Medical Image Analysis 18 (2014) 834–842

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

The benefit of tree sparsity in accelerated MRI

Chen Chen, Junzhou Huang *

Department of Computer Science and Engineering, University of Texas at Arlington, 500 UTA Boulevard, Arlington, TX 76019, United States

ARTICLE INFO

Article history:
Received 5 April 2013
Received in revised form 4 December 2013
Accepted 5 December 2013
Available online 16 December 2013

Keywords:
Compressive sensing MRI
SparseMRI
Tree sparsity
Structured sparsity

ABSTRACT

The wavelet coefficients of a 2D natural image are not only approximately sparse with a large number of coefficients tend to be zeros, but also yield a quadtree structure. According to structured sparsity theory, the required measurement bounds for compressive sensing reconstruction can be reduced to $\mathcal{O}(K+\log(N/K))$ by exploiting the tree structure rather than $\mathcal{O}(K+K\log(N/K))$ for standard K-sparse data. In this paper, we proposed two algorithms with convex relaxation to solve the tree-based MRI problem. They are based on FISTA and NESTA frameworks respectively to validate the benefit of tree sparsity in accelerated MRI, both of which have very fast convergence rate and low computational cost. These properties make them be comparable with the fastest algorithms with standard sparsity. Extensive experiments are conducted to validate how much benefit it can bring by tree sparsity. Experimental results show that the proposed methods can definitely improve existing MRI algorithms, although with gaps to the theory.

1. Introduction

According to compressive sensing (CS) theory (Donoho, 2006; Candès et al., 2006), only $\mathcal{O}(K+K\log(N/K))$ sampling measurements are enough to recover K-sparse data with length N. Applying this theory in magnetic resonance imaging (MRI), the scanning time can be significantly reduced (Lustig et al., 2007). 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 can reconstruct MR images with good quality from approximate 20% sampling (Lustig et al., 2007; Ma et al., 2008; Yang et al., 2010; Huang et al., 2011b). They have a general model for the MRI problem:

$$\hat{x}=\arg\min_x\left\{\frac{1}{2}\|Rx-b\|^2+\alpha\|x\|_{TV}+\beta\|\Phi x\|_1\right\} \quad (1)$$

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 ℓ_1 norms are non-smooth which makes this problem have no closed form solution. Classical conjugate gradient decent method is first used to solve this problem Lustig et al. (2007). TVCMRI Ma et al. (2008) and RecPF Yang et al. (2010) use an operator-splitting method and a variable splitting method to solve this problem respectively. FCSA Huang et al. (2011b) decomposes the original problem into two easy subproblems and separately solve each of them with FISTA (Beck and Teboulle, 2009b; Beck and Teboulle, 2009a). These are the state-of-the-art algorithms for CS-MRI. Other methods tried to reconstruct compressed MR images by performing L_p-quasinorm ($p<1$) regularization optimization (Ye et al., 2007; Chartrand, 2007; Chartrand, 2009). However, they all assume that the wavelet coefficients are only approximately sparse, while no structured information has been exploited other than sparsity. When the sampling ratio is below the minimum bound proved in CS theory, there is no guarantee for the performance of standard sparsity based algorithms.

Recent works show that the required measurements can be further reduced by exploiting some structured prior information of the original data (Huang et al., 2011c; Baraniuk et al., 2010). Specially, only $\mathcal{O}(K+\log(N/K))$ sampling measurements are needed for tree sparse data, which widely appear on the wavelet coefficients of nature images. This property has been studied in image compression (Manduca and Said, 1996; Said and Pearlman, 1996; Crouse et al., 1998). Based on this observation, some algorithms have been proposed to improve standard CS recovery by utilizing the tree structure of wavelet coefficients, which can be classified in: greedy algorithms (Baraniuk et al., 2010; La and Do, 2006; Huang et al., 2011c), convex optimization (Rao et al., 2011), Bayesian learning (He and Carin, 2009; He et al., 2010) and AMP (Som and Schniter, 2012). Due to the high computational complexity of greedy algorithms and Bayesian learning, they are difficult to be introduced in CS-MRI. AMP algorithm assumes the sampling

* Corresponding author. Tel.: +1 817 272 9596.
E-mail address: jzhuang@uta.edu (J. Huang).

1361-8415/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.media.2013.12.004

matrix is random Gaussian matrix (Donoho et al., 2009), while in MRI it is a partial Fourier transform that may result in inaccurate reconstruction. Convex optimization method (Rao et al., 2011) is preferred to improve CS-MRI. However, it applies SpaRSA (Wright et al., 2009) to solve their model, which makes it still hard to be comparable with existing fastest CS-MRI methods due to the relatively slow convergence rate of SpaRSA.

In order to investigate the benefit of tree sparsity in CS-MRI and study the feasibleness of improving previous method, in this paper, we propose two tree-based algorithms based on FISTA (Beck and Teboulle, 2009b) and NESTA (Becker et al., 2011) frameworks to validate this. The ℓ_1 norms in the original methods are extended to $\ell_{2,1}$ norms with overlapping groups which encourage tree sparsity (Jacob et al., 2009; Bach et al., 2011; Bach, 2011). Each group contains the wavelet coefficient and its ancestor (s). Due to the overlapping of $\ell_{2,1}$ norm in our implementation, an auxiliary variable is introduced in order to solve the FISTA typed problem, while the dual problem in NESTA typed algorithm is solved directly. Inherent the fast convergence property, the proposed algorithms then can outperform the exiting fastest algorithms with standard sparsity. Extensive experiments are conducted to validate whether tree sparsity has the benefit and how much is the benefit in practical MRI. The results show that existing methods can be further improved by exploiting the wavelet tree structure, although there is gap to theories.

The conference version of this submission has appeared in MICCAI Workshop'12 (Chen and Huang, 2012). This submission has undergone substantial revisions and offers new contributions in the following aspects:

1. The introduction and related work sections are rewritten to provide an extensive review of relevant works and to make our contributions clear.
2. Besides the tree-based FISTA algorithm, we develop another fast tree-based algorithm based on NESTA framework to validate the algorithmic benefit of tree sparsity, which further confirms the conclusion in this work.
3. The experiment section is substantially extended by adding numerous comparisons, such as new data, CPU time, SNR and sample ratio. The extensive comparisons further confirm the conclusion in this article.

2. Related work

2.1. Theoretical benefit of wavelet tree structure

The wavelet coefficients for natural data (signals or images) are often approximately 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. Fig. 1 shows the wavelet quadtree structure of an MR image. If this structure can be utilized, the result will be better as more prior information exploited.

Based on structured sparsity theories, only $\mathcal{O}(K+\log(N/K))$ measurements is needed to recover tree-sparse data rather than $\mathcal{O}(K+K\log(N/K))$ for standard K-sparse data (Huang et al., 2011c; Huang, 2011; Baraniuk et al., 2010). Once the location of a non-zero element is fixed, all its ancestors should be non-zeros. Therefore the number of the solution subspaces for $Rx=b$ is significantly limited. The wavelet coefficients also tend to have this good property. If a parent coefficient has a large/small value, its children also tend to be large/small. By exploiting wavelet tree structure, significant improvement can be achieved especially when the data is very compressible ($K \ll N$).

