Weighted-CS for reconstruction of highly under-sampled dynamic MRI sequences

Dornoosh Zonoobi and Ashraf A. Kassim*

Dept. Electrical and Computer Engineering
National University of Singapore, Singapore
* E-mail: ashraf@nus.edu.sg Tel: +65-65162277

***Abstract*—This paper investigates the potential of the new Weighted-Compressive Sensing approach which overcomes the major limitations of other compressive sensing and outperforms current state-of-the-art methods for low-rate reconstruction of sequences of MRI images. The underlying idea of this approach is to use the image of the previous time instance to extract an estimated probability model for the image of interest, and then use this model to guide the reconstruction process. This is motivated by the observation that MRI images are hugely sparse in Wavelet domain and the sparsity changes slowly over time.**

I. Introduction

Magnetic Resonance Imaging (MRI) is an essential medical imaging tool burdened by its intrinsic slow data acquisition process. Since data acquisition is sequential in MR imaging modalities, the scan time (time to get enough data to accurately reconstruct one frame) is reduced if fewer measurements are needed for reconstruction. Goal of many researchers is therefore to employ a smaller set of samples than normally required to reconstruct the original images. However, when k-space is under sampled, the Nyquist criterion is violated, and conventional Fourier reconstructions exhibit aliasing artifacts.

Compressed sensing (CS) [1]–[4], on the other hand, aims to reconstruct signals and images from significantly fewer measurements than were traditionally thought necessary. CS has already had noteworthy impacts in this field [5], [6] which enables the reconstruction of MRI images from much smaller set of measurements than the number of Nyquist-rate samples. It was shown that MRI has two key features that makes the application of CS successful; first is that MRI images are naturally compressible by sparse coding in an appropriate transform domain (e.g., by wavelet transform [7]); second is that MRI scanners acquire sample images in 2D Fourier transform, which is known to be incoherent w.r.t. sparse domain, rather than acquiring direct pixel samples [8].

Much of the current CS-based works in the literature are, mostly concerned with reconstruction of static MRI images. However, in many important and challenging MRI applications, such as volumetric (3D) MRI or real-time MRI (refers to the continuous monitoring of moving objects in real time), instead of just one image slice, we are dealing with sequences of MRI images which are closely related to each other. Real-time MRI, for example, currently is only possible with low image quality or limited rate of slices per second due to the time-consuming scanning process [9]. Therefore developing CS-based methods, specifically adopted for dealing with dynamic sequences of MRI images, can greatly benefit these applications. To the best of our knowledge, only a few works have developed reconstruction methods for sequences of MRI images (e.g. [8], [9]).

II. The Background

A. Notations

Throughout the paper, vectors are denoted by boldface letters (e.g. $\mathbf{f}, \mathbf{F}$) and f_i is the i-th element of the vector $\mathbf{f}$. Scalars are shown by small regular letters (e.g. n, k) and matrices are denoted by bold capital Greek letters ($\mathbf{\Psi}, \mathbf{\Phi}$). Superscript (t) added to a vector/matrix refers to that of time t. We use the notation $\mathbf{\Phi}|_{\mathcal{S}}$ to denote the sub-matrix containing the columns of $\mathbf{\Phi}$ with indices belonging to $\mathcal{S}$. For a vector, the notation $\mathbf{y}|_{\mathcal{S}}$ forms a sub-vector that contains elements with indices in $\mathcal{S}$.

B. Problem Formulation

In its noiseless formulation, the problem can be posed as follows: let $\mathbf{f}^{(t)} \in \mathcal{R}^n$ be the slowly time-varying MRI signal at epoch t, which is sparse in a transform domain ($\mathbf{F}^{(t)} = \mathbf{\Psi}\mathbf{f}^{(t)}$). $\mathcal{S}^{(t)} := \{k \in \{1, \cdots, n\} : \mathbf{F}^{(t)}|_k \neq 0\}$ denote the support of the signal in the transform domain. At each slice, the image is measured using a sampling matrix ($\mathbf{\Phi}$ of size $m \times n$) and $\mathbf{y}^{(t)} = \mathbf{\Phi}\mathbf{f}^{(t)}$ is the observation vector. The problem, at each time t, is then to recover the original image, $\mathbf{f}^{(t)}$, from the corresponding compressive samples ($\mathbf{y^{(t)}}$), assuming that the image of interest is sparse. The most straightforward application of CS for reconstruction of the dynamic signal, would be to use the simple CS at each time frame separately [9]. This means that at each time frame, we search for a signal that satisfies the observations and is sparsest in the transform domain. Equivalently, at each time instant, t, we find a $\mathbf{g}$ which is an estimate of $\mathbf{f}^{(t)}$, as a solution to the following optimization problem:

$$\mathbf{f}^{(t)} = \arg\min\{\left\|\mathbf{G}\right\|_{\ell_1} \text{ subject to } \mathbf{\Phi}\mathbf{g} = \mathbf{y}^{(t)}\}. \quad (1)$$

where $\mathbf{G} = \mathbf{\Psi}\mathbf{g}$.

