The Benefit of Tree Sparsity in Accelerated MRI

Chen Chen and Junzhou Huang

Department of Computer Science and Engineering,
The University of Texas at Arlington , TX, USA 76019

Abstract. The wavelet coefficients of a 2D natural image are not only very sparse with only a small number of coefficients have large values, but also yield a quadtree structure. According to structured sparsity theory, the required measurement bounds for compressive sensing reconstruction can be reduced to $\mathcal{O}(K + \log n)$ by exploiting this tree structure rather than $\mathcal{O}(K + K \log n)$ for standard K-sparse data. In this paper, we proposed a new model to validate how much the wavelet tree structure can help to accelerate Magnetic Resonance Imaging (MRI). This model is decomposed to two subproblems. The first subproblem has closed form solution. For the other one, we apply FISTA to solve it, which guarantees this subproblem can be solved with the similar convergence to the existing fastest MRI algorithms. Numerous experiments are conducted to validate how much benefit it can bring by tree sparsity. Experimental results show that the proposed method can definitely improve existing MRI algorithms, although with gaps to the theory.

1 Introduction

According to compressive sensing (CS) theory [1] [2], only $\mathcal{O}(K + K \log n)$ sampling measurements are enough to recover K-sparse data with length n. Applying this theory in Magnetic Resonance Imaging (MRI), the MR scanning time can be significantly reduced [3]. Suppose x is a MR image and R is a partial Fourier transform, the sampling measurement b of x is defined as $b = Rx$. Recent methods all can reconstruct MR images with good quality from approximate 20% sampling [3] [4] [5] [6]. They have a general model for the MRI problem:

$$\hat{x} = \arg\min_{x} \{ \frac{1}{2} \|Rx - b\|^2 + \alpha \|x\|_{TV} + \beta \|\Phi x\|_1 \} \tag{1}$$

where α and β are two positive parameters; and Φ denotes a wavelet transform. It is based on the fact that smooth MR images of organs should have relatively small total variations and these images can be sparsely represented in the wavelet domain. TV is defined as: $\|x\|_{TV} = \sum_i \sum_j \sqrt{(\nabla_1 x_{ij})^2 + (\nabla_2 x_{ij})^2}$, where ∇_1 and ∇_2 denote the forward finite difference operators on the first and second coordinates.

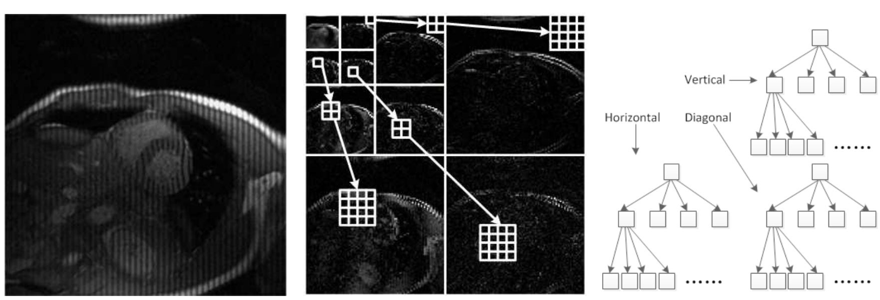
In this model, both TV and $L1$ norms are nonsmooth which makes this problem has no closed form solution. Classical conjugate gradient decent method is first used to solve this problem [3]. TVCMRI [4] and RecPF [5] use an operator-splitting method and a variable splitting method to solve this problem respectively. FCSA [6] decomposes the original problem into two easy subproblems and separately solve each of them with FISTA [7] [8]. These are the state-of-the-art algorithms for CS-MRI.

Recent works show that the required measurements can be further reduced by exploiting the structure of the sparse prior [9] [10]. Specially, only $\mathcal{O}(K + \log n)$ sampling measurements are needed for tree sparse data. Some algorithms have been proposed to improve standard CS recovery by utilizing the tree structure of wavelet coefficients [11] [12] [13] [14]. Although their experiments show the improvement by tree sparsity, none of them conducts experiments on MR images to validate the practical benefit for accelerated MRI. In addition, for an image with 256×256 pixels, there are big gaps between their improvements and the theory.

