IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 22, 2019, accepted October 14, 2019, date of publication October 17, 2019, date of current version October 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948017

Real Time Dynamic Magnetic Resonance Imaging via Dictionary Learning and Combined Fourier Transform

YANG WANG[1,2], NING CAO[2], YUELI CUI[1], AND QIANG ZHOU[1]
[1]School of Electronic and Information Engineering, Taizhou University, Taizhou 318000, China
[2]School of Computer and Information, Hohai University, Nanjing 210098, China

Corresponding author: Ning Cao (caoning@vip.163.com)

This work was supported by the Natural Science Foundation of Zhejiang Province, China, under Grant LY18F010005.

ABSTRACT Real time dynamic magnetic resonance imaging (dMRI) requires that the image acquisition and reconstruction are carried out simultaneously and the reconstruction speed catches up with imaging speed. In this paper, a novel compressed sensing (CS) reconstruction algorithm for real time dynamic MRI is proposed. The first frame with more k-space measurements is reconstructed precisely as the reference image. Different from previous methods who start their reconstructions from zero-filled k-space measurements, a Combined Fourier Transform (CFT) algorithm is implemented in our method, which can dynamically aggregate the k-space measurements from previous sampled frames to create a highly accurate predictive image for the current frame. We then combine the CFT algorithm with a 3D path-based dictionary leaning algorithm, which is named as DLCFT in our work for fast real time dMRI reconstruction. The proposed algorithm is compared with four state-of-the-art online and offline methods on two real and complex perfusion MR sequences and a real functional brain MR sequence. Experimental results show that the proposed algorithm outperforms these methods with faster convergence and higher reconstruction accuracy.

INDEX TERMS Dynamic magnetic resonance imaging, compressed sensing, combined Fourier transform, dictionary leaning.

I. INTRODUCTION

Magnetic resonance imaging (MRI) is a widely used imaging modality in hospitals for diagnosis due to its superiority in non-invasion and high contrast resolution of soft tissue. However, current MRI facilities have several shortcomings, one of which is the relatively slow imaging speed, particularly in the field of dynamic magnetic resonance imaging (dMRI). A dMRI sequence with high spatiotemporal resolution could be time consuming due to the nuclear physics mechanism [1]. Slow imaging speed will cause motion artifacts, which still constitutes the main challenge to dMRI, such as dynamic cardiac cine. Patients are requested to hold their breath in case of artifacts during the scan. However, many patients are unable to sustain breath-hold for a long time, so the data scanning process must be accelerated.

Undersampling in k-space to reduce the number of measurements is a feasible way to improve imaging speed, but it often leads to aliasing artifacts under the Nyquist sampling criterion. How to reconstruct high-resolution dynamic MR images from undersampled k-space measurements is an underdetermined problem that is still in challenging at present.

It is well known that most natural signals are sparse in the transform domain that can be expressed by a small number of parameters or features. For example, in Fourier transform domain, an image can be represented by a small number of transform coefficients; in a video sequence, adjacent frames can be represented by a reference image and a residual image using motion compensation techniques. dMRI sequences contain abundant redundancy in spatiotemporal dimensions, which promotes the compressed sensing (CS) technology wildly used in MRI reconstruction [2]–[10]. According to the theory of compressed sensing, a signal of length n can be accurately reconstructed from a small number of incoherent observations ($k + k \log n$, k is the sparsity of the signal) by a compressed sensing and reconstruction framework, as long as the measurement matrix is incoherent with the transform base and satisfies restricted isometry property (RIP) [11], [12].

The associate editor coordinating the review of this manuscript and approving it for publication was Fang Yang.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

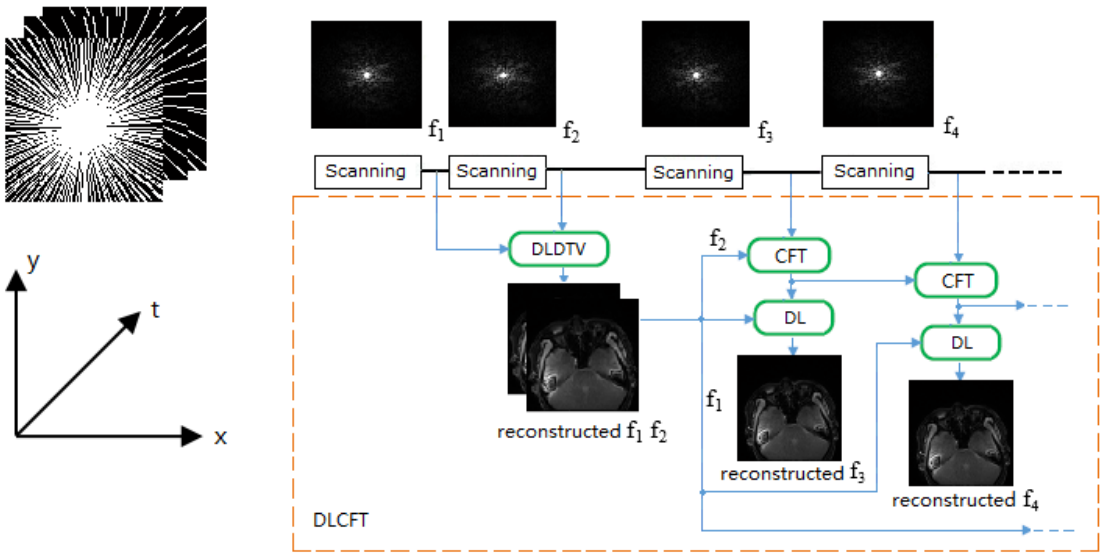

FIGURE 1. Schematic of the proposed algorithm.

CS dMRI reconstruction methods can be implemented in offline mode or online mode. Offline methods need to obtain the whole frames before reconstruction, which takes up a lot of time. Typical offline methods such as motion correction [13]–[15], low rank approximation [16]–[18], and dictionary learning [19]–[21], utilize the information and sparsity characteristics of entire dataset for high-resolution recovery, while their computational time is relatively long, which limits their applications in real time situations. Online methods implement their programs only on the previously sampled data or reconstructed images, which saves the computing time. Online mode is more suitable for real time dynamic MRI, while the reconstruction accuracy cannot be guaranteed due to the lack of prior information of the whole dataset.

Online methods can implement their programs in parallel or serial mode. In serial mode, a frame of image is recovered using the prior information from adjacent frames, which is a common strategy among exiting online schemes. However, it will lead to error accumulation on a long dMRI sequence. Chen *et al.* [22] propose a Dynamic Total Variation (DTV) method to solve the error accumulation problem, who uses the first frame with more measurements as a reference frame, and reconstructs the later frames one by one in parallel mode. Wang *et al.* [23] extend this parallel mode to arbitrary *n*-frame structure flexibly using Dictionary Learning and Dynamic Total Variation (DLDTV) as spatiotemporal regularizations, which greatly improves the accuracy of reconstruction.

All these CS-dMRI strategies focus on the sparsity from the measurements in adjacent frames or the whole dMRI sequence, while these measurements are scattered randomly in the whole dMRI sequence. If a frame to be reconstructed can aggregate the k-space measurements from sampled images and give a highly accurate predictive image before the CS reconstruction, it will greatly shorten the reconstruction time.

In this paper, we proposed a novel algorithm named Dictionary Learning and Combined Fourier Transform (DLCFT) for compressed sensing dMRI reconstruction (see fig. 1). The 1^{st} frame is sampled with more measurements and can be reconstructed by any online method. In the 1^{st} step, we choose DLDTV algorithm to reconstruct the first two frames together, which will be explained in the experiment part. In the 2^{nd} step, as soon as the t^{th} frame ($t > 2$) is sampled, we add its k-space measurements to the reconstructed 2^{nd} frame and use a Combined Fourier Transform (CFT) algorithm to generate a highly accurate predictive image before reconstruction. In the 3^{rd} step, a 3D path-based dictionary leaning algorithm is implemented on the subset of the 1^{st} image and the predictive image for fast reconstruction. Numerical experiments conducted on two real and complex perfusion MR sequences and a real brain MR sequence validate the effectiveness and efficiency of the proposed method.

The rest of this paper is organized as follows. Section 2 reviews the background and some sparsity models in CS dMRI. Section 3 details the DLCFT algorithm. Section 4 gives experimental procedures and compares our method with four state-of-the-art online and offline methods. Section 5 concludes the whole paper.

