Anthony G. Christodoulou and Sajan Goud Lingala

Accelerated Dynamic Magnetic Resonance Imaging Using Learned Representations

A new frontier in biomedical imaging

Dynamic magnetic resonance imaging (MRI) can be used to scan a wide range of dynamic processes within the body, including the motion of internal organs, tissue-level nuclear magnetic resonance (NMR) relaxation, and dynamic contrast enhancement (DCE) of dye agents. The ability of MRI to safely provide unique soft-tissue contrast and comprehensive functional information has made dynamic MRI central to a number of imaging exams for cardiac, interventional, vocal tract, cancer, and gastrointestinal applications, among others. Unfortunately, MRI is a notoriously slow imaging modality due to fundamental physical and physiological limitations. These limitations result in tradeoffs between spatial and temporal resolutions, spatial coverage, and the signal-to-noise ratio and have made dynamic MRI a challenging technical goal.

During the past decade there was a surge of several compressed sensing (CS) MRI approaches (for example, [1] and [2]) that have made substantial headway toward addressing this challenge. At its core, CS-MRI exploits the redundancy of the images in predetermined transform domains (for example, Fourier and wavelet). However, one drawback has been that fixed transforms often do not provide the most efficient representation of images, thereby limiting the maximum achievable acceleration. To address these shortcomings, several adaptive, learning-based methods have been proposed. These schemes have revolutionized dynamic MRI for many applications, offering exciting new capabilities in biomedical imaging.

Here we offer a unified view of several different approaches to dynamic imaging using learned representations, focusing on the signal processing aspects that make each class of learning methods so powerful. Included are low-rank (LR) [3]–[5], blind CS (BCS) [6], higher-order multidynamic [7]–[10], explicit motion estimation and compensated recovery [11], [12], and manifold-regularized recovery methods [13], [14] as well as deep-learning approaches [15]–[17].

Problem formulation

We consider a complex image, $\tilde{x}(\mathbf{r}, \tau_1(t), \tau_2(t), \ldots, \tau_L(t))$, that is a scalar function of spatial location, $\mathbf{r} = [x, y, z]^T$, and

Digital Object Identifier 10.1109/MSP.2019.2942180
Date of current version: 15 January 2020

$L \geq 1$ time-varying independent variables, $\{\tau_\ell(t)\}_{\ell=1}^{L}$, each representing a different physical or physiological dynamic process, such as motion and NMR relaxation or simply the passage of time, t. The dynamic MR signal from this image can be observed from noisy, spatially encoded discrete measurements in $(\mathbf{k}, t)$ space:

$$b_c(\mathbf{k}_m, t_n) = \int_{\mathbf{r}} x(\mathbf{r}, t_n) s_c(\mathbf{r}, t_n) \exp(-j\mathbf{k}_m^T \mathbf{r}) d\mathbf{r} + \eta_c(\mathbf{k}_m, t_n) \tag{1}$$

$$x(\mathbf{r}, t) = \tilde{x}(\mathbf{r}, \tau_1(t), \tau_2(t), \ldots, \tau_L(t)), \tag{2}$$

where $s_c(\mathbf{r}, t_n)$ denotes the coil-sensitivity profile of the cth channel (which may vary with time), $(\mathbf{k}_m, t_n)$ indicates the (m, n)th sampling location in $(\mathbf{k}, t)$ space, and η denotes additive complex-valued Gaussian noise. Specific dynamic processes can be targeted by holding the unwanted τ's constant during signal acquisition. For example, the ℓth dynamic process can be isolated by enforcing $\tau_q(t) = \kappa_q \forall q \neq \ell$ during acquisition for some set of constant values, $\{\kappa_q\}_{q \neq \ell}$. Practically speaking, this means either physically holding $\{\tau_q(t)\}_{q \neq \ell}$ constant (for instance, pausing respiratory motion by having patients hold their breath) or waiting to acquire data only when the condition $\tau_q(t) = \kappa_q \forall q \neq \ell$ is met [such as removing the appearance of cardiac motion by synchronizing the data acquisition to a monitored electrocardiogram (ECG) signal]. The decision of which dynamic processes to image and which to suppress is application dependent.