In this paper, we propose a new model with wavelet tree sparsity instead of the wavelet sparsity in FCSA, to validate how much it can improve the result of standard sparsity. The tree structure is modeled as an overlapping group regularization. The original problem can be decomposed to two subproblems and solved very efficiently by existing techniques. Numerous experiments are conducted to validate whether tree sparsity has the benefit and how much is the benefit in practical MRI. The results show that the existing methods can be further improved by exploiting the wavelet tree structure.

2 Related Work

2.1 Theoretical Benefit of Wavelet Tree Structure



(a) (b) (c)

Fig. 1: Wavelet quadtree structure: a) A cardiac MR image; (b)(c) The corresponding tree structure of the wavelet coefficients.

The wavelet coefficients for natural data (signals or images) are often very sparse, with only a small number of the coefficients have large values and a large fraction of them are approximate zeros. Apart from this, the wavelet coefficients also yield a quadtree structure for a 2D image. The coefficients in the coarsest scale can be seen as the root nodes and the coefficients in the finest scale are the leaf nodes. Each coefficient (non leaf) has four children in the finer scale below it. Figure 1 shows the wavelet quadtree structure of a MR image.

Besides the sparsity of wavelet coefficients, this tree structure also provides another good prior for compressive sensing recovery. If a parent coefficient has a large/small value, its children also tend to be large/small. If this prior is fully exploited, only $\mathcal{O}(K + \log n)$ measurements is needed to recover tree-sparse data rather than $\mathcal{O}(K + K \log n)$ for standard K-sparse data [9] [10]. The improvement by exploiting tree structure can be significant when n is large.

2.2 Algorithmic Benefit of Wavelet Tree Structure

Rao et al. consider the wavelet tree structure as group lasso regularization [11]:

$$\min_x \{F(x) = \frac{1}{2}\|A\theta - b\|_2^2 + \beta \sum_{g \in \mathcal{G}} ||\theta_g||_2\} \tag{2}$$

where θ is the wavelet coefficients. $A = R\Phi^T$ for MR image reconstruction problem, Φ^T is an inverse wavelet transform. β is positive parameter, $\mathcal{G}$ denotes the all parent-child groups and g is the group index. When θ is recovered, it can be transferred to the image by an inverse wavelet transform.

They replicate the overlapped elements and propose two convex models OGL and OGLR with only non-overlapping terms. These tree-based group lasso methods have shown great benefit when comparing with standard lasso.

In statistical learning, Turbo AMP [14] models the wavelet coefficients as conditionally Gaussian with hidden Markov tree states. Both the wavelet coefficients and states are propagated on a factor graph. It assumes the observed measurements are linearly containing with noise:

$$y = Ax + w = A\Phi^T\theta + w \tag{3}$$

x is the original image to be reconstructed and w is Gaussian white noise. MCMC [12], and VB [13] also solve (3) with probabilistic inference. They model the coefficients with hierarchical tree graph. The posterior probability of coefficients are derived from the value of their parents. All these algorithms with tree structure show superiority to standard sparse methods. However, none of them have been validated on real MR images.

3 Algorithm

We also model the wavelet tree structure as overlapping group regularization (2). We do not introduce OGL and OGLR [11] into MRI problem as following reasons: a) OGLR needs to replicate the matrix A, which brings much inconvenience for the partial Fourier transform and will slow down the whole algorithm; b) the parent-child relationship in OGL is hard to track; c) they apply SpaRSA [15] to solve their models, with a relative slow convergence rate. Instead, we extend the overlapping term to nonoverlapping with a sparse matrix G. Then the original problem (2) is transferred to:

$$\hat{x} = \arg\min_{x,z}\{\frac{1}{2}\|Rx - b\|_2^2 + \beta \sum_{g_i=1}^{s} ||z_{g_i}||_2 + \frac{\lambda}{2}||z - G\Phi x||_2^2\} \tag{4}$$

where β and λ are positive parameters and Φ denotes a wavelet transform. g_i denotes one of the parent-child groups and s is the total number of groups. G is a sparse binary matrix indicates the grouping index with only one of 1 in each row. z is the extended vector of wavelet coefficients x without overlapping.

