Fast TVL1-L2 MR Image Reconstruction using Variable Splitting and Accelerated Alternating Direction Method with Adaptive Restart

Shoulie Xie, Cuntai Guan, Weimin Huang, and Zhongkang Lu
Neuro & Biomedical Technology Department, Institute for Infocomm Research, Singapore
e-mail: {slxie,ctguan,wmhuang,zklu}@i2r.a-star.edu.sg

***Abstract*—This paper presents a fast algorithm for magnetic resonance (MR) image reconstruction from undersampled k-space measurements. The underlying MR image reconstruction is formulated as solving a TVL1-L2 minimization problem whose objective function consists of total variation (TV) regularizer, wavelet-based ℓ_1-norm regularizer and ℓ_2 data fidelity. Our approach is based on a variable splitting strategy and an accelerated alternating direction method of multiplier (ADMM) with restart. This paper shows that our proposed algorithm is fast and efficient for solving the TVL1-L2 MR image reconstruction problem. More precisely, a variable splitting method is used to split the variable into three variables and obtain an equivalent constrained optimization formulation, which is then addressed with an accelerated ADMM with adaptive restart. This ADMM algorithm is acceleration because the next iterate is computed by employing two previous computed iterates, and the restart rule is employed to enforce monotonicity and convergence in solving weakly convex TVL1-L2 optimization. Moreover thanks to intrinsic spatial-frequency encoding in MRI data, the inverse of regularized Hessian matrix can perform efficiently by exploiting fast Fourier transform (FFT) and fast wavelet transform (or tight frame). Experimental examples also demonstrate that the proposed algorithm is fast and efficient compared to the classical ADMM in TVL1-L2 MR image reconstruction.**

I. Introduction

Magnetic resonance imaging (MRI) has been widely used in medical diagnosis because of its non-invasive nature and glorious depiction of soft tissue changes. But due to physical and physiological limitations, most MRI scanners take long time to generate an MR image. Recent studies in compressive sensing (CS) theory and its applications to MRI ([1]– [2]) indicate that MR images can be accurately reconstructed from highly undersampled *k*-space data by employing nonlinear reconstruction schemes. That is to say, CS-based MRI has the potential to reduce the scanning time considerably.

The objective of MR image reconstruction is to recover the unknown true MR image $u \in \mathbb{R}^n$ from a noisy undersampled *k*-space measurement $y \in \mathbb{C}^m$ ($m \ll n$) modeled by

$$y = Hu + \nu \tag{1}$$

where $H = SF$ is linear sampling operator with $F \in \mathbb{C}^{n\times n}$ being discrete Fourier transform and $S \in \mathbb{R}^{m\times n}$ sampling pattern containing m rows of identity matrix of order n, $\nu \in \mathbb{C}^m$ is assumed to be a circular complex white Gaussian noise with mean zero and variance σ^2.

It is well known that the MR image reconstruction is an ill-posed inverse problem due to noise and undersampling. But in order to obtain acceptable solution to this inverse problem, variational regularization approaches are often used to find approximate ones by exploiting signal prior information. In general, MR images have some blocky structures and sparse representations under certain wavelet orthogonal basis or redundant frame, so total-variation (TV) and wavelet-based sparsity are often exploited as the prior information. Thus the regularization approach for the CS-based MR image reconstruction is described by the following unconstrained TVL1-L2 optimization problem:

$$\min_u \frac{1}{2}\|y - Hu\|_2^2 + \gamma\|Wu\|_1 + \tau TV(u) \tag{2}$$

where W is the discrete wavelet orthogonal transform or redundant frame, TV is the total-variation operator, $\gamma > 0$ and $\tau > 0$ are regularization parameters that weight the influence of the fidelity and the regularization terms.

The nonlinear conjugate gradient method (NLCG) was first used in [1] to tackle this regularization problem, where a smooth function was employed to approximate the nonsmooth part. Then the classical ADMM algorithm was adopted to solve this problem in [3], [4], where the regularization problem (2) was decoupled into several subproblems that can be solved efficiently. Specially it was shown in [4] that ADMM-based algorithm is faster than NLCG method for solving (2).

