

Appendix A: CDLMRI

We provide more details in relation to Section 2.2 CDLMRI in this appendix.

Stage 1) Coupled Dictionary Learning (details). The dictionary update step is to solve the optimization problem:

$$\begin{aligned} & \underset{\Psi_c, \Psi, \Phi_c, \Phi}{\text{minimize}} \sum_{ij} \left\{ \|\mathbf{R}_{ij} \mathbf{x}^{(1)} - (\Psi_c \mathbf{z}_{ij} + \Psi \mathbf{u}_{ij})\|_2^2 \right. \\ & \quad \left. + \|\mathbf{R}_{ij} \mathbf{x}^{(2)} - (\Phi_c \mathbf{z}_{ij} + \Phi \mathbf{v}_{ij})\|_2^2 \right\} \\ & \text{subject to} \quad \left\| \begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix} \right\|_2^2 \leq 1, \|\psi_k\|_2^2 \leq 1, \|\phi_k\|_2^2 \leq 1, \forall k \end{aligned} \quad (9)$$

Given a subset of the patches to constitute the training dataset $\mathbf{X}^{(1)} = [\dots, \mathbf{x}_{ij}^{(1)}, \dots]$ and $\mathbf{X}^{(2)} = [\dots, \mathbf{x}_{ij}^{(2)}, \dots]$ in Stage 1), the optimization problem (9) is equivalent to:

$$\begin{aligned} & \underset{\Psi_c, \Psi, \Phi_c, \Phi}{\text{minimize}} \left\| \begin{bmatrix} \mathbf{X}^{(1)} - \Psi \mathbf{U} \\ \mathbf{X}^{(2)} - \Phi \mathbf{V} \end{bmatrix} - \begin{bmatrix} \Psi_c \\ \Phi_c \end{bmatrix} \mathbf{Z} \right\|_F^2 \\ & \text{subject to} \quad \left\| \begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix} \right\|_2^2 \leq 1, \|\psi_k\|_2^2 \leq 1, \|\phi_k\|_2^2 \leq 1, \forall k \end{aligned}$$

where, $\mathbf{Z} = [\dots, \mathbf{z}_{ij}^{(1)}, \dots]$, $\mathbf{U} = [\dots, \mathbf{u}_{ij}^{(1)}, \dots]$, $\mathbf{V} = [\dots, \mathbf{v}_{ij}^{(1)}, \dots]$. Taking the dictionary update of Ψ_c and Φ_c for example, we update the atom pairs one by one. For the k -th atom pair ψ_c and ϕ_c , we can immediately establish that

$$\begin{aligned} \mathbf{d}_k \leftarrow \min_{\mathbf{d}} \left\| \begin{bmatrix} \mathbf{X}^{(1)} - \Psi \mathbf{U} \\ \mathbf{X}^{(2)} - \Phi \mathbf{V} \end{bmatrix} - \begin{bmatrix} \Psi_c \\ \Phi_c \end{bmatrix} \mathbf{Z} + \begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix} \mathbf{z}^k - \mathbf{d} \mathbf{z}^k \right\|_F^2 \\ \text{s.t.} \quad \|\mathbf{d}\|_2^2 \leq 1, \end{aligned}$$

By expanding the Frobenius norm and removing the constant term, it turns out that the above problem is equivalent to the optimization problem

$$\begin{aligned} \min_{\mathbf{d}} \quad \frac{1}{2} \|\mathbf{d} \mathbf{z}^k\|_F^2 - \mathbf{d}^T \left(\begin{bmatrix} \mathbf{X}^{(1)} - \Psi \mathbf{U} \\ \mathbf{X}^{(2)} - \Phi \mathbf{V} \end{bmatrix} - \begin{bmatrix} \Psi_c \\ \Phi_c \end{bmatrix} \mathbf{Z} + \begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix} \mathbf{z}^k \right) \mathbf{z}^{kT} \\ \text{s.t.} \quad \|\mathbf{d}\|_2^2 \leq 1, \end{aligned}$$

