Efficient Preconditioning in Joint Total Variation Regularized Parallel MRI Reconstruction*

Zheng Xu[1], Yeqing Li[1], Leon Axel[2], and Junzhou Huang[1,**]

[1] Department of Computer Science and Engineering,
University of Texas at Arlington, Arlington TX 76019, USA
[2] Department of Radiology, New York University, New York, NY 10016, USA
jzhuang@uta.edu

Abstract. Parallel magnetic resonance imaging (pMRI) is a useful technique to aid clinical diagnosis. In this paper, we develop an accelerated algorithm for joint total variation (JTV) regularized calibrationless Parallel MR image reconstruction. The algorithm minimizes a linear combination of least squares data fitting term and the joint total variation regularization. This model has been demonstrated as a very powerful tool for parallel MRI reconstruction. The proposed algorithm is based on the iteratively reweighted least squares (IRLS) framework, which converges exponentially fast. It is further accelerated by preconditioned conjugate gradient method with a well-designed preconditioner. Numerous experiments demonstrate the superior performance of the proposed algorithm for parallel MRI reconstruction in terms of both accuracy and efficiency.

1 Introduction

Parallel MR imaging is a powerful method that uses multiple receiver coils for reducing scanning time in MRI[5,6]. Based on the way in utilizing the sensitivity information and local kernel in k-space, these methods are classified broadly into two main types. Reconstruction techniques such as SENSE[11] and CSSENSE[8] expect accurate estimation of reception profiles from each coil element to optimally reconstruct undersampled MR image. However, it is often very difficult to accurately and robustly measure the sensitivities and even small errors can result in inconsistencies that lead to visible artifacts in the image. These disadvantages therefore motivate the other type of methods, termed auto-calibrating methods, e.g. GRAPPA[4] and SPIRiT[9], that derive sensitivity information from auto-calibration signals (ACSs) and thus avoid side effects brought by the difficult and inaccurate sensitivity map estimation. However, it is often limiting or totally infeasible to acquire sufficient ACSs. For example, for non-Cartesian imaging, ACS acquisition requires much longer time and can probably lead to artifacts due to off-resonance. To overcome these shortcomings, several calibrationless methods have been proposed recently, e.g. CaLMMRI[10], FISTA_JTV[2] and SAKE[15].

* This work was partially supported by U.S. NSF IIS-1423056, CMMI-1434401, CNS-1405985.

** Corresponding author.

N. Navab et al. (Eds.): MICCAI 2015, Part II, LNCS 9350, pp. 563–570, 2015.
DOI: 10.1007/978-3-319-24571-3_67

Among these methods, joint total variation (JTV) model has been demonstrated as a powerful tool for calibrationless parallel MR image reconstruction. The JTV model is designed based on the observation of the gradient sparsity of each coil image and the cross-channel similarity of parallel MR images. Previous attempt to solve this model is shown in [2] which is based on FISTA_JTV algorithms. The numerical experiments exhibit its effectiveness and efficiency on the parallel MR images with real measurements. However, real world parallel MR images are often sampled in complex measurements. For complex parallel MR images, the FISTA_JTV algorithm usually fails to solve the model efficiently as it requires more inner loop iterations to converge.

Here, we propose a novel optimization scheme to solve the joint total variation model more efficiently based on the iteratively reweighted least squares (IRLS) framework. It preserves the fast convergence speed of traditional IRLS which converges exponentially fast. Since it requires solving a linear inverse subproblem in each IRLS step, we propose a new pseudo-diagonal preconditioner to significantly accelerate this process with preconditioned conjugate gradient method. Extensive experiment results show that it impressively outperforms previous state-of-art methods for the MR image reconstruction in terms of both reconstruction accuracy and computational complexity.

2 Joint Total Variation Regularized Model

Based on the assumption that the gradients of aliased images from all the coils are jointly sparse, the formulation of joint total variation (JTV) model is designed as follows[2]:

$$\min_x \frac{1}{2}\|\mathcal{R}\mathcal{F}x - b\|^2 + \lambda\|[\nabla_h x, \nabla_v x]\|_{2,1}. \tag{1}$$

where $x \in \mathbb{C}^{m \times n \times c}$, is the c-channel parallel MR image with each coil size $m \times n$. $\mathcal{R}$ is the subsampling operator in frequency domain, and $\mathcal{F}$ is the Fourier transform operator. λ is the non-negative tunning parameter balancing the data fitting and JTV regularization. The JTV regularization term $\|[\nabla_h x, \nabla_v x]\|_{2,1} = \sum_{i=1}^{m}\sum_{j=1}^{n}\sqrt{\sum_{k=1}^{c}(\nabla_h x_{i,j,k})^2 + (\nabla_v x_{i,j,k})^2}$ and ∇_h and ∇_v are the horizontal and vertical discrete gradient operators (i.e. $\nabla_h x_{i,j,k} = x_{i,j+1,k} - x_{i,j,k}$, $\nabla_v x_{i,j,k} = x_{i+1,j,k} - x_{i,j,k}$). The JTV regularization sums up the horizontal and vertical discrete gradients across all data channels. Therefore, minimizing the JTV regularization can lead to the joint gradient sparse solution for pMRI reconstruction.