2.2. Algorithmic benefit of wavelet tree structure

Rao et al. (Rao et al., 2011) consider the wavelet tree structure as overlapping group lasso regularization (Jacob et al., 2009):

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

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 one of such groups. When θ is recovered, it can be transferred to the image by an inverse wavelet transform. By geometric interpretation, ℓ_1 norm ball has some singular values at the axes, which only encourages sparseness with no constrains on the selection of axes. The singular values appears on $\ell_{2,1}$ norm ball only when all coordinates in the same group are zeros. Intuitively, overlapping group inducing norm ball encourages overlapping group sparsity (Jacob et al., 2009). Our work is motivated by Rao's method (Rao et al., 2011), however, we do not introduce this method to CS-MRI, because the replicating of the sampling matrix in their algorithm is not preferred. Moreover, its solver SpaRSA (Wright et al., 2009) only achieves the convergence rate $\mathcal{O}(1/k)$ in function value, which is unable to be comparable with the fastest ones with $\mathcal{O}(1/k^2)$.

2.3. Difference with previous work

The wavelet tree sparsity has been used to improve FCSA (Huang et al., 2010; Huang et al., 2011b; Huang et al., 2011a;

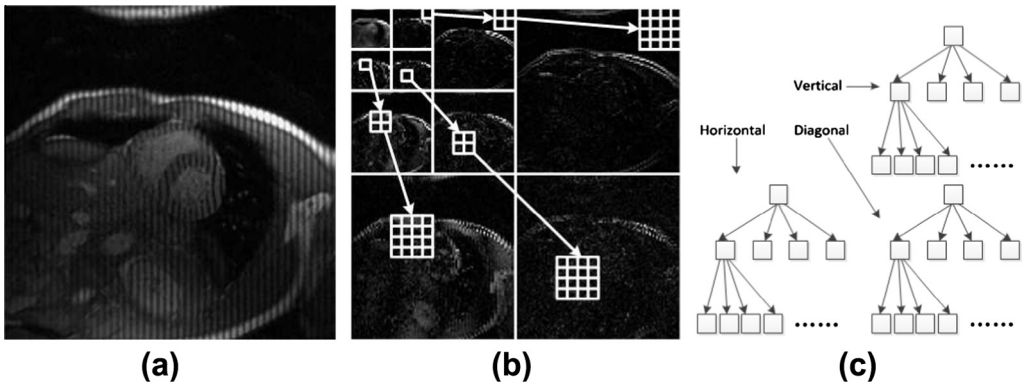

Fig. 1. Wavelet quadtree structure: (a) A cardiac MR image. (b) and (c) The corresponding tree structure of the wavelet coefficients.

Huang et al., 2010) in our previous work (Chen and Huang, 2012). Although both of works study the wavelet tree sparsity in CS-MRI, this work differs from the previous one in the following aspects: (a) the previous work is on the application level, where TV, ℓ_1 norm are all combined to improve FCSA. This work is on the algorithmic level, where we only study the improvements of tree-sparsity inducing norm compared with ℓ_1 norm. (b) Besides the FISTA based algorithm, we propose another tree-based algorithm under NESTA framework here, to validate the benefit of tree sparsity under multiple schemes. (c) The experiments here are conducted to compared the proposed tree-based algorithms with the corresponding ℓ_1 norm based algorithms. This work also studies the impact of different overlapping group size on MRI. The purpose of which is totally different from that of previous work.

3. Algorithm

To validate the benefit of tree sparsity in accelerated MRI, we propose two algorithms to efficiently solve the constrained and unconstrained tree-based CSMRI problems respectively. The tree structure in MR images is approximated as overlapping groups (Jacob et al., 2009; Rao et al., 2011). The unconstrained problem is solved in FISTA (Beck and Teboulle, 2009b) framework and the constrained problem is solved in NESTA (Becker et al., 2011) framework. Both FISTA and NESTA have the optimal convergence rate for first order methods, that is, $\mathcal{O}(1/k^2)$ in function value where k is the iteration number (Nesterov, 1983).

3.1. Unconstrained tree-based MRI

Following overlapping group sparsity algorithms (Jacob et al., 2009; Rao et al., 2011), The unconstrained MRI problem with tree sparsity can be formulated as:

$$\hat{x} = \min_x \left\{ \frac{1}{2}\|Rx - b\|_2^2 + \beta \sum_{g\in\mathcal{G}} \|\Phi x\|_2 \right\} \tag{3}$$

where x is the MR image to be reconstructed, R is the partial Fourier transform, b is the measurement vector, Φ denotes the wavelet transform, β is a positive parameter need to be tuned. Here, g denotes one of the groups that encourages tree sparsity (e.g. one node and its parent) and $\mathcal{G}$ denotes the set of all such groups. Due to the non-smoothness and non-separability of the overlapping group penalty, it is not easy to solve the problem directly. Instead, we introduce a variable z to constrain the problem:

$$\hat{x} = \arg\min_{x,z} \left\{ \frac{1}{2}\|Rx - b\|_2^2 + \beta \sum_{g\in\mathcal{G}} \|z_g\|_2 + \frac{\lambda}{2}\|z - G\Phi x\|_2^2 \right\} \tag{4}$$

where λ is another positive parameter, G is a binary matrix to duplicate the overlapped entries. z is the extended vector of wavelet coefficients x without overlapping.

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

$$z_g = \arg\min_{z_g} \left\{ \beta\|z_g\|_2 + \frac{\lambda}{2}\|z_g - (G\Phi x)_g\|_2^2 \right\}, \qquad g \in \mathcal{G} \tag{5}$$

It has closed form solution by soft thresholding:

$$z_g = \max\left(\|r\|_2 - \frac{\beta}{\lambda}, 0\right)\frac{r}{\|r\|_2}, \qquad g \in \mathcal{G} \tag{6}$$

where $r = (G\Phi x)_g$. We denote this step by $z = shrinkgroup(G\Phi x, \frac{\beta}{\lambda})$ for convenience. For the x-subproblem:

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

This is a combination of two quadratic terms and has closed form solution: $x = (R^TR + \lambda\Phi^TG^TG\Phi)^{-1}(R^Tb + \Phi^TG^Tz)$. However, the inverse of $R^TR + \lambda\Phi^TG^TG\Phi$ 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, which called FISTA_Tree. Here $\nabla f(r^k) = R^T(Rr^k - b) + \lambda\Phi^TG^T(G\Phi r^k - z)$. R^T and Φ^T denote the inverse partial Fourier transform and the inverse wavelet transform.

Algorithm 1. FISTA_Tree

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

Computational complexity. Note that $G \in \mathbb{R}^{N'\times N}$ is a sparse matrix with each row containing only one non-zero element 1. Therefore, the multiplication by G only cost $\mathcal{O}(N') = \mathcal{O}(N)$ with our group configuration. Suppose 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.