In this method, however, the fact that all image slices are closely related is not exploited. The authors of [10], propose a method based on the fact that the MRI images

change slowly over time. They dynamically update the solution of the above minimization, without directly solving it. The proposed dynamic update scheme systematically breaks down the solution update into a small number of linear steps. However, this approach merely uses past reconstructions to speed up the current optimization and does not improve the reconstruction error and therefore the number of samples needed is equal to the conventional CS [11]. Authors of [12], discussed the problem of reconstructing a signal when some priori information exists about the signal (we refer to their method as priori-CS). Their method can be applied to the problem of sparse sequence reconstruction as follows:

$$\mathbf{f}^{(t)} = \arg\min\{\left\|\mathbf{G}\right\|_{\ell_1} \text{ subject to } \mathbf{\Phi g} = \mathbf{y}^{(t)}, \|\mathbf{G} - (\mathbf{f}^{(t-1)})\|_{\ell_1} \leq \epsilon\}. \quad (2)$$

The assumption is that ℓ_1 of difference between the current signal and the previous signal, $(\mathbf{G} - \mathbf{f}^{(t-1)}\|_{\ell_1})$ would be small, if the support is changing slowly over time. However, this assumption is not always valid as values and locations of the non-zero elements (spikes) of a sparse signal will typically change over time. Recently Vaswani and Lu [8], [11], proposed the modified-CS which uses the support of the previous time instant ($\mathcal{S}^{(t-1)}$) as an estimated support of the signal of interest ($\mathbf{F}^{(t)}$) at current time and then use this estimate for reconstruction of $\mathbf{F}^{(t)}$, by finding a signal which satisfies the observations and is sparsest outside $\mathcal{S}^{(t-1)}$. This is equivalent to solving the following optimization problem:

$$\min\{\left\|\mathbf{G}|_{\hat{\mathcal{S}}^{(t-1)}}\right\|_{\ell_1} \text{ such that } \mathbf{\Phi g} = \mathbf{y}^{(t)}\}. \quad (3)$$

where $\mathbf{G}|_{\hat{\mathcal{S}}^{(t-1)}} := \{G_i : i \in \hat{\mathcal{S}}^{(t-1)}\}$ and $\hat{\mathcal{S}}^{(t-1)} := \{k \in \{1, \cdots, n\} : \mathbf{F^{(t-1)}}|_{\mathbf{k}} = \mathbf{0}\}$, is the complement of $\mathcal{S}^{(t-1)}$. It is shown in [11], that under fairly general conditions, the number of samples needed would be less than the conventional CS.

In this paper, we apply a CS-based method called "Weighted-CS" [13] to the problem of reconstruction of sequences of MRI images . Through extensive experimental results we show that Weighted-CS is able to achieve exact reconstruction from even fewer number of samples than the modified-CS by extracting more priori information from the previous reconstructed image than just the estimated support. The paper is organized as follows: next section presents the Details of Weighted-CS algorithm as applied to sequences of MRI images. We present and analyze our experimental results in section IV before providing the concluding remarks in section V.

III. WEIGHTED-CS

To guide the reconstruction process, our approach makes use of critical a priori knowledge including the estimated support of the signal of interest in the transform domain. The idea is based on the observation that in sequences of MRI images, nearby slices are closely related to each other (see figure 3). Therefore, the conjecture is that we should be able to extract some priori information about sparsity of $\mathbf{F}^{(t)}$ from $\mathbf{F}^{(t-1)}$. More specifically, we try to estimate the probability of each element of $\mathbf{F}^{(t)}$ having a non-zero value, from $\mathbf{F}^{(t-1)}$. For the signal of interest at time t ($\mathbf{F}^{(t)}$), let p_i be the probability of its i^{th} element having a non-zero value, $p_i := P(F_i > 0)$ and let $\mathbf{p} := [p_1, p_2, \cdots, p_n] \in \mathcal{R}^n$ be the sparsity probability vector. In our proposed method, we first estimate this probability vector $\mathbf{p}$ for the current signal ($\mathbf{F}^{(t)}$) from the reconstructed signal of the previous time instant ($\mathbf{F}^{(t-1)}$) and then use $\mathbf{p}$ to aid the reconstruction of $\mathbf{F}^{(t)}$. This is discussed in detail in the following sub-sections.