where \mathbf{z}^k denotes the k -th row of \mathbf{Z} .¹ We compute the derivative of the objective w.r.t. \mathbf{d} , leading to a norm equation:

$$\mathbf{d}_k \leftarrow \frac{1}{\mathbf{z}^k \mathbf{z}^{kT}} \left(\begin{bmatrix} \mathbf{X}^{(1)} - \Psi \mathbf{U} \\ \mathbf{X}^{(2)} - \Phi \mathbf{V} \end{bmatrix} - \begin{bmatrix} \Psi_c \\ \Phi_c \end{bmatrix} \mathbf{Z} \right) \mathbf{z}^{kT} + \begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix}$$

Then, we apply the ℓ_2 norm constraint.

$$\begin{bmatrix} \psi_{ck} \\ \phi_{ck} \end{bmatrix} = \mathbf{d}_k \leftarrow \frac{\mathbf{d}_k}{\max(\|\mathbf{d}_k\|_2, 1)}$$

The dictionary update of Ψ and Φ is performed in a similar way. In order to accelerate the training, the proposed algorithm can be updated to online training version without difficulty.

Stage 3) k -space Consistency Enforcing (details). In this stage, we aim to enforce consistency between the denoised image and its measurements in the k -space domain. In particular, given the estimated

¹Note that \mathbf{z}^k is a row vector resulting from the derivative w.r.t the k -th atom pair, while \mathbf{z}_{ij} is a column vector corresponding to the ij -th patch pair.

patches $\hat{\mathbf{x}}_{ij}^{(1)}$ from Stage 2), this step is formulated as a least square problem:

$$\min_{\mathbf{x}^{(1)}} \sum_{ij} \left\| \mathbf{R}_{ij} \mathbf{x}^{(1)} - \hat{\mathbf{x}}_{ij}^{(1)} \right\|_2^2 + \nu_1 \left\| \mathbf{F}_{u1} \mathbf{x}^{(1)} - \mathbf{y}^{(1)} \right\|_2^2, \quad (10)$$

which admits an analytical solution satisfying the normal equation

$$\left(\sum_{ij} \mathbf{R}_{ij}^H \mathbf{R}_{ij} + \nu_1 \mathbf{F}_{u1}^H \mathbf{F}_{u1} \right) \mathbf{x}^{(1)} = \sum_{ij} \mathbf{R}_{ij}^H \hat{\mathbf{x}}_{ij}^{(1)} + \nu_1 \mathbf{F}_{u1}^H \mathbf{y}^{(1)}, \quad (11)$$

where the superscript $()^H$ denotes the Hermitian transpose operation. The term $\sum_{ij} \mathbf{R}_{ij}^H \mathbf{R}_{ij} \in \mathbb{C}^{N \times N}$ is a diagonal matrix where each diagonal entry is the number of overlapping patches at the corresponding pixel location in $\mathbf{x}^{(1)}$. Assuming that patches wrap around at image boundaries, the number of overlapping patches at each pixel is the same, denoted by β .² Thus, the term $\frac{1}{\beta} \sum_{ij} \mathbf{R}_{ij}^H \hat{\mathbf{x}}_{ij}^{(1)}$ represents the denoised image $\hat{\mathbf{x}}^{(1)}$, where the intensity value of each pixel is the average of all the overlapping patches that cover this pixel. Multiplying by the normalized full Fourier transform matrix \mathbf{F} on the both sides of equation (11) leads to

$$\begin{aligned} & \left(\mathbf{F} \sum_{ij} \mathbf{R}_{ij}^H \mathbf{R}_{ij} \mathbf{F}^H + \nu_1 \mathbf{F} \mathbf{F}_{u1}^H \mathbf{F}_{u1} \mathbf{F}^H \right) \mathbf{F} \mathbf{x}^{(1)} \\ & = \mathbf{F} \sum_{ij} \mathbf{R}_{ij}^H \hat{\mathbf{x}}_{ij}^{(1)} + \nu_1 \mathbf{F} \mathbf{F}_{u1}^H \mathbf{y}^{(1)}. \end{aligned} \quad (12)$$

