac{\beta_2}{2} \sum_{i=1}^{3} \|\nabla_i X - V_i - B_i\|_F^2, \end{aligned} \tag{7}$$

where β_1 and β_2 are two positive parameters, A, B_1, B_2 and B_3 represent four Lagrange multipliers. The problem of minimizing $\mathcal{L}(X, U, V_i, A, B_i)$ can be iteratively and alternatively solved by the following step **a)** to step **d)**:

a) The X-subproblem is given as follows

$$\begin{aligned} \min_X \frac{L}{2}\|X - Y\|_F^2 + \frac{\beta_1}{2}\|\nabla X - \nabla D(P) - U - A\|_F^2 \\ + \frac{\beta_2}{2}\omega_i \|\nabla_i X - V_i - B_i\|_F^2, \end{aligned} \tag{8}$$

which has the following closed-form solution by least squares method [18],

$$X = \mathcal{F}^{-1}\left(\frac{L\mathcal{F}(Y) + Q_1 + Q_2}{R}\right), \tag{9}$$

where

$$\begin{aligned} Q_1 &= \beta_1 \mathcal{F}^{-1}(\nabla)\mathcal{F}(\nabla D(P) + U + A), \\ Q_2 &= \beta_2 \sum_{i=1}^{3} \mathcal{F}^{-1}(\nabla_i)\mathcal{F}(V_i + B_i), \\ R &= L + \beta_1 \mathcal{F}^{-1}(\nabla)\mathcal{F}(\nabla) + \beta_2 \sum_{i=1}^{3} \mathcal{F}^{-1}(\nabla_i)\mathcal{F}(\nabla_i), \end{aligned}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and the inverse Fourier transform.

b) The U-subproblem is shown as follows

$$\min_U \lambda \|U\|_{2,1} + \frac{\beta_1}{2}\|\nabla X - \nabla D(P) - U - A\|_F^2, \tag{10}$$

which also has the following closed-form solution [19],

$$U_{i,j,:} = \mathbf{Shrink}_1((\nabla X - \nabla D(P) - A)_{i,j,:}, \frac{\lambda}{\beta_1}), \tag{11}$$

where $\mathbf{Shrink}_1(r, \alpha) = \frac{r}{\|r\|_2} * max(\|r\|_2 - \alpha, 0)$.

c) The V_i-subproblem is given by minimizing the following function

$$\min_{V_i} \omega_i \|V_i\|_1 + \frac{\beta_2}{2}\|\nabla_i X - V_i - B_i\|_F^2. \quad (12)$$

The closed-form solution of V_i-subproblem by soft-thresholding strategy [20] is as follows

$$V_i = \mathbf{Shrink}_2(\nabla_i X - B_i, \frac{\omega_i}{\beta_2}), \ \ i = 1, 2, 3, \quad (13)$$

where $\mathbf{Shrink}_2(r, \alpha) = \frac{r}{|r|_1} . * max(|r|_1 - \alpha, 0)$, and $.*$ denotes component-wise multiplication.

d) Update the Lagrange multipliers as follows

$$\begin{aligned} A^{k+1} &= A^k - (\nabla X^{k+1} - \nabla D(P) - U^{k+1}), \\ B_i^{k+1} &= B_i^k - (\nabla_i X^{k+1} - V_i^{k+1}). \end{aligned} \quad (14)$$

The following is the summarized iterative algorithm based on the FISTA framework and ADMM method to solve the proposed model (5).