Due to the non-smoothness of JTV regularization, it's hard to optimize the JTV objective function efficiently. Although FISTA_JTV[2] has been proposed as an accelerated algorithm and experimentally proven as a fast method for parallel MR images with real number measurements, it remains challenging to efficiently reconstruct parallel MR images in complex measurements. This situation therefore motivates us to develop a faster algorithm, especially for complex measurements, to solve JTV model in the next section.

3 Algorithm

3.1 Iteratively Reweighted Least Squares Framework

In this section, we first briefly review the iteratively reweighted least squares (IRLS) method[1,3] and show how to fit JTV model into the IRLS framework. The key idea of the IRLS method is to approximate the ℓ_1 regularization by a weighted ℓ_2 regularization, making the loss function strongly convex. The formulation is then able to be optimized in a linear convergence rate as shown in Theorem 1.

Theorem 1. *(Theorem 6.1 in [3]) Let $\{x^t\}$ be the sequence generated by IRLS method, x^* be the ℓ_1-minimizer with $\mu \in (0,1)$, thus*

$$\|x^t - x^*\|_1 \leq \mu^t \|x^0 - x^*\|_1. \tag{2}$$

Using the techniques introduced in [7], we have the weighted ℓ_2 form for the JTV model (1).

$$\min_x \frac{1}{2}\|\mathcal{R}\mathcal{F}x - b\|^2 + \frac{\lambda}{2}\langle \nabla_h x, W^t \nabla_h x\rangle + \frac{\lambda}{2}\langle \nabla_v x, W^t \nabla_v x\rangle. \tag{3}$$

where $\langle \nabla_h x, W^t \nabla_h x\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{c} W^t_{i,j,k}(\nabla_h x_{i,j,k})^2$, and similarly $\langle \nabla_v x, W^t \nabla_v x\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{c} W^t_{i,j,k}(\nabla_v x_{i,j,k})^2$. Here $W^t_{i,j,k} = (\sqrt{\sum_{p=1}^{c}(\nabla_h x^t_{i,j,p})^2 + (\nabla_v x^t_{i,j,p})^2} + \varepsilon)^{-1}$ for $k = 1, 2, \ldots, c$ and ε is a real infinitesimal added to avoid $W^t_{i,j,k}$ to be infinite.

By solving the Euler-Lagrange equation of (3), we have the least squares subproblem in each IRLS iteration

$$(\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F} + \lambda\langle \nabla_h, W^t\nabla_h\rangle + \lambda\langle \nabla_v, W^t\nabla_v\rangle)x = \mathcal{F}^H\mathcal{R}^H b. \tag{4}$$

where the superscript H denotes the conjugate transpose of a linear bounded operator. However, it is not feasible to calculate exact matrix inverse since it requires $\mathcal{O}(pq^2)$ time for a $p \times q$ matrix. Another option is to use the classical iterative methods, e.g. Jacobian, Gauss-Seidel iteration whose convergence are not guaranteed. Therefore, it is challenging to design an efficient algorithm with clear theoretical convergence justification to solve the subproblem (4).

3.2 Preconditioned Conjugate Gradient Descent

In the sequel, we study the preconditioned conjugate gradient (PCG) method aiming at solving subproblem (4) efficiently. We first show, in Theorem 2, the conjugate gradient(CG) algorithm is able to solve the least squares problem in a linear convergence rate.

Theorem 2. *(Section 9.2 in [14]) Let $\{x^t\}$ be the sequence generated by conjugate gradient iteration, x^* be the optimal solution for $b = Ax$. Let κ be the condition number of A, $\|x\|_A = \langle x, Ax\rangle$ be the energy norm, we have*

$$\|x^t - x^*\|_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^t \|x^0 - x^*\|_A. \tag{5}$$

The practical convergence speed of CG highly depends on the condition number κ of A. When κ is relatively large, the coefficient of convergence $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ approaches 1, leading to the arbitrarily slow convergence of CG. However, in practices, the A is usually with high condition number. It thus motivates the preconditioning techniques which is typically related to reducing condition number of A by designing a transformation P. The design of preconditioner is problem-dependent and the preconditioner not only needs to be as close as possible to the original system matrix A but also be able to be inverted efficiently.