Recently a fast ADMM algorithm was proposed in [5] where iterations could be accelerated by using two previous computed iterates, which borrowed the idea of fast iterative shrinkage thresholding algorithm (FISTA) [6]. The fast convergence rate of the accelerated ADMM was provided for strongly convex optimization problems, but can not be guaranteed for weakly convex problems. The authors handled weakly convex problems by incorporating a simple restart rule to enforce convergence, the numerical examples in [5] showed that the accelerated ADMM with this restart was superior to the classical ADMM for solving the weakly convex problems when the regularization was either ℓ_1-norm or anisotropic TV. This inspires us to use the accelerated ADMM with restart to solve the weakly convex TVL1-L2 MR image reconstruction with both complex isotropic TV and ℓ_1-norm regularizers.

The proposed approach in this paper is based on a variable splitting technique and the accelerated ADMM with adaptive restart. First the variable splitting technique is used to split the variable u into three variables $u,\ v_1$ and v_2, each variable serves as the argument of each function, thus an equivalent constrained optimization formulation is obtained, which is then addressed using the accelerated ADMM with restart.

Furthermore, application of the accelerated ADMM with restart to TVL1-L2 regularization problem (2) involves solving a linear system whose size is often very large. This seems like an unsurmountable obstacle due to the required huge computations for the regularized Hessian matrix and its inverse. But since MRI scanners acquire data in the spatial-frequency encoded domain, we can show that the regularized Hessian matrix and its inverse perform quite efficiently by exploiting fast Fourier transform (FFT) and fast wavelet algorithms. Thus the second-order information in Hessian matrix can be used, which speeds up the convergence of the proposed algorithms compared with other first-order gradient algorithms. Therefore the proposed algorithm can solve efficiently the TVL1-L2 MR image reconstruction problem (2).

II. Proposed Approach

Using variable splitting, the problem (2) can be rewritten as an equivalent constrained optimization problem:

$$\begin{aligned}\min_{u,v_1,v_2} \frac{1}{2}\|Hu-y\|_2^2+\gamma\|v_1\|_1+\tau TV(v_2)\\ \text{subject to } v_1=Wu \text{ and } v_2=u.\end{aligned} \tag{3}$$

Obviously, the variable splitting scheme splits the variable u in (2) into three variables $u,\ v_1$ and v_2, each variable serves as the argument of each function.

Generally $f(u)=\frac{1}{2}\|Hu-y\|_2^2$ is not the strong convex due to under-sampling and noise, and $g(v)=\gamma\|v_1\|_1+\tau TV(v_2)$ is also weakly convex. In order to accelerate the classical ADMM algorithm, the idea of fast ADMM with restart rule [5] is extended here to deal with the constrained TVL1-L2 optimization problem (3) with both complex isotropic TV and ℓ_1-norm regularizers. First, the classical ADMM technique leads to the following iterative scheme for solving (3):

$$\begin{aligned}u_k=\arg\min_u \|Hu-y\|_2^2+\rho\|Wu-\hat{v}_{1,k}+\hat{\eta}_{1,k}\|_2^2\\ +\rho\|u-\hat{v}_{2,k}+\hat{\eta}_{2,k}\|_2^2,\end{aligned} \tag{4}$$

$$\begin{aligned}(v_{1,k},\ v_{2,k})=\arg\min_{v_1,v_2}\gamma\|v_1\|_1+\tfrac{\rho}{2}\|Wu_k-v_1+\hat{\eta}_{1,k}\|_2^2\\ +\tau TV(v_2)+\tfrac{\rho}{2}\|u_k-v_2+\hat{\eta}_{2,k}\|_2^2,\end{aligned} \tag{5}$$

$$\eta_{1,k}=\hat{\eta}_{1,k}+(Wu_k-v_{1,k}),\quad \eta_{2,k}=\hat{\eta}_{2,k}+(u_k-v_{2,k}), \tag{6}$$

and then an adaptive restart rule is adopted for accelerating its convergence as follows: Calculate

$$Err_k=\|\eta_k-\hat{\eta}_k\|^2+\|v_k-\hat{v}_k\|^2,$$

if $Err_k<\epsilon Err_{k-1}$ then

$$\alpha_{k+1}=\tfrac{1}{2}(1+\sqrt{1+4\alpha_k^2}),\quad \beta_k=\tfrac{\alpha_k-1}{\alpha_{k+1}},$$
$$\hat{v}_{k+1}=v_k+\beta_k(v_k-v_{k-1}),\quad \hat{\eta}_{k+1}=\eta_k+\beta_k(\eta_k-\eta_{k-1}),$$

else

$$\alpha_{k+1}=1,\ \hat{v}_{k+1}=v_{k-1},\ \hat{\eta}_{k+1}=\eta_{k-1},\ Err_k=\epsilon^{-1}Err_{k-1}$$

where $\epsilon\in(0,1)$, $v_k=(v_{1,k},\ v_{2,k})$ and $\eta_k=(\eta_{1,k},\ \eta_{2,k})$.