II. NOTATIONS AND RELATED WORK

A. IMAGING MODEL

Suppose a dMRI sequence is denoted as $x=[x_1, x_2, \ldots, x_T]$ where x_T is the T^{th} frame image, and its k-space partial sampled Fourier measurements is denoted as $y=[y_1, y_2, \ldots, y_T]$. $X \in R^{N\times T}$ and $y \in C^{M\times T}$, where M is the number of k-space measurements and N is the number of pixels in each frame. The vector form of the imaging system can be

written as:

$$y = F_u \mathrm{x} + e \tag{1}$$

where e is the k-space sampling noise, $F_u \in C^{MT \times NT}$ is the sensing matrix of entire sequence. Equation (1) can be unfold frame by frame as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} F_{\Omega_1} & 0 & \cdots & 0 \\ 0 & F_{\Omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{\Omega_T} \end{bmatrix} \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \\ \vdots \\ \mathrm{x}_T \end{bmatrix} + \begin{bmatrix} \mathrm{e}_1 \\ \mathrm{e}_2 \\ \vdots \\ \mathrm{e}_T \end{bmatrix} \tag{2}$$

where $F_{\Omega i}$ denotes an operator who only computes Fourier coefficients at sampled locations indexed by Ω_i. Ω_i is randomly selected at each frame to satisfy the incoherence request of the CS reconstruction system.

B. COMPRESSED SENSING STRUCTURE

The compressed sensing dMRI reconstruction process can be formulated as the following l_0 regularized least-squares problem:

$$\min_{\mathrm{x}} \frac{1}{2} \|F_u \mathrm{x} - \mathrm{y}\|_2^2 + \lambda \|\Psi \mathrm{x}\|_0 \tag{3}$$

where the first term keeps the data fidelity towards the measurements, the second term keeps the solution to be sparse in the transformed domain Ψ, and λ is a positive regularization parameter that makes a trade-off between two terms. The l_0 optimization problem needs to list all possible linear combinations of non-zero positions in x, which is non-convex and NP hard [24]. Many greedy algorithms have been exploited to solve this problem effectively, such as Basis Pursuit (BP) and Orthogonal Matching Pursuit (OMP) [25], [26]. Another appropriate solution is to replace the l_0 norm with l_1 norm because l_1 is the optimal convex approximation of l_0 with many efficient solutions, and l_1 can also guarantee the solution to be sparse with the transform Ψ [27]. Equation (3) can be rewritten as a convex optimization problem based on l_1 norm:

$$\min_{\mathrm{x}} \frac{1}{2} \|F_u \mathrm{x} - \mathrm{y}\|_2^2 + \lambda \|\Psi \mathrm{x}\|_1 \tag{4}$$

where $\|\mathrm{x}\|_1 = \sum |x_i|$ calculates the absolute sum of each element in x. Equation (4) can be solved by a lot of fast and efficient methods such as Fourier transform, wavelet transform and Total Variation (TV) regularization. However, these methods with fixed sparse basis often fail to capture characteristics of the entire sequence, which will emerge blocking artifacts as the sampling ratio decreases.

Sparse representation over adaptive dictionary learning has been widely used in many applications such as image processing and machine learning. Many researchers have explored dictionary learning methods to reconstruct medical images. Bai *et al.* [28] propose a CBCT algorithm based on 3-D dictionary learning for volumetric CT image reconstruction. Liu *et al.* [29] propose a KSVD dictionary learning technique for 3D PACT reconstruction. Jiang *et al.* [30] propose a dictionary training algorithm on patches with two different resolutions for super-resolution CT imaging. Chen *et al.* [31] propose a dictionary learning penalized imaging method for PET reconstruction. Wang and Ying [32] propose a spatiotemporal dictionary learning combined with 3D TV method for compressed sensing dynamic cardiac cine MRI reconstruction. All these DL methods work in offline mode and their reconstruction speed are relatively slow. Nguyen-Duc and Jeong [33] build a multi-scale 3D dictionary with elastic net regularization for dynamic MRI reconstruction, which can be implemented on GPU for efficient computation.

In adaptive dictionary learning methods, the images of a sequence need to be divided into overlap patches before the training. Suppose the patch size is $\sqrt{N} \times \sqrt{N}$ and the patch matrix is denoted as $X = \{R_i \mathrm{x}\}_{i=1}^{l}$, R_i is the operator taking the overlap patches from images, l is the total number of patches, $R_i \mathrm{x}$ is the vector form of chosen patches. The objective function of patch-based dictionary learning algorithms can be formulized as:

$$\min_{D,\alpha} \sum_i \|R_i \mathrm{x} - D\alpha_i\|_2^2 \ \ s.t.\ \forall i, \quad |\alpha_i|_0 \le T_0 \tag{5}$$

where D is an overcomplete dictionary of size $N \times M (M >> N)$, which means the dictionary has more atoms (columns) than rows. $\alpha = \{\alpha_i\}_{i=1}^{l}$ is the sparse coefficient matrix, and T_0 is the sparsity constraint. X is sparsely represented in the dictionary D. Equation (5) can be rewritten as an unconstrained mode:

$$\min_{\mathrm{D},\alpha} \frac{1}{2} \sum_i \|R_i \mathrm{x} - D\alpha_i\|_2^2 + \lambda \sum_i \|\alpha_i\|_0 \tag{6}$$

where λ is the regularization parameter related to sparsity. Researchers have exploited a lot of effective dictionary learning algorithms, such as orthogonal dictionary learning, convolutional dictionary learning and K-SVD dictionary learning. Here we chose the K-SVD mode in our work since it is regarded as a standard algorithm for dictionary learning and can be converged quickly, which is vital in real time reconstructions.

III. PROPOSED METHOD

A. ALGORITHMS FOR THE REFERENCE IMAGE

In real time dMRI reconstruction, a highly accurate reference image is vital, especially for the first few frames who are lack of adequate information for reconstruction. dMRI sequences contain much redundancy in temporal dimension, which implicates that once the first frame is captured and accurately reconstructed, the following images to be reconstructed will be very close to it. The first frame can be used as a reference frame for the whole sequence. The first frame needs to be acquired with more sample data, as is a common strategy in online methods. Many CS methods can be implemented on this single frame image efficiently, such as total variation, wavelet transform and dictionary learning.

Some online algorithms exploit both the spatial and temporal sparsity between the first frame and later frames, such as dynamic total variation (DTV) and 3D Dictionary Learning

(3D DL). The DTV method in [22] is formulated as the following Lagrange relaxed form:

$$\min_{z}\left\{\frac{1}{2}\left\|Az-y\right\|_2^2+\lambda\left\|z\right\|_{TV}\right\} \tag{7}$$

where $z = x_t - x_1$ is the residual image of the t^{th} frame, A is spatiotemporal sensing matrix. The authors use a reweighted least squares algorithm to reconstruct the dMRI sequence frame by frame in parallel.

DL methods can be easily applied to dMRI reconstruction by changing the patches into 3D blocks extracted from adjacent frames. A novel algorithm combining Dictionary Learning and Dynamic Total Variation (named as DLDTV) has been exploited in our prior work [23], who takes full advantage of the spatiotemporal sparsity of dMRI sequences, and can be implemented in parallel mode for fast reconstruction. DLDTV aims to solve the sparsity-based optimization problem expressed as follows:

$$\begin{aligned}&\min_{x,D,\alpha}\sum_i\left\|R_ix-D\alpha_i\right\|_2^2+\lambda_1\left\|F_ux-y\right\|_2^2+\lambda_2\left|\nabla_tx\right|_1\\&s.t.\ \forall i,\quad \left|\alpha_i\right|_0\le T_0\end{aligned} \tag{8}$$

where the first term corresponds to dictionary sparse representation, the second term is relevant to data fidelity, and the third term $\nabla_t x$ exploits temporal gradient sparsity on l_1 norm. λ_1 and λ_2 are penalty parameters.

All these online methods take full advantage of the prior information in the 1st frame and temporal sparsity between the first frame and later frames, but they cannot utilize the global sparsity of the whole dMRI sequence, which will lead to a decrease in the accuracy of reconstruction. It motivates us to integrate these algorithms into our method and exploit a new algorithm to utilize the global prior information in a parallel online scheme.