For least squares subproblem (4), we design a novel preconditioner to accelerate the CG method. Observing $\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}$ is diagonal dominant, we can discard the non-diagonal elements in $\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}$ without bringing in large error. In this way, we can design the following preconditioner for solving subproblem (4)

$$P = \overline{\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}}I + \lambda\langle\nabla_h, W^t\nabla_h\rangle + \lambda\langle\nabla_v, W^t\nabla_v\rangle. \tag{6}$$

where $\overline{\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}}$ denotes the mean of diagonal elements of $\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}$ and I is the identity matrix. The $\overline{\mathcal{F}^H\mathcal{R}^H\mathcal{R}\mathcal{F}}I$ is bounded and the parameter λ is usually small (e.g. 10^{-6}) hence successfully suppressing the condition number of the preconditioner P. Moreover, this preconditioner is observed as a penta-diagonal matrix, whose inverse can be evaluated in linear time by [12]. Therefore, the proposed preconditioner is able to solve problem (4) more accurately, while, in the meantime, it does not increase time complexity.

The proposed algorithm is summarized in Algorithm 1. We denoted this algorithm as *PRIM* which is short for PReconditioned Iterative reweighted Method for parallel MRI reconstruction. Although the proposed algorithm has inner loop, we observe that usually ten PCG iterations are sufficient to obtain a solution very close to the optimal one for the parallel MRI reconstruction. This is because both the inner and outer loops have linear convergence rates. The theoretically fast convergence constitutes a key feature of the proposed method.

Another key feature of our method is the cost of each iteration is only $\mathcal{O}(cmn\log(mn))$. The step of updating $W^t_{i,j,k}$ and P^t requires $\mathcal{O}(cmn)$ time. Updating S^t requires $\mathcal{O}(cmn\log(mn))$ time since in each step it requires to evaluate Fast Fourier Transformation. The inverse of P^t can be calculated in $\mathcal{O}(cmn)$ time as shown in [12]. As a result, the time complexity of each iteration in PRIM is $\mathcal{O}(cmn\log(mn))$.

With these two key features, the PRIM efficiently solves the compressive sensing parallel MR image reconstruction model regularized by joint total variation. The experiment results in the next section demonstrate its superior performance compared with all previous state-of-art methods for pMRI reconstruction.

Algorithm 1. PRIM

Input: $\mathcal{R}$, b, x^0, λ, $t = 0$, ε
while not meet the stopping criterion **do**
  $W^t_{i,j,k} := 1/(\sqrt{\sum_{p=1}^{c} (\nabla_h x^t_{i,j,p})^2 + (\nabla_v x^t_{i,j,p})^2 + \varepsilon})$
  $S^t := \mathcal{F}^H \mathcal{R}^H \mathcal{R} \mathcal{F} + \lambda \langle \nabla_h, W^t \nabla_h \rangle + \lambda \langle \nabla_v, W^t \nabla_v \rangle$
  $P^t := \overline{\mathcal{F}^H \mathcal{R}^H \mathcal{R} \mathcal{F}} I + \lambda \langle \nabla_h, W^t \nabla_h \rangle + \lambda \langle \nabla_v, W^t \nabla_v \rangle$
  while not meet the PCG stopping criterion **do**
    Update x^{t+1} by PCG for $S^t x = (\mathcal{R}\mathcal{F})^H b$ with preconditioner P^t
  end while
  $t := t + 1$
end while

4 Experiments

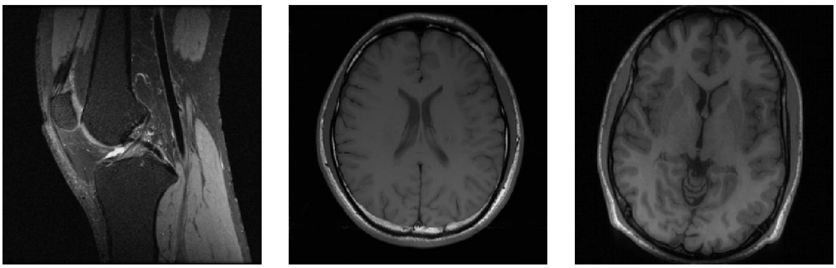
4.1 Experiment Setup

The experiments are conducted on three parallel MRI datasets in Figure 1.