3.2. Constrained tree-based MRI

NESTA (Becker et al., 2011) solves the constrained problem of standard sparsity:

$$\min_\theta \|\theta\|_1, \quad s.t. \quad \|b - A\theta\|_2 \leqslant \epsilon \tag{8}$$

where θ denotes the set of wavelet coefficients with $\theta = \Phi x$, $A = R\Phi^T$, Φ^T denotes the inverse wavelet transform, ϵ is a small constant. It reaches the optimal convergence rate for first order methods. Similar as the previous subsection, we extend it to solve the tree-based MRI problem:

$$\min_\theta \|G\theta\|_{2,1}, \quad s.t. \quad \|b - A\theta\|_2 \leqslant \epsilon \tag{9}$$

where $\|G\theta\|_{2,1} = \sum_{g\in\mathcal{G}}\|(G\theta)_g\|_2$, and $g, \mathcal{G}$ are the same as those in Algorithm 1. Recall $\ell_{2,1}$ norm also have the form:

$$\|G\theta\|_{2,1} = \max_{u\in\mathcal{Q}} < u, Gx > \tag{10}$$

where the dual feasible set is:

$$\mathcal{Q} = \{u : \|u\|_{2,\infty} \leqslant 1\} = \{u : \max_{g\in\mathcal{G}}\|u_g\|_2 \leqslant 1\} \tag{11}$$

We relax the non-smooth $\ell_{2,1}$ norm to smooth function with:

$$f_\mu(\theta) = \max_{u\in\mathcal{Q}}(< u, G\theta > - \frac{\mu}{2}\|u\|_2^2) \tag{12}$$

where μ is a small fixed number.

Note that $(G\theta)_g = G_g\theta$ where G_g the rows of G correspond to group g. The first order gradient of $f_\mu(\theta)$ with Lipschitz constant L_μ is given by:

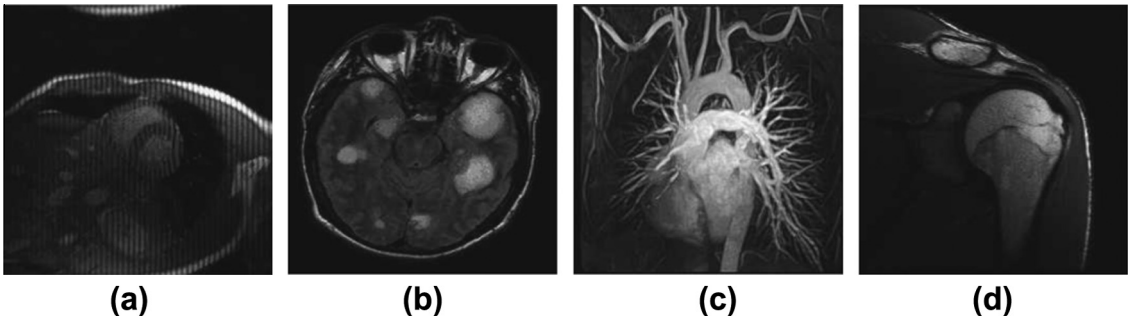

Fig. 2. MR images: (a) cardiac; (b) brain; (c) chest and (d) shoulder.

$$\nabla f_\mu(\theta)_g = \begin{cases} \mu^{-1} G_g^T G_g \theta, & \|G_g\theta\|_2 < \mu \\ G_g^T G_g \theta / \|G_g\theta\|_2, & \texttt{otherwise} \end{cases} \tag{13}$$

NESTA assumes the rows of the sampling matrix A are orthogonal, that is, $AA^T = I$ where I denotes the identical matrix. Fortunately, the partial Fourier transform in compressive sensing MRI satisfies this assumption: $AA^T = R\Phi^T\Phi R^T = RR^T = I$, where R^T denotes the inverse operator of R. The whole algorithm based in NESTA (Becker et al., 2011) framework is given in Algorithm 2.

Algorithm 2. NESTA_Tree

Input: $\theta_0, \epsilon, k = 1, L_\mu, \mu$
while not meet the stopping criterion **do**
1. Compute $\nabla f_\mu(\theta)$
2. Compute y^k
   $q = \theta^k - L_\mu^{-1} \nabla f_\mu(\theta)$
   $\lambda_\epsilon = \max(0, \epsilon^{-1}\|b - Aq\|_2 - L_\mu)$
   $y^k = (I - \frac{\lambda_\epsilon}{\lambda_\epsilon + L_\mu} A^T A)(\frac{\lambda_\epsilon}{L_\mu} A^T b + q)$
3. Compute z^k
   $\alpha^k = 1/2(k+1)$
   $q = x_0 - L_\mu^{-1} \sum_{i \leqslant k} \nabla \alpha_i f_\mu(\theta)$
   $\lambda_\epsilon = \max(0, \epsilon^{-1}\|b - Aq\|_2 - L_\mu)$
   $z^k = (I - \frac{\lambda_\epsilon}{\lambda_\epsilon + L_\mu} A^T A)(\frac{\lambda_\epsilon}{L_\mu} A^T b + q)$
4. Update θ^k
   $\tau^k = 2(k+3)$
   $\theta^k = \tau^k z^k + (1 - \tau^k) y^k$
5. $k = k + 1$
end while

Computational complexity. As shown in Algorithm 2, the complexity of the proposed algorithm the same as the original NESTA algorithm (Becker et al., 2011). It is $6\mathcal{C} + \mathcal{O}(N)$, where $\mathcal{C}$ denotes the complexity of applying A or A^T. In CSMRI, $\mathcal{C} = \mathcal{O}(N \log N)$ if fast Fourier transform (FFT) is applied. Therefore, the total computational complexity is $\mathcal{O}(N \log N)$ for each iteration, the same as that of Algorithm 1.

If we compare the two types of algorithms, the parameters can be manually set in the unconstrained algorithm to determine how sparse the data is. Or the weights between sparseness and the least square fitting can be controlled. However, the constrained algorithm always seeks for the sparsest solution that satisfy the constrain. In the application of MRI, we find that if good parameters can be tuned, the unconstrained algorithm (Algorithm 1) performs better, or vice versa. In contrast, the constrained algorithm (Algorithm 2) has the convenience without tuning the parameter.

Table 1
Comparisons of SNR (db) on different group sizes for tree sparsity.

β\group size	1	2	3	4
5×10^{-2}	17.30	17.94	16.45	15.33
10^{-2}	16.49	16.99	16.95	16.53
5×10^{-3}	16.36	16.62	16.66	16.48
10^{-3}	16.21	16.27	16.29	16.27

Table 2
Comparisons of computational costs (s) on different group sizes for tree sparsity.

β\group size	1	2	3	4
5×10^{-2}	0.69	0.99	1.11	1.17
10^{-2}	0.70	0.95	1.09	1.15
5×10^{-3}	0.72	0.97	1.07	1.11
10^{-3}	0.70	0.97	1.07	1.11

4. Experiments

4.1. Experiment setup

We compare the unconstrained algorithm FISTA_Tree with CG (Lustig et al., 2007), TVCMRI (Ma et al., 2008), RecPF (Yang et al., 2010), FCSA (Huang et al., 2011b) and compare the constrained algorithm NESTA_Tree with the original NESTA (Becker et al., 2011) algorithm for CSMRI. For fair comparisons, all code are downloaded from the authors' websites and we carefully follow their experiment setup. We apply all these methods on four real-valued MR images: cardiac, brain, chest and shoulder respectively (shown in Fig. 2). In addition, a complex valued MR brain image[1] is added to validate the benefit of tree sparsity on complex-valued data. Suppose R is a partial Fourier transform with M rows and N columns. The sampling ratio is defined as M/N. For simulations with real-valued images, we follow the sampling strategy of previous works (Ma et al., 2008; Huang et al., 2011b), which randomly choose more Fourier coefficients from low frequency and less on high frequency. For complex-valued data, the radial sampling mask is used (Yang et al., 2010), which is more feasible in practical.