Algorithm 1: The summarized algorithm for model (5)

Input: $\mathbf{Y_{MS}}, \mathbf{P}, L, \lambda, \omega_1, \omega_2, \omega_3, \beta_1, \beta_2$
Output: Fused image $\hat{\mathbf{X}}$
Initialize: $t^{(0)} = 1, X^{(0)} = \mathbf{0}$
1) Calculate $Y^{(0)} = \psi^T Y_{MS}$
For $k = 1 : \tau$
  2) Calculate $Y = Y^{(k-1)} - \psi^T(\psi X - Y_{MS})/L$
  // Solve the problem (6) by ADMM
  While *not converged* **do**
    3) Update X by (9)
    4) Update U by (11)
    5) Update V_i by (13), $i = 1, 2, 3$
    6) Update the multipliers by (14)
  EndWhile
  7) Update $X^{(k)}$ by step 3
  8) Calculate $t^{(k)} = [1 + \sqrt{1 + 4(t^{(k-1)})^2}]/2$
  9) Calculate $Y^{(k)} = X^{(k)} + \frac{t^{(k-1)}-1}{t^{(k)}}(X^{(k)} - X^{(k-1)})$
EndFor
Return: $\hat{\mathbf{X}} = X^{(\tau)}$

where τ is the number of maximum iterations defined empirically, and L is set as 1 in the experiments.

3. RESULTS

In this section, we compare the proposed method with some competitive image fusion methods on a real pansharpening dataset (WV-2, $1024 \times 1024 \times 4$) acquired by the WorldView-2 sensor [1] and a color image fusion example ($512 \times 512 \times 3$) from CAVE dataset [2]. The scale factors are both set as 4. In particular, we use bicubic interpolation as the downsampling operator to simulate experimental low spatial-resolution multispectral (color) image. The experiments are implemented in MATLAB(R2013b) on a computer of 4Gb RAM and Intel(R) Core(TM) i5-2450M CPU E5504: @2.50 GHz, 2.50 GHz.

For the parameters in the proposed method, we set $\lambda = 5$, $\omega_1 = \omega_2 = 1$, $\omega_3 = 0.1$ and $\beta_1 = \beta_2 = 0.1$ for WV-2 dataset, while set $\lambda = 2.5$, $\omega_1 = \omega_2 = 0.1$, $\omega_3 = 0.01$ and $\beta_1 = \beta_2 = 0.1$ for CAVE dataset. We use Quality Index for four spectral bands (**Q4**), Spectral Angle Mapper (**SAM**), $Erreur\ Relative\ Globale\ Adimensionnelle\ de\ Synthèse$ (**ERGAS**), Peak Signal-to-Noise Ratio (**PSNR**) and Mean Structural Similarity (**MSSIM**) as indices to estimate the performance of different methods (refer to [2, 21] for details, the larger Q4, PSNR, MSSIM and smaller SAM, ERGAS, the better performance).

Table 1. The quantitative results of different methods for WV-2 and CAVE datasets. Q4 is for 4 bands data, thus it is meaningless for color CAVE dataset (**Bold**: the best).

	Q4	SAM	ERGAS	PSNR	MSSIM
WV-2 dataset					
GS	0.8763	4.5897	4.4285	28.343	0.8918
CNMF	0.9599	3.9412	2.8910	31.032	0.9825
AWLP	0.9601	3.8280	3.2536	30.277	0.9831
DGS	0.9540	3.2945	3.2130	31.207	0.9904
Proposed	**0.9718**	**3.1295**	**2.4908**	**32.708**	**0.9944**
CAVE dataset					
GS	-	3.5832	2.0545	38.161	0.9775
CNMF	-	13.3755	12.6076	24.117	0.8474
AWLP	-	4.0938	2.7952	32.795	0.9683
DGS	-	2.2638	1.7802	36.966	0.9890
Proposed	-	**2.1123**	**1.1316**	**42.960**	**0.9944**

We evaluate the proposed method against some competitive CS and MRA methods, *i.e.*, GS [5] and AWLP [6], two variational methods, *i.e.*, CNMF [7] and DGS [8] [3]. For each of comparing methods, we used the parameters suggested in their respective papers and codes. From Fig. 1 that shows the visual results of WV-2 remote sensing data, we know that the proposed method obtains more spatial image details and well preserves the spectral information of multispectral image, while GS and AWLP loss the spectral information significantly. The two state-of-the-art variational methods yield excellent spatial performance, however they can not preserve more spectral information than our method. Fig. 2 exhibits the visually fused results of a high spatial-resolution gray image and a RGB color natural image, which is a special example from smartphone with dual camera. Our method still performs best both on preserving spatial and spectral informa-

[1] http://www.openremotesensing.net/index.php/codes/11-pansharpening

[2] http://www.cs.columbia.edu/CAVE/databases/multispectral/

[3] Codes of GS, AWLP and CNMF [1], DGS's on the author's homepage.