The matrix $\mathbf{F} \mathbf{F}_{u1}^H \mathbf{F}_{u1} \mathbf{F}^H$ is a diagonal matrix consisting of ones (corresponding to sampling locations in k -space) and zeros. Under the "wrap around" assumption, $\mathbf{F} \sum_{ij} \mathbf{R}_{ij}^H \mathbf{R}_{ij} \mathbf{F}^H = \beta \mathbf{I}_P$. Thus, the matrix pre-multiplying $\mathbf{F} \mathbf{x}^{(1)}$ in (12) is diagonal and trivially invertible. The vector $\mathbf{F} \mathbf{F}_{u1}^H \mathbf{y}^{(1)}$ represents the zero-filled Fourier measurements. Dividing both sides of (12) by the constant β to obtain

$$\tilde{\mathbf{y}}_{pq}^{(1)} = \begin{cases} (\mathbf{F} \hat{\mathbf{x}}^{(1)})_{pq}, & (p, q) \notin \Omega^{(1)} \\ \frac{1}{1 + \tilde{\nu}_1} \left(\mathbf{F} \hat{\mathbf{x}}^{(1)} + \tilde{\nu}_1 \mathbf{F} \mathbf{F}_{u1}^H \mathbf{y}^{(1)} \right)_{pq}, & (p, q) \in \Omega^{(1)} \end{cases}$$

where $\tilde{\nu}_1 = \nu_1 / \beta$, $\hat{\mathbf{x}}^{(1)} = \frac{1}{\beta} \sum_{ij} \mathbf{R}_{ij}^H \hat{\mathbf{x}}_{ij}^{(1)}$ denotes the denoised image. We denote by $\Omega^{(1)}$ the subset of k -space that has been sampled and by $\tilde{\mathbf{y}}_{pq}^{(1)}$ the updated value at location (p, q) in the k -space. Note that (8) uses the dictionaries that were learned in Stage 1) to interpolate the non-sampled Fourier frequencies, and update the sampled frequencies. Then, we immediately obtain the solution:

$$\hat{\mathbf{x}}^{(1)} = \mathbf{F}^H \tilde{\mathbf{y}}^{(1)} \quad (13)$$

where \mathbf{F}^H denotes the conjugate Fourier transform matrix. $\tilde{\mathbf{y}}^{(1)}$ denotes the estimated k -space samples as in (8). In other words, the estimation $\hat{\mathbf{x}}^{(1)}$ is obtained by inverse DFT of $\tilde{\mathbf{y}}^{(1)}$. Then the process returns to the Stage 1). The whole process is shown in Algorithm 1.

²In particular, $\beta = n$ when the overlap stride $r = 1$, where the *overlap stride* is defined as the distance in pixels between corresponding pixel locations in adjacent image patches.

Appendix B: More Experiments

Tissue can be characterized by two different relaxation times – T1 (longitudinal relaxation time) and T2 (transverse relaxation time).³ T1-weighted and T2-weighted pair of MRI scans are two basic types of multi-contrast data, where the former is produced by using short TE and TR times and conversely the latter is produced by using longer TE (Time to Echo) and TR (Repetition Time) times. In general, T1-weighted MRI images results in highlighted/bright fat tissue, such as subcutaneous fat (SC fat) and bone marrow, and suppressed/dark water-based tissue, such as Cerebrospinal fluid (CSF). In contrast, T2-weighted MRI images highlight both fat tissue and water-based tissue. Therefore, the correlation of T1-weighted and T2-weighted is complex, instead of simple reverse mapping relationship.