B. COMBINED FOURIER TRANSFORM

Current offline methods are always too complex and slow to implement their programs on massive data from the entire sequence, while online methods cannot always guarantee fairly high reconstruction accuracy due to the lack of entire prior information. How to keep data concise while maintaining the information rich seems to be a dilemma. Considering the data measurement is done in k-space, it motivates us to design a novel algorithm to aggregate all the k-space data into an artificial MR image. The artificial image should be highly similar to the current frame to be reconstructed, then this dilemma can be easily solved. This dynamic data aggregation algorithm is formulated as follows:

$$\begin{cases}\hat{x}_{pre_i}^{k+1}(kx,ky)=\begin{cases}\dfrac{\hat{x}_{pre_i}^{k}(kx,ky)+\nu\hat{x}_{zf_i}(kx,ky)}{1+\nu}, & (kx,ky)\in\Omega_i\\ \hat{x}_{pre_i}^{k}(kx,ky), & (kx,ky)\notin\Omega_i\end{cases}\\ x_{pre_i}^{k+1}(x,y)=\left|IFFT(\hat{x}_{pre_i}^{k+1}(kx,ky))\right|\end{cases} \tag{9}$$

where x_{pre_i} is a combined predictive image for the i^{th} frame, $\hat{x}_{pre_i}$ is its Fourier transform value in k-space, the superscript k denotes the number of iteration. $\hat{x}_{zf_i}$ is the i^{th} frame k-space acquisitions with zero-filled values at unsampled locations and Ω_i is the set of sampled locations. *IFFT* is the inverse fast Fourier transform. Taking sampling noise into account, the parameter ν acts as a counterweight to balance the predictive image and current undersampled zero-filled image at sampled locations. Parameter ν is a constant related to the standard variance ($\nu = q/\sigma$) of additive white Gauss noise, where q plays the key role under noisy sampling condition and will be discussed in experiments. Under noise free condition, $\sigma \to 0$ and $\nu \to \infty$, $\hat{x}_{pre_i} = \hat{x}_{zf_i}$, which means the sampled values $\hat{x}_{zf_i}$ should be fully trusted.

Equation (9) implements an alternating forward and backward Fourier transform on the predictive image and the current frame for data aggregation. Since the k-space data in each iteration is a combination of the current frame and the predictive image, we name it Combined Fourier Transform (CFT) algorithm. $x_{pre_3}^1 = x_2$, $x_{pre_i}^1 = x_{pre_{i-1}}^k$ (see Fig.1), where x_2 is the reconstructed 2^{nd} frame using online methods listed in sections 3.1. Algorithm 1 gives the flowchart of CFT algorithm. The function *PSNR* computes the peak signal to noise ratio of the reconstructed image, which is used as stopping criterion for each frame.

Algorithm 1 CFT

Input: x–ground-truth dMRI sequence, x_2–reconstructed 2^{nd} frame Ω–sampling mask, ε–stopping threshold
Output: x_{pre}—predictive images for the whole sequence
Initialization: set default iterations n, *seqlen*=size(x,3)

```
for i = 3 to seqqlen
    if i == 3
      x_temp(i, 1) = x_2;
    else
      x_temp(i, 1) = x_pre(i − 1);
    end
  for j = 1 to n
      implement forward CFT to get x̂_temp(i, j + 1);
      implement back CFT to get x_temp(i, j + 1);
      if (PSNR(x_temp(i, j + 1)) -PSNR(x_temp(i, j))) < ε
          x_pre(i) = x_temp(i, j + 1)
          break;
      end
end
```

To our knowledge, it is the first time such a CFT dynamic aggregation algorithm has been proposed for dMRI imaging. From Fig.1, one can see that the CFT is actually a data preprocessing operation independent of the CS methods, which means that it can be combined with any online CS schemes. A highly accurate predictive image can save the computing time tremendously, as is validated in our experiments.

C. DICTIONARY LEARNING

How to build an efficient over-complete dictionary D to exploit the optimal sparse representation of images is the key to reconstruction. Here we utilize an adaptive 3D dictionary learning method for fast parallel reconstruction. The dictionary learning extracts the overlapped patches from the original dMRI sequence, which best match the unique structure characteristics of the images themselves. The objective function of the reconstruction algorithm is as follows:

$$\min_{\mathrm{x},D,\alpha} \sum_i \|R_i\mathrm{x}-D\alpha_i\|_2^2+\lambda\,\|F_u\mathrm{x}-y\|_2^2 \;\; s.t.\;\forall i, \quad |\alpha_i|_0 \le T_0 \tag{10}$$

Equation (10) is a simplified vision of (8) without the temporal gradient constraint $\nabla_t\mathrm{x}$. It is because only the first frame and a predictive image are used in our parallel scheme, and these two images may not be sparse enough along the temporal gradient when the interval between them increases. Equation (10) is nonconvex and NP hard because the dictionary D is over-complete. Typical greedy algorithms, such as Orthogonal Matching Pursuit (OMP), Method of Optimal Directions (MOD), and K-Singular Value Decomposition (K-SVD) adopt two-step alternating optimization to update dictionary D and sparse coefficients α_i [19]. Here we use K-SVD to solve equation (10) for its simple structure, fast convergence speed and flexible match to other sparse coding methods such as OMP, MOD and MAP. Equation (10) can be decomposed into two-step alternating iterations:

(1) Dictionary training and sparse coding

In this step the x is fixed, D and α are varying. Equation (10) turns into solving the following subproblem:

$$\min_{D,\alpha} \sum_i \|R_i\mathrm{x}-D\alpha_i\|_2^2 \; s.t. \;\; \|d_i\|_2=1 \quad \forall i, \;\; |\alpha_i|_0 \le T_0 \tag{11}$$

where d_i is the overlapping sampling step length. The minimum $\|d_i\|_2 = 1$ is used in our method for the best learning effect. The dictionary training and sparse coding problem can be easily resolved by the standard K-SVD algorithm. K-SVD employs a two-stage iterative update to optimize this nonconvex problem.

In the first stage, a DCT dictionary is used as the initial dictionary. The dictionary is fixed and α_i is the only variable, which can be solved by the following equation:

$$\alpha_i^k = \arg\min_{\alpha_i} \sum_i \left\|R_i\mathrm{x}^{k-1}-D^{k-1}\alpha_i\right\|_2^2 \forall i, \quad |\alpha_i|_0 \le T_0 \tag{12}$$

where the superscript k denotes the iterations. This problem can be efficiently solved by standard OMP algorithms. Rubinstein *et al.* [34] explore a faster Batch-OMP method to accelerate the sparse coding programs.

In the second stage, the sparsity coefficients α_i are fixed and the dictionary is updated column by column using these patches whose sparse coefficients related to the current column (atom), then the problem can be formulated as a simple rank-1 approximation:

$$D^k = \arg\min_D \sum_i \left\|E_k - d_k\alpha_i^k\right\|_2^2 \; s.t.\;\forall i, \quad |\alpha_i|_0 \le T_0 \tag{13}$$

where $E_k = R_i\mathrm{x}^{k-1} - \sum_{j\neq k} d_j\alpha_i^k$ is the residual matrix without the k^{th} atom. The problem can be solved efficiently via k-SVD decomposition. In the proposed method, the dictionary is adaptively trained via K-SVD and coded via Batch-OMP.

(2) Image reconstruction

In this step, D and α are fixed, the image x is reconstructed by solving the following least squares problem:

$$\mathrm{x}^k = \arg\min_{\mathrm{x}} \sum_i \left\|R_i\mathrm{x} - D^k\alpha_i^k\right\|_2^2+\lambda\,\|F_u\mathrm{x}-y\|_2^2 \tag{14}$$

By taking the derivative of the only variable x in (14) and setting it equal to 0, we get the solution to x:

$$\mathrm{x}^k = \frac{\sum_i R_i^T D^k\alpha_i^k + \lambda F_u^H\mathrm{y}}{\sum_i R_i^T R_i + \lambda F_u^H F_u} \tag{15}$$

where F_u^H is the Hermitian transpose of F_u, R^T is the inverse matrix of R, $\sum R_i^T R_i = nI$ is an identity matrix and n is the patch size in vector dimension. In order to reduce the matrix operation complexity, the operation process is transformed from image domain to Fourier domain. Equation (15) can be rewritten as follows:

$$F\mathrm{x}^k = \frac{F\sum_i R_i^T D^k\alpha_i^k + \lambda FF_u^H\mathrm{y}}{F\sum_i R_i^T R_i F^H + \lambda FF_u^H F_u F^H} \tag{16}$$

where $F_uF_u^H$ are k-space acquisitions with zero-filled values in unsampled locations, which is denoted as $\hat{x}_{zf}$, $FF_u^HF_uF^H$ is a $p \times p$ diagonal matrix (p is the dimension of the whole MR images) with a 1 in the diagonal at the sampled locations or a 0 otherwise. Equation (16) can be simplified as:

$$\hat{x}_{recon_i}^k(kx,ky) = \begin{cases} \dfrac{\hat{x}_{DL}^k(kx,ky) + \frac{\lambda}{n}\hat{x}_{zf}(kx,ky)}{1+\frac{\lambda}{n}}, \\ \quad (kx,ky) \in \Omega_i \\ \hat{x}_{DL}^k(kx,ky), \\ \quad (kx,ky) \notin \Omega_i \end{cases} \tag{17}$$

where $\hat{x}_{recon_i}^k$ is the reconstrued i^{th} frame, $\hat{x}_{DL}^k$ is the k-space values reconstructed by dictionary learning. Here we define $\lambda/n = q/\sigma = \nu$ which has the same meaning as the parameter in (9). A uniform parameter setting could simplify the complexity of our algorithm.

