

Compressed Sensing Dynamic MRI Reconstruction using GPU-accelerated 3D Convolutional Sparse Coding

Tran Minh Quan and Won-Ki Jeong

E-mail: {quantm, wkjeong}@unist.ac.kr
Ulsan National Institute of Science and Technology (UNIST) *

Abstract. In this paper, we introduce a fast alternating method for reconstructing highly undersampled dynamic MRI data using 3D convolutional sparse coding. The proposed solution leverages Fourier Convolution Theorem to accelerate the process of learning a set of 3D filters and iteratively refine the MRI reconstruction based on the sparse codes found subsequently. In contrast to conventional CS methods which exploit the sparsity by applying universal transforms such as wavelet and total variation, our approach extracts and adapts the temporal information directly from the MRI data using compact shift-invariant 3D filters. We provide a highly parallel algorithm with GPU support for efficient computation, and therefore, the reconstruction outperforms CPU implementation of the state-of-the art dictionary learning-based approaches by up to two orders of magnitude.

1 Introduction

Dynamic cardiac MRI is considered as the gold standard among several imaging modalities in heart function diagnosis. However, due to its long acquisition time, its clinical application has been limited to non-time-critical ones. Recent research advances in Compressed Sensing (CS) [?] have been successfully applied to MRI [8] to reduce acquisition time. Nevertheless, CS-MRI poses a new challenge – the reconstruction time also increases because it needs to solve an in-painting inverse problem in the frequency domain (i.e., k-space). Therefore, accelerating the reconstruction process is a top priority to adopt CS framework to fast MRI diagnosis.

Conventional CS-MRI reconstruction methods have exploited the sparsity of signal by applying universal sparsifying transforms such as Fourier (e.g., discrete Fourier transform (DFT) or discrete cosine transform (DCT)), Total Variation (TV), and Wavelets (e.g., Haar, Daubechies, etc.). This research direction has focused on accelerating the sparsity-based energy minimization problem, with [10] or without hardware supports [7]. Some strategies were designed to accelerate the minimization process such as using TV plus nuclear norm [14] or proposed

* Corresponding author: wkjeong@unist.ac.kr

the solver in other sparsity domain such as low-rank technique [9, 12]. More recently, the other approaches leveraging the state-of-the-art data-driven method, i.e., dictionary learning [2] (DL), have been proposed to further enhance the reconstruction quality [3, 5, 6, 11]. However, the existing learning-based methods suffer from the drawback of patch-based dictionary (i.e., redundant atoms and longer running times).

Convolutional sparse coding (CSC) is a new learning-based sparse representation that approximates the input signal with a superposition of sparse feature maps convolved with a collection of filters. This advanced technique replaces the patch-based dictionary learning process with an energy minimization process using a convolution operator on the image domain, which leads to an element-wise multiplication in frequency domain, derived within Alternating Direction Method of Multiplier (ADMM) framework [4], and later its direct inverse problem is introduced by Wohlberg [13]. CSC can generate much compact dictionaries due to its *shift-invariant* nature of filters, and the pixel-wise computation in Fourier domain maps well to parallel architecture. However, such advanced machine learning approaches have not been fully exploited in CS-MRI literature yet. Therefore, in this paper, we propose a novel CS dynamic MRI reconstruction that exploits the compactness and efficiency of 3D CSC. The proposed 3D CSC directly encodes both spatial and temporal features from dynamic cardiac 2D MRI using a compact set of 3D atoms (i.e., filters) without regularizers enforcing temporal coherence (e.g., total variation along the time axis). We also show that the proposed method maps well to data-parallel architecture, such as GPUs, for further accelerating its running time significantly, up to two orders of magnitude faster compared to the state-of-the-art CPU implementation of CS-MRI using patch-based dictionary learning. To the best of our knowledge, this is the first CS-MRI reconstruction method based on GPU-accelerated 3D CSC.

2 Method

Fig. 1 is a pictorial description of the proposed method. If the inverse Fourier transform is directly applied to undersampled MRI k-space data (Fig. 1a $\times 4$ undersampling), the reconstructed images will suffer from artifacts (Fig. 1b). The zero-filling reconstruction will serve as an initial guess for our iterative reconstruction process with randomly initialized filters, e.g., a collection of 16 atoms of size $9 \times 9 \times 9$ as shown in Fig. 1d. Then the image and filters are iteratively updated until they converge as shown in Fig. 1c and e, f.