A. Reconstruction using the sparsity probability model

In this sub-section, we discuss the problem of recovering a signal $\mathbf{F}$ using its sparsity probability model. More specifically, let $\mathbf{p} := [p_1, p_2, \cdots, p_n] \in \mathcal{R}^n$ be the priori knowledge of the signal of interest's sparsity, where $p_i := P(F_i > 0)$ is the probability of F_i having a non-zero value. It is clear that if $p_i = 0$, then F_i is always 0, while if $p_i = 1$, then it is known beforehand that $\mathbf{F}$ has a spike at location i (though its value is unknown). $p_i = 0.5$ basically means that no priori information of F_i sparsity is available and it is as likely as not to have a non-zero value.

In order to incorporate the probability model of the signal into the process of reconstruction, we propose to minimize a weighted ℓ_1 norm (4), where the weights are adjusted according to the probability of each entry being non-zero:

$$\min ||\mathbf{WG}||_{\ell_1} \text{ subject to } \mathbf{\Phi g} = \mathbf{y}. \quad (4)$$

where $\mathbf{W} = diag([w_1, w_2, \cdots, w_n])$. Intuitively, a smaller weight should be given to those entries with higher probability of being non-zero while those elements with small probability should be penalized with larger weights. Naturally, we want to reward and penalize the elements uniformly using a linear function. Thus, the choice of the weight for each element is:

$$w_i = 2(1 - p_i) \quad (5)$$

Figure 1 shows the chosen weight with respect to the value of the probability. It can be seen that as the probability of an element being non-zero, increases, its weight decreases accordingly. Notice that if $p_i = 0.5$ then $w_i = 1$ and $w_i < 1$ when $p_i > 0.5$ (similarly if $p_i < 0.5$ then $w_i > 1$). It should be noted that the weighted ℓ_1 approach (4), could be seen as a generalized ℓ_1 minimization, since when no priori information of the sparsity is available (which means for all elements $p_i = 0.5$), it reduces to the conventional ℓ_1 minimization as all weights would be equal.

B. Estimation of sparsity probability vector

In this sub-section, we discuss estimation of the sparsity probability model for a signal at time t, from the reconstructed signal at time $t-1$. Based on the assumption that sparsity changes smoothly with time, given a spike in the signal at time $t-1$, there is a good chance that either it remains in the same location, or shifts to some point in the same vicinity in the next time frame (t). Similarly, at time t it is expected that zeros appear in the vicinity of zeros at time

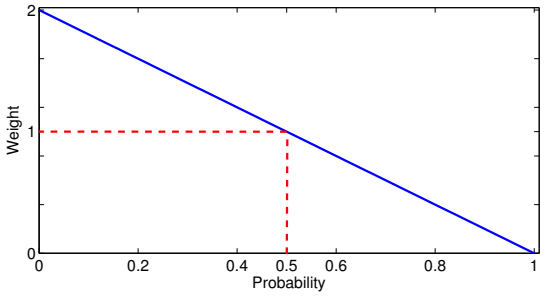
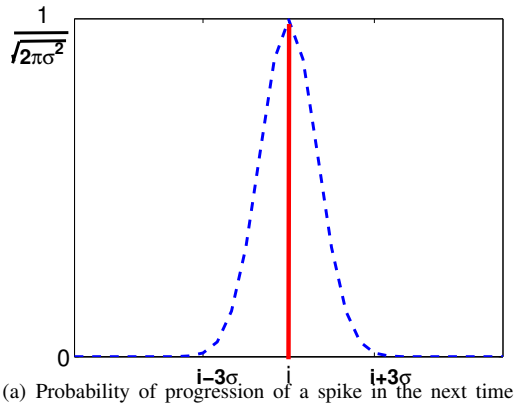
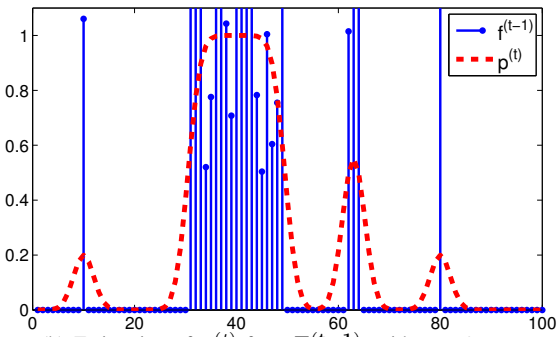


Fig. 1. Chosen weight (w_i) vs. probability (p_i).