The DL algorithm implements these two iterative steps on a dynamic subset denoted as $\mathrm{x}_s^{k+1} = [\mathrm{x}_1, \mathrm{x}_{DL}^k, \mathrm{x}_{pre_i}]$ for a faster and better reconstruction of i^{th} frame ($\mathrm{x}_s^1 = [\mathrm{x}_1, \mathrm{x}_{pre_i}]$), x_{DL}^k is the k^{th} reconstructed image of the i^{th} frame. Each subset only contains the first frame and the predictive image of the current frame; therefore, the proposed algorithm runs in parallel. On the other hand, a smaller subset and dictionary

size will speed up the reconstruction, which is vital in real time imaging.

The major steps of the proposed optimization algorithm are presented in algorithm 2. The proposed algorithm can be decomposed into three independent parts. Step 1 reconstructs the first two frames in high quality for reference to the later frames, where the optimal DLDTV algorithm is selected based on our previous work. Step 2 figures out a highly accurate predictive image for each frame by the CFT algorithm, which is independent of the DL algorithm. Step 3 utilize a parallel 3D DL scheme to reconstruct the dMRI sequence frame by frame.

Algorithm 2 DLCFT

Input: x–ground-truth dMRI sequence, Ω–sampling mask, ε–stopping threshold in each step, *i, j*–default iterations.

Output: reconstructed dMRI images x_{recon}

Initialization: k=1, $\mathrm{x}_{recon} = \mathrm{x}_{zf}$

Repeat:

Step1: reconstructing first two frames $[\mathrm{x}_1, \mathrm{x}_2]$ using Eq. (8);

Step2: computing x_{pre} for each frame using Eq. (9);

Step3: computing x_{recon} using Eqs. (11) and (15).

stopping criterion for i^{th} frame in step2 and step3:

$(\mathrm{PSNR}(\mathrm{x}^{k+1}_{pre_i}) - \mathrm{PSNR}(x^{k}_{pre_i})) < \varepsilon$,

$(\mathrm{PSNR}(\mathrm{x}^{k+1}_{recon_i}) - \mathrm{PSNR}(x^{k}_{recon_i})) < \varepsilon$.

IV. EXPERIMENTS AND COMPARISONS

A. EXPERIMENT SETUP

All experiments and numerical comparisons are evaluated using MATLAB on a laptop with 2.9GHz Intel core i7 3520M CPU and 12G DDR3 RAM. We compared our scheme with three online methods: TV and Dynamic Total Variation (DTV) [22], Dictionary Learning and Dynamic Total Variation (DLDTV) [23], and one offline method: k-t SLR [18]. We use the default settings of these methods from each author's homepage.

Two real and complex cardiac perfusion sequences (of 3D size 192 × 192 × 30) and a real functional brain sequence (of 3D size 128 × 128 × 40) are selected here to evaluate algorithm performances. The pixel values (real) or modulus (complex) in each sequence are normalized to (0,1) before reconstruction. We prefer the real sequences in our experiments if not specifically mentioned.

Pseudo radial sampling masks are used here to simulate the undersampled scheme. In each frame, the initial phase angles of radial lines are generated randomly to get more prior information from different frames, which is vital in our method. The first frame is set to a 50% sampling ratio if not specifically mentioned in all experiments.

Quantitative analyses are presented to compare the reconstructed performance among different models using three criteria: peak signal to noise ratio (PSNR), root mean squared error (RMSE), and structural similarity index (SSIM), which are formulated as:

$$PSNR(\hat{\mathrm{x}}) = 20\lg(1/RMSE(\hat{\mathrm{x}})) \tag{18}$$

$$RMSE(\hat{\mathrm{x}}) = \left\|\hat{\mathrm{x}} - \mathrm{x}\right\|_2^2 / N \tag{19}$$

$$SSIM(\hat{\mathrm{x}}) = \frac{(2\mu_{\hat{\mathrm{x}}}\mu_{\mathrm{x}} + c_1)(2\sigma_{\hat{\mathrm{x}}\mathrm{x}} + c_2)}{(\mu_{\hat{\mathrm{x}}}^2 + \mu_{\mathrm{x}}^2 + c_1)(\sigma_{\hat{\mathrm{x}}}^2 + \sigma_{\mathrm{x}}^2 + c_2)} \tag{20}$$

where $\hat{\mathrm{x}}$ denotes the reconstructed image and x is the ground truth image, μ and σ are the mean and standard deviation of a frame image, $\sigma_{\hat{x}x}$ is the covariance of $\hat{\mathrm{x}}$ and x. c_1 and c_2 are small constants used to avoid the denominators equal to zero. For the whole dMRI sequence, we simply take their mean values for performance evaluation.

All experiments are conducted in a noise-free environment if not specifically mentioned.

B. RECONSTRUCTION OF THE FIRST TWO FRAMES

Since the 1^{st} frame is used as the reference image for later frames, its reconstruction precision plays a key role in our method. Many online methods can be implemented here, such as TV, DTV [20], DL [22], and DLDTV [23]. All these methods can make full use of the spatiotemporal sparsity between the first two frames for better performance. We use their default settings and applicate them on the first two frames. While in the DLDTV method, only two frames are implemented in each iteration, we cannot use the default patch size [3, 3, 3]. Table 1 compares the reconstructed PSNR (dB) values vs. patch size in this method. One can see that [3, 3, 2] is the optimum selection for both frames.

TABLE 1. Reconstructed PSNR of the first two frames vs. patch size by DLDTV method.

Patch size	[2,2,2]	**[3,3,2]**	[4,4,2]	[5,5,2]
1^{st} frame	42.21	**42.50**	42.41	42.36
2^{nd} Frame	37.98	**38.21**	38.14	38.12

Table 2 compares the reconstructed PSNRs by different methods. It's obviously that DLDTV outperforms other algorithms on both frames, especially the 1^{st} frame, which is used as reference image for later frames. DLDTV is the best choice to reconstruct the first two frames in our method.

TABLE 2. Reconstructed PSNR of the first two frames by different methods on the perfusion sequence.

Algorithms	1^{st} frame (dB)	2^{nd} Frame (dB)
TV	39.65	33.48
DTV	41.21	38.06
DL	41.08	37.54
DLDTV	**42.50**	**38.21**

C. RECONSTRUCTION ACCURACY COMPARISONS

Once the first two frames have been reconstructed, the rest of these frames can be implemented in our DLCFT method. Many parameters should be set optimally. In the CFT step,

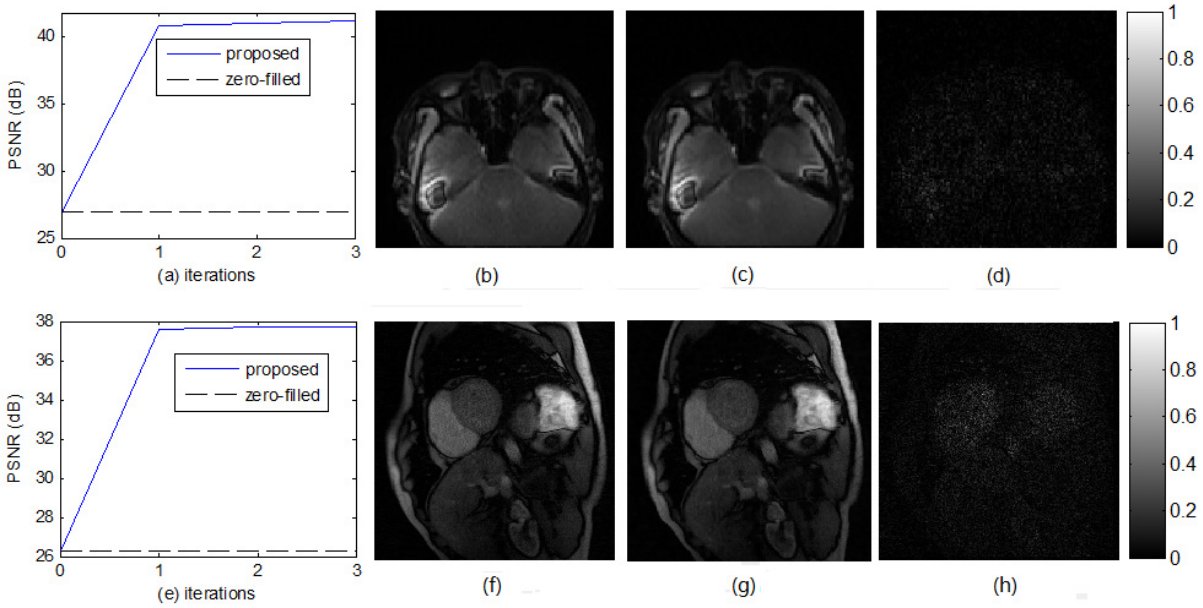

FIGURE 2. Reconstruction quality evaluated by PSNR and algorithm convergence on the brain sequence. (a) Reconstruction convergence. (b) The fully-sampled image. (c) Reconstructed image. (d) Errors × 5. (e), (f), (g), (h) are corresponding items on perfusion sequence.

