

SPATIO-TEMPORAL MRI RECONSTRUCTION BY ENFORCING LOCAL AND GLOBAL REGULARITY VIA DYNAMIC TOTAL VARIATION AND NUCLEAR NORM MINIMIZATION

Cagdas Ulas ^{†¶*} Pedro A. Gómez ^{†¶} Jonathan I. Sperl [¶] Christine Preibisch [‡] Bjoern H. Menze ^{†§}

[†] Department of Computer Science, TU München, Munich, Germany

[§] Institute for Advanced Study, TU München, Munich, Germany

[‡] Neuroradiology, Klinikum rechts der Isar der TU München, Munich, Germany

[¶] GE Global Research, Munich, Germany

ABSTRACT

In this paper, we propose a new spatio-temporal reconstruction scheme for the fast reconstruction of dynamic magnetic resonance imaging (dMRI) data from undersampled k -space measurements. To utilize both spatial and temporal redundancy in dMRI sequences, our method investigates the potential benefits of enforcing local spatial sparsity constraints on the difference to a reference image for each frame and additionally exploiting the low-rank property of global spatio-temporal signal via nuclear norm (NN) minimization. We present here an iterative algorithm that solves the convex optimization problem in an alternating fashion. The proposed method is tested on in-vivo 3D cardiac MRI and dynamic susceptibility contrast (DSC)-MRI brain perfusion datasets. In comparison to two state-of-the-art methods, numerical experiments demonstrate the superior performance of our method in terms of reconstruction accuracy.

Index Terms— compressed sensing, dynamic MR imaging, low-rank approximation, total variation, nuclear norm

1. INTRODUCTION

Dynamic magnetic resonance imaging (dMRI) is an important medical imaging technique that enables the visualization of anatomical and functional changes of internal body structures through time, resulting in a spatio-temporal signal. Although MRI is a non-invasive, non-ionizing technology and provides an unmatched quality in soft tissue contrast, physical and physiological limitations on scanning speed makes this an inherently slow process [1]. Besides, there is a trade-off between the spatial and temporal resolution. The reason is that acquiring fewer k -space samples than those dictated by the Nyquist criterion accelerates the process significantly, but exhibits aliasing artifacts in image space. Fortunately, dynamic MR sequences usually provide redundant information

in both spatial and temporal domains, which allows the reduction of acquisition time by using compressed sensing (CS) approaches [2, 3]. More recently, CS theory has been applied to MRI enabling highly accurate reconstructions from fewer k -space measurements depending on the assumption of sparsity of the reconstructed data under some transform domain [4].

In recent years, researchers have proposed sophisticated CS-based reconstruction methods that exploit both spatial and temporal redundancies of the entire dataset, such as spatio-temporal total variation [5], dictionary learning [6], and low-rank approximation and sparsity [5, 7]. In general, dynamic MR images are temporally correlated due to slow changes of the same organ(s) through the whole image sequence, and such high correlation in the temporal domain has been successfully investigated based on a sparsity constraint in the temporal domain for dMRI reconstruction [6]. As an extension of the conventional spatial total variation (TV), a new sparsity inducing norm called dynamic total variation (dTV) [8] has been recently introduced to utilize both spatial and temporal correlations in online reconstruction.

In this paper, we make an attempt to integrate two fundamentally different approaches for CS-based reconstruction: we enforce local coherences at the pixel-level via dynamic total variation (dTV) and global regularity in the full spatio-temporal domain via a nuclear norm (NN) minimization constraint. We present the dTV/NN optimization in a joint formal framework which allows us to rely on an iterative minimization algorithm. The joint minimization problem is solved iteratively by utilizing an alternating minimization strategy. The proposed method is validated on two different dynamic MR sequences with comparisons to state-of-the-art methods.

Our main contributions can be summarized as follows: We propose a novel reconstruction scheme that iteratively enforces not only the local (spatial) regularity in every single frame but also the global (spatio-temporal) regularity of a full sequence. To this end, we introduce a reconstruction model that is jointly using dTV sparsity and nuclear norm penalties, exploiting both the sparsity of inter-frame differences and the low-rank structure of the dynamic MR sequences in the full

This research has received funding from the European Union's H2020 Framework Programme (H2020-MSCA-ITN-2014) under grant agreement no 642685 MacSeNet.

* Corresponding author. E-mail: cagdas.ulas@tum.de

spatio-temporal space. Our approach also employs, for the first time, the dTV sparsity inducing norm in an offline reconstruction scheme.