The expression in (1) can be more compactly written in matrix-vector form as $\mathbf{b} = \mathcal{A}(\mathbf{X}) + \boldsymbol{\eta}$, where the matrix $\mathbf{X}$ has elements $[\mathbf{X}]_{mn} = x(\mathbf{r}_m, t_n)$, and the linear operator $\mathcal{A}$ models the coil-sensitivity encoding as well as the Fourier encoding on an arbitrary sampling trajectory (for example, Cartesian or non-Cartesian). The goal of accelerated dynamic MR image reconstruction is typically to recover $x(\mathbf{r}, t)$ from undersampled $(\mathbf{k}, t)$-space measurements through the use of low-dimensional learned representations of the underlying image. The recovery of $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L)$ is also possible and is discussed in the "Multidynamic Models" section.

Spatiotemporal decomposition models

The general partial separability (PS) model proposed by Liang [3] can be used to interpret several fixed and learned spatiotemporal decomposition models in a unified manner. The PS model utilizes a global signal model to represent the signal-time profile at each voxel, $(x(\mathbf{r}, t))$, as

$$x(\mathbf{r}, t) = \sum_{i=1}^{R} x_i(\mathbf{r}, t) = \sum_{i=1}^{R} u_i(\mathbf{r}) v_i(t). \tag{3}$$

The decomposition implies that $x_i(\mathbf{r}, t)$ is separable as functions of $\mathbf{r}$ and t; that is, it can be factored as a product of the spatial model coefficients $u_i(\mathbf{r})$ and the temporal basis functions $v_i(t)$. Here, R denotes the total number of basis functions (or the model order). Equation (3) implies that $x(\mathbf{r}, t)$ can be efficiently represented in either of two low-dimensional feature spaces, $\text{span}(\{u_i(\mathbf{r})\}_{i=1}^{R})$ and $\text{span}(\{v_i(t)\}_{i=1}^{R})$, providing an avenue for representation learning to recover $x(\mathbf{r}, t)$ from undersampled measurements.

Both $\text{span}(\{u_i(\mathbf{r})\}_{i=1}^{R})$ and $\text{span}(\{v_i(t)\}_{i=1}^{R})$ are R-dimensional linear subspaces (that is, linear manifolds intersecting the origin), so we adopt a linear algebraic framework to understand (3). In matrix form, (3) states that the dynamic Casorati matrix $\mathbf{X} \in \mathbb{C}^{M \times N}$ can be decomposed as a product of a spatial coefficient matrix, $\mathbf{U} \in \mathbb{C}^{M \times R}$, and a matrix, $\mathbf{V} \in \mathbb{C}^{R \times N}$, containing temporal functions:

$$\underbrace{\begin{pmatrix} x(\mathbf{r}_1, t_1) & \ldots & x(\mathbf{r}_1, t_N) \\ x(\mathbf{r}_2, t_1) & \ldots & x(\mathbf{r}_2, t_N) \\ \vdots & \ldots & \vdots \\ x(\mathbf{r}_M, t_1) & \ldots & x(\mathbf{r}_M, t_N) \end{pmatrix}}_{\mathbf{X}_{M \times N}} = \underbrace{\begin{pmatrix} u_1(\mathbf{r}_1) & \ldots & u_R(\mathbf{r}_1) \\ u_1(\mathbf{r}_2) & \ldots & u_R(\mathbf{r}_2) \\ \vdots & \ldots & \vdots \\ u_1(\mathbf{r}_M) & \ldots & u_R(\mathbf{r}_M) \end{pmatrix}}_{\mathbf{U}_{M \times R}} \underbrace{\begin{pmatrix} v_1(t_1) & \ldots & v_1(t_N) \\ v_2(t_1) & \ldots & v_2(t_N) \\ \vdots & \ldots & \vdots \\ v_R(t_1) & \ldots & v_R(t_N) \end{pmatrix}}_{\mathbf{V}_{R \times N}}. \tag{4}$$

Here, M represents the total number of voxels, and N denotes the number of time frames. Based on the constraints imposed on the matrices, $\mathbf{U}$ and $\mathbf{V}$, and the model order, R, the general PS model in (4) can help to classify various models, as laid out in Figure 1.