TABLE 3. PSNR and computing time (in seconds) vs. patch size and sparse level T_0 on the perfusion sequence by proposed method.

Patch size	T_0=1	T_0=2	T_0=3	T_0=4
[2,2,2]	**37.61 (24.3s)**	37.59 (23.8s)	37.55 (23.4s)	37.54 (23.6s)
[3,3,2]	37.40 (26.4s)	37.51 (34.4s)	37.52 (33.0s)	37.51 (33.5s)
[4,4,2]	37.30 (50.0s)	37.43 (47.6s)	37.48 (47.9s)	37.49 (45.9s)

TABLE 4. Reconstruction accuracy comparisons of different methods on two sequences.

Sequence	Method	Mean PSNR	Mean RMSE	Mean SSIM
perfusion	TV	32.56	0.0238	0.8274
	DTV	35.83	0.0163	0.9036
	k-t SLR	36.02	0.0159	0.9065
	DLDTV	37.45	0.0135	0.9298
	DLCFT	**37.59**	**0.0132**	**0.9449**
brain	TV	32.83	0.0233	0.8919
	DTV	39.22	0.0112	0.9679
	k-t SLR	38.97	0.0133	0.9663
	DLDTV	40.88	0.0091	0.9734
	DLCFT	**41.61**	**0.0084**	**0.9761**

the maximum iteration is set 10 or reaching the stop criterion $(\text{PSNR}(i+1) - \text{PSNR}(i) < 0.002)$ to balance reconstruction accuracy and computing time. For later frames, the temporal gradient sparsity regularization in DLDTV cannot be used because the later frames and the 1st frame may not sparse enough in temporal dimension. Here we only use the DL algorithm and CFT predictive algorithm to reconstruct later frames, so the parameters of DL need to be reset properly.

Table 3 gives the reconstructed results vs. the patch size and sparse level T_0 by the proposed method. One can see that the optimal selection of the combined parameters is $[n_x, n_y, n_t] = [2, 2, 2]$ and $T_0 = 1$, weighing the accuracy and speed of the reconstruction.

Other parameters in DL step are set as follows: the selected patch numbers for online training n=5000, the overcompleteness of dictionary is set $k \geq 4$, which is common strategy in traditional DL method, so the dictionary size of the proposed method is fixed at 8×32 and only trained once in each DL iteration for fast reconstruction. Our method works well with these settings and the reconstructed values are robust to minor changes in these settings.

Fig. 2 shows the results and algorithm convergence by the proposed method at 15% sampling ratio on both sequences. The reconstructed PSNRs rise rapidly and tend to be stable within only three iterations, which indicates that our method converges quickly. Compared with the fully-sampled image, the reconstructed errors of both images are minimal even multiplied by 5, which indicates that the visual effects are consistent with numerical values.

The reconstructed results of a frame are presented in Figs. 3 and 4 for the two sequences respectively. DLDTV and our method provide higher quality images, especially around the ventricles and lesion sites in brain. while TV method lags behind a lot with obviously artifacts because it only reveals the spatial sparsity in both sequences. From the mean RMSEs (see Table 4) and the visual comparisons, our method is even better than the state-of-the-art offline methods.

Table 4 compares the reconstructed PSNRs, RMSEs and SSIMs by different methods. One can see that the proposed DLCFT leads others a lot in PSNR (0.73 dB), RMSE (0.0007) and SSIM (0.0027) on the cine sequence, which also validates the superiority of our algorithm.

Fig. 5 compares the reconstructed RMSEs frame by frame on both sequences. From the results, one can see that the

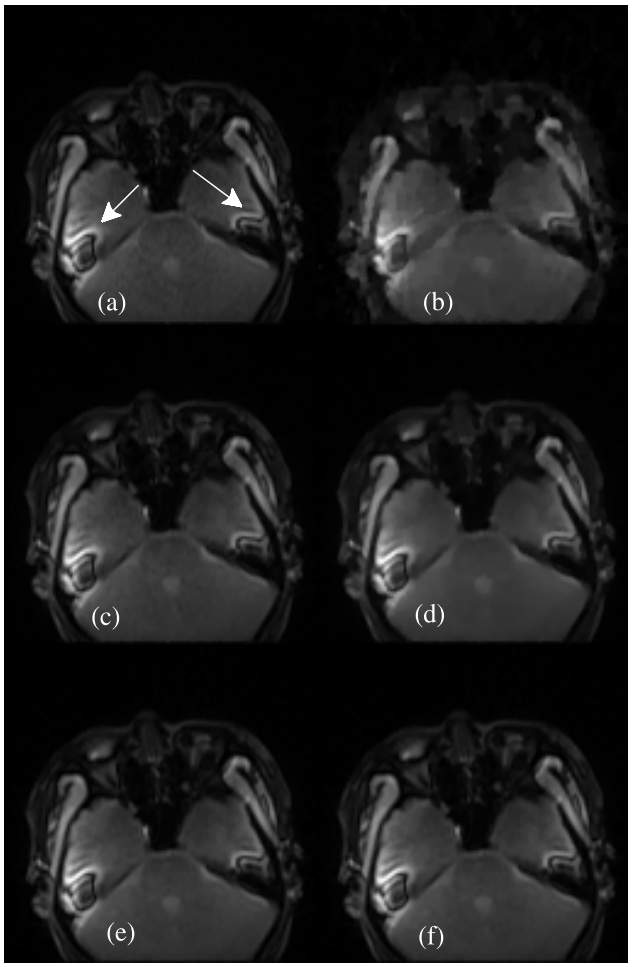

FIGURE 3. Results of a frame on brain sequence. (a) The fully-sampled image. The rest images are reconstructed by (b) TV; (c)DTV; (d) k-t SLR; (e) DLDTV; (f) proposed method.

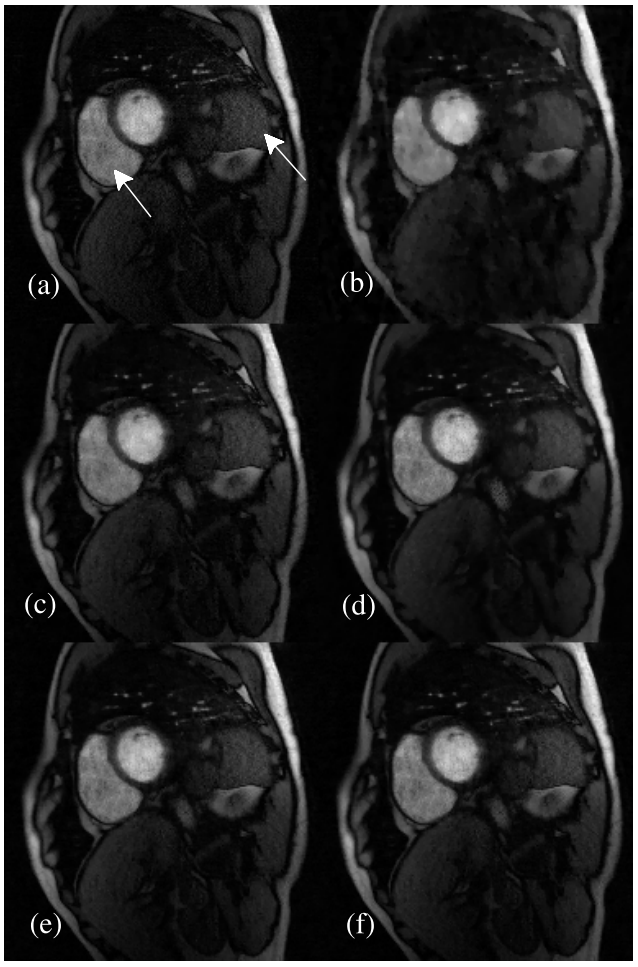

FIGURE 4. Results of a frame on perfusion sequence. (a) The fully-sampled image. The rest images are reconstructed by (b) TV; (c)DTV; (d) k-t SLR; (e) DLDTV; (f) proposed method.

proposed method gets the smallest reconstruction error overall. However, DLCFT lags slightly behind our previous work DLTV in the first few frames but still superior to other online methods. It is because the CFT algorithm can only use limited k-space measurements in the first few frames and the DL algorithm is executed frame by frame in parallel in the proposed method, while our previous work DLDTV can use the adjacent 3 or 4 frames as a subset (with more measurements) for each parallel reconstruction at the cost of computing time. Detailed comparisons of computing time will be discussed in the following subsection F. As the number of frames increases and more k-space measurements are gathered, the superiority of our method becomes obvious (see the later frames in Fig. 5), especially compared to those who only use inter-frame prediction and tend to be hampered by error accumulation.