The proposed CS-MRI reconstruction algorithm is a process of finding s (i.e., a stack of 2D MR images for a given time duration) in the energy minimization problem defined as follows:

$$\begin{aligned} \min_{d, x, s} \quad & \frac{\alpha}{2} \left\| s - \sum_k d_k * x_k \right\|_2^2 + \lambda \sum_k \|x_k\|_1 \\ \text{s.t.} : \quad & \|R\mathcal{F}_2 s - m\|_2^2 < \varepsilon^2, \|d_k\|_2^2 \leq 1 \end{aligned} \quad (1)$$

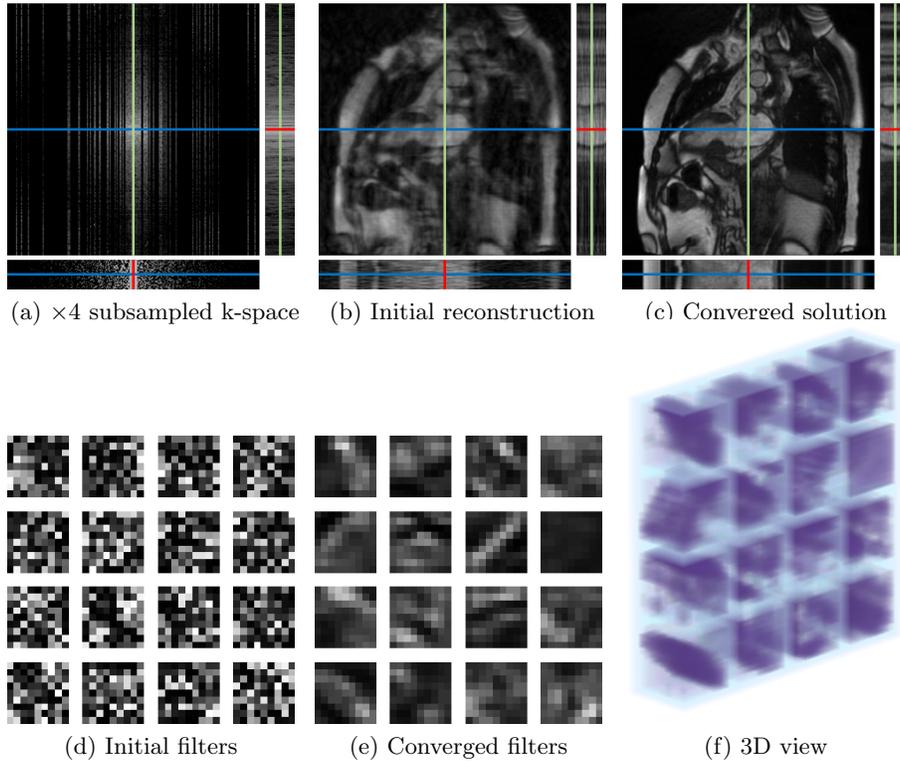


Fig. 1: An overview of CS-MRI reconstruction using 3D CSC method.

where d_k is the k -th filter (or atom in the dictionary) and x_k , is its corresponding sparse code for s . In Eq. 1, the first term measures the difference between s and its sparse approximation $s - \sum_k d_k * x_k$, weighted by α . The second term is the sparsity regularization of x_k using an ℓ_1 norm with a weight λ instead of an ℓ_0 norm as used in [2, 5, 6]. The rest of the equation is the collection of constraints - the first constraint enforces the consistency between undersampled measurement m and the undersampled reconstructed image using the mask R and the Fourier operator \mathcal{F} , and the second constraint restricts the Frobenius norm of each atom d_k within a unit length. In the following discussion, we will use a simplified notation without indices k and replace the result of Fourier transform of a given variable by using the subscript f (for example, d_f is the simplified notation for $\mathcal{F}d$ in 3D domain and s_{f_2} is the simplified notation for \mathcal{F}_2s in 2D spatial domain) to derive the solution of Eq. 1. Therefore, problem 1 can be rewritten using auxiliary variables y and g for x and d as follows:

$$\begin{aligned} & \min_{d,x,g,y,s} \frac{\alpha}{2} \left\| s - \sum d * x \right\|_2^2 + \lambda \|y\|_1 \\ \text{s.t. : } & x - y = 0, \|R\mathcal{F}_2 s - m\|_2^2 < \varepsilon^2, g = \mathcal{P}roj(d), \|g\|_2^2 \leq 1 \end{aligned} \quad (2)$$

where g and d are related by a projection operator as a combination of a truncated matrix followed by a padding-zero matrix in order to make the dimension of g same as that of x . Since we will leverage Fourier transform to solve this problem, g should be zero-padded to make its size same as g_f and x_f . The above constrained problem can be rebuilt in an unconstrained form with dual variables u, h , and further regulates the measurement consistency and the dual differences with γ, ρ , and σ , respectively:

$$\begin{aligned} & \min_{d,x,g,y,s} \frac{\alpha}{2} \left\| s - \sum d * x \right\|_2^2 + \lambda \|y\|_1 + \frac{\gamma}{2} \|R\mathcal{F}_2 s - m\|_2^2 \\ & + \frac{\rho}{2} \|x - y + u\|_2^2 + \frac{\sigma}{2} \|d - g + h\|_2^2 \text{ s.t. : } g = \mathcal{P}roj(d), \|g\|_2^2 \leq 1 \end{aligned} \quad (3)$$

Then we can solve problem 3 by iteratively finding the solution of independent smaller problems, as described below:

Solve for x :

$$\min_x \frac{\alpha}{2} \left\| \sum d * x - s \right\|_2^2 + \frac{\rho}{2} \|x - y + u\|_2^2 \quad (4)$$

If we apply the Fourier transform to the 4, it becomes:

$$\min_{x_f} \frac{\alpha}{2} \left\| \sum d_f x_f - s_f \right\|_2^2 + \frac{\rho}{2} \|x_f - y_f + u_f\|_2^2 \quad (5)$$

Then the minimum solution of 5 can be found by taking the derivative of 5 with respect to x_f and setting it to zero as follows:

$$(\alpha D_f^H D_f + \rho I) x_f = D_f^H s_f + \rho (y_f - u_f) \quad (6)$$

Note that the notation D_f stands for the concatenated matrix of all diagonalized matrices d_{fk} as follows: $D_f = [\text{diag}(d_{f1}), \dots, \text{diag}(d_{fk})]$ and D_f^H is the complex conjugated transpose of D_f .

Solve for y :

$$\min_y \lambda \|y\|_1 + \frac{\rho}{2} \|x - y + u\|_2^2 \quad (7)$$

y for ℓ_1 minimization problem can be found by using a shrinkage operation:

$$y = \mathcal{S}_{\lambda/\rho}(x + u) \quad (8)$$

Update for u : The update rule for u can be defined as a fixed-point iteration with the difference between x and y (u converges when x and y converge each other) as follows:

$$u = u + x - y \quad (9)$$

Solve for d :

$$\min_d \frac{\alpha}{2} \left\| \sum d * x - s \right\|_2^2 + \frac{\sigma}{2} \|d - g + h\|_2^2 \quad (10)$$

Similar to x , d can be solved in the Fourier domain:

$$\min_{d_f} \frac{\alpha}{2} \left\| \sum d_f x_f - s_f \right\|_2^2 + \frac{\sigma}{2} \|d_f - g_f + h_f\|_2^2 \quad (11)$$

$$(\alpha X_f^H X_f + \sigma I) d_f = X_f^H s_f + \sigma (g_f - h_f) \quad (12)$$

where X_f stands for the concatenated matrix of all diagonalized matrices x_{fk} as follows: $X_f = [\text{diag}(x_{f1}), \dots, \text{diag}(x_{fk})]$ and X_f^H is the complex conjugated transpose of X_f .

Solve for g :

$$\min_g \frac{\sigma}{2} \|d - g + h\|_2^2 \quad \text{s.t. :} \quad g = \text{Proj}(d), \|g\|_2^2 \leq 1 \quad (13)$$

g can be found by taking the inverse Fourier transform of d_f . This projection should be constrained by suppressing the elements which are outside the filter size d_k , and followed by normalizing its ℓ_2 -norm to a unit length.

Update for h : Similar to u , the update rule for h can be defined as follows:

$$h = h + d - g \quad (14)$$

Solve for s :

$$\min_s \frac{\alpha}{2} \left\| s - \sum d * x \right\|_2^2 + \frac{\gamma}{2} \|R\mathcal{F}_2 s - m\|_2^2 \quad (15)$$

The objective function of 15 can be transformed into 2D Fourier domain:

$$\min_{s_{f_2}} \frac{\alpha}{2} \left\| s_{f_2} - \mathcal{F}_t^H \sum d_f x_f \right\|_2^2 + \frac{\gamma}{2} \|R s_{f_2} - m\|_2^2 \quad (16)$$

Since d_f and x_f obtained previously in 3D Fourier domain, we need to bring it onto the same space by applying an inverse Fourier transform along time-axis \mathcal{F}_t^H . Then s_{f_2} can be found by solving the following linear system:

$$(\gamma R^H R + \alpha I) s_{f_2} = \gamma R^H m + \alpha \mathcal{F}_t^H \sum d_f x_f \quad (17)$$

Note that the efficient solutions of 6, 12 and 17 can be determined via the Sherman-Morrison formula for independent linear systems as shown in [13]. To this end, after the iteration process, s will be the results of applying a 2D inverse Fourier transform \mathcal{F}_2^H on s_{f_2} .