2. METHODS

2.1. Problem formulation

Here, we denote \mathbf{X}^{3D} as a dMRI sequence to be represented as a spatio-temporal 3D volume of size $P = N \times N \times T$, i.e., the images are of size $N \times N$ and T is the total number of frames in the sequence. Let X_t denote the MR image matrix at the t th frame, y_t is the k -space data for the t th frame and $\mathbb{T} = \{1, 2, \dots, T\}$ is the set of frame number indices. The main objective here is to reconstruct all X_t 's, $t \in \mathbb{T}$, from the collected k -space measurements y_t 's. The MR imaging equation for each frame is formulated as

$$y_t = \mathcal{F}_t x_t + \eta \quad (1)$$

where \mathcal{F}_t denotes the undersampling 2D Fourier operator for frame t , i.e., $\mathcal{F}_t = R_t \mathcal{F}_{2D}$, where $R_t \in \mathbb{R}^{m \times N}$, $m \ll N$, is the undersampling mask to acquire only a subset Ω of k -space, x_t denotes the MR image vector formed by row/column concatenation of X_t and $\eta \in \mathbb{C}^m$ is additive Gaussian noise in k -space. We stack the data for all the frames of the MR sequence as columns and denote them as follows: $Y = [y_1 | y_2 | \dots | y_T]$, $X = [x_1 | x_2 | \dots | x_T]$, and $\mathcal{F}_u = \text{diag}\{\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_T\}$.

We propose solving the following optimization problem for the reconstruction of dMRI sequences:

$$\begin{aligned} \min_X \quad & \nu_1 \|X\|_* + \nu_2 (dTV(X, \bar{x})) \\ \text{s.t.} \quad & \|\mathcal{F}_u X - Y\|_2^2 \leq \epsilon \end{aligned} \quad (2)$$

where ν_1 and ν_2 are respective regularization parameters for the two terms, and $\|X\|_*$ denotes the nuclear norm of X and is calculated as

$$\|X\|_* = \sum_i \sigma_i(X)$$

where $\sigma_i(X)$ represents the i^{th} singular value of X . For an image x_t with N^2 pixels, $dTV(X, \bar{x})$ can be defined as

$$dTV(X, \bar{x}) = \sum_{t \in \mathbb{T}} \sum_{n=1}^{N^2} \sqrt{(\nabla_x(x_t - \bar{x})_n)^2 + (\nabla_y(x_t - \bar{x})_n)^2}$$

where \bar{x} is the reference image calculated by averaging all the frames in the sequence, ∇_x and ∇_y represent the finite-difference matrices along the x and y dimensions respectively.

Let us introduce new variables $z_t = x_t - \bar{x}$ and $b_t = y_t - \mathcal{F}_t \bar{x}$, then the problem (2) can be reformulated as follows:

$$\begin{aligned} \min_{X, z} \quad & \nu_1 \|X\|_* + \nu_2 \sum_{t \in \mathbb{T}} \|z_t\|_{TV} \\ \text{s.t.} \quad & \begin{cases} \|\mathcal{F}_u X - Y\|_2^2 \leq \epsilon \\ \|\mathcal{F}_t z_t - b_t\|_2^2 \leq \epsilon, \forall t \end{cases} \end{aligned} \quad (3)$$

where $z = [z_1, \dots, z_T]$ and $\|z_t\|_{TV} = \|[D_1 z_t, D_2 z_t]\|_{2,1}$, where D_1 and D_2 are two $N^2 \times N^2$ first order finite difference matrices in vertical and horizontal directions, and $\ell_{2,1}$ norm is the summation of the ℓ_2 norm of each row, $[\mathbf{a}_1, \mathbf{a}_2]$ denotes concatenating two vectors \mathbf{a}_1 and \mathbf{a}_2 horizontally.

2.2. Image reconstruction algorithm

The optimization problem (3) is convex and we choose to split it into two simpler subproblems that can be efficiently solved with greedy algorithms. Alternating the solution of these two subproblems iteratively will give an approximate solution to Eq. (3). In this approach, an approximate generic solution is refined towards a better solution.