Keep the sampling ratio of the first frame unchanged, we switch the sampling ratios of the rest frames from 15% to 40%. Fig. 6 compares the mean reconstructed errors of the whole sequence by different methods. TV still performs worst at any sampling ratio. DTV and k-t SLR perform moderate and get very close values. Our method outperforms other methods at low sampling ratios but very close to DLDTV at high sampling ratios. Table 5 gives the numerical comparisons of mean RMSEs vs. sampling ratios by different methods. The proposed method leads quite a lot at low sampling ratios but slightly inferior to DLDTV at 40% sampling on the perfusion sequence, which is consistent with Fig. 6. However, a lower sampling rate is necessary in rapid dMRI, that is where our algorithm leads.

TABLE 5. Numerical comparisons of reconstructed mean RMSEs vs. sampling ratios (15%-40%) by different methods.

	Method	15%	20%	25%	30%	40%
Perfusion	TV	0.0238	0.0194	0.0168	0.0149	0.0120
	DTV	0.0163	0.0146	0.0134	0.0123	0.0103
	k-t SLR	0.0159	0.0140	0.0128	0.0125	0.0105
	DLDTV	0.0135	0.0119	0.0106	0.0095	**0.0075**
	proposed	**0.0132**	**0.0117**	**0.0105**	**0.0095**	0.0077
Brain	TV	0.0233	0.0175	0.0126	0.0099	0.0062
	DTV	0.0112	0.0095	0.0078	0.0068	0.0051
	k-t SLR	0.0113	0.0097	0.0083	0.0073	0.0062
	DLDTV	0.0091	0.0079	0.0067	0.0057	0.0044
	proposed	**0.0084**	**0.0074**	**0.0064**	**0.0056**	**0.0043**

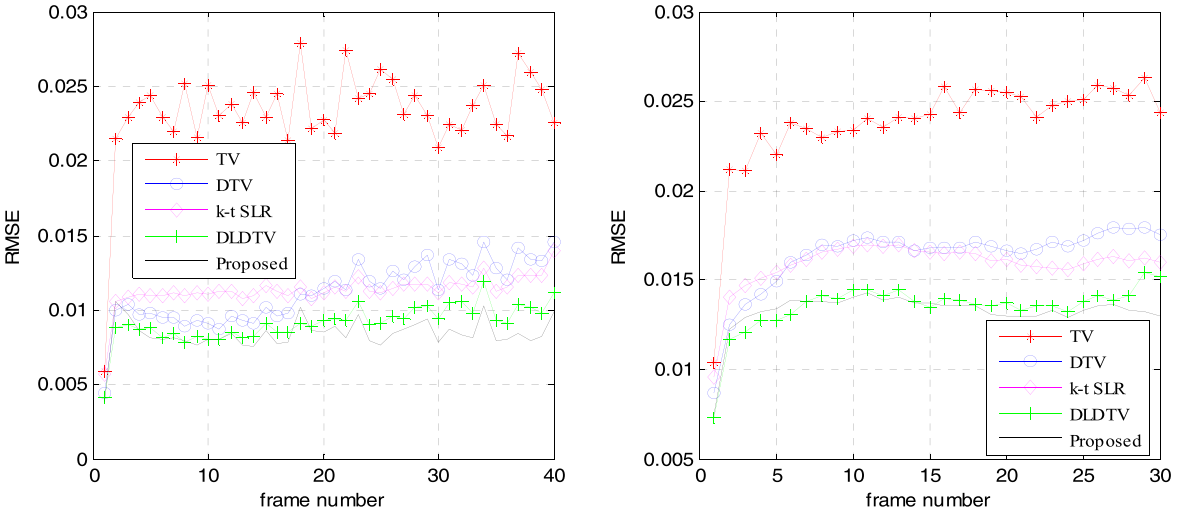

FIGURE 5. Comparisons of reconstructed RMSE. (left) On brain sequence. (right) On perfusion sequence.

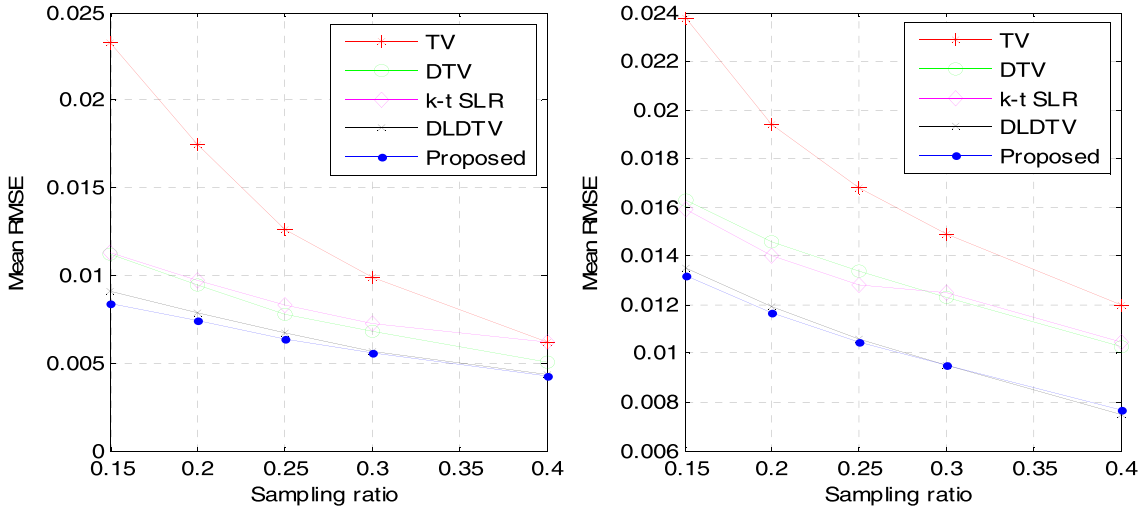

FIGURE 6. Comparisons of mean RMSEs vs. sampling ratios. (left) On brain sequence. (right) On perfusion sequence. The sampling ratio of the first frame is fixed at 50% on both datasets.

As a typical online method, only two frames are conducted in each iteration, our method relies heavily on the reconstruction quality of the first frame, which is proportional to the k-space sampling rate on it. Fig. 7 shows the reconstructed RMSEs vs. this parameter. One can see that the RMSEs of the first few frames vary dramatically when the sampling ratios of the 1^{st} frame changes. It is because less k-space measurements are available for the first few frames and they are more dependent on the first frame for reference. For the later frames, these gaps are subtle as more measurements are aggregated into the predictive image by the CFT algorithm who plays the key role in reconstruction quality. A fully-sampled 1^{st} frame is the optimal choice if allowed.

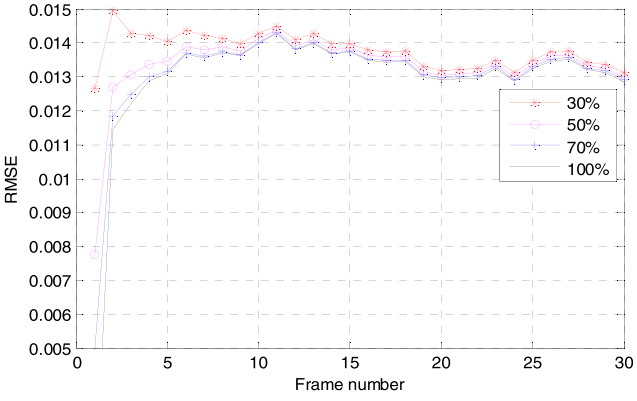

FIGURE 7. Comparisons of RMSEs vs. sampling ratios of the first frame by proposed method on perfusion sequence. The sampling ratio of the remaining frames is 15%.