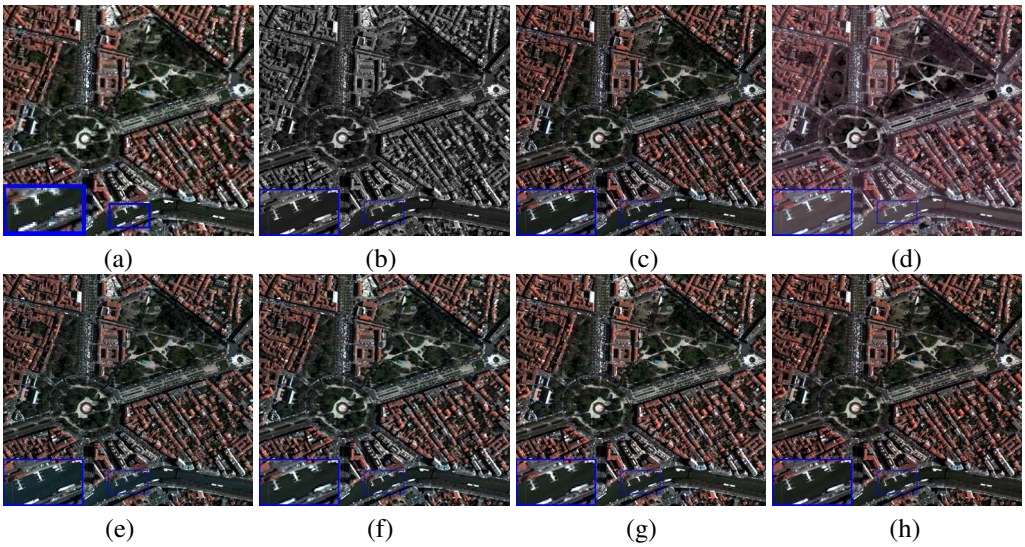

Fig. 1. WV-2 dataset: (a) Low spatial-resolution multispectral image; (b) PAN image; (c) Reference image; (d) GS [5]; (e) CNMF [7]; (f) AWLP [6]; (g) DGS [8]; (h) Proposed. Readers are recommended to zoom in all figures for better visibility.

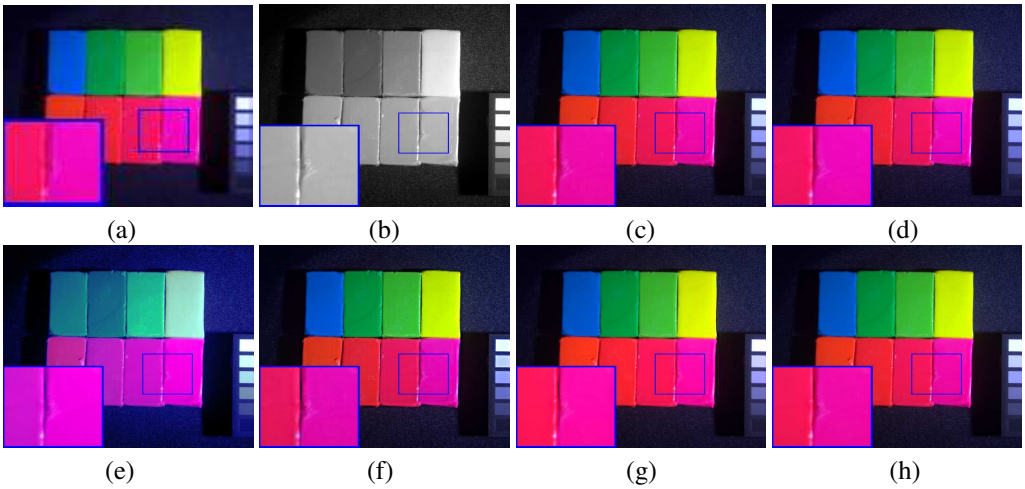

Fig. 2. CAVE dataset: (a) Low spatial-resolution color image; (b) High spatial-resolution gray image; (c) Reference image; (d) GS [5]; (e) CNMF [7]; (f) AWLP [6]; (g) DGS [8]; (h) Proposed.