All terms in our model are convex. For the z subproblem:

$$z_{g_i} = \arg\min_{z_{g_i}}\{\beta||z_{g_i}||_2 + \frac{\lambda}{2}||z_{g_i} - (G\Phi x)_{g_i}||_2^2\}, i = 1, 2, ..., s \tag{5}$$

It has closed form solution by soft thresholding:

$$z_{g_i} = \max(||r_i||_2 - \frac{\beta}{\lambda}, 0)\frac{r_i}{||r_i||_2}, i = 1, 2, ..., s \tag{6}$$

where $r_i = (G\Phi x)_{g_i}$. We denote this step by $z = shrinkgroup(G\Phi x, \frac{\beta}{\lambda})$ for convenience.

For the x-subproblem:

$$x = \arg\min_{x}\{\frac{1}{2}\|Rx - b\|_2^2 + \frac{\lambda}{2}||z - G\Phi x||_2^2\} \tag{7}$$

This is a combination of two quadratic terms and has closed form solution: $x = (R^T R + \lambda\Phi^T G^T \Phi G)^{-1}(R^T b + \Phi^T G^T z)$. However, the inverse of $R^T R + \lambda\Phi^T G^T \Phi G$ is not easily obtained. In order to validate the benefit of tree structure, we apply FISTA to solve the x subproblem, which can match the convergence rate of FCSA. Let $f(x) = \frac{1}{2}\|Rx - b\|_2^2 + \frac{\lambda}{2}||z - G\Phi x||_2^2$, which is a convex and smooth function with Lipschitz L_f, and $g(x) = 0$. Then our algorithm can be summarized in Algorithm 1.

The proximal map is defined for any scaler $\rho > 0$:

$$prox_\rho(g)(x) := \arg\min_{u}\{g(u) + \frac{1}{2\rho}\|u - x\|^2\} \tag{8}$$

and $\nabla f(r^k) = R^T(Rr^k - b) + \lambda\Phi^T G^T(G\Phi r^k - z)$. R^T and Φ^T denote the inverse partial Fourier transform and the inverse wavelet transform. Note that G is a sparse matrix with each row containing only one nonzero element 1. Suppose

Algorithm 1 TreeMRI

Input: $\rho = 1/L_f$, $r^1 = x^0$, $t^1 = 1, \beta, \lambda, N$
for $k = 1$ **to** N **do**
  $z = shrinkgroup(G\Phi x^{k-1}, \beta/\lambda)$
  $x^k = r^k - \rho \nabla f(r^k)$
  $t^{k+1} = [1 + \sqrt{1 + 4(t^k)^2}]/2$
  $r^{k+1} = x^k + \frac{t^k - 1}{t^{k+1}}(x^k - x^{k-1})$
end for

x is an image with n pixels. The *shrinkgroup* step can implemented in only $\mathcal{O}(n \log n)$ time and the gradient step also takes $\mathcal{O}(n \log n)$. We can find the total time complexity in each iteration is still $\mathcal{O}(n \log n)$, the same as that of TVCMRI, RecPF and FCSA. This good feature guarantees the proposed algorithm could be comparable with the fastest MRI algorithms in terms of execution speed.

4 Experiments

4.1 Experiment Setup

For fair comparisons, we follow the experiment setup used in previous works [4] [6] and download codes from their websites. Suppose R is a partial Fourier transform with m rows and n columns. The sampling ratio is defined as m/n. Lower sampling ratio needs less MR scanning time to acquire. We follow the sampling strategy of previous works [4] [6], which randomly choose more Fourier coefficients from low frequency and less on high frequency.

All experiments are on a laptop with 2.5GHz Intel core i5 2530M CPU. Matlab version is 7.8(2009a). All measurements are added with 0.01 Gaussian white noise. Signal-to-Noise Ratio (SNR) is used for result evaluation. We mainly compare the proposed algorithm with the classical method CG [3] and several fastest MRI algorithms TVCMRI [4], RecPF [5] and FCSA [6]. In order to show the benefit of tree structure, we remove the TV term in all algorithms.We use the same setting $\beta = 0.035$ in previous works and $\lambda = 0.2 \times \beta$ for our model.

For convenience, all test images are resized to 256×256 and the wavelet decomposition level is set to 4. To perform fair comparisons, all methods run 50 iterations except that the CG runs only 8 iterations due to its higher computational complexity.

