4999

A Fast Double Stochastic Proximal Method for CS-MRI Reconstruction with Multiple Wavelets

Tao Hong[1], Luis Hernandez-Garcia[1], and Jeffrey Fessler[2]
[1]Department of Radiology, University of Michigan, Ann Arbor, Ann Arbor, MI, United States, [2]Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, Ann Arbor, MI, United States

Synopsis

Keywords: Image Reconstruction, Sparse & Low-Rank Models

Motivation: Gu et al. [3] showed one can obtain comparable performance as the physics-guided deep learning (PG-DL) networks [4] for CS-MRI reconstruction by using multiple wavelets as the regularizers.

Goal(s): Develop an efficient numerical algorithm for CS-MRI reconstruction with multiple wavelets.

Approach: Study a fast double stochastic proximal method (FDSPM) for compressed sensing MRI (CS-MRI) reconstruction.

Results: Our experiments demonstrate that FDSPM converges in less CPU time than classical CS algorithms for image reconstruction.

Impact: Exploring efficient algorithms for multiple regularizers CS-MRI reconstruction can motivate new efficient network structures that are easy to train.

Introduction

The CS-MRI reconstruction with $R > 1$ coils and $Q > 1$ different regularizers can be formulated as the following minimization problem [3]:

$$\arg\min_{x\in\mathbb{C}^N} F(x) \equiv \frac{1}{R}\sum_{r=1}^{R} \underbrace{\frac{1}{2}\|A_r x - b_r\|_2^2}_{f_r(x)} + \sum_{g=1}^{Q} g_q(T_q x), \qquad (1)$$

where $A_r :\in \mathbb{C}^{M\times N} = PFS_r$ denotes the forward model defining a mapping from the signal x to the acquired data b_r. P, F, $\{S_r\}_r$, and $\{T_q\}_q$ represent the downsampling mask, the nonuniform FFT, the sensitivity mapping, and the (e.g., non-orthogonal) wavelet transform, respectively. Here, we focus on $g_q(x) = \lambda_q\|x\|_1$.

Methods

Denote by

$$\mathrm{prox}_h(v) = \arg\min_{x\in\mathbb{C}^N} \frac{1}{2}\|x - v\|^2 + h(x).$$

At kth iteration, FDSPM needs to compute

$$x_{k+1} = \mathrm{prox}_h\left(x_k - \frac{1}{L|\Box_k|}\sum_{r\in\Box_k}\nabla f_r(x_k)\right), \qquad (2)$$

where $h(x) = \sum_{q=1}^{Q} g_q(T_q x) = \sum_{q=1}^{Q} \lambda_q\|T_q x\|_1$, L denotes the Lipschitz constant of $\frac{1}{R}\sum_r f_r$, and $\Box_k$ a randomly chosen subset of the whole $\{A_r\}_r$. Define $\Box(x) = [T_1 x; T_2 x; \cdots; T_Q x]$. Then the adjoint of $\Box$ is $\Box^{\Box}(y) = \sum_{q=1}^{Q} T_q^{\Box} y_q$ with $y = [y_1; y_2; \cdots; y_Q]$. With the definition of $\Box$, we rewrite (1) as

$$\mathrm{prox}_h(u_k) = \arg\min_{x\in\mathbb{C}^N} \underbrace{\frac{1}{2}\|x - u_k\|_2^2}_{F(x)} + G(\Box(x)), \qquad (3)$$

where $G(y) = \sum_q g_q(y_q)$ and $u_k = \left(x_k - \frac{1}{L|\Box_k|}\sum_{r\in\Box_k}\nabla f_r(x_k)\right)$. Since $G(y)$ is nonsmooth, we solve (3) via its dual formulation which is

$$\min_y F^*\left(\Box^{\Box}(y)\right) + G^*(-y), \qquad (4)$$

where F^* and G^* are the convex conjugate functions of F and G, respectively. Since Q can be much larger than 1, we use the randomized block proximal gradient method (RBPGM) for (4) that the computation at each iteration is independent of the number of Q. By using the the Moreau decomposition property ($\mathrm{prox}_{\lambda h}(x) + \lambda\mathrm{prox}_{\lambda^{-1}h^*}(x/\lambda) = x$), we can write the primal sequence representation of RBPGM for (4) as described in Figure 1. The main computation at each iteration of Figure 1 is to apply one time T_q and its adjoint since we only need to update one y_q.

Results

All experiments are implemented in SigPy [5] and the brain and knee images from [7] are used as our test image. Figures 1-4 show the results and experimental details.

Conclusion

We propose a FDSPM method for CS-MRI reconstruction using multiple wavelet regularizers.The computation at each iteration of FDSPM is independent of the number of coils R and the number of used wavelets Q. Gu et al. [3] proposed an unroll network based on the alternating direction method of multipliers (ADMM) [1] to solve (1) by only learning $\{\lambda_q\}_q$ and stepsizes. Moreover, [3] showed that their approach yields comparable performance as the PG-DL networks [4] which need to learn millions of parameters.One of the interesting applications of FDSPM is to efficiently train the model proposed in [3] by unrolling FDSPM instead of ADMM; one may also use FDSPM to accelerate the testing stage of the network proposed in [3].

Acknowledgements

This work was funded by National Institutes of Health grant R01NS112233.

References

[1] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization andstatistical learning via the alternating direction methodof multipliers. Foundations and Trends® in Machine Learning, 3(1):1–122, 2011.

[2] Laurent Condat and Peter Richtarik. Randprox: Primal-dual optimization algorithms with randomized proximal updates. arXiv preprint arXiv:2207.12891, 2022.