Implementation details: Since the above derivation consists only Fourier transform and element-wise operations, it maps well to data-parallel architecture, such as GPUs. We used MATLAB to implement the proposed method using the GPU. We set $\alpha = 1$, $\gamma = 1$, $\lambda = 0.1$, $\rho = 10$, $\sigma = 10$ and keep refining the filter banks as well as the reconstruction iteratively until they converge.

3 Result

In order to assess the performance of the proposed method, we compared our algorithm with the stage-of-the-art dictionary learning-based CS reconstruction from Caballero et. al. [5], and the conventional CS reconstruction using wavelet and total variation energy from Quan et. al. [10]. We used three cardiac MRI datasets from The Data Science Bowl [1] – 2 chamber view (2ch), 4 chamber view (4ch), and short axis view (sax). Each dataset consists of 30 frames of a 256×256 image across the cardiac cycle of a heart. In the experiment, we used 3D atoms of size $9 \times 9 \times 9$ and CS-undersampling factor was set to $\times 4$.

Running time evaluation: In order to make this direct performance comparison of learning-based methods between the proposed one and Caballero et al. [5], we measured wall clock running time of both methods on a PC equipped with an Intel i7 CPU with 16 GB main memory and an NVIDIA GTX Geforce Titan X GPU. Our prototype code is written in MATLAB 2015b including GPU implementation, and we used the author-provided MATLAB code for Caballero et al. [5]. As shown in Table 1, we observed that our CPU-based method is about $54 \times$ to $73 \times$, or about two orders of magnitude, faster than the stage-of-the-art DL-based CS-MRI reconstruction method for 100 epochs (i.e., the number of learning iterations). In addition, our GPU-based accelerated implementation also outperforms the CPU version about $1.25 \times$ to $3.82 \times$, which is greatly reduced to a level closer to be ready for clinical application. We expect that the performance of our method can improve further by using CUDA C/C++ without MATLAB.

Table 1: Reconstruction times of learning-based methods (100 epochs)

	2ch	4ch	sax
Caballero et al. [5]	558 Min	475 Min	427 Min
Our method (CPU)	7.67 Min	7.73 Min	7.94 Min
Our method (GPU)	6.14 Min	2.70 Min	2.08 Min

Quality evaluation: Fig. 2 visualizes the reconstruction errors compared to the full reconstruction of each method, respectively. As can be seen, our approach generated less error compared to the stage-of-the-art method of [5] and conventional CS-reconstruction using wavelet and TV energy [10]. Their glitches on the temporal profile are clearly observed since total variation along time axis may smooth out the temporal features that move quickly, especially near the heart boundary. In our case, the learned atoms are in 3D with larger supports, which helps to capture the time trait better even under fast motion and reduces errors in the reconstructed images. In addition, shift-invariance of CSC helps to generate more compact filters compared to the patch-based method.

Fig. 3 shows the achieved Peak Signal-To-Noise-Ratios (PSNRs) measured between the CS-reconstruction results and the full reconstruction. As shown in this figure, our method requires more iterations (epochs) to converge to the steady state, but the actual running time is much faster than the others due to GPU acceleration. In the mean time, our method can reach much higher PSNRs.

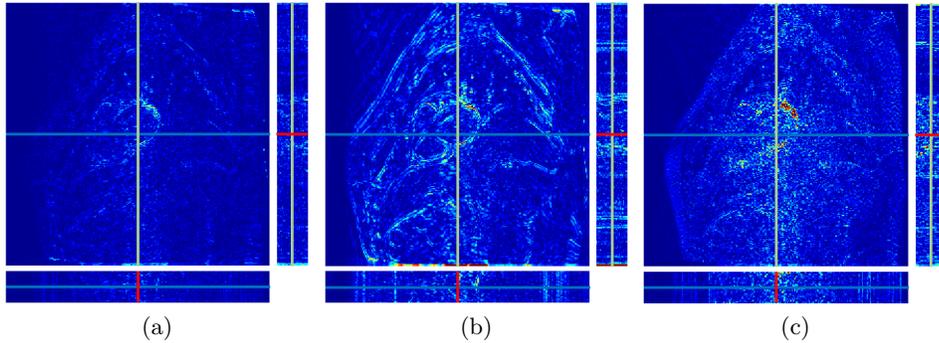


Fig. 2: Error plots (red: high, blue: low) between full reconstruction and the result from (a) the proposed method, (b) the DL-based method [5], and (c) wavelet and total variation energy method [10].