***3T Knee**[13].* The MR image shown in figure 1(a) is a 3D FSE CUBE sequence with proton density weighting scanned on a GE 3T whole body scanner (TE=25ms, TR=1550ms, FOV=160mm, 320×320 matrix). This dataset is open access to public in `http://mridata.org`.

***3T Brain**[8].* Figure 1(b) shows an image scanned from a GE 3T commercial scanner with an eight-channel head coil using a two-dimensional T1-weighted spin echo protocol (TE = 11ms, TR = 700ms, FOV = 22cm, 256×256 pixels).

***Signa-Excite 1.5T Brain**[9].* Figure 1(c) shows a T1-weighted image from spoiled gradient echo (SPGR) sequence, scanned on a GE Signa-Excite 1.5-T scanner with an eight-channel receive coil (TE = 8ms, TR=17.6 ms, FOV = 20cm, 200×200 pixels).



(a) 3T Knee (b) 3T Brain (c) Signa-Excite 1.5T Brain

Fig. 1. Three MR images used in the experiments.

We implemented the proposed method for problem (1) and apply them to MRI k-space data with complex measurements. All experiments are conducted on a PC with Intel i7-4770 @ 3.40 GHz CPU and 16 GB RAM. We compare the proposed method with the state-of-art methods GRAPPA[4], CGSPIRiT[9], and CSSENSE[8]. Moreover, we compare them with SAKE[15] which is another calibrationless method proposed recently. We also compare the proposed method with FISTA_JTV[2] that solves the same JTV model. For fair comparisons, we download the codes from their websites and follow their default parameter settings carefully.

4.2 Numerical Results

Figure 2 shows the reconstruction results on 3T Knee MR image at a reduction factor $R = 4$, together with the ground-truth image. Quantitative comparison results for 3T Knee, 3T Brain and Signa-Excite 1.5T Brain datasets are shown in Table 1, 2 and 3, respectively. Compared to all other methods, the proposed method always preserves most details and suppress most noise (as shown in the zoomed region of interest). It does make sense because the JTV regularization acquires the prior knowledge that the gradients of MR images are typically sparse. The reconstruction results of GRAPPA, CGSPIRiT and SAKE are much more noisy in that their models do not include that assumption. Moreover, the gradient of each image coil(channel) is assumed similar in the JTV model which is the main reason for the better performance than CSSENSE. In this way, the JTV regularization makes the reconstruction result purer and more recognizable compared with other methods.

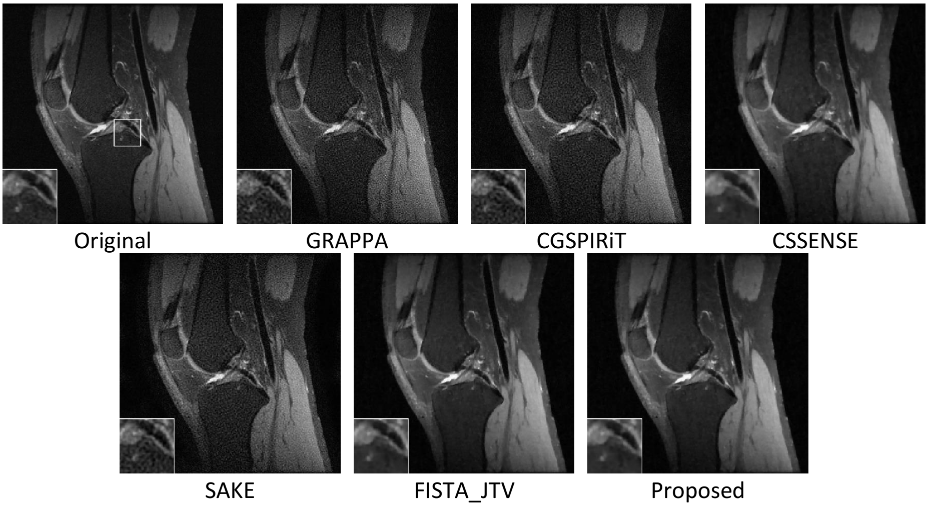


Fig. 2. Visual results of different methods compared with ground-truth. Best viewed in $\times 2$ pdf.

Table 1. Quantitative comparison on 3T Knee dataset.

	GRAPPA	CGSPIRiT	CSSENSE	SAKE	FISTA_JTV	Proposed
RMSE ($\times 10^{-4}$)	11.7595	9.7159	6.9781	9.5650	6.3886	**6.3323**
SNR (db)	9.7149	11.3729	14.2479	11.5089	15.0145	**15.0914**
Time (s)	171.81	9.85	193.21	27.35	115.80	**9.74**

Table 2. Quantitative comparison on 3T Brain dataset.