- Subproblem 1 : Enforcing local (spatial) regularity

$$\min_{z_t} \quad \frac{1}{2} \|\mathcal{F}_t z_t - b_t\|_2^2 + \nu_2 \|z_t\|_{TV}, \quad \forall t \quad (4)$$

For each frame x_t in the sequence, we solve the optimization problem (4) to reconstruct each frame individually given a reference image \bar{x} . This guarantees that the sum of TV norms in Eq. (3) is also minimized. The problem (4) can be efficiently solved by the fast iteratively reweighted least squares (FIRLS) algorithm [9] based on preconditioned conjugate gradient. This algorithm provides fast convergence and low computational cost by designing a new preconditioner which can be accurately approximated using the properties of the Fourier transform and diagonally dominant structure of the $\mathcal{F}_t^H \mathcal{F}_t$ matrix, where H denotes the conjugate transpose. We refer the reader to [9] for more details on FIRLS.

- Subproblem 2 : Enforcing spatio-temporal regularity

$$\min_X \quad \frac{1}{2} \|\mathcal{F}_u X - Y\|_2^2 + \nu_1 \|X\|_* \quad (5)$$

The spatio-temporal signal representation of a dMRI sequence can be arranged as a 2D matrix of X , where each column represents a vectorized image frame. Due to the repetitive structure of the dMRI sequence between consecutive frames, and the resulting high correlation between each column of X , this matrix can be generally approximated to be low-rank, i.e., X has only a few significant singular values.

By exploiting the low-rank property of X , we can solve a low-rank matrix recovery problem using convex nuclear norm as a prior. In this way, we pose low-rank matrix recovery as a nuclear norm regularized linear least squares problem as stated in (5). This problem can be solved iteratively through an accelerated proximal gradient (APG) algorithm [10]. The algorithm provides a computationally efficient way of recovering low-rank matrices iteratively and consists of two main steps: a first order update and a proximal projection of the penalty that is solved via the singular value thresholding operator, i.e., $S_\alpha(\mathbf{G}) = \mathbf{U} \text{diag}\{(\Sigma - \alpha)_+\} \mathbf{V}^H$, where \mathbf{U} , Σ , \mathbf{V} are obtained from singular value decomposition of \mathbf{G} .

Our proposed scheme follows an iterative refinement of an

Algorithm 1 NNMdTV reconstruction

```
1: Input:  $Y, \nu_1, \nu_2, \mathcal{F}_u, IterNo$ 
2: Output:  $X$ 
3: Initialize:  $X = X_0 = \mathcal{F}_u^H Y, \bar{x} = \bar{x}_0 = mean(\mathcal{F}_u^H Y)$ 
4: for  $i = 1$  to  $IterNo$  do
5:   for each  $t \in \{1, 2, \dots, T\}$  do
6:      $\hat{z}_t \leftarrow \arg \min_{z_t} \frac{1}{2} \|\mathcal{F}_t z_t - b_t\|_2^2 + \nu_2 \|z_t\|_{TV}$ 
7:      $x_t \leftarrow \hat{z}_t + \bar{x}_i$ 
8:   end for
9:   Form updated  $X_i = [x_1 | x_2 | \dots | x_T]$ 
10:   $X_i \leftarrow \arg \min_X \frac{1}{2} \|\mathcal{F}_u X - Y\|_2^2 + \nu_1 \|X\|_*$ 
11:   $\bar{x}_i \leftarrow mean(X_i)$ 
12: end for
```

initial solution. First, we start with zero-filled sequence and iteratively improve the previous reconstruction by first solving the Subproblem 1 for each frame and refining this solution by solving the Subproblem 2 as a following step. Second, in each iteration we update the reference image that is used for solving Subproblem 1, providing a better reference image given as the input to the problem (4), thus yielding more accurate reconstructions. Throughout the paper we will simply term our proposed method as NNMdTV. Algorithm 1 summarizes the steps of the NNMdTV algorithm.

3. EXPERIMENTS AND RESULTS

3.1. Experimental setup

We evaluate our method on two different types of dynamic MR sequences. We use an in-vivo breath-hold cardiac perfusion sequence [8] of size $192 \times 192 \times 40$ and a dynamic susceptibility contrast (DSC)-MRI brain perfusion sequence of size $128 \times 128 \times 60$ with normalized intensities. Both sequences are artificially corrupted by multiplying its corresponding k -space representation with a binary undersampling mask and subsequently adding complex Gaussian white noise with a standard deviation σ . A radial sampling mask is used to simulate undersampling. The same undersampling mask is used for all frames in our experiments.