Linear models with complex exponential bases

The earliest adaptive model in dynamic MRI, known as *dynamic imaging by model estimation* (*DIME*), was proposed by Liang et al. [18]. It focuses on the recovery of a quasi-periodic dynamic signal as a linear combination of a limited number of complex exponential temporal basis functions. Therefore, the generalized PS model in (3) and (4) reduces to

$$x(\mathbf{r}, t) = \sum_{i=1}^{R} u_i(\mathbf{r}) e^{j2\pi f_i t}; \quad R < N. \tag{5}$$

The frequencies $\{f_i\}_{i=1}^{R}$ can be obtained from training data during the scan, which are typically composed of low spatial but high temporal-resolution dynamic data. In a second step, the coefficients $\{u_i(\mathbf{r})\}_{i=1}^{R}$ or $\mathbf{U}$ are determined by the method of least-squares-fitting to the acquired *k-t* data ($\mathbf{b}$) as

$$\hat{\mathbf{U}} = \underset{\mathbf{U}}{\text{argmin}} \| \mathcal{A}(\mathbf{U}\mathbf{V}) - \mathbf{b} \|_2^2; \quad V_{\text{in}} = e^{j2\pi f_i t_n}. \tag{6}$$

In addition, to improve the conditioning of (6), it is common to include l_2 regularization on $\mathbf{U}$. The idea was extended by several researchers who reformulated the problem as a spatial-spectral filtering issue; the goal was to design *k-t* lattice-sampling patterns to minimize the coherent aliasing in the spatial-spectral space (such as in [19]). These schemes were particularly developed to model quasi-periodic temporal patterns as observed with cardiac cine MRI and task-based functional MRI.

CS

CS under the synthesis-basis representation can be interpreted by the generalized PS model as having sparse coefficients, $\mathbf{U}$, of a fixed temporal basis, $\mathbf{V}$, with $R \geq N$, where $\mathbf{V}$ is determined a priori. The temporal basis $\{v_i(t)\}_{i=1}^{R}$ belongs to an overcomplete dictionary that is chosen either off the shelf (such as wavelets and complex exponential bases) or learned from a physical model describing the evolution of the time series (for instance, via dictionary learning [20]). CS assumes the model coefficients $u_i(\mathbf{r})$ to be sparse and estimates them from the measured $(\mathbf{k}, t)$-space data via the following minimization:

$$\hat{\mathbf{U}} = \underset{\mathbf{U}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{UV}) - \mathbf{b} \right\|_2^2 + \lambda \left\| \operatorname{vec}(\mathbf{U}) \right\|_1; \quad \text{for predetermined } \mathbf{V} \text{ and } R \geq N. \tag{7}$$

An alternative to (7) is to use an analysis formulation [1],

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{X}) - \mathbf{b} \right\|_2^2 + \lambda \left\| \psi(\mathbf{X}) \right\|_1, \tag{8}$$

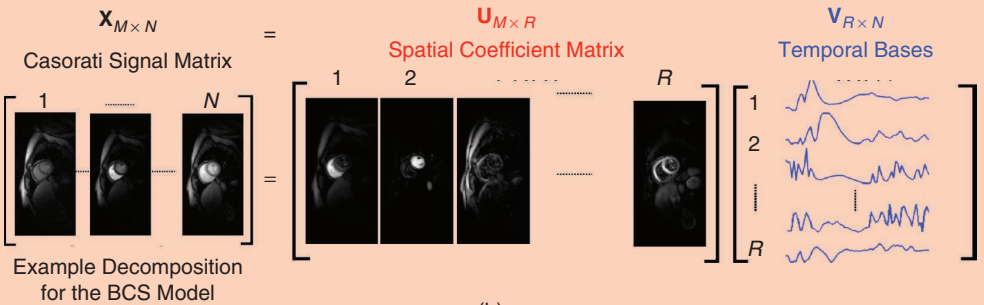
where ψ is an appropriate sparsity-inducing operator. Note that (7) and (8) produce the same reconstruction when $\psi(\mathbf{X}) = \mathbf{XV}^{-1}$. CS eliminates the need for training scans and, instead, relies on incoherent sampling. The performance of CS is heavily dependent on the specific dictionary or sparsifying operator. For example, Fourier dictionaries may be ideally suited for breath-held cine applications due to the pseudo-periodicity of cardiac motion. However, for applications involving complex temporal dynamics (such as free-breathing and contrast-enhanced MRI), the use of Fourier dictionaries may be suboptimal (as demonstrated in [5]).