tion. Table 1 reports the quantitative results of the comparing approaches, it also indicates the superiority of our method. Note that CNMF method for the CAVE dataset performs not good maybe due to the weakness of CNMF method for color image example (only 3 bands).

4. CONCLUSIONS

In this paper, we proposed a variational image fusion model based on the combination of DGS and ASSTV. A FISTA and ADMM based algorithm was designed for solving the proposed model. We compared our method with some CS and MRA methods and two recent state-of-the-art variational methods. By the experimental results, the proposed method preserved more image information than the comparing approaches, especially on preserving spectral information which may resulted by the ASSTV term. In the next work, we'll enhance spatial image details via iterative strategy.

5. ACKNOWLEDGMENT

Thank 973 Program (2013CB329404), NSFC (61370147, 61402082) and Fundamental Research Funds for the Central Universities (ZYGX2016KYQD142) for funding.

6. REFERENCES

[1] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Transactions on geoscience and remote sensing*, vol. 40, no. 10, pp. 2300–2312, 2002.

[2] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pansharpening algorithms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2565–2586, 2015.

[3] W. J. Carper, "The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 457–467, 1990.

[4] P. Kwarteng and A. Chavez, "Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens*, vol. 55, pp. 339–348, 1989.

[5] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," Jan. 4 2000, US Patent 6,011,875.

[6] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. application to wavelet-based methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2376–2385, 2005.

[7] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.

[8] C. Chen, Y. Li, W. Liu, and J. Huang, "Image fusion with local spectral consistency and dynamic gradient sparsity," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2760–2765.

[9] Xiyan He, Laurent Condat, José M Bioucas-Dias, Jocelyn Chanussot, and Junshi Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 4160–4174, 2014.

[10] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1852–1866, 2015.

[11] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[12] L.-J. Deng, H. Guo, and T.-Z. Huang, "A fast image recovery algorithm based on splitting deblurring and denoising," *Journal of Computational and Applied Mathematics*, vol. 287, pp. 88–97, 2015.

[13] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120–145, 2011.

[14] L.-J. Deng, W. Guo, and T.-Z. Huang, "Single image super-resolution via an iterative reproducing kernel Hilbert space method," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, pp. 2001–2014, 2016.

[15] L. J. Deng, W. Guo, and T. Z. Huang, "Single image super-resolution by approximate Heaviside functions," *Information Sciences*, vol. 348, pp. 107–123, 2016.

[16] Xi-Le Zhao, Fan Wang, Ting-Zhu Huang, Michael K Ng, and Robert J Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 4045–4058, 2013.

[17] Xi-Le Zhao, Fan Wang, and Michael K Ng, "A new convex optimization model for multiplicative noise and blur removal," *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 456–475, 2014.

[18] X.-L. Zhao, W. Wang, T.-Y. Zeng, T.-Z. Huang, and M. K. Ng, "Total variation structured total least squares method for image restoration," *SIAM Journal on Scientific Computing*, vol. 35, no. 6, pp. B1304–B1320, 2013.

[19] W. Deng, W. Yin, and Y. Zhang, "Group sparse optimization by alternating direction method," *SPIE Optical Engineering+ Applications. International Society for Optics and Photonics*, 2013.

[20] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, pp. 613–627, 1995.

[21] Z. Wang and A. C. Bovik, "Modern image quality assessment," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 2, pp. 1–156, 2006.