3.2. Evaluation

For quantitative evaluation, we adopt the Peak Signal-to-Noise Ratio (PSNR) as the metric in our experiments. We compare our method with two state-of-the-art methods: k-t SLR [5] and dynamic total variation (dTV) [8]. The codes of dTV and k-t SLR reconstruction methods are downloaded from each author's website and for k-t SLR we use the default parameter settings in the package for all experiments. For dTV reconstruction, we use the first frame as the reference frame with 1/4 sampling rate and 1/6 sampling rate for the remaining frames. The sampling rate for all frames is also

set to 1/6 for NNMdTV and k-t SLR. To ensure fair comparison, the parameters settings used in dTV reconstruction are also used in our NNMdTV method for all experiments. For the NNMdTV method, we set $\nu_1 = 5 \times 10^{-8}$, $\nu_2 = 0.001$ and $IterNo = 5$ for both sequences. The noise standard deviation is set to $\sigma = 10^{-5}$ for all reconstruction methods.

3.3. Experimental results

In Figs. 1 and 2, we present qualitative results for the DSC brain and cardiac perfusion datasets respectively. Fig. 1 shows the temporal profile of the DSC brain data along a fixed row. From the error maps (see Fig. 1(d, f, h)), it is clearly visible that NNMdTV reconstructs better result compared to other two methods. The red arrows in Fig. 1(h) show the regions where the reconstruction is improved with NNMdTV. A frame of the reconstructed cardiac sequence is shown in Fig. 2. Visible artifacts can clearly be seen on the images reconstructed by k-t SLR. In contrast, compared to the dTV, the reconstruction result of NNMdTV is more similar to the fully-sampled frame, and less noisy (see Fig. 2(h)).

Quantitative results of different methods on two perfusion datasets are shown in Fig. 3. From the figure, we can clearly observe that the proposed NNMdTV achieves the highest PSNR for each iteration and for all frames of the sequences. The graphs at the top of Fig. 3 mainly validate the fact that iteratively updated mean reference image in NNMdTV enables better reconstruction accuracy.

4. CONCLUSION

In this paper, we have proposed a new CS-based reconstruction model for dynamic MRI based on the joint minimization of local differences in each frame and global differences in the full spatio-temporal space and developed an iterative reconstruction algorithm to solve this minimization problem.

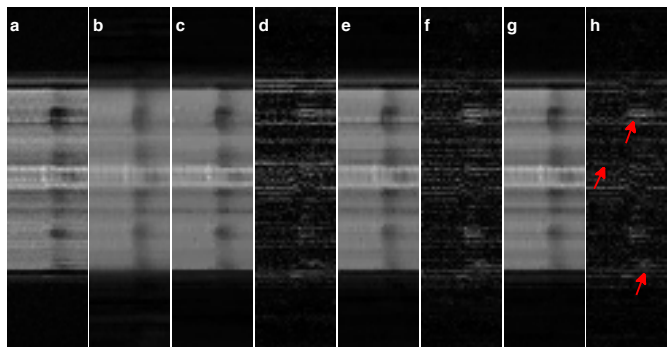


Fig. 1. Temporal profile of row 75 in the original DSC brain dataset (a), its undersampled by 6 zero-filled version (b), and reconstructions using k-t SLR (c), dTV (e), and NNMdTV (g) with their respective errors multiplied by 3 (d, f, h).