	GRAPPA	CGSPIRiT	CSSENSE	SAKE	FISTA_JTV	Proposed
RMSE ($\times 10^{-4}$)	4.0159	3.5716	5.7170	3.3266	3.2563	**3.1963**
SNR (db)	21.9516	22.9699	18.8838	23.5871	23.7726	**23.9341**
Time (s)	172.69	9.79	193.40	28.34	113.81	**9.36**

Table 3. Quantitative comparison on Signa-Excite 1.5T Brain dataset.

	GRAPPA	CGSPIRiT	CSSENSE	SAKE	FISTA_JTV	Proposed
RMSE ($\times 10^{-4}$)	6.7155	5.5824	5.5891	4.7867	4.5506	**4.5155**
SNR (db)	17.2415	18.8467	18.8363	20.1823	20.6218	**20.6889**
Time (s)	173.03	9.82	193.47	28.49	116.07	**9.49**

It has also been noticed that our method always consumes least time on each MR image. GRAPPA requires more time for calibration when using random mask. SAKE calculates Singular Value Decomposition (SVD) of the pMRI data tensor in each iteration. The SVD takes $\mathcal{O}(n^3)$ time, resulting in the less efficiency compared with the proposed method. FISTA_JTV requires more inner loop iterations to converge in complex measurements. Overall, the proposed method is able to outperform the other methods in computational performance due to its superior convergence property and lower per-iteration computational cost.

5 Conclusion

We have proposed an novel algorithm PRIM to solve the joint total variation model for parallel MRI reconstruction. It is based on the iteratively reweighted least squares framework and preconditioned conjugate gradient method. Moreover, we have designed a novel preconditioner to strengthen its converge property as it requires to efficiently solve the least squares subproblem. The efficiency of PRIM is theoretically guaranteed and also exhibited in extensive experiments. With the joint gradient sparsity assumption in JTV model, the proposed method is able to provide much more accurate reconstruction result than other state-of-art methods with less time. All these benefits lead us closer to the calibrationless real-time parallel MRI reconstruction than ever before.

References

1. Chen, C., Huang, J., He, L., Li, H.: Preconditioning for accelerated iteratively reweighted least squares in structured sparsity reconstruction. In: CVPR 2014, pp. 2713–2720. IEEE (2014)
2. Chen, C., Li, Y., Huang, J.: Calibrationless parallel MRI with joint total variation regularization. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part III. LNCS, vol. 8151, pp. 106–114. Springer, Heidelberg (2013)
3. Daubechies, I., DeVore, R., Fornasier, M., Güntürk, C.S.: Iteratively reweighted least squares minimization for sparse recovery. Comm. on Pure and Applied Math. 63(1), 1–38 (2010)
4. Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn. Reson. Med. 47(6), 1202–1210 (2002)
5. Huang, J., Zhang, S., Li, H., Metaxas, D.: Composite splitting algorithms for convex optimization. Computer Vision and Image Understanding 115(12), 1610–1622 (2011)
6. Huang, J., Zhang, S., Metaxas, D.: Efficient MR image reconstruction for compressed MR imaging. Medical Image Analysis 15(5), 670–679 (2011)
7. Hunter, D.R., Lange, K.: A tutorial on MM algorithms. The Amer. Stat. 58(1), 30–37 (2004)
8. Liang, D., Liu, B., Wang, J., Ying, L.: Accelerating SENSE using compressed sensing. Magn. Reson. Med. 62(6), 1574–1584 (2009)
9. Lustig, M., Pauly, J.M.: SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space. Magn. Reson. Med. 64(2), 457–471 (2010)
10. Majumdar, A., Ward, R.K.: Calibration-less multi-coil mr image reconstruction. Magn. Reson. Imag. 30(7), 1032–1045 (2012)
11. Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., et al.: SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42(5), 952–962 (1999)
12. Saad, Y.: Iterative methods for sparse linear systems. Siam (2003)
13. Sawyer, A.M., Lustig, M., Alley, M., Uecker, P., Virtue, P., Lai, P., Vasanawala, S., Healthcare, G.: Creation of fully sampled MR data repository for compressed sensing of the knee (2013)
14. Shewchuk, J.R.: An introduction to the conjugate gradient method without the agonizing pain (1994)
15. Shin, P.J., Larson, P.E., Ohliger, M.A., Elad, M., Pauly, J.M., Vigneron, D.B., Lustig, M.: Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. Magn. Reson. Med. 72(4), 959–970 (2014)