It should be noted that the scaled multipliers are used here, thus the combined primal and dual error Err_k is easy to be calculated without requirement of multiplying the penalty parameter ρ, which is unlike the calculation in [5] where the multiplications of ρ and ρ^{-1} were required.

On the other hand, a simple restart rule is incorporated to enforce monotonicity and convergence in this algorithm, although the convergence rate of $O(1/k^2)$ can not be guaranteed for the weekly convex objective terms.

The above minimization problem (5) can further be separated into the following two optimal subproblems:

$$v_{1,k} = \arg\min_{v_1}\gamma\|v_1\|_1+\frac{\rho}{2}\|Wu_k-v_1+\hat{\eta}_{1,k}\|_2^2, \tag{7}$$

$$v_{2,k} = \arg\min_{v_2}\tau TV(v_2)+\frac{\rho}{2}\|u_k-v_2+\hat{\eta}_{2,k}\|_2^2. \tag{8}$$

Then the minimization problem (7) can be solved by the soft thresholding method [7], which has a closed form:

$$v_{1,k}=\text{soft}(Wu_k+\hat{\eta}_{1,k},\gamma/\rho), \tag{9}$$

where $\text{soft}(x,\tau)=\frac{(|x|-\tau)}{|x|}x$ for $|x|>\tau$, and zero for $|x|\le\tau$.

The minimization problem (8) is pure isotropic TV-based denoising problem, which does not have a closed-form solution but can be solved efficiently and accurately with numerical methods such as Chambolle's algorithm [8]. The numerical solution is expressed as

$$v_{2,k}=\text{prox}_{\tau TV/\rho}(u_k+\hat{\eta}_{2,k}) \tag{10}$$

where $\text{prox}_{\tau TV}$ is the Moreau proximity operator defined by

$$\text{prox}_{\tau TV}(x)=\arg\min_v \tau TV(v)+\frac{1}{2}\|v-x\|_2^2. \tag{11}$$

Finally we need compute the u_k in (4). Note that the minimization problem with respect to u is a strictly convex quadratic, it can be reduced to the following linear system:

$$u_k=\mathcal{A}^{-1}(H^Ty+\rho W^T(\hat{v}_{1,k}-\hat{\eta}_{1,k})+\rho(\hat{v}_{2,k}-\hat{\eta}_{2,k})) \tag{12}$$

where

$$\mathcal{A}=H^TH+\rho W^TW+\rho I. \tag{13}$$

The matrix $\mathcal{A}$ can be seen as a regularized version of the Hessian of function $f(u)=1/2\|Hu-y\|_2^2$ by adding the terms ρI and ρW^TW. In general, the computations of this matrix and its inverse are not affordable for large-size matrix H, we may take gradient-based algorithms which lead to the linear system being solved inexactly. However, for underlying MR image reconstruction, the matrix $H=SF$ has a special undersampling Fourier structure. The undersampling pattern $S\in\mathbb{R}^{m\times n}$ can be formed by taking a subset of rows of an identity matrix. Hence we have $SS^T=I$. In additional, for an orthogonal or a redundant tight wavelet frame, we have $W^TW=I$. Using matrix inversion formula, we can obtain

$$\mathcal{A}^{-1} = \frac{1}{2\rho}\left(I-\frac{1}{1+2\rho}F^TS^TSF\right). \tag{14}$$

Therefore, in view of (14), the computation of u_k in (12) only involves matrix-vector products with fast Fourier and wavelet

algorithms. Since S^TS is equal to an identity matrix with some zeros in the diagonal, it is a binary mask and can be computed with $O(n)$ cost. Then the products by F^T and F can be computed with $O(n \log n)$ cost using the fast Fourier transform [9]. Furthermore, for an orthogonal wavelet basis or a redundant tight wavelet frame, any matrix-vector multiplications can be performed by the fast $O(n \log n)$ implementations [10]. In conclusion, u_k can be computed with $O(n \log n)$ cost using the fast wavelet transform W and fast Fourier transform F. And the second-order information of function f is used in our proposed algorithm so that it is faster than the existing gradient-like algorithms that only use first-order information.