In order to study the benefit of tree structure in CSMRI, we remove the TV term in all algorithms. The parameters for real-valued images and complex-valued image are tuned separately. There is no continuation step (Becker et al., 2011) in the NESTA_Tree algorithm. All experiments are on a desktop with 3.4 GHz Intel core i7 3770 CPU. Matlab version is 7.8 (2009a). Measurements are added by Gaussian white noise with 0.01 standard deviation. Signal-to-Noise Ratio (SNR) is used for result evaluation:

[1] http://www.eecs.berkeley.edu/~mlustig/CS.html

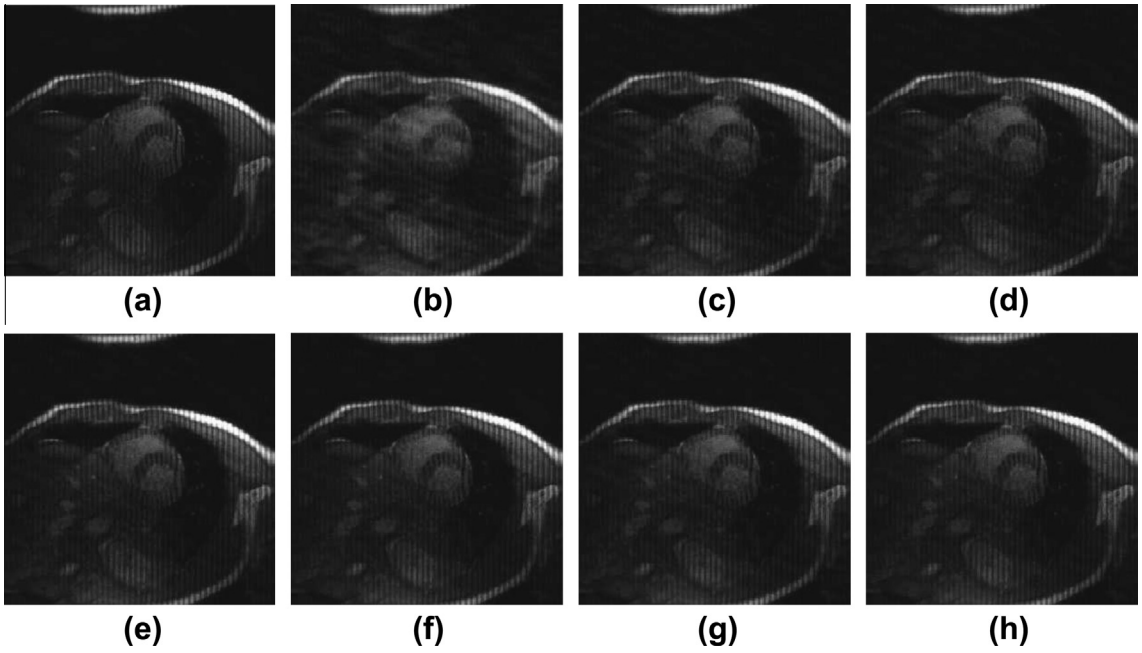

Fig. 3. Cardiac MR image reconstruction from 20% sampling. (a) The original image. Recovered by: (b) CG (Lustig et al., 2007); (c) TVCMRI (Ma et al., 2008); (d) RecPF (Yang et al., 2010); (e) FCSA (Huang et al., 2011b); (f) FISTA_Tree; (g) NESTA (Becker et al., 2011); (h) NESTA_Tree. All algorithms are without total variation regularization. Their SNR are 9.86, 14.70, 15.14, 17.31, 17.93, 16.31 and 16.96 respectively. Their CPU time costs are 1.34 s, 1.12 s, 1.25 s, 0.67 s, 0.85 s, 0.88 s and 1.05 s.

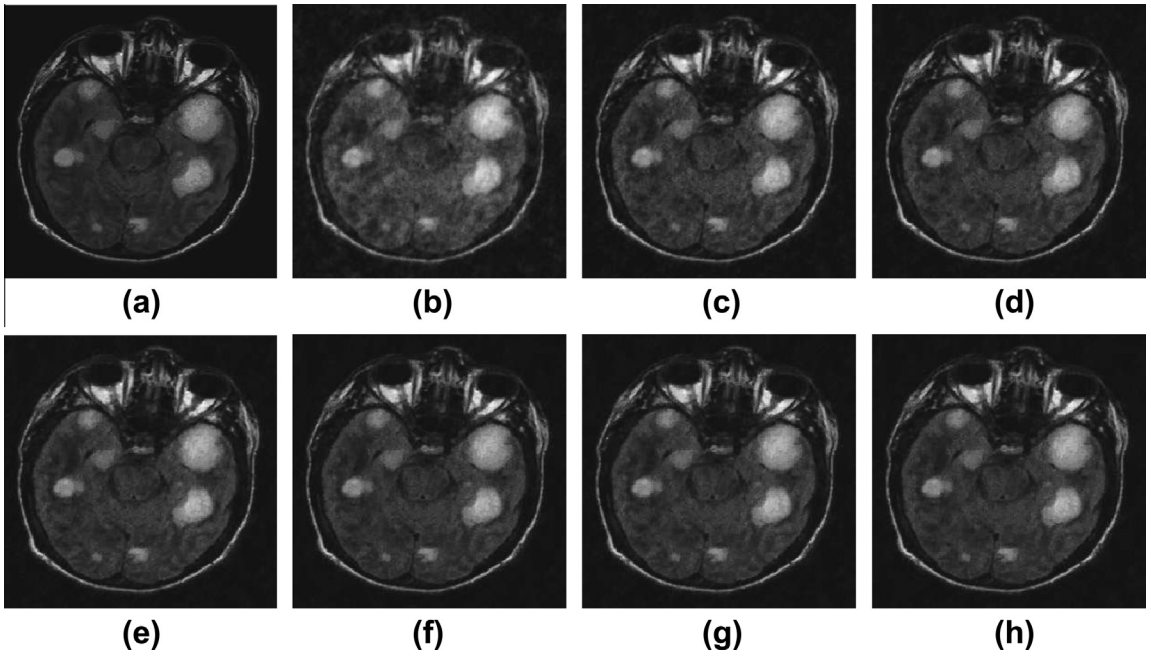

Fig. 4. Brain MR image reconstruction from 20% sampling. (a) The original image. Recovered by: (b) CG (Lustig et al., 2007); (c) TVCMRI (Ma et al., 2008); (d) RecPF (Yang et al., 2010); (e) FCSA (Huang et al., 2011b); (f) FISTA_Tree; (g) NESTA (Becker et al., 2011); (h) NESTA_Tree. All algorithms are without total variation regularization. Their SNR are 10.25, 13.81, 14.22, 15.65, 16.13, 15.05 and 15.52 respectively. Their CPU time costs are 1.36 s, 1.11 s, 1.17 s, 0.71 s, 1.02 s, 0.91 s and 1.03 s.

$$SNR = 10\log_{10}(V_s/V_n) \tag{14}$$

where V_n is the Mean Square Error between the original image x_0 and the solution x; V_s denotes the variance of the values in x_0.