$t-1$. More specifically, suppose there is a spike in the i^{th} location of $\mathbf{F}^{(t-1)}$. In the next time instant, there is a very high probability this spike remains in the same location but also some possibility that it moves to some other point in the vicinity. Thus, the probability of the spike appearing at the each location decreases as we get farther from (i). This motivated us to use a Gaussian distribution (figure 2(a)) to provide an estimate of the probability of the progression of a spike in the next time frame.



(a) Probability of progression of a spike in the next time instant



(b) Estimation of $\mathbf{p}^{(t)}$ from $\mathbf{F}^{(t-1)}$, with $\sigma = 2$

Fig. 2. Illustration of estimation of sparsity probability.

Thus the probability of j^{th} element of $\mathbf{F}^{(t)}$ being non-zero ($p_j^{(t)}$) is the accumulated probability of the spikes of $\mathbf{F}^{(t-1)}$ moving to location j at time t as follows:

$$p_j^{(t)} = \sum_{i \in \mathcal{S}^{(t-1)}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(j-i)^2/2\sigma^2} \qquad (6)$$

where $\mathcal{S}^{(t-1)}$ is the support of $\mathbf{F}^{(t-1)}$ and σ^2 is the variance of the Gaussian distribution which is set proportional to the signal's rate of change with time. Figure 2(b) shows a syntetic signal $\mathbf{F}^{(t-1)}$ together with the probability of the elements of $\mathbf{F}^{(t)}$ being non-zero in the next time frame ($\mathbf{p}^{(t)}$) which is the dashed red line. From this figure and equation (6), it can be seen that the maximum value of $p_j^{(t)}$ is 1. This coincides with the locations where in the previous time instant, $F_j^{(t-1)}$ and all elements in its vicinity ($[j-3\sigma, j+3\sigma]$) are non-zero. Similarly, $p_j^{(t)}$ is 0, if there is not any spike in this span.

The above formulation can be extended for 2D images using a 2D Gaussian distribution.

It should be mentioned that if no priori knowledge on the sparsity of the signal of interest is available, including at $t = 1$ where there is no previous time instant to estimate the probability from, we set all $p_i = 0.5$ and basically we solve a basic ℓ_1 minimization.

C. Reconstruction of the time sequences of sparse signals

In the earlier sub-sections, we explained the methods for estimating the sparsity probability model and its use for guiding the reconstruction. Using these methods, the algorithm for reconstruction of sequences of MRI images is summarized in Algorithm 1 below. It should be noted that at time $t = 1$, if no priori knowledge of the image is available, we solve using ℓ_1 minimization thus requiring more samples to achieve a perfect reconstruction.

Algorithm 1 Reconstruction of a sequences of sparse signals varying with time using Weighted-CS.

Input: $\mathbf{\Phi}$ and $\mathbf{y}^{(t)}$.
Output: $\mathbf{g}^{(t)}$.

1) If $t = 1$ then
   $\mathbf{g}^{(t)} = \arg\min\{\|\mathbf{G}\|_{\ell_1}$ subject to $\mathbf{\Phi}^{(1)}\mathbf{g} = \mathbf{y}^{(1)}\}$;
2) $t = t + 1$;
3) $\mathcal{S}^{(t-1)} := \{k \in \{1, \cdots, n\} : G_k^{(t-1)} \neq 0\}$
4) Compute $\mathbf{p}^{(t)} := [p_1^{(t)}, \cdots, p_n^{(t)}]$ from (6).
5) Compute $\mathbf{W}^{(t)}$ from (5).
6) $\mathbf{g}^{(t)} = \arg\min\{\|\mathbf{W}^{(t)}\mathbf{G}\|_{\ell_1}$ subject to $\mathbf{\Phi}\mathbf{g} = \mathbf{y}^{(t)}\}$;
7) Go to step 2

IV. Experimental Results

We tested our algorithm on 5 sequences of MRI images of the foot, knee, ankle, neck (all of size $512 \times 512 \times 20$) and skull base ($512 \times 512 \times 40$), obtained from [14]. Figure 3 shows some of these MRI images.