In this experiment, we use under-sampled T1-weighted MRI as the target contrast and corresponding fully-sampled T2-weighted as the guidance contrast to replicate the same scenario as in [3]. Similar to previous approaches [3, 9, 12, 28–30], the data acquisition was simulated by retrospectively under-sampling the 2D discrete Fourier transform of clinical magnitude MR images.⁴ The sampling masks include Cartesian 1D and 2D random sampling. We compare the proposed approach with DLMRI [9] to show the benefits of integrating guidance information into the MRI reconstruction task. We also compare with SVTMRI [3] which uses the structure-guided total variation to integrate the guidance contrast to aid the reconstruction of the target one. Figure 3 and 4 show reconstruction results for the scenario where a variable density Cartesian mask is employed for under-sampling on the target T1-weighted contrast, with a fully sampled T2-weighted MRI for guidance contrast.

³T1 (longitudinal relaxation time) is a measure of the time taken for excited spinning protons to realign with the external magnetic field and return to equilibrium. T2 (transverse relaxation time) is a measure of the time taken for excited spinning protons to lose phase coherence among the nuclei spinning perpendicular to the main field.

⁴After using the Fourier transform to transform measured k-space data into image space, the image data is of complex type, which is then manipulated for different clinical utility. In clinical practice, magnitude images are nearly exclusively used for diagnosis as it maximizes the signal-to-noise ratio (SNR). Phase-images are occasionally generated in clinical MRI for the depiction of flow and characterization of susceptibility-induced distortions. Therefore, from the perspective of diagnosis, we focus on the magnitude images