In view of (9), (10), (12) and (14), we can obtain the following fast algorithm to solve the TVL1-L2 regularization problem (2).

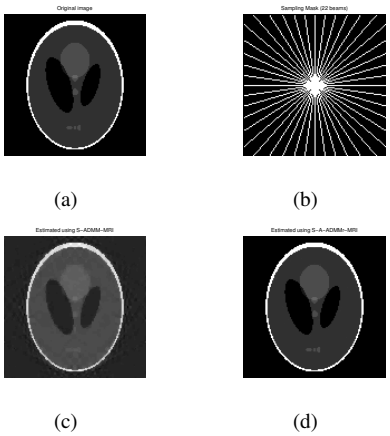
Algorithm Split-accelerated ADMM with restart for TVL1-L2 MR image reconstruction *(S-A-ADMMr-MRI)*:

1) Set $k = 0$, choose $\rho > 0$, $\hat{v}_{1,0}$, $\hat{\eta}_{1,0}$, $\hat{v}_{2,0}$, $\hat{\eta}_{2,0}$, $\alpha_1 = 1$ and $\epsilon \in (0, 1)$
2) *repeat*
3) $r_k = H^T y + \rho W^T(\hat{v}_{1,k} - \hat{\eta}_{1,k}) + \rho(\hat{v}_{2,k} - \hat{\eta}_{2,k})$
4) $u_k = \frac{1}{2\rho}(r_k - \frac{1}{1+2\rho} F^T S^T S F r_k)$
5) $v_{1,k} = \text{soft}(Wu_k + \hat{\eta}_{1,k}, \frac{\gamma}{\rho})$
6) $v_{2,k} = \text{prox}_{\tau/\rho TV}(u_k + \hat{\eta}_{2,k})$
7) $\eta_{1,k} = \hat{\eta}_{1,k} + (Wu_k - v_{1,k})$
8) $\eta_{2,k} = \hat{\eta}_{2,k} + (u_k - v_{2,k})$
9) $Err_k = \|\eta_{1,k} - \hat{\eta}_{1,k}\|^2 + \|\eta_{2,k} - \hat{\eta}_{2,k}\|^2 + \|v_{1,k} - \hat{v}_{1,k}\|^2 + \|(v_{2,k} - \hat{v}_{2,k}\|^2$
10) if $Err_k < \epsilon Err_{k-1}$ then
11) $\quad \alpha_{k+1} = \frac{1}{2}(1 + \sqrt{1 + 4\alpha_k^2}), \quad \beta_k = \frac{\alpha_k - 1}{\alpha_{k+1}}$
12) $\quad \hat{v}_{1,k+1} = v_{1,k} + \beta_k(v_{1,k} - v_{1,k-1})$
13) $\quad \hat{v}_{2,k+1} = v_{2,k} + \beta_k(v_{2,k} - v_{2,k-1})$
14) $\quad \hat{\eta}_{1,k+1} = \eta_{1,k} + \beta_k(\eta_{1,k} - \eta_{1,k-1})$
15) $\quad \hat{\eta}_{2,k+1} = \eta_{2,k} + \beta_k(\eta_{2,k} - \eta_{2,k-1})$
16) else
17) $\quad \alpha_{k+1} = 1, \hat{v}_{1,k+1} = v_{1,k-1}, \hat{\eta}_{1,k+1} = \eta_{1,k-1}$
18) $\quad \hat{v}_{2,k+1} = v_{2,k-1}, \hat{\eta}_{2,k+1} = \eta_{2,k-1}$
19) $\quad Err_k = \epsilon^{-1} Err_{k-1}$
20) end if
21) $k \leftarrow k + 1$
22) *until stopping criterion is satisfied.*

III. Experimental Results

In this section, numerical results are reported to illustrate performance of our proposed fast S-A-ADMMr-MRI algorithm for TVL1-L2 MR image reconstruction from undersampled measurements. In [3], [4], [11], classical ADMM algorithm was shown to be more efficient than the existing algorithms such as NLCG and FISTA. Hence, we only compare our proposed fast S-A-ADMMr-MRI algorithm with classical ADMM algorithm for MR image reconstruction. Our simulations are written in MATLAB and performed on a Dell computer with Intel Xeon CPU 2.66GHz and 4GB of RAM. Our test sets are generated from three images: the Shepp-Logan phantom, a real human brain and a real rat brain images. Moreover, in our experiments, the transform W is a redundant Haar wavelet frame with four levels.