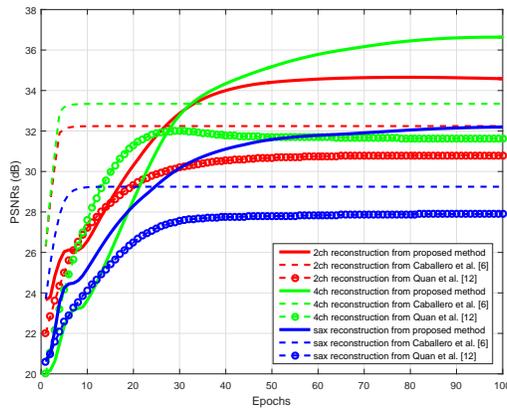


Fig. 3: Convergence rate evaluation based on PSNRs.

4 Conclusion

In this paper, we introduced an efficient CS-MRI reconstruction method based on pure 3D convolutional sparse coding where shift-invariant 3D filters can represent the temporal features of the MRI data. The proposed numerical solver is derived under the ADMM framework by leveraging the Fourier convolution theorem, which can be effectively accelerated using GPUs. As a result, we achieved faster running time and higher PSNRs compared to the state-of-the-art CS-MRI reconstruction methods, such as using a patch-based dictionary learning and conventional wavelet and total variation energy. In the future, we plan to conduct a proper controlled-study of tuning-parameters and assess its feasibility in clinical applications.

Acknowledgements. This work was partially supported by the R&D program of MOTIE/KEIT (No. 10054548, the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- 2014R1A1A2058773) and the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government, MSIP(NRF-2015M3A9A7029725).

References

1. Data science bowl cardiac challenge data (2015), <https://www.kaggle.com/c/second-annual-data-science-bowl/data>
2. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing* 54(11), 4311–4322 (Nov 2006)
3. Awate, S., DiBella, E.: Spatiotemporal dictionary learning for undersampled dynamic MRI reconstruction via joint frame-based and dictionary-based sparsity. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI). pp. 318–321 (May 2012)
4. Bristow, H., Eriksson, A., Lucey, S.: Fast Convolutional Sparse Coding. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 391–398 (Jun 2013)
5. Caballero, J., Price, A., Rueckert, D., Hajnal, J.: Dictionary Learning and Time Sparsity for Dynamic MR Data Reconstruction. *IEEE Transactions on Medical Imaging* 33(4), 979–994 (Apr 2014)
6. Caballero, J., Rueckert, D., Hajnal, J.V.: Dictionary Learning and Time Sparsity in Dynamic MRI. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 256–263. No. 7510 (2012)
7. Jung, H., Sung, K., Nayak, K.S., Kim, E.Y., Ye, J.C.: k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. *Magnetic Resonance in Medicine* 61(1), 103–116 (2009)
8. Lustig, M., Donoho, D., Santos, J., Pauly, J.: Compressed sensing MRI. *IEEE Signal Processing Magazine* 25(2), 72–82 (2008)
9. Otazo, R., Cands, E., Sodickson, D.K.: Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magn. Reson. Med.* 73(3), 1125–1136 (Mar 2015)
10. Quan, T.M., Han, S., Cho, H., Jeong, W.K.: Multi-GPU Reconstruction of Dynamic Compressed Sensing MRI. In: *Medical image computing and computer-assisted intervention (MICCAI)*, pp. 484–492 (2015)
11. Ravishanker, S., Bresler, Y.: MR Image Reconstruction From Highly Undersampled k-Space Data by Dictionary Learning. *IEEE Transactions on Medical Imaging* 30(5), 1028–1041 (May 2011)
12. Tremouilhac, B., Dikaios, N., Atkinson, D., Arridge, S.R.: Dynamic MR Image Reconstruction ;Separation From Undersampled (k-t)-Space via Low-Rank Plus Sparse Prior. *IEEE Transactions on Medical Imaging* 33(8), 1689–1701 (Aug 2014)
13. Wohlberg, B.: Efficient convolutional sparse coding. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 7173–7177 (May 2014)
14. Yao, J., Xu, Z., Huang, X., Huang, J.: Accelerated Dynamic MRI Reconstruction with Total Variation and Nuclear Norm Regularization. In: *Medical image computing and computer-assisted intervention (MICCAI)*, pp. 635–642 (Oct 2015)