4.2. Group configuration for tree sparsity

In all previous works, the tree structure are approximated as overlapping groups (Rao et al., 2011; Chen and Huang, 2012). In additional, all of them only consider each wavelet coefficient and its parent are assigned into one group. However, the relationship between one coefficient and its grandparent is not exploited. We first conduct an experiment to validate the influence of the group size to the reconstruction result. Four group sizes are compared: (a) each group one contains one coefficient, which is the same case as standard sparsity; (b) each group contains a coefficient and its parent, which is the same as previous works (Rao et al., 2011; Chen and Huang, 2012); (c) each group contains a coefficient, its parent

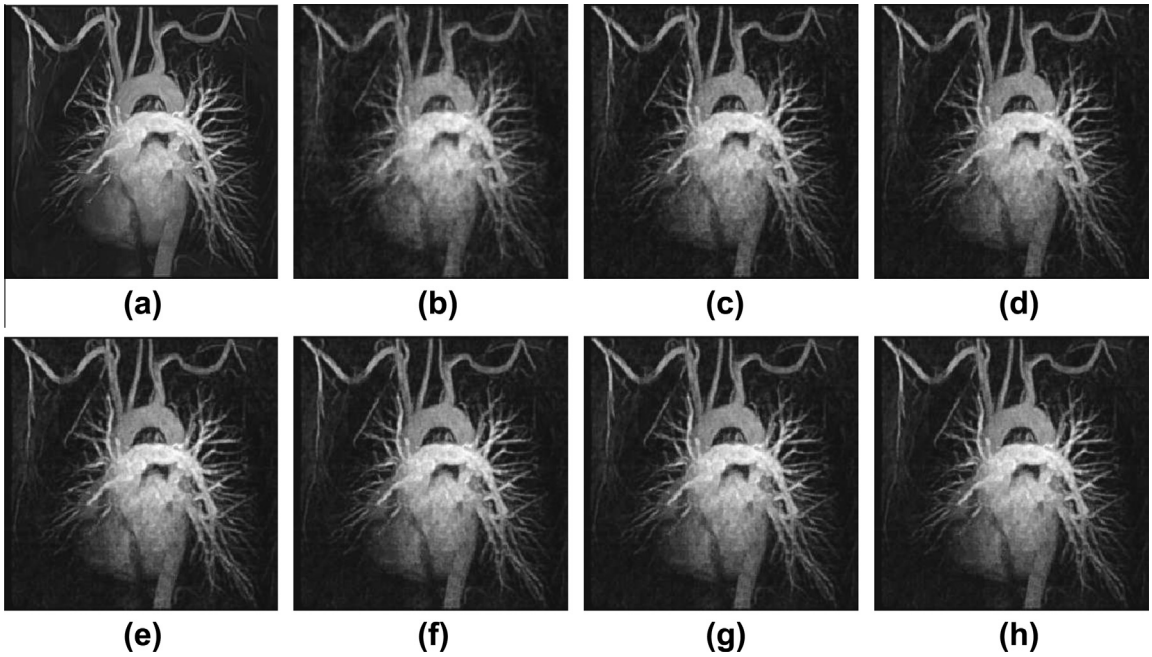

Fig. 5. Chest MR image reconstruction from 20% sampling. (a) The original image. Recovered by: (b) CG (Lustig et al., 2007); (c) TVCMRI (Ma et al., 2008); (d) RecPF (Yang et al., 2010); (e) FCSA (Huang et al., 2011b); (f) FISTA_Tree; (g) NESTA (Becker et al., 2011); (h) NESTA_Tree. All algorithms are without total variation regularization. Their SNR are 11.82, 15.09, 15.36, 15.98, 16.35, 15.91 and 16.30 respectively. Their CPU time costs are 1.28 s, 1.12 s, 1.23 s, 0.67 s, 0.96 s, 0.84 s and 1.07 s.

and its grandparent; (d) each group contains 4 coefficients, where the grandparent's parent is also assigned in the same group.

With these group configurations, we test their performance on the FISTA_Tree algorithm, except the standard sparsity case is performed on FISTA. The parameter β determines how strong the tree sparsity assumption is. Tables 1 and 2 show the average SNRs and CPU time on the four MR images with various parameter settings. With smaller parameters, the third group configuration performs the best, while the second group configuration is the best with larger parameters. The computational time increase monotonously as the size of group becomes bigger. Due to the above two reasons, we encourage the use of the second group configuration on CS-MRI.

4.3. Visual comparisons

We compare proposed tree-based algorithms with the fastest MRI algorithms to validate how much the tree structure can improve existing results. To perform fair comparisons, all methods

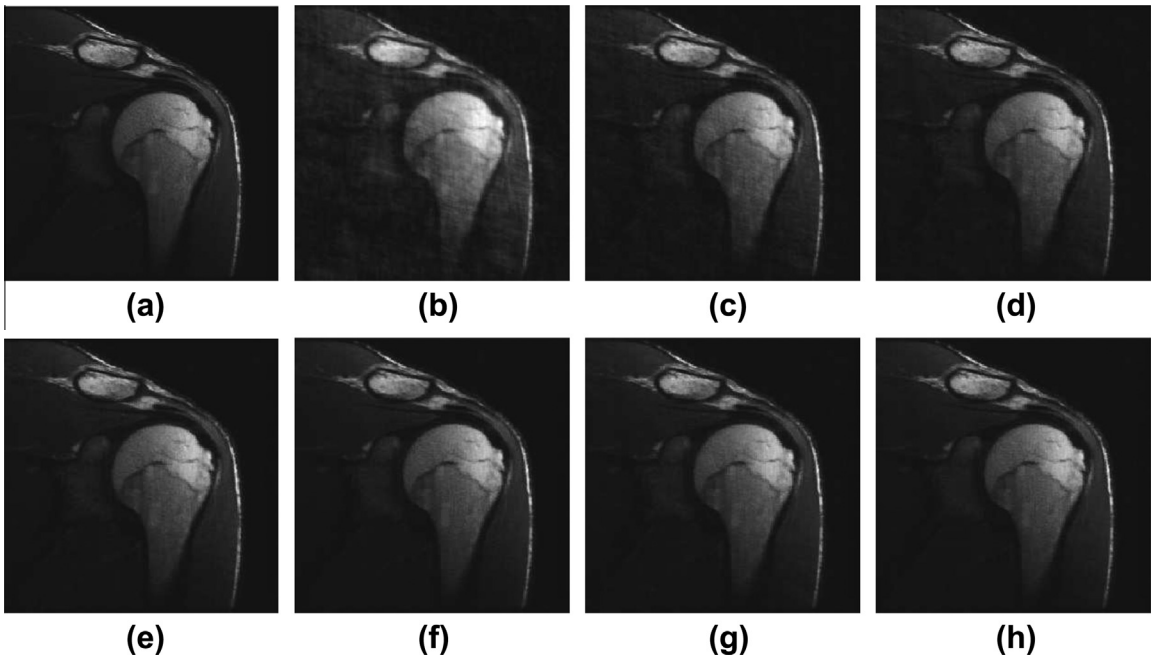

Fig. 6. Shoulder MR image reconstruction from 20% sampling. (a) The original image. Recovered by: (b) CG (Lustig et al., 2007); (c) TVCMRI (Ma et al., 2008); (d) RecPF (Yang et al., 2010); (e) FCSA (Huang et al., 2011b); (f) FISTA_Tree; (g) NESTA (Becker et al., 2011); (h) NESTA_Tree. All algorithms are without total variation regularization. Their SNR are 12.31, 16.80, 17.90, 20.77, 21.04, 20.17 and 20.62 respectively. Their CPU time costs are 1.36 s, 1.07 s, 1.25 s, 0.67 s, 0.95 s, 0.82 s and 1.07 s.