In the first experiment, we consider the well-known Shepp-Logan phantom with size 128×128 shown in Fig. 1(a). The undersampled k-space data is generated from 22 radial lines of its 2-D discrete Fourier transform (sampling ratio 16.27%) as shown in Fig. 1(b). The projections are also corrupted by circular complex Gaussian noise with variance $\sigma^2 = 0.5 \times 10^{-6}$. In our simulations, we empirically choose $\gamma = 10^{-5}$, $\tau = 0.5\gamma$, $\rho = 0.01$ and $\epsilon = 0.999$. Figs. 1(c)



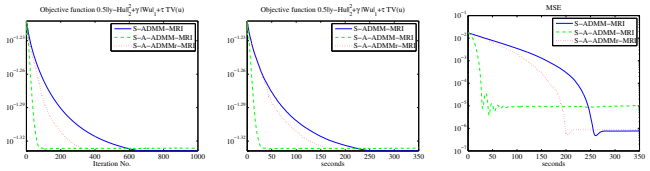
(a) (b) (c) (d)

Fig. 1. (a) Shepp-Logan phantom with size 128×128. (b) Sampling mask with 22 radial lines. Phantom MR image reconstruction with 400 iterations: (c) Reconstructed image using variable splitting and classical ADMM method (S-ADMM-MRI). (d) Reconstructed image using variable splitting and accelerated ADMM with restart (S-A-ADMMr-MRI).

and 1(d) plot the reconstructed images at the 400th iterate by using the proposed the proposed S-A-ADMMr-MRI and the classical ADMM approaches, respectively. It can be seen that artifacts occur around edges in the reconstructed image using the classical ADMM algorithm. The proposed S-A-ADMMr-MRI algorithm can lead to much better visual quality.

To visually illustrate the relative speed and performance of the algorithms, Fig. 2 gives the evolutions of the objective function and MSE ($= \frac{1}{N}\|u_k - u\|^2$ with u being the true image of size N). We can conclude from these figures that the accelerated ADMM without restart (S-A-ADMM-MRI) can only convergent into a neighbor region of the optimal objective values and MSE values, although it is faster than the accelerated ADMM with restart. Therefore, the proposed S-A-ADMMr-MRI provides a good trade off between convergence speed and accuracy while compared with the accelerated ADMM without restart and classical ADMM algorithms.

For both human brain and rat brain images with sized 128×128, the undersampled k-space measurements are generated from 66 radial lines of their 2-D Fourier transforms (sampling ratio is 44.22%). The undersampled measurements are also corrupted by Gaussian noise with $\sigma^2 = 0.5 \times 10^{-6}$. In these two cases, we empirically choose $\gamma = 10^{-4}$, $\tau = 0.2\gamma$, $\rho = 5\gamma$. In running algorithms, the stopping criterion is set as $|\text{obj}(k+1) - \text{obj}(k)|/\text{obj}(k) \leq \text{Tol}$, where $\text{obj}(k)$ is the



(a) (b) (c)

Fig. 2. Evolution of the objective function and MSE (redundant wavelets): (a) Objective value versus iteration number. (b) Objective value over time. (c) MSE over time.

objective value at kth iteration, and Tol is the tolerance and chosen as 5×10^{-5} here. Furthermore the same parameter values are chosen for both human brain and rat brain image reconstructions, which are used to illustrate robustness of the proposed S-A-ADMMr-MRI algorithm.