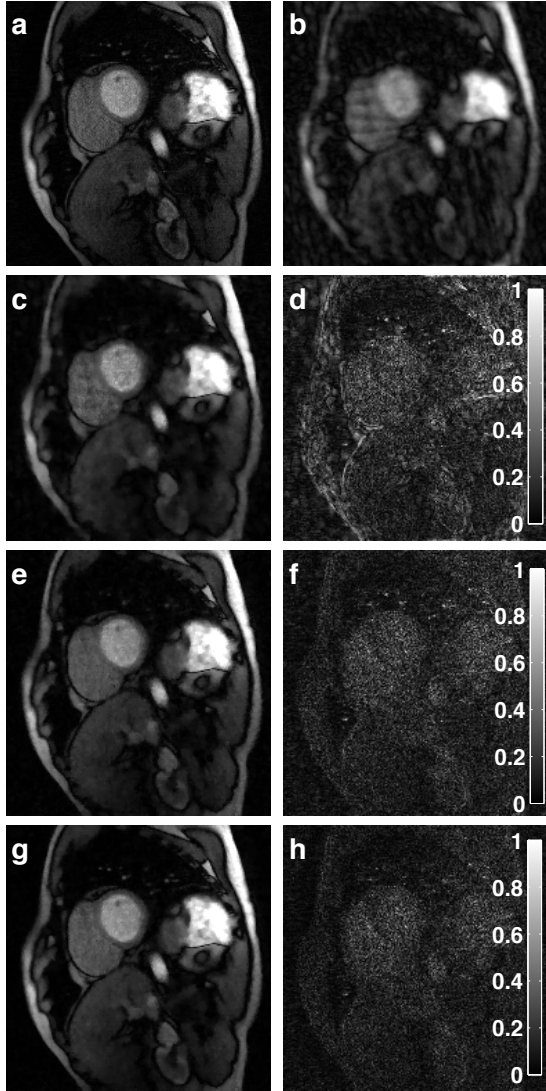


Fig. 2. Visual comparison of a fully sampled frame of cardiac dataset (a), its undersampled by 6 zero-filled version (b), and reconstructions using k-t SLR (c), dTV (e), and NNMdTV (g) with their respective errors magnified by 4 (d, f, h).

Experiments on two different perfusion datasets have demonstrated the effectiveness of our method over the state-of-the-art. Future work will aim at extending our method with the use of patch-wise redundancies of spatio-temporal neighborhoods in adjacent frames and making it more robust to noisy scenarios and large inter-frame motion.

5. REFERENCES

[1] P. A. Gomez et al., “Learning a spatiotemporal dictionary for magnetic resonance fingerprinting with compressed sensing,” in *1st Int. Work. on Patch-based Techniques in Medical Imaging, MICCAI*, October 2015.

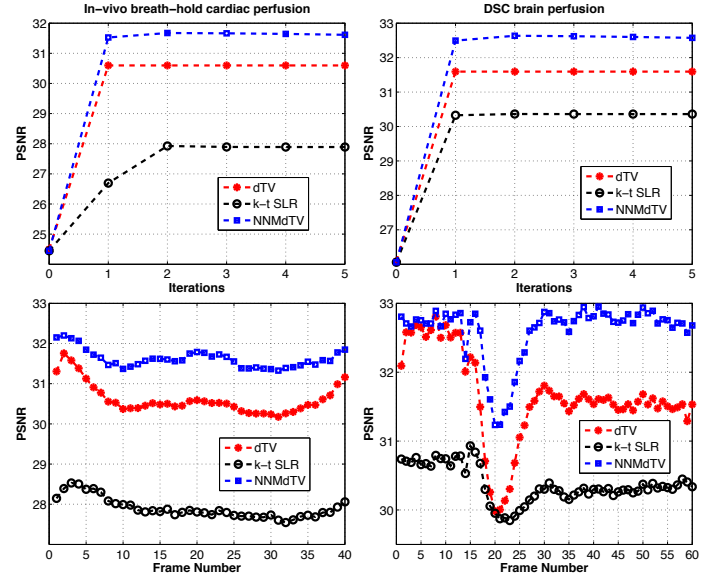


Fig. 3. PSNR comparisons of different reconstruction methods. Cardiac dataset (left), DSC brain dataset (right).

[2] M. Lustig, D. Donoho, and J. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[3] U. Gamper, P. Boesiger, and S. Kozerke, “Compressed sensing in dynamic MRI,” *Magn. Reson. Med.*, vol. 59, no. 2, pp. 365–373, 2008.

[4] N. Vaswani and W. Lu, “Modified-CS: Modifying compressive sensing for problems with partially known support,” *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, 2011.

[5] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, “Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR,” *IEEE Trans. on Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.

[6] J. Caballero, A.N. Price, D. Rueckert, and J. Hajnal, “Dictionary learning and time sparsity for dynamic MR data reconstruction,” *IEEE Trans. on Med. Imag.*, vol. 33, no. 4, pp. 979–994, April 2014.

[7] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z.-P. Liang, “Image reconstruction from highly undersampled (k,t)-space data with joint partial separability and sparsity constraints,” *IEEE Trans. on Med. Imag.*, vol. 31, no. 9, pp. 1809–1820, September 2012.

[8] C. Chen, Y. Li, L. Axel, and J. Huang, “Real time dynamic MRI with dynamic total variation,” in *Proc. 17th Int. Conf. MICCAI*, 2014, LNCS, pp. 138–145.

[9] C. Chen, J. Huang, L. He, and H. Li, “Preconditioning for accelerated iteratively reweighted least squares in structured sparsity reconstruction,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014, pp. 2713–2720.

[10] K. C. Toh and S. Yun, “An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems,” *Pacific Journal of Optimization*, vol. 6, pp. 615–640, 2010.