For the first image in each sequence, $t = 1$, since no priori knowledge is available, 30% of samples are taken along 150 radial lines in the Fourier domain (see sampling mask in figure 4(a)) while in the successive frames, only 10% of samples are taken along 50 radial lines (figure 4(b)) . The reconstruction is carried out in the Wavelet domain which is assumed to be the sparse domain.

Figure 5(a)-(e) compares the reconstruction performance of the Weighted-CS with σ set to 12 for MRI sequences

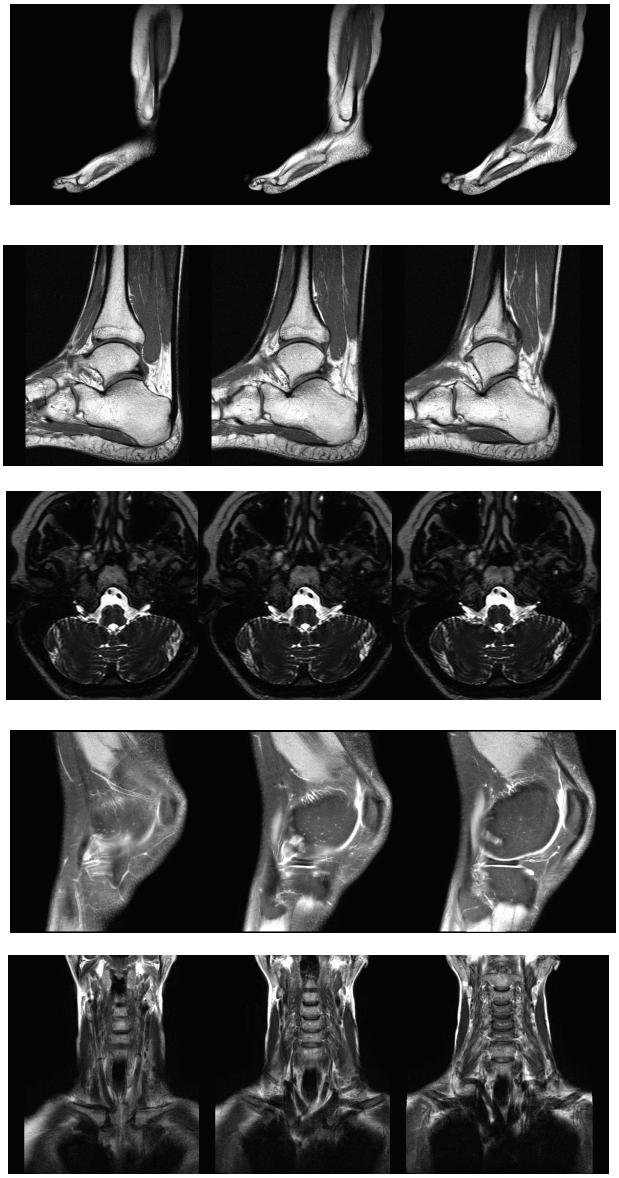
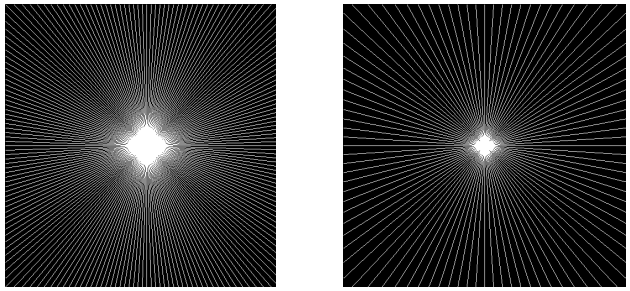


Fig. 3. Sample images of MRI sequences.

with those obtained using the modified-CS, priori-CS and ℓ_1 minimization (CS). It can be seen that our method outperforms the other methods significantly in terms of the improved PSNR.

V. Conclusion

We have employed a Weighted-CS based method to reconstruct sequences of MRI images from a lower that conventionally possible number of samples. Our method uses the image of the previous time instance to efficiently reconstruct the current image of interest using a weighted-ℓ_1 minimization. Our experimental results on MRI datasets, has shown that our algorithm can achieve a significant reduction in the number of samples needed compared to the other state-of-art CS algorithms. One possible extension of our work would be to use different measures of sparsity such as [15], [16] or [17], to further reduce the number of samples needed.