run 50 iterations except that the CG runs only 8 iterations due to its higher computational complexity. Total variation terms are removed in all algorithms, as we only want to validate how much benefit the wavelet tree sparsity can bring compared to standard wavelet sparsity. In this case, FCSA (Huang et al., 2011b) is similar as FISTA (Beck and Teboulle, 2009b). Fig. 3–6 shows the visual results on the four MR images with 20% sampling. It can be found that the proposed unconstrained algorithm FISTA_Tree is always better than CG (Lustig et al., 2007), TVCMRI (Ma et al., 2008), RecPF (Yang et al., 2010) and FCSA (Huang et al., 2011b). These results are consistent with previous observations (Huang et al., 2011b). Compared the proposed NESTA_Tree with NESTA, our method is still much better. These results are reasonable because no structured prior information has been exploited in previous algorithms other than sparsity, while the tree structure in our algorithms is utilized. Any coefficient that disobeys the tree structure will be penalized in our algorithms, which makes the results closer to the original ones.

4.4. SNRs and CPU time

Fig. 7 gives the performance comparisons between different methods in terms of SNR with 50 iterations. Due to the faster convergence rate of FISTA and NESTA, they always outperforms CG, TVCMRI and RecPF. Moreover, the tree-based algorithms approximated by overlapping group sparsity are always better than those with standard sparsity. Table 3 shows all computational costs of different algorithms. CG has the highest computational complexity.

Table 3
Comparisons of average computational costs (s) on different MR images with 20% sampling.

	Cardiac	Brain	Chest	Shoulder
CG	9.14	8.87	9.21	9.17
TVCMRI	1.12	1.11	1.12	1.07
RecPF	1.25	1.17	1.23	1.25
FCSA	0.67	0.71	0.67	0.67
FISTA_Tree	0.85	1.02	0.96	0.95
NESTA	0.88	0.91	0.84	0.82
NESTA_Tree	1.05	1.03	1.07	1.07

TVCMRI and RecPF is much faster than CG and slower than FCSA. It is to be expected that tree based algorithms FISTA_Tree and NESTA_Tree are slower than FISTA and NESTA respectively, since the overlapping structure needs more time for computing than non-overlapping structure. However, applying the wavelet transform and the Fourier transform is still the dominant cost, which is the same for all algorithms. As a result, FISTA_Tree and NESTA_Tree are comparable to the corresponding standard sparsity algorithms in term of reconstruction speed, and bring much more improvement on accuracy.

4.5. Sampling ratios

All algorithms are compared under different sampling ratios on the four MR images. Since we have shown that the CG method is far less efficient than other methods, we do not include it in this

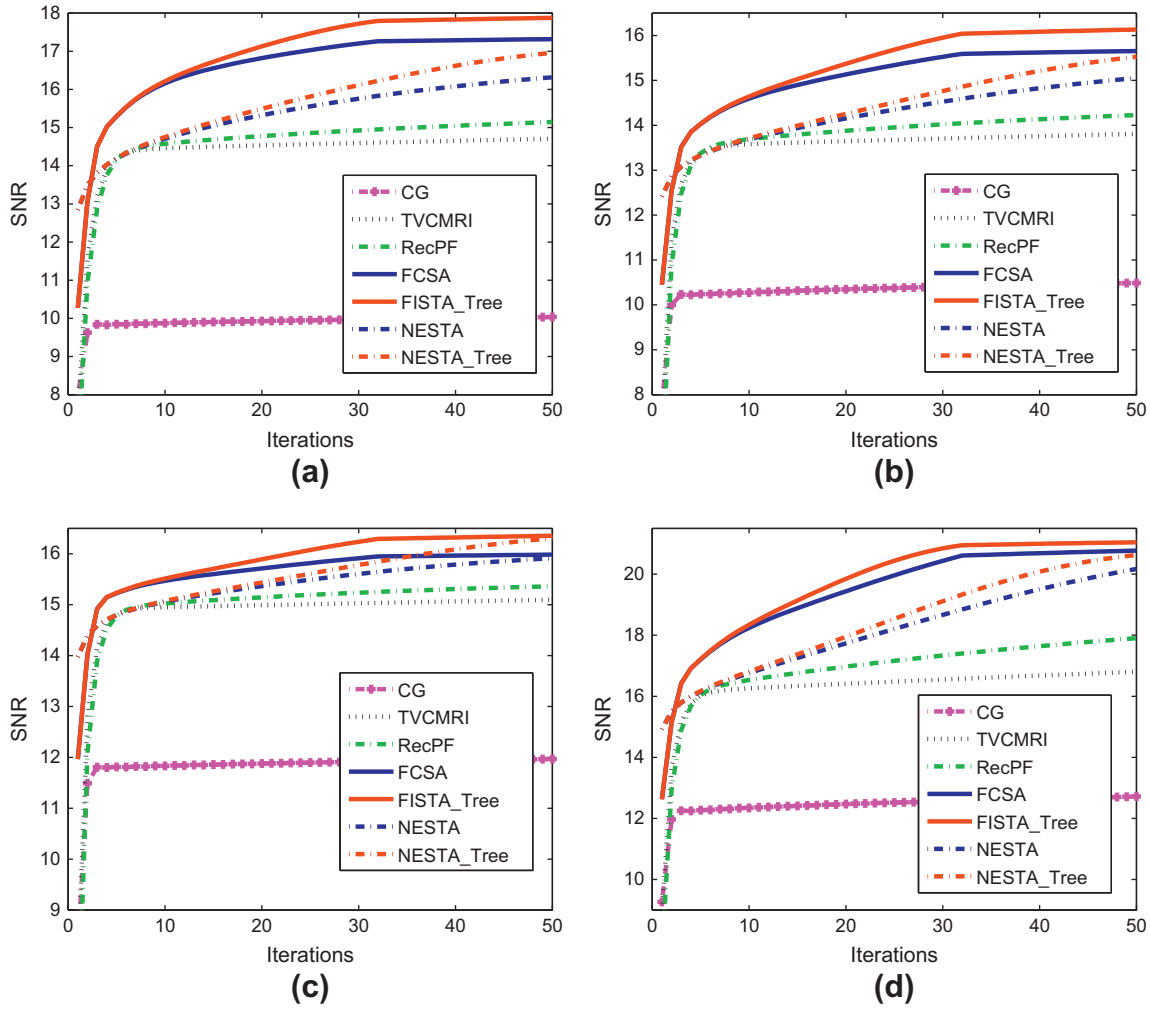

Fig. 7. Performance comparisons (SNRs) on different MR images: (a) cardiac image; (b) brain image; (c) chest image and (d) shoulder image.