D. COMPLEX DMRI RECONSTRUCTION

Since the real MR images are often complex image domain, we use a complex perfusion sequence of size $192 \times 192 \times 30$ to validate the efficiency of the proposed method. Some parameters should be adjusted accordingly. For the first two frames, we still select the optimal DLDTV algorithm. In the 1^{st} step, the real and imaginary parts of complex images reconstructed by KSVD and batch-OMP independently. In the 2^{nd} step,

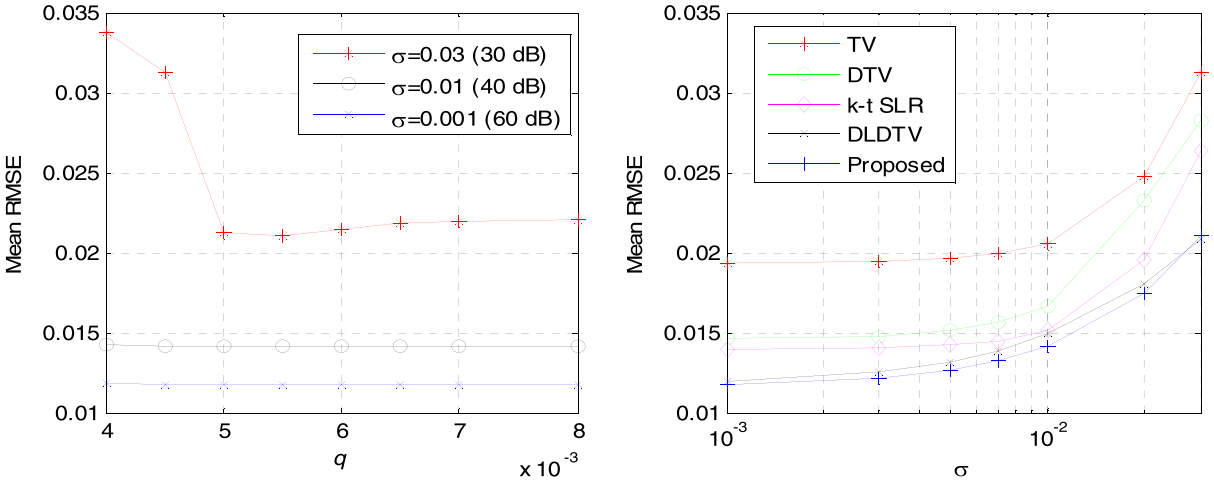

FIGURE 8. Noise robustness comparison on perfusion sequence at 20% sampling ratio, the AWGN noise level is evaluated by its standard deviation σ in k-space or PSNR (dB) in image domain. (left) Reconstructed RMSE vs. q by proposed method. (right) robustness vs. σ by different methods.

the modulus of images instead of their complex pixel values are used in DTV subprogram. For more details of this strategy we refer to [21]. In the CFT subprogram used for later frames, the predictive image for each frame should also be complex. Equation (9) can be rewritten as:

$$\begin{cases} \hat{x}_{pre_i}(kx, ky) = \begin{cases} \dfrac{\hat{x}_{pre_i}(kx, ky) + \nu \hat{x}_{zf_i}(kx, ky)}{1+\nu}, \\ \quad (kx, ky) \in \Omega_i \\ \hat{x}_{pre_i}(kx, ky), \\ \quad (kx, ky) \notin \Omega_i \end{cases} \\ \mathrm{x}_{pre_i}(x, y) = IFFT(\hat{x}_{pre_i}(kx, ky)) \end{cases} \tag{21}$$

where the combined forward and backward Fourier transform is executed only once in complex mode (using complex x_{pre_i} instead of its modulus), and the default $\mathrm{x}_{pre_3} = \mathrm{x}_2$, $\mathrm{x}_{pre_i} = \mathrm{x}_{pre_{i-1}}$.

Table 6 compares the results on the complex perfusion sequence by different methods. One can see that although DLDTV has a narrow lead over the proposed method, it runs slowly and takes more than 400 seconds, which means it is not fit for real time complex dynamic MR imaging. The CFT algorithm in the proposed method give a highly accurate prediction to each frame, which greatly accelerate convergence. The proposed method is optimal considering both reconstruction speed and performance.

TABLE 6. Reconstruction comparisons by different methods on a complex perfusion sequence.

Algorithms	Mean PSNR	Mean RMSE	Mean SSIM	Computing time (s)
TV	30.35	0.0308	0.7605	**32.4**
DTV	34.56	0.0188	0.8767	34.7
k-t SLR	35.49	0.0169	0.8834	250.9
DLDTV	**36.01**	**0.0159**	**0.9001**	417.4
proposed	35.78	0.0162	0.8999	50.7

E. NOISE ROBUSTNESS COMPARISONS

If the data acquired in k-space is polluted by the noise, it will greatly affect the accuracy of reconstruction. Assume a noisy dMRI measurements obtained in k-space described as $\hat{x}_n = F_u \mathrm{x} + n$, where n is the additive white Gaussian noise of power σ^2 added in k-space, which can also be measured by PSNR (dB) in image domain. The constant q in both CFT and DL algorithms plays an important role to reconstruction quality in noisy environments.

Fig. 8 (left) reveals mean RMSEs vs. q in various noise environments. It's easy to find that the mean RMSE is insensitive to q in a medium-noise or low-noise environment (e.g., $\sigma = 0.001$, PSNR=60dB). While in a high-noise environment (e.g., $\sigma = 0.03$, PSNR=30dB), the results depend heavily on it. Here we fixed $q = 5.5 \times 10^{-3}$ to minimize the reconstructed errors at different noise levels.

Fig. 8 (right) compares robustness to added noise calibrated by mean RMSEs among different methods. The perfusion sequence of 20% sampling ratio is used here to testify this performance. One can see that only slight deterioration of RMSEs occurs in a low-noise case (40-60dB), which is easily encountered in actual situation. While in high-noise case (30dB), only the proposed method and DLDTV can maintain relatively low error. In general, the proposed method has the best noise robustness.

F. ALGORITHM SPEED COMPARISONS

Algorithm speed is vital for real time imaging. The computational complexity of the proposed method mainly depends on the dictionary learning process. The process of dictionary learning includes dictionary training and sparse coding and the main calculation work occurs in the sparse coding step. An efficient batch based OMP algorithm explored by Rubinstein is implemented in our method for sparse coding, who is 5-7 times faster than standard OMP algorithm. On the other hand, Fig. 2 points out that the proposed method converges quickly and it is also acceptable to run the DL iteration only

once. Table 7 compares the results using DLCFT and its two simplified versions (using the DL iteration only once for each frame's reconstruction, which is named as simplified DLCFT, or just CFT itself without the DL iteration, which is named as CFT) on the perfusion sequence. One can see that the simplified version is still comparable to DLCFT while its speed is much faster. CFT without DL is obviously the fastest while the reconstructed values lags behind a lot. Table 7 demonstrates that the CFT algorithm pays the key role in our method. The proposed method can be implemented flexibly in different modes to meet different needs.

TABLE 7. Comparisons of proposed method and its simplified versions.

Algorithms	Mean PSNR	Mean RMSE	Mean SSIM	Computing time (s)
DLCFT	**37.59**	**0.0132**	**0.9449**	24.3
Simplified DLCFT	37.47	0.0135	0.9344	9.8
CFT	37.03	0.0142	0.9259	**5.7**

Table 8 compares the computing time of different method on both sequences. The average reconstruction speed of our method is 0.81 (24.3/30) seconds per frame on the perfusion sequence and 0.67 seconds per frame on the brain sequence. The simplified DLCFT who uses the DL algorithm only once for each frame lead a lot in reconstruction speed. The speed of our method is calculated in MATLAB implementation, which can be further accelerated by GPU parallel computing algorithms because our method is implemented in parallel.

TABLE 8. Computing time (in seconds) of different methods on both sequences.

Algorithms	perfusion	brain
TV	26.8	16.0
DTV	26.6	16.0
k-t SLR	294.8	220.2
DLDTV	113.2	106.4
DLCFT	24.3	26.8
simplified DLCFT	**9.8**	**8.1**

In our method, once a frame of measurements is obtained, the predictive image can be calculated immediately since the CFT is a fast-linear transformation. Then the DL algorithm can reconstruct the current frame immediately only using the 1^{st} frame and the predictive image. In short, our method can be called a real-time method.