4.2 Visual Comparisons

We compare proposed tree-based algorithm with the fastest MRI algorithms to validate how much the tree structure can improve existing results. Total variation terms are all removed, leaving only wavelet sparsity and wavelet tree sparsity to be compared. We conduct experiments on MR images that used in previous work [6]. Figure 2 shows the visual results on a Cardiac MR image. It can be found that the SNR has been improved by the proposed tree-based algorithm. The image reconstructed by the proposed algorithm has detailed textures and is the closest to the original one. Experiments on other MR images are also conducted but not shown here due to the page limitation. All results show that the wavelet tree sparsity obtain significantly better reconstruction over standard sparsity.

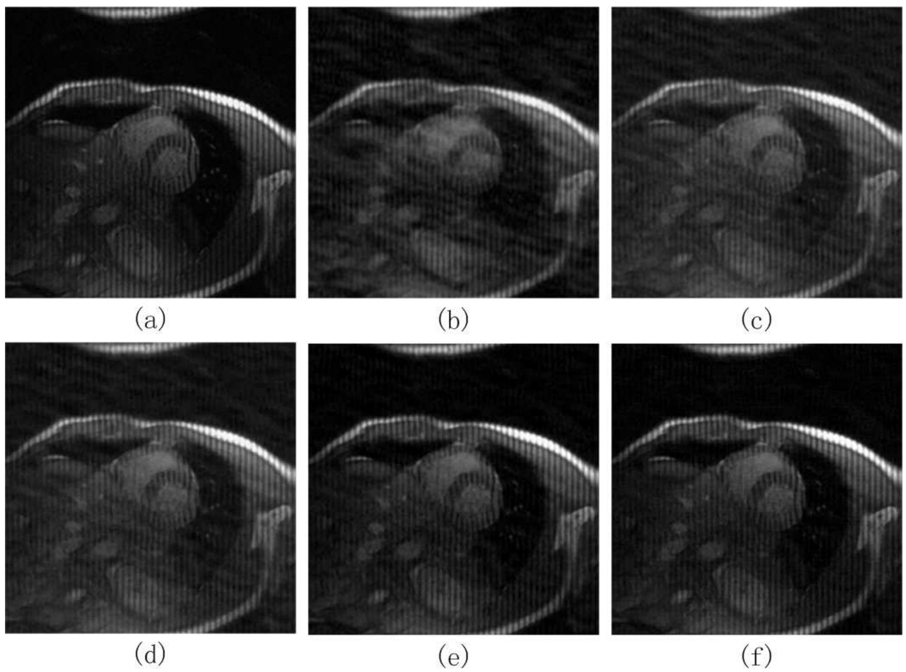


Fig. 2: Cardiac MR image reconstruction from 20% sampling. a) The original image Recovered by : b) CG [3]; (b) TVCMRI [4]; d) RecPF [5]; e) FCSA [6]; f) the proposed algorithm. All algorithms are without total variation regularization. Their SNR are 9.86, 14.31, 15.14, 17.31 and 17.93.

4.3 CPU Time and SNRs

In the last subsection, our experiment confirms the conclusion in [6] that the FCSA [6] is better than TVCMRI [4] and RecPF [5] and far better than the classical CG [3]. We give the performance comparisons between different methods in terms of the CPU time over SNR in this subsection. Besides the Cardiac image, we also show the SNR performance in terms of both iterations and execution time on the Chest image used in previous work [6]. In Figure 3, the SNR of the proposed tree-based algorithm always is the highest in each iteration. When considering CPU execution time, the superiority is weakened by its higher time cost. Although the benefit obtained by wavelet tree structure is far away from the conclusion in theory, it is reasonable because we only model the tree structure as overlapping group sparsity. Only parent-child grouping is contained in the model but not the whole tree graph structure.