[3] Hongyi Gu, Burhaneddin Yaman, Steen Moeller, Jutta Ellermann, Kamil Ugurbil, and Mehmet Akcakaya. Re-visiting ℓ_1-wavelet compressed-sensing MRI in the era of deep learning. Proceedings of the National Academy of Sciences, 119(33):e2201062119, 2022.

[4] Vishal Monga, Yuelong Li, and Yonina C Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signaland image processing. IEEE Signal Processing Magazine, 38(2):18–44, 2021.

[5] Frank Ong and Michael Lustig. SigPy: a python packagefor high performance iterative reconstruction. In Pro-ceedings of the ISMRM 27th Annual Meeting, Montreal,Quebec, Canada, volume 4819, 2019.

[6] Yao-Liang Yu. Better approximation and faster algorithmusing the proximal average. Advances in Neural Informa-tion Processing Systems, 26, 2013.

[7] Jure Zbontar, Florian Knoll, Anuroop Sriram et al. fastMRI: An open dataset and benchmarks for accelerated MRI. arXiv preprint arXiv:1811.08839, 2018.

Figures

Algorithm 1 Randomized block proximal gradient method

Initialization:
$\{L_q = \|\boldsymbol{T}_q\|^2\}_q$, $\mathbf{y}^0 = \mathbf{0}$, and the maximal number of iterations Max_Iter.

Output: $\boldsymbol{x}_{k+1} = \arg\max_{\boldsymbol{v}} \{< \boldsymbol{v}, \mathcal{A}^T(\mathbf{y}^{k+1}) > -F(\boldsymbol{v})\}$.

$k \leftarrow 0$
for all $k \leq$ Max_Iter **do**
  $\boldsymbol{v}^k = \arg\max_{\boldsymbol{v}} \{< \boldsymbol{v}, \mathcal{A}^T(\mathbf{y}^k) > -F(\boldsymbol{v})\}$
  % $\boldsymbol{v}^k = \mathcal{A}^T(\mathbf{y}^k) + \boldsymbol{u}^k$
  Pick $i_k \in \{1, 2, \cdots, Q\}$
  $\mathbf{y}_{i_k}^{k+1} = \mathbf{y}_{i_k}^k - \frac{1}{L_{i_k}} \boldsymbol{T}_{i_k} \boldsymbol{v}^k + \frac{1}{L_{i_k}} \mathrm{prox}_{L_{i_k} g_{i_k}} \left(\boldsymbol{T}_{i_k} \boldsymbol{v}^k - L_{i_k} \mathbf{y}_{i_k}^k \right)$
  $k \leftarrow k + 1$
end for

Figure 1. The Randomized block proximal gradient method for (4).

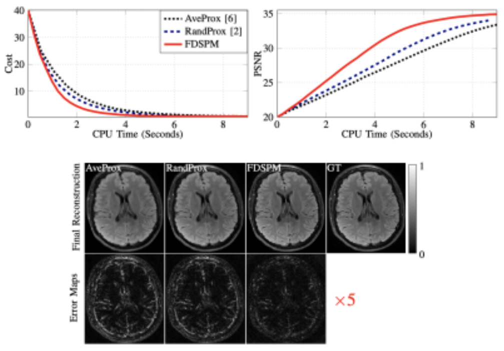

Figure 2. Performance on the brain image with $Q = 9$ different wavelets, i.e., 'haar', 'db2', 'db3', and 'db4' with 4 levels and 'db10', 'sym3', 'sym5', 'sym8', and 'sym9' with 3 levels. Acquisition: spiral trajectory with 32 interleaves 1688 readout points and $R = 12$ coils. Matrix size $= 256 \times 256$. FDSPM settings: $|\square_k| = 4$ and Max_Iter = 6. First row: the cost and PSNR versus the CPU time; Second row: the final reconstructed images and the ground truth; Third row: the corresponding error maps $\times 5$.

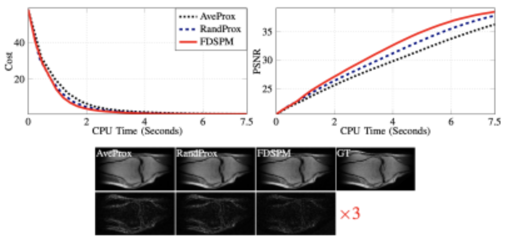

Figure 3. Performance on the knee image. Same setting as Figure 2.

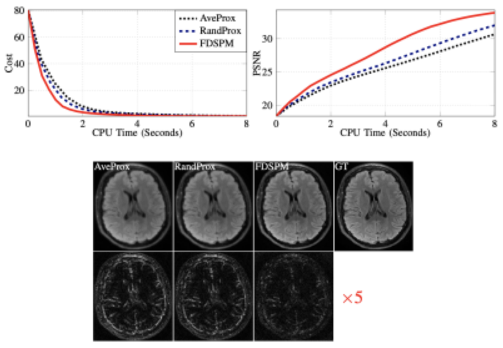

Figure 4. Performance on the brain image with $Q = 9$ different wavelets. Acquisition: radial trajectory with 96 spokes 512 readout points and $R = 12$ coils. Matrix size $= 256 \times 256$. FDSPM settings: $|\square_k| = 4$ and Max_Iter = 6.

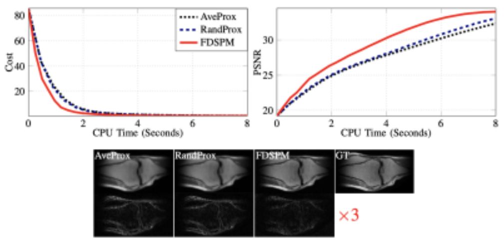

Figure 5. Performance on the knee image. Same setting as Figure 4.

Proc. Intl. Soc. Mag. Reson. Med. 32 (2024)
4999