V. CONCLUSION

we have proposed a novel reconstruction algorithm (DLCFT) for real time dynamic MRI, which uses a combined Fourier transform to dynamically aggregate k-space measurements and creates a highly accurate predictive image for reference, was demonstrated to successfully reconstruct three real and complex dynamic MR sequences with different accelerators. It provided images with superior spatiotemporal resolutions compared to four state-of-art methods. The source of improvement is the dynamic k-space data aggregation mode who uses the whole sampled frames, instead of conventional models who only use the inter-frame sparsity with respect to a single reference frame. Our contribution provides a fast and flexible method to implement CS reconstruction on real time dynamic MRI. The leading superiority of our method is not obvious in reconstruction quality and speed while using the complex dMRI sequences. It is because the DL algorithm converge slowly in complex mode without other sparse regularizations. In the future, we plan to combine our method with other efficient sparse regularizations and exploit a more accurate prediction algorithm from both image domain and Fourier domain to accelerating the algorithm further. The design of CFT scheme can also be extended to other CS dynamic MR imaging applications. Further accelerating the running time using high-performance computing systems such as GPU parallel computing is another future direction to explore.

REFERENCES

[1] 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. Med. Imag.*, vol. 31, no. 9, pp. 1809–1820, Sep. 2012.
[2] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.
[3] J. Huang, S. Zhang, and D. Metaxas, "Efficient MR image reconstruction for compressed MR imaging," *Med. Image Anal.*, vol. 15, no. 5, pp. 670–679, Oct. 2011.
[4] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic MRI," *Magn. Reson. Med.*, vol. 59, no. 2, pp. 365–373, Feb. 2008.
[5] Y. Zhang, Z. Dong, P. Phillips, S. Wang, G. Ji, and J. Yang, "Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging," *Inf. Sci.*, vol. 322, pp. 115–132, Nov. 2015.
[6] Q. Liu, S. Wang, L. Ying, X. Peng, Y. Zhu, and D. Liang, "Adaptive dictionary learning in sparse gradient domain for image recovery," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4652–4663, Dec. 2013.
[7] D. Guo and X. Qu, "Improved reconstruction of low intensity magnetic resonance spectroscopy with weighted low rank Hankel matrix completion," *IEEE Access*, vol. 6, pp. 4933–4940, 2018.
[8] Z. Zhan, J.-F. Cai, D. Guo, Y. Liu, Z. Chen, and X. Qu, "Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1850–1861, Sep. 2016.
[9] N. He, R. Wang, and Y. Wang, "Dynamic MRI reconstruction exploiting blind compressed sensing combined transform learning regularization," *Neurocomputing*, to be published. doi: 10.1016/j.neucom.2018.12.087.
[10] Y. Liu, S. Wu, X. Huang, B. Chen, and C. Zhu, "Hybrid CS-DMRI: Periodic time-variant subsampling and omnidirectional total variation based reconstruction," *IEEE Trans. Med. Imag.*, vol. 36, no. 10, pp. 2148–2159, Oct. 2017.
[11] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[12] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[13] Q. Lin, Z. Liang, C. Duan, J. Ma, H. Li, C. Roque, J. Yang, G. Zhang, H. Lu, and X. He, "Motion correction for MR cystography by an image processing approach," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 9, pp. 2401–2410, Sep. 2013.
[14] B. B. Mehta, D. Ma, E. Y. Pierre, Y. Jiang, S. Coppo, and M. A. Griswold, "Image reconstruction algorithm for motion insensitive MR fingerprinting (MRF): MORF," *Magn. Reson. Med.*, vol. 80, no. 6, pp. 2485–2500, Dec. 2018.
[15] M. S. Asif, L. Hamilton, M. Brummer, and J. Romberg, "Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI," *Magn. Reson. Med.*, vol. 70, no. 3, pp. 800–812, 2013.

[16] S. Xiao, H. Deng, C. Duan, J. Xie, H. Zhang, X. Sun, C. Ye, and X. Zhou, "Considering low-rank, sparse and gas-inflow effects constraints for accelerated pulmonary dynamic hyperpolarized ^{129}Xe MRI," *J. Magn. Reson.*, vol. 290, pp. 29–37, May 2018.
[17] J. Yao, Z. Xu, X. Huang, and J. Huang, "An efficient algorithm for dynamic MRI using low-rank and total variation regularizations," *Med. Image Anal.*, vol. 44, pp. 14–27, Feb. 2018.
[18] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: K-t SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.
[19] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[20] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, May 2011.
[21] J. Caballero, A. N. Price, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity for dynamic MR data reconstruction," *IEEE Trans. Med. Imag.*, vol. 33, no. 4, pp. 979–994, Apr. 2014.
[22] C. Chen, Y. Li, L. Axel, and J. Huang, "Real time dynamic MRI by exploiting spatial and temporal sparsity," *Magn. Reson. Imag.*, vol. 34, no. 4, pp. 473–482, 2016.
[23] Y. Wang, N. Cao, Z. Liu, and Y. Zhang, "Real-time dynamic MRI using parallel dictionary learning and dynamic total variation," *Neurocomputing*, vol. 238, pp. 410–419, May 2017.
[24] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.
[25] M. Usman, C. Prieto, F. Odille, D. Atkinson, T. Schaeffter, and P. G. Batchelor, "A computationally efficient OMP-based compressed sensing reconstruction for dynamic MRI," *Phys. Med. Biol.*, vol. 56, no. 7, pp. N99–N144, Apr. 2011.
[26] M. Niu, S. Salari, I.-M. Kim, F. Chan, and S. Rajan, "Recovery probability analysis for sparse signals via OMP," *IEEE Trans. Aerosp. Elect. Syst.*, vol. 51, no. 4, pp. 3475–3479, Oct. 2015.
[27] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[28] T. Bai, H. Yan, X. Jia, S. Jiang, G. Wang, and X. Mou, "Z-index parameterization for volumetric CT image reconstruction via 3-D dictionary learning," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2466–2478, Dec. 2017.
[29] F. Liu, X. Gong, L. V. Wang, J. Guan, L. Song, and J. Meng, "Dictionary learning sparse-sampling reconstruction method for *in-vivo* 3D photoacoustic computed tomography," *Biomed. Opt. Express*, vol. 10, no. 4, pp. 1660–1777, Apr. 2019.
[30] C. Jiang, Q. Zhang, R. Fan, and Z. Hu, "Super-resolution CT image reconstruction based on dictionary learning and sparse representation," *Sci. Rep.*, vol. 8, Jun. 2018, Art. no. 8799. [Online]. Available: https://www.nature.com/articles/s41598-018-27261-z.pdf
[31] S. Chen, H. Liu, P. Shi, and Y. Chen, "Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography," *Phys. Med. Biol.*, vol. 60, no. 2, pp. 807–823, Jan. 2015.
[32] Y. Wang and L. Ying, "Compressed sensing dynamic cardiac cine MRI using learned spatiotemporal dictionary," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 4, pp. 1109–1120, Apr. 2014.
[33] T. Nguyen-Duc and W.-K. Jeong, "Compressed sensing dynamic MRI reconstruction using multi-scale 3D convolutional sparse coding with elastic net regularization," in *Proc. 15th Int. Symp. Biomed. Imag. (ISBI)*, Washington, DC, USA, Apr. 2018, pp. 332–335.
[34] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," CS Technion., Tech. Rep. CS-2008-08, 2008. [Online]. Available: https://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-info.cgi/2008/CS/CS-2008-08

YANG WANG received the B.S. degree in electronic engineering from the Zhejiang University of Technology, China, in 2002, and the M.S. degree in signal and information processing from Shanghai University, China, in 2005. He is currently pursuing the Ph.D. degree with Hohai University, China. He is also a Lecturer with the Department of Electric Engineering, Taizhou University, China. His research interests focus on computer-aided diagnosis and biomedical image processing.

NING CAO received the B.S. and M.S. degrees in signal and information processing from Southeast University, China, in 1980 and 1987, respectively. He is currently a Professor and the Director of the Department of Information and Communication Engineering, Hohai University, China. His major research interests include signal estimation and detection, cognitive radio, image processing, and transmission.

YUELI CUI received the B.S. degree in electronics and communication engineering from Zhejiang University City College, China, in 2006, and the M.S. degree in electronics and communication engineering from Hebei University, China, in 2009. He is currently pursuing the Ph.D. degree with Ningbo University. He is also a Lecturer with the Department of Electronics and Information Engineering, Taizhou University, China. His research interests include image processing, pattern recognition, and information safety.

QIANG ZHOU received the B.S. degree in electronic information engineering form Zhejiang Gongshang University, China, in 2005, and the M.S. degree in electronic and communication engineering form the Hangzhou University of Electronic Science and Technology. He is currently an Experimenter with the Department of Electric Engineering, Taizhou University, China. His major research interests include sensor application and detection technology.