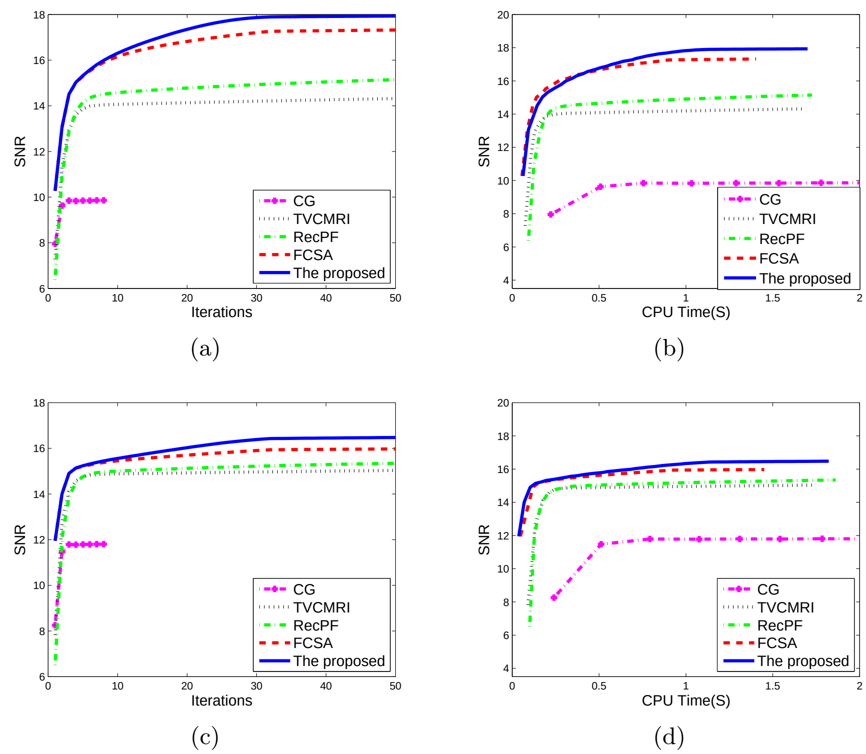


Fig. 3: Performance comparisons on: a)b) The Cardiac image; c)d) The Chest image. All algorithms are without total variation regularization.

5 Conclusion and Future Work

In order to validate the benefit of wavelet tree structure in MR image reconstruction, we propose an overlapping group sparse model for CS-MRI and compare with the good model used in recent papers based on standard sparsity. All total variation terms are removed to emphasize the benefit of tree sparsity. Numerous experiments are conducted to show the practical improvement of the proposed tree-based algorithm on MR images. The results tell that the benefit of the proposed algorithm far from the conclusion in structured sparsity theory. That is because we model the tree structure as overlapping group sparsity which only utilizes every parent-child relationship, while not considering the whole tree-graph structure. The future work will focus on designing efficient algorithms to exploit the whole wavelet tree graph.

References

1. Donoho, D.: Compressed sensing. IEEE Transactions on Information Theory **52**(4) (2006) 1289–1306
2. Candes, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory **52** (2006) 489–509
3. Lustig, M., Donoho, D., Pauly, J.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine **58** (2007) 1182–1195
4. Ma, S., Yin, W., Zhang, Y., Chakraborty, A.: An efficient algorithm for compressed mr imaging using total variation and wavelets. In: Proceedings of CVPR. (2008)
5. Yang, J., Zhang, Y., Yin, W.: A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. IEEE Journal of Selected Topics in Signal Processing, Special Issue on Compressive Sensing **4**(2) (2010) 288–297
6. Huang, J., Zhang, S., Metaxas, D.: Efficient MR Image Reconstruction for Compressed MR Imaging. Medical Image Analysis **15**(5) (2011) 670–679
7. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2**(1) (2009) 183–202
8. Beck, A., Teboulle, M.: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Transaction on Image Processing **18**(113) (2009) 2419–2434
9. Huang, J., Zhang, T., Metaxas, D.: Learning with structured sparsity. Journal of Machine Learning Research **12** (2011) 3371–3412
10. Baraniuk, R.G., Cevher, V., Duarte, M.F., Hegde, C.: Model-based compressive sensing. IEEE Transactions in Information Theory **56** (2010) 1982–2001
11. Rao, N., Nowak, R., Wright, S., Kingsbury, N.: Convex approaches to model wavelet sparsity patterns. In: IEEE international Conference on Image Processing. (2011)
12. He, L., Carin, L.: Exploiting structure in wavelet-based bayesian compressive sensing. IEEE Transactions on Signal Processing **57**(9) (2009) 3488–3497
13. He, L., Chen, H., Carin, L.: Tree-structured compressive sensing with variational bayesian analysis. IEEE Signal Processing Letters **17**(3) (2010) 233–236
14. Som, S., Potter, L., Schniter, P.: Compressive imaging using approximate message passing and a markov-tree prior. Proc. Asilomar Conference on Signals, Systems, and Computers **52** (2010) 489–509
15. Wright, S., Nowak, R., Figueiredo, M.: Sparse reconstruction by separable approximation. IEEE Transactions on Signal Processing **57** (2009) 2479–2493