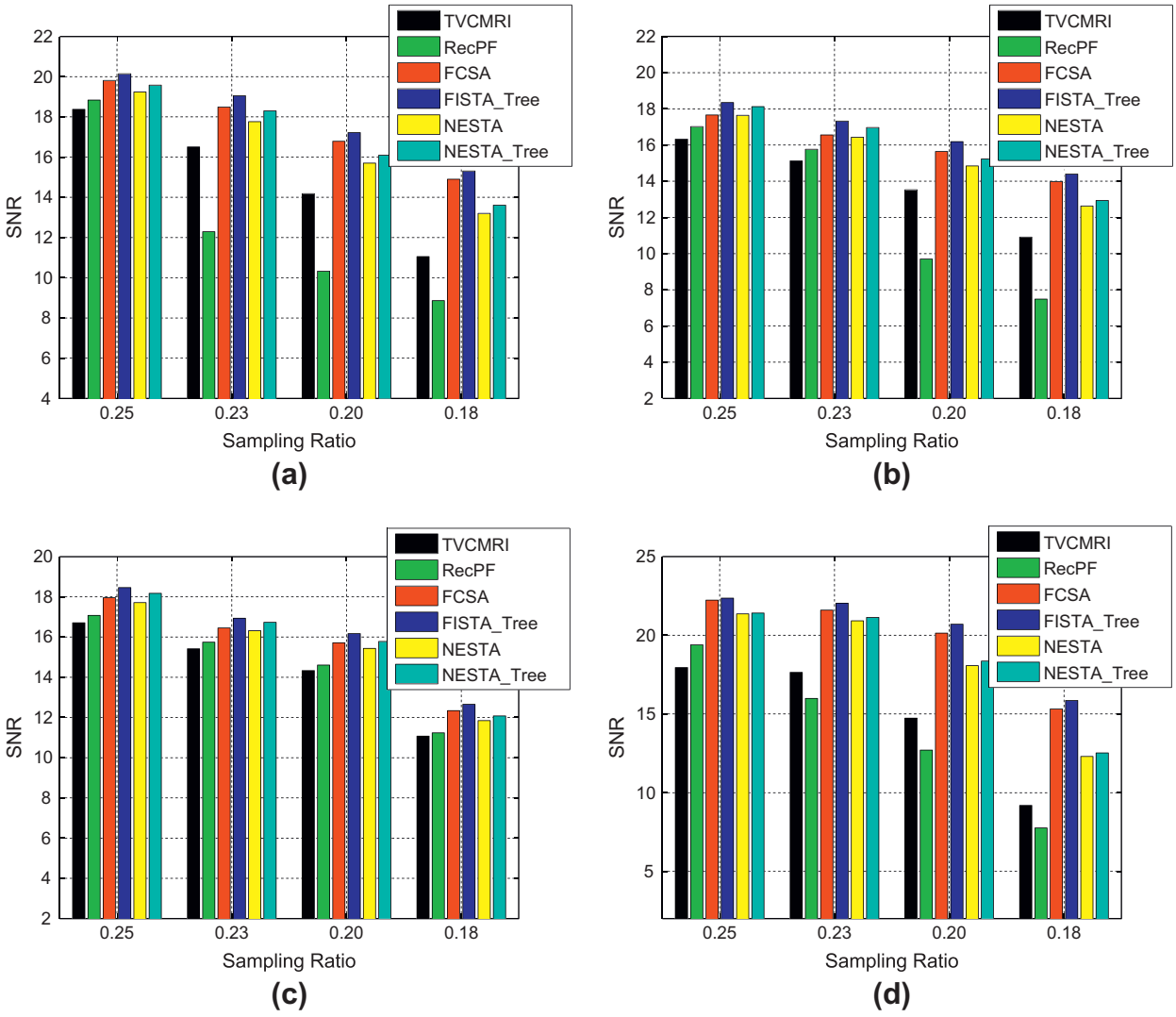

Fig. 8. Performance comparisons on the four MR images with different sampling ratios: (a) cardiac image; (b) brain image; (c) chest image and (d) shoulder image.

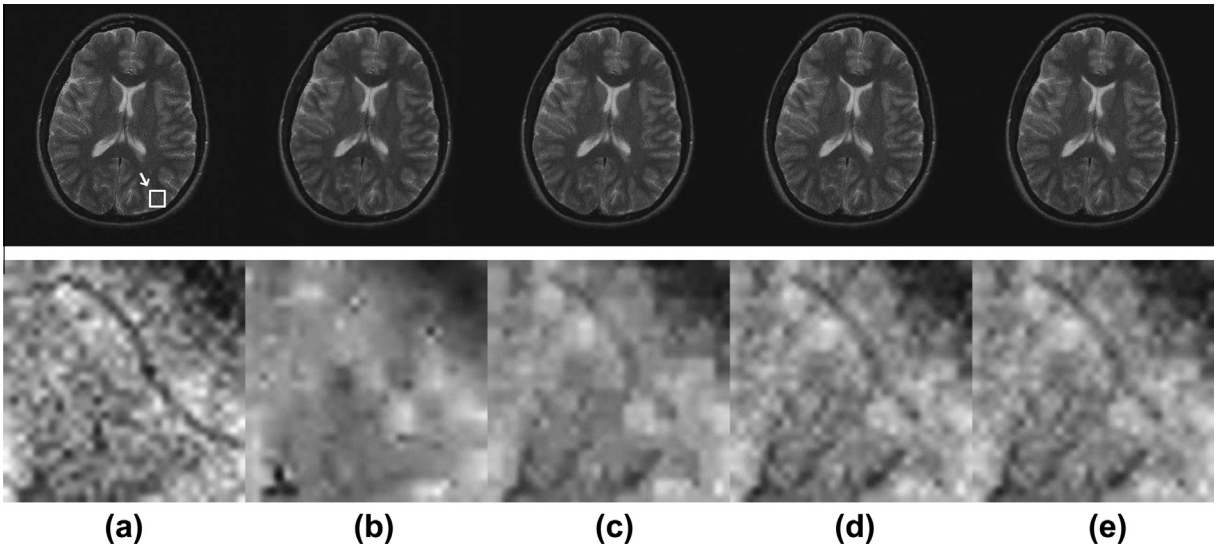

Fig. 9. Reconstruction on complex-valued MR image with 20% sampling. The first shows the visual results of by different algorithm. The second row shows the zoomed in areas indicated by the white boxes. (a) Inverse FFT with full sampling. (b) CG. (c) FCSA. (d) FISTA_Tree. (e) NESTA_Tree. Their SNRs are 15.31, 16.32, 16.65, 16.68 respectively.

experiment. To reduce the randomness, we run each experiments 100 times to obtain the average results of each method. The sampling ratio ranges from 17% to 25%. Fig. 8 shows these results on the four images. We could observe that TVCMRI and RecPF are not comparable to recent algorithms with fast convergence rate. Under the same framework and with similar convergence rate, the tree-based algorithms (i.e. FISTA_Tree and NESTA_Tree) are always better than the corresponding standard sparsity algorithms (i.e. FCSA and NESTA) respectively. These results further demonstrate the benefit of tree sparsity in accelerated MRI.