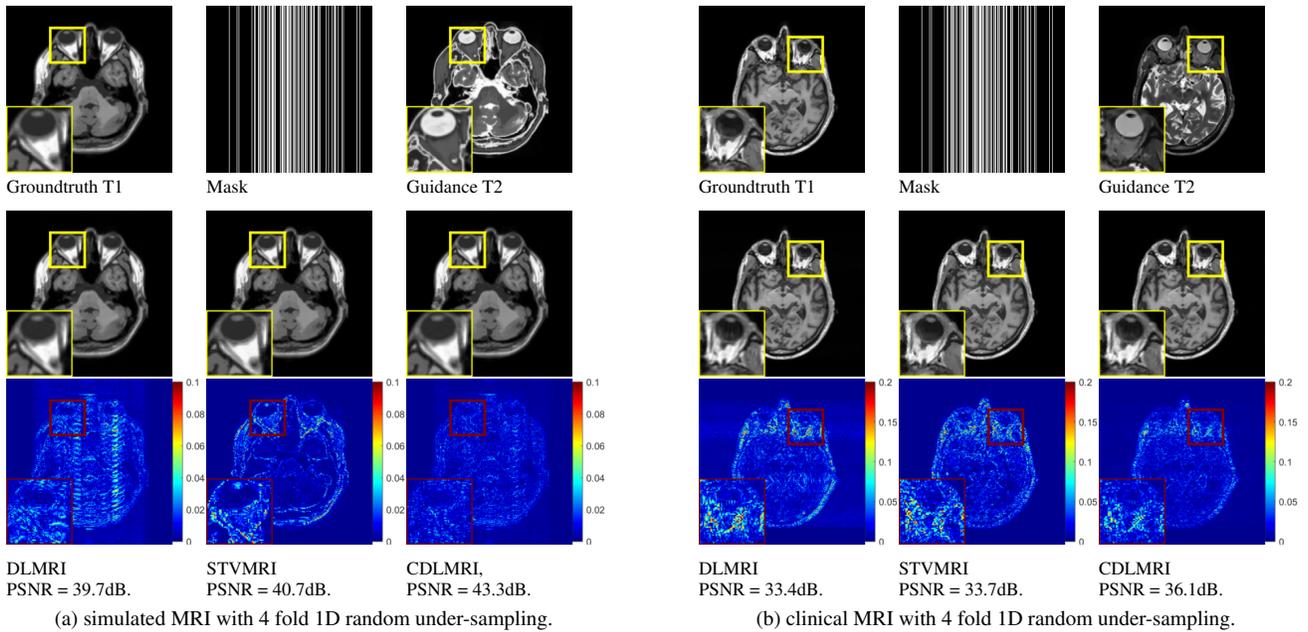


Fig. 3. Reconstruction for T1-weighted MRI, with fully-sampled T2-weighted version as reference using 4 fold Cartesian 1D random under-sampling. The first row shows the groundtruth T1-weighted contrast, sampling mask and guidance T2-weighted contrast. The second and third rows show the reconstructed images and the corresponding residual error from DLMRI [9], STVMRI [3], and the proposed CDLMRI. It can be seen that the proposed approach reliably reconstructs fine details and substantially suppresses aliasing, noise and artifacts, leading to the smallest residual error.

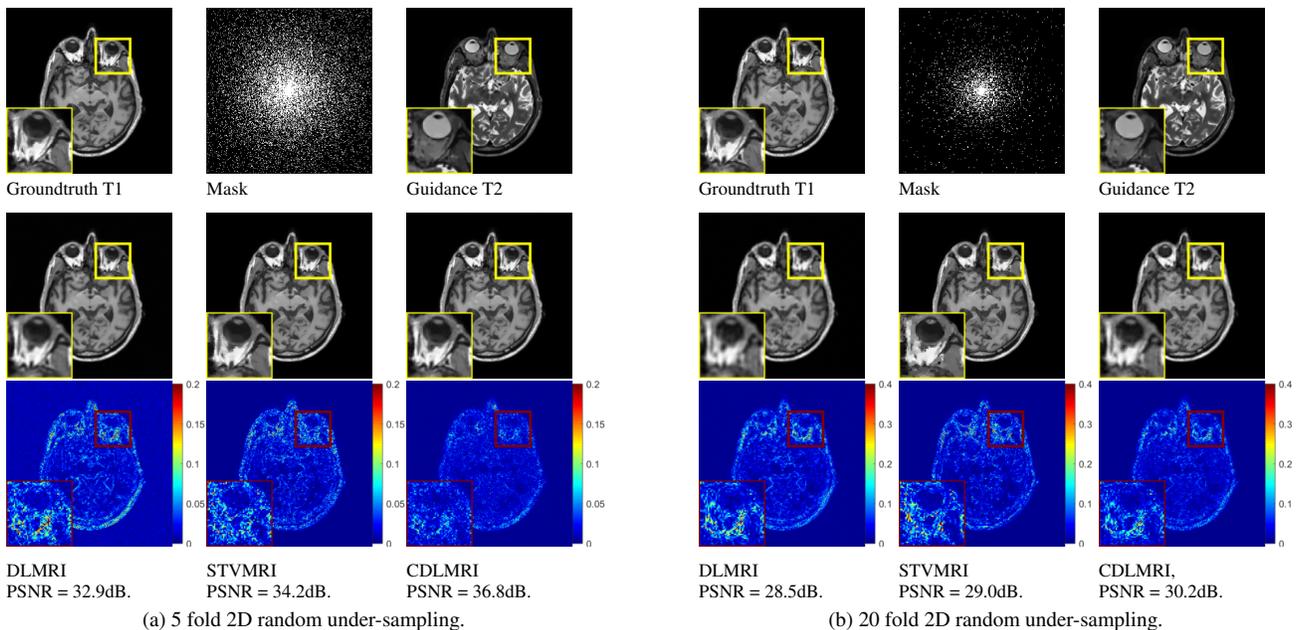


Fig. 4. Reconstruction for T1-weighted MRI, with fully-sampled T2-weighted version as reference using 5 fold and 20 fold 2D random under-sampling, using DLMRI [9], STVMRI [3], and the proposed CDLMRI. The first row shows the groundtruth T1-weighted, sampling mask and guidance modality T2-weighted. The second and third rows show the reconstructed images and the corresponding residual error from DLMRI [9], STVMRI [3], and the proposed CDLMRI. It can be seen that the proposed approach outperform the competing approaches, leading to the smallest residual error.