Blind linear or LR models

Blind linear models can be thought of as a generalization of adaptive dynamic imaging by model estimation to nonperiodic dynamic data sets; instead of assuming the temporal basis functions in $\mathbf{V}$ to be periodic exponentials, they are derived from data, such as principal component analysis (PCA), singular value decomposition (SVD), Karhunen–Loève theorem (KLT), and others. These models are LR because they capitalize the position redundancy of the Casorati matrix, $\mathbf{X}$. Several models (including k-t PCA [4], a variant of PS [3]) utilize a two-step strategy to reconstruct $x(\mathbf{r}, t)$. The temporal bases $v_i(t)$ are estimated from low spatial-resolution but high temporal-resolution data. Subsequently, the spatial coefficients $u_i(\mathbf{r})$ are

Model	Constraint on **U** (Spatial Coefficients)	Constraint on **V** (Choice of Temporal Bases)	Number of Bases (R)
Linear Model (Complex Exponential Bases)	Nonsparse	Orthogonal	Few
CS	Sparse	Predetermined and Nonorthogonal	Many (Overcomplete)
Blind Linear Model (Low Rank)	Nonsparse	Learned and Orthogonal	Few
BCS	Sparse	Learned and Nonorthogonal	Many (Overcomplete)

(a)



(b)

FIGURE 1. The generalized PS model provides a unified view of various spatiotemporal decomposition models. (a) The models based on choices of constraints on the spatial coefficient matrix, **U**, the temporal matrix, **V**, and the model order, R. (b) An example pictorial decomposition of **U** and **V** for the BCS model on a myocardial perfusion data set.

estimated by fitting the model in (3) to the acquired *k-t* data in a least-squares sense that is similar to (6).

One limitation of the two-step strategy is that it requires sufficient training data to reliably approximate the true PCA/SVD/KLT basis functions. It demands a tradeoff between the time spent to collect the training data and the undersampled high-spatial-frequency measurements. To address this, single-step recovery schemes have been proposed that jointly estimate the spatial weights and temporal bases from the undersampled data itself [5]. The reconstruction problem can be reformulated as a rank-constrained optimization:

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{X}) - \mathbf{b} \right\|_2^2 \quad \text{such that rank}\,(\mathbf{X}) \le R < N. \tag{9}$$

Several researchers have proposed alternate regularizers as a surrogate to the rank, including the convex nuclear norm ($p = 1$) and the nonconvex Schatten p norm ($p < 1$). In addition, methods that jointly exploit the low rank and transform sparsity of dynamic time series have been developed. For example, the method of *k-t* sparsity and low rank (a joint exploitation of the finite difference sparsity and low rank) can be formulated [5] as

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{X}) - \mathbf{b} \right\|_2^2 + \lambda_1 \underbrace{\left\| \mathbf{X} \right\|_p^p}_{\text{Schatten } p \text{ norm}} + \lambda_2 \underbrace{\left\| \sqrt{\left|\nabla_x(\mathbf{X})\right|^2 + \left|\nabla_y(\mathbf{X})\right|^2 + \left|\nabla_t(\mathbf{X})\right|^2} \right\|_1}_{\text{spatiotemporal total variation}}. \tag{10}$$