The simulation results on the human brain and rat brain image reconstructions are reported in Table I. The number of iterations (Iters), CPU time (Time) in seconds, the final objective function value (Obj) and MSE tabulated for each experiment are the average values over 10 instances. We can see that for the TVL1-L2 human brain MR image reconstruction, the S-A-ADMMr-MRI algorithm was able to reconstruct the image in 124 iterations with 55.58 seconds and achieved an MSE 20.09, while the S-ADMM-MRI algorithm took 186 iterations with 69.11 seconds to achieve an MSE value of 20.46. For the rat brain MR image reconstruction, the S-A-ADMMr-MRI algorithm took 62 iterations with 26.92 seconds to achieve an MSE value of 28.48, but the S-ADMM-MRI took 106 iterations with 31.75 seconds to obtain an MSE 28.90. Clearly, the proposed S-A-ADMMr-MRI algorithm is more efficient than the classical ADMM algorithm.

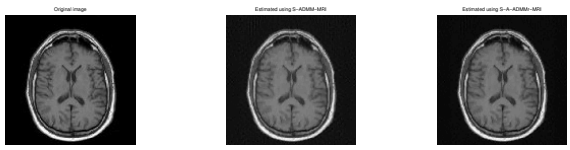
TABLE I
COMPARISON OF THE BRAIN MR IMAGE RECONSTRUCTIONS

Images	Algorithm	Iters	Time	Obj	MSE
Human brain	S-ADMM-MRI	186	69.11	153.721	20.46
	S-A-ADMMr-MRI	124	55.58	153.694	20.09
Rat brain	S-ADMM-MRI	106	31.75	173.695	28.90
	S-A-ADMMr-MRI	62	26.92	173.664	28.48

Moreover, the reconstructed images produced by these algorithms are shown in Figs. 3 and 4. These experimental results show that the reconstructed images by our proposed fast S-A-ADMMr-MRI algorithm have higher quality than those by the classical ADMM algorithm for TVL1-L2 MR image reconstruction under our experimental conditions.

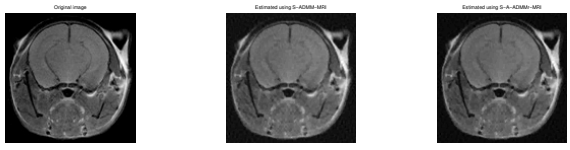
IV. CONCLUSIONS

An efficient algorithm for solving TVL1-L2 MR image reconstruction from undersampled k-space data has been presented by using variable splitting and accelerated ADMM with restart. Experimental results have also shown that for the wavelet frame-based MR image reconstruction with TVL1-L2 regularization from undersampled measurements, the proposed S-A-ADMMr-MRI algorithm is clearly faster than the classical ADMM-based algorithm while obtaining better reconstructed image quality. The experimental results for the underlying human and rat brain MR image reconstructions also show that the proposed S-A-ADMMr-MRI algorithm is robust and efficient.



(a) (b) (c)

Fig. 3. Human brain MR image reconstruction with 120 iterations: (a) Original human brain image. (b) Reconstructed image using variable splitting and classical ADMM method (S-ADMM-MRI, MSE: 32.85). (c) Reconstructed image using variable splitting and accelerated ADMM with restart (S-A-ADMMr-MRI, MSE: 20.63).



(a) (b) (c)

Fig. 4. Rat brain MR image reconstruction with 60 iterations: (a) Original rat brain image. (b) Reconstructed image using variable splitting and classical ADMM method (S-ADMM-MRI, MSE: 37.51). (c) Reconstructed image using variable splitting and accelerated ADMM with restart (S-A-ADMMr-MRI, MSE: 28.41).

REFERENCES

[1] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magn. Resonance Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[2] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact frequency reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Theory*, vol.52, no. 2, pp. 489–509, 2006.

[3] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 288–297, 2010.

[4] J. Aelterman, H. Q. Luong, B. Goossens, A. Pizurica, and W. Philips, "Augmented Lagrangian based reconstruction of non-uniformly sub-Nyquist sampled MRI data," *Signal Process.*, vol. 91, pp. 2731–2742, 2011.

[5] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk, "Fast alternating direction optimization methods," Technical Report CAM12-35, UCLA, May 2012.

[6] A. Beck and M. Teboulle, "Fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[7] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM J. Multiscale Model. Sim.*, vol. 4, pp. 1168–1200, 2005.

[8] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math Imag. Vis.*, vol. 20, no. 1-2, pp. 89-97, 2004.

[9] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer-Verlag, 2006.

[10] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 2009.

[11] S. Xie and S. Rahardja, "An Alternating Direction Method for Balanced Image Restoration," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4557–4567, 2012.