4.6. Complex-valued with radial sampling mask

We have observed the superior performance for tree-based MRI algorithms from numerical simulations. In this subsection, we validate their performance on a complex-valued MR image with 512×512 pixels. The sampling mask is radial mask, which is more feasible than the random sampling mask in practical. With previous results, we only compare the classical method CG (Lustig et al., 2007) and the fastest algorithm FCSA (Huang et al., 2011b) with the proposed tree-based algorithms.

Fig. 9 presents the visual results reconstructed by difference methods. The image with full sampling is used as reference image. We could observe that tree-based algorithms FISTA_Tree and NESTA_Tree achieve higher SNR than standard sparsity algorithms CG and FCSA. Due to the relative slow convergence rate of CG, it is still not converged after 50 iterations. That is why it has inferior performance to FCSA. This data is scanned with noise. Therefore, we also compare image quality besides SNR. From the zoomed in areas, image details are lost in the image reconstructed by CG and blurred in that reconstructed by FCSA. However, both tree-based algorithms can preserve significant features on the MR image even with a low sampling ratio.

5. Conclusion and future work

In order to validate the benefit of wavelet tree sparsity in MR image reconstruction, we propose two tree-based algorithms for CS-MRI and compare them with the state-of-the-art algorithms based on standard sparsity. In order to observe the benefit of tree sparsity more clearly, total variation terms are removed in all algorithms. Extensive 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 is far from the conclusion in structured sparsity theory. That is because the tree structure is not as strictly as the structured sparsity theories assumed on practical data. In addition, the approximation in our implementations may be not precise enough. To bridge the gap between the theories and practice, future work will weight the wavelet coefficients on different levels differently, but not treat them equally.

References

Bach, F., 2011. Optimization with sparsity-inducing penalties. Found. Trends Machine Learn. 4, 1–106.

Bach, F., Jenatton, R., Mairal, J., Obozinski, G., 2011. Structured sparsity through convex optimization. arXiv:1109.2397.

Baraniuk, R., Cevher, V., Duarte, M., Hegde, C., 2010. Model-based compressive sensing. IEEE Trans. Inform. Theory 56, 1982–2001.

Beck, A., Teboulle, M., 2009a. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Trans. Image Process. 18, 2419–2434.

Beck, A., Teboulle, M., 2009b. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imag. Sci. 2, 183–202.

Becker, S., Bobin, J., Candès, E., 2011. NESTA: a fast and accurate first-order method for sparse recovery. SIAM J. Imag. Sci. 4, 1–39.

Candès, E., Romberg, J., Tao, T., 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory 52, 489–509.

Chartrand, R., 2007. Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process. Lett. 14, 707–710.

Chartrand, R., 2009. Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data. In: Proceedings of ISBI.

Chen, C., Huang, J., 2012a. The benefit of tree sparsity in accelerated MRI. In: MICCAI Workshop on Sparsity Techniques in Medical Imaging.

Chen, C., Huang, J., 2012b. Compressive Sensing MRI with Wavelet Tree Sparsity. In: Proceedings of the Annual Conference on Advances in Neural Information Processing Systems (NIPS), pp. 1124–1132.

Crouse, M.S., Nowak, R.D., Baraniuk, R.G., 1998. Wavelet-based statistical signal processing using hidden markov models. IEEE Trans. Signal Process. 46, 886–902.

Donoho, D., 2006. Compressed sensing. IEEE Trans. Inform. Theory 52, 1289–1306.

Donoho, D., Maleki, A., Montanari, A., 2009. Message-passing algorithms for compressed sensing. Proc. Natl. Acad. Sci. 106, 18914–18919.

He, L., Carin, L., 2009. Exploiting structure in wavelet-based bayesian compressive sensing. IEEE Trans. Signal Process. 57, 3488–3497.

He, L., Chen, H., Carin, L., 2010. Tree-structured compressive sensing with variational bayesian analysis. IEEE Signal Process. Lett. 17, 233–236.

Huang, J., 2011. Structured sparsity: Theorems, Algorithms and Applications. Ph.D. thesis, Rutgers University.

Huang, J., Zhang, S., Li, H., Metaxas, D., 2011a. Composite splitting algorithms for convex optimization. Comput. Vis. Image Understand. 115, 1610–1622.

Huang, J., Zhang, S., Metaxas, D., 2010a. Efficient MR image reconstruction for compressed MR imaging. In: Proceedings of the 13th International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI), pp. 135–142.

Huang, J., Zhang, S., Metaxas, D., 2010b. Fast optimization for mixture prior models. In: European Conference on Computer Vision (ECCV), pp. 607–620.

Huang, J., Zhang, S., Metaxas, D., 2011b. Efficient MR image reconstruction for compressed MR imaging. Med. Image Anal. 15, 670–679.

Huang, J., Zhang, T., Metaxas, D., 2011c. Learning with structured sparsity. J. Machine Learn. Res. 12, 3371–3412.

Jacob, L., Obozinski, G., Vert, J., 2009. Group lasso with overlap and graph lasso. In: Proceedings of the 26th Annual International Conference on Machine Learning (ICML), pp. 433–440.

La, C., Do, M.N., 2006. Tree-based orthogonal matching pursuit algorithm for signal reconstruction. In: IEEE International Conference on Image Processing (ICIP), pp. 1277–1280.

Lustig, M., Donoho, D., Pauly, J., 2007. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magnet. Reson. Med. 58, 1182–1195.

Ma, S., Yin, W., Zhang, Y., Chakraborty, A., 2008. An efficient algorithm for compressed mr imaging using total variation and wavelets. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8.

Manduca, A., Said, A., 1996. Wavelet compression of medical images with set partitioning in hierarchical trees. In: Proceedings of the SPIE Symposium on Medical Imaging.

Nesterov, Y., 1983. A method for unconstrained convex minimization problem with the rate of convergence $\mathcal{O}(1/k^2)$. Doklady AN USSR (translated as Soviet Math. Docl.) 269, 543–547.

Rao, N., Nowak, R., Wright, S., Kingsbury, N., 2011. Convex approaches to model wavelet sparsity patterns. In: Proceedings of the 18th IEEE International Conference on Image Processing (ICIP), pp. 1917–1920.

Said, A., Pearlman, W.A., 1996. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans. Circ. Syst. Video Technol. 6, 243–250.

Som, S., Schniter, P., 2012. Compressive imaging using approximate message passing and a markov-tree prior. IEEE Trans. Signal Process. 60, 3439–3448.

Wright, S., Nowak, R., Figueiredo, M., 2009. Sparse reconstruction by separable approximation. IEEE Trans. Signal Process. 57, 2479–2493.

Yang, J., Zhang, Y., Yin, W., 2010. A fast alternating direction method for tvl1-l2 signal reconstruction from partial fourier data. IEEE J. Select. Topics Signal Process. 4, 288–297.

Ye, J., Tak, S., Han, Y., Park, H., 2007. Projection reconstruction MR imaging using focuss. Magnet. Reson. Med. 57, 764–775.