(a) (b)

Fig. 4. Sampling mask for (a) t=1 (b) subsequent frames.

Another interesting challenge would be to develop fast iterative schemes for real-time implementation of Weighted-CS on MRI scanners.

VI. Acknowledgments

This work is supported by a grant (TDSI/11-014/1A) from the Temasek Defence Systems Institute (TDSI), Singapore.

References

[1] E. Candes, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, IEEE Trans. on Information Theory 52 (12) (2006) 5406 – 25.

[2] E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. on Information Theory 52 (2) (2006) 489 – 509.

[3] E. Candes, M. Wakin, An introduction to compressive sampling, IEEE Signal Processing Magazine 25 (2) (2008) 21 – 30.

[4] D. Donoho, Compressed sensing, IEEE Trans. on Information Theory 52 (4) (2006) 1289 – 1306.

[5] J.Trzasko, A. Manduca, Highly undersampled magnetic resonance image reconstruction via homotopic l_0-minimization, IEEE Trans. on Medical Imaging 28 (1).

[6] M. Lustig, D. Donoho, J. M. Pauly, Sparse mri: The application of compressed sensing for rapid mr imaging, Magnetic Resonance in Medicine 58 (6) (2007) 1182–1195.

[7] A. A. Kassim, N. Yan, D. Zonoobi, Wavelet packet transform basis selection method for set partitioning in hierarchical trees., Electronic Imaging 17 (3) (2008) 033007.

[8] W. Lu, N. Vaswani, Modified compressive sensing for real-time dynamic mr imaging, in: ICIP, 2009, pp. 3045–3048.

[9] U. Gamper, P. Boesiger, S. Kozerkey, Compressed sensing in dynamic mri, Magnetic Resonance in Medicine 59 (2) (2008) 365–373.

[10] M. Asif, J. Romberg, Dynamic updating for sparse time varying signals, in: Information Sciences and Systems, 2009. CISS 2009. 43rd Annual Conference on, 2009, pp. 3 –8.

[11] N. Vaswani, W. Lu, Modified-cs: Modifying compressive sensing for problems with partially known support, IEEE Transactions on Signal Processing 58 (9) (2010) 4595 –4607.

[12] Y. Zhang, On theory of compressive sensing via l1 minimization: Simple derivations and extensions, Caam technical report tr08-11, Rice University (2008).

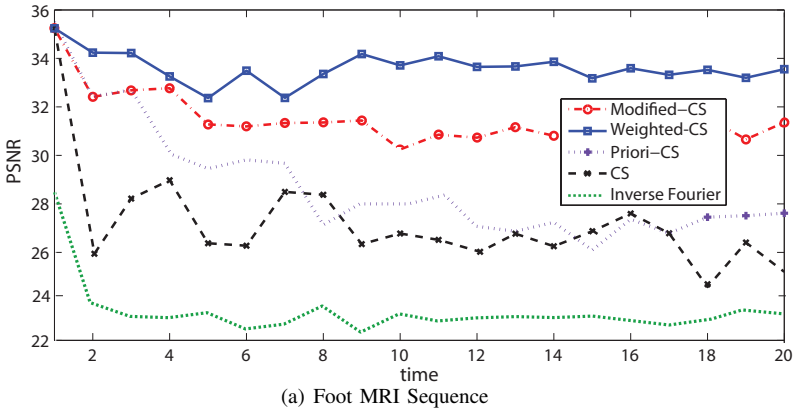
[13] D. Zonoobi, A. A. Kassim, On the reconstruction of sequences of sparse signals- the Weighted-cs, Journal of Visual Communication and Image Representation.