The two-step recovery scheme of the blind linear model has also been improved by utilizing additional sparsity constraints [21]. These improve the recovery of the $u_i(\mathbf{r})$ problem for (3) as

$$\hat{\mathbf{U}} = \underset{\mathbf{U}}{\operatorname{argmin}} \left\| \mathcal{A}\Big(\mathbf{U} \underbrace{\mathbf{V}}_{\text{from PCA/KLT/SVD}}\Big) - \mathbf{b} \right\|_2^2 + \lambda \underbrace{\left\| \psi(\mathbf{UV}) \right\|_1}_{\text{sparsity regularizer}}, \tag{11}$$

where the choice of ψ in (11) is motivated by the diffusion-MRI application. It is worth noting that, since $\mathbf{V}$ is explicitly estimated in the two-step recovery scheme, it has the advantage of using less memory to store $\mathbf{U}$ and $\mathbf{V}$ as opposed to storing the entire $\mathbf{X}$ matrix, which is a key contributing factor when translated to the multidynamic setting, as we discuss in the "Multidynamic Models" section.

The BCS model

The BCS model [6] shares similarities with the CS and blind linear models. Similar to CS, the voxel-intensity profiles are modeled as a sparse linear combination of the basis functions in a dictionary. However, instead of assuming a fixed dictionary, $\mathbf{V}$, BCS estimates the dictionary from the undersampled measurements themselves. BCS differs from blind linear models by assuming the sparsity of the coefficients, $\mathbf{U}$, and using $R \ge N$ temporal basis functions, which are not necessarily orthogonal.

BCS reconstruction involves the joint estimation of $u_i(\mathbf{r})$ and $v_i(t)$ in (3) from undersampled *k-t* measurements, ($\mathbf{b}$), with a sparsity constraint on $u_i(\mathbf{r})$ and a dictionary constraint on $v_i(t)$, which is required to avoid scale ambiguity in the product of $u_i(\mathbf{r})v_i(t)$. The choices for the sparsity and dictionary constraints can range from the convex l_1 norm or the nonconvex l_0, l_p; $(0 < p < 1)$ norms on $u_i(\mathbf{r})$ to the unit-column and Frobenius norm constraints in the dictionary. An example BCS reconstruction with the l_1 coefficient sparsity and unit-column-norm dictionary constraint can be formulated as the following constrained optimization problem:

$$\{\hat{\mathbf{U}}, \hat{\mathbf{V}}\} = \underset{\mathbf{U},\mathbf{V}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{UV}) - \mathbf{b} \right\|_2^2 + \lambda \left\| \operatorname{vec}(\mathbf{U}) \right\|_1 \quad \text{such that } \left\| v_i(t) \right\|_2^2 < 1 \forall i = 1, 2, \ldots, R. \tag{12}$$

It is worth noting that various other formulations of the BCS scheme have been proposed that cannot be described by the generalized PS model. For instance, [22] considers overlapping patches of $\mathbf{X}$ in the $\mathbf{r}$ and t dimensions and exploits the sparsity in a learned spatiotemporal overcomplete dictionary from the patches. Convolutional sparse coding assumes $\mathbf{X}$ to be a superposition of the sparse feature images convolved with a collection of filters, where both the feature images and filters are acquired from the undersampled data [23].

Multidynamic models

An alternative to imaging $x(\mathbf{r}, t)$ is to perform multidynamic imaging [7]–[10] (Figure 2), which preserves the multiple independent variables $\{\tau_\ell\}_{\ell=1}^{L}$ and applies multidimensional signal modeling to image the underlying $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L)$. This image, $\tilde{x}$, is said to have L time dimensions.

The multidynamic imaging of $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L)$ has several benefits over imaging $x(\mathbf{r}, t)$. First, it is no longer necessary to enforce $\tau_q(t) = \kappa_q \forall q \ne \ell$ during acquisition to isolate a single τ_ℓ: The desired dynamic image can simply be retroactively extracted as a temporal slice of $\tilde{x}$, that is, $\tilde{x}(\mathbf{r}, \kappa_1, \kappa_2, \ldots, \tau_\ell, \ldots, \kappa_L)$. Second, if multiple τ's are of interest they can be imaged without collapsing them into a single time dimension, t, where they would otherwise confound each other during analysis. This allows multipurpose imaging, such as cine imaging and T_1 mapping, in the same scan and multiparameter mapping (for example, T_1–T_2 mapping). Finally, there are additional benefits to the image analysis: When multiple processes are quantified from the same scan, there is no need for image registration. This is especially relevant to the image fusion of the parameter maps acquired in separate breath-holds since patients typically do not reproduce their exact respiratory position during successive scans (Figure 3).

Imaging the entire $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L)$ presents its own challenges, of course. These primarily stem from the curse of dimensionality, wherein the size of $\tilde{x}$ grows geometrically with the number of dimensions. Let $\{\mathbf{r}_m\}_{m=1}^{M}$ be the set of M voxels in the image and $\{\tau_{\ell,n_\ell}\}_{n_\ell=1}^{N_\ell}$ be a set of N_ℓ discrete values along the ℓth time dimension, τ_ℓ. The number of elements in $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L)$ is, therefore, $M\prod_{\ell=1}^{L} N_\ell$, revealing geometric growth as L increases. As the number of elements in $\tilde{x}$ rises, so do the sampling (for instance, the scan time) and memory/storage requirements, presenting major practical challenges.

Fortunately, this challenge is balanced by the blessing of dimensionality, wherein the signals in high-dimensional spaces are often highly structured and can be very efficiently represented with the appropriate signal modeling [24]. This provides an opportunity for the multidimensional extensions in (3) to represent $\tilde{x}$ with fewer degrees of freedom than $M\Pi_{\ell=1}^{L} N_\ell$, thereby reducing the scan time and/or storage requirements. By explicitly separating the sources of the image dynamics into multiple time dimensions, each dynamic in $\tilde{x}$ can be afforded its own temporal model.

To interpret various models in a unified manner, we revisit (3) in the context of multidynamic imaging:

$$\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_L) = \sum_{i=1}^{R} u_i(\mathbf{r}) \tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L), \tag{13}$$

which simply replaces the $\{v_i(t)\}_{i=1}^{R}$ with the multidimensional functions $\{\tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L)\}_{i=1}^{R}$. We notate $\tilde{x}$ in multidimensional array/tensor form [25] as the $(L+1)$-way tensor $\mathcal{X} \in \mathbb{C}^{M \times N_1 \times N_2 \times \ldots \times N_L}$ with elements $[\mathcal{X}]_{mn_1n_2\ldots n_L} = \tilde{x}(\mathbf{r}_m, \tau_{1,n_1}, \tau_{2,n_2}, \ldots, \tau_{L,n_L})$. In this form, (13) becomes

$$\mathcal{X} = \mathcal{V} \times_1 \mathbf{U}, \tag{14}$$

where $\mathcal{V} \in \mathbb{C}^{R \times N_1 \times \cdots \times N_L}$ has the elements $[\mathcal{V}]_{in_1n_2\ldots n_L} = \tilde{v}_i(\tau_{1,n_1}, \tau_{2,n_2}, \ldots, \tau_{L,n_L})$ and the definition of $\mathbf{U}$ is unchanged from (4). As before, the constraints imposed on $\mathbf{U}$ and $\mathcal{V}$ and the value of R help to classify various multidynamic-image models.

Multidimensional CS

Multidimensional CS for multidynamic imaging was first described by Feng et al. as the extra-dimensional golden angle radial sparsity reconstruction (XD-GRASP) method [7], which demonstrated cardiac- and respiratory-resolved imaging of the heart as well as respiratory- and DCE-resolved abdominal imaging. Cheng et al. expanded upon the

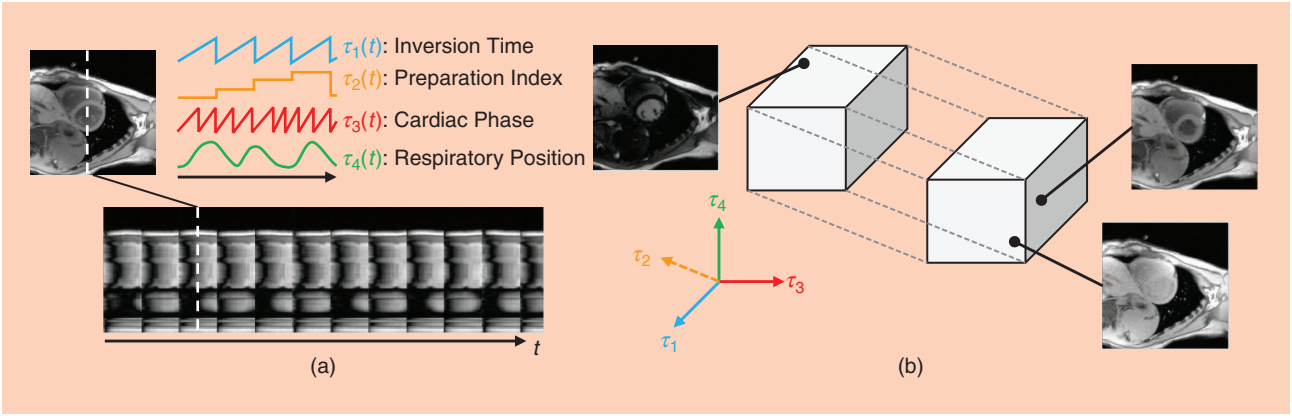


FIGURE 2. An illustration of the multidynamic imaging concept. Four dynamic processes occur during the scan: T_1 relaxation (a function of inversion time), T_2 relaxation (a function of T_2 preparation-module duration), cardiac motion, and respiratory motion. (a) All four processes overlap in the real-time dynamic image $x(\mathbf{r}, t)$, complicating the analysis of any individual dynamic process. (b) Multidynamic imaging separates the processes into different dimensions of $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \tau_3, \tau_4)$.

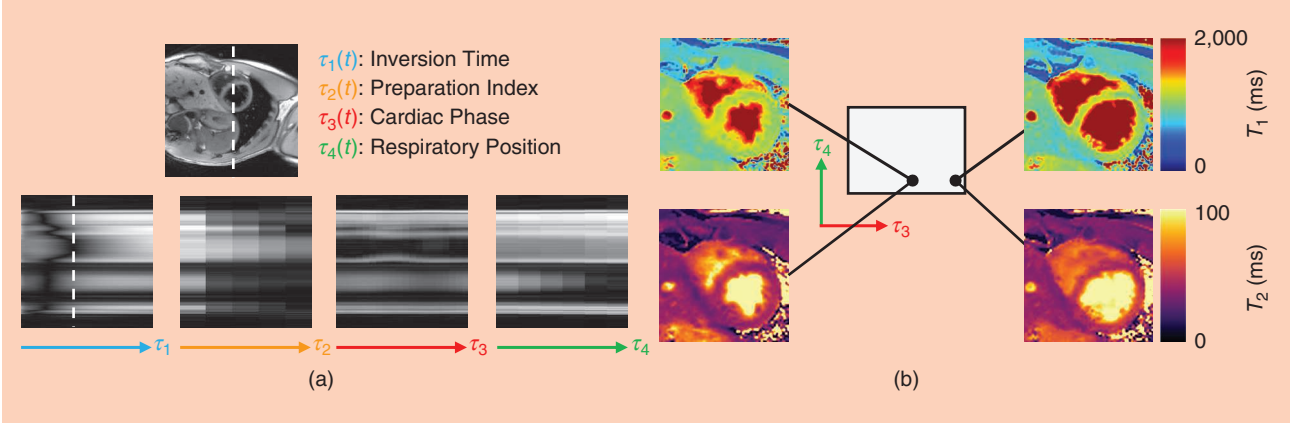


FIGURE 3. Some of the capabilities of multidynamic imaging, as performed using MR multitasking [10]. (a) The multidynamic image $\tilde{x}(\mathbf{r}, \tau_1, \tau_2, \tau_3, \tau_4)$ can be sliced along any of its four time dimensions to retrospectively isolate individual dynamic processes. Note that, by performing cardiac- and respiratory-resolved imaging, the example here did not require ECG gating or subject breath-holding. (b) Because relaxation processes