[14] Magnetic resonance technology information portal.
URL http://www.mr-tip.com

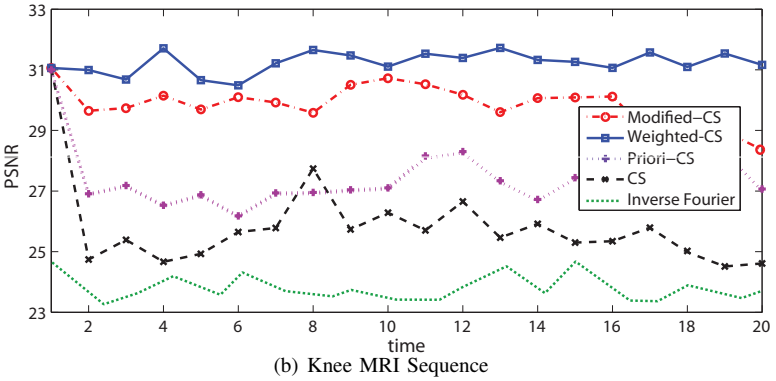
[15] R. Chartrand, Exact reconstruction of sparse signals via nonconvex minimization, IEEE Signal Processing Letters 14 (10) (2007) 707–710.

[16] Y. V. Venkatesh, A. A. Kassim, D. Zonoobi, Medical image reconstruction from sparse samples using simultaneous perturbation stochastic optimization, in: Proc. of the ICIP, 2010, pp. 3369–3372.

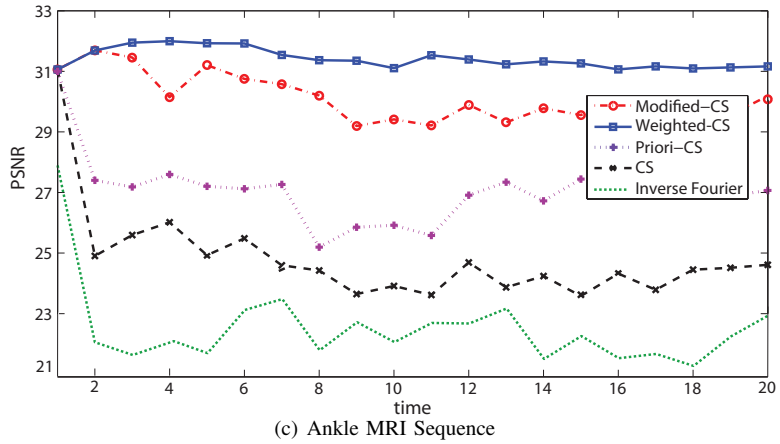
[17] D. Zonoobi, A. A. Kassim, Y. Venkatesh, Gini index as sparsity measure for signal reconstruction from compressive samples, IEEE Journal of Selected Topics in Signal Processing 5 (5) (2011) 927 –932.



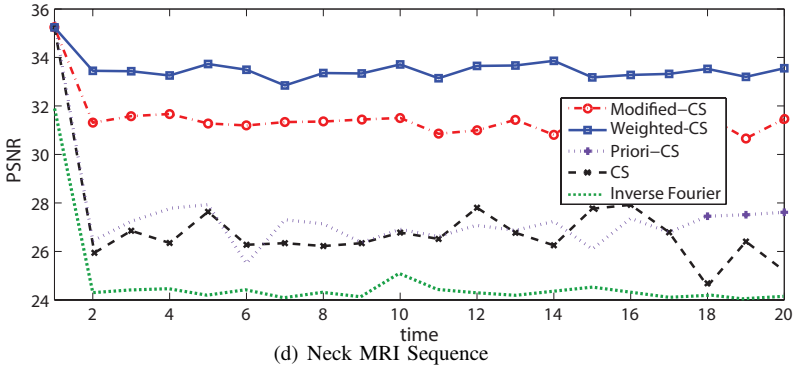
(a) Foot MRI Sequence



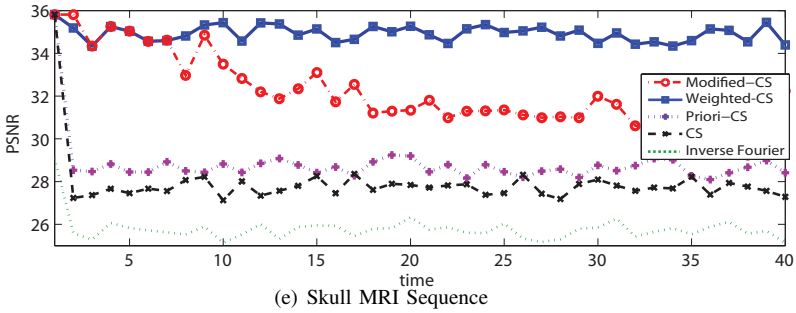
(b) Knee MRI Sequence



(c) Ankle MRI Sequence



(d) Neck MRI Sequence



(e) Skull MRI Sequence

Fig. 5. PSNR of the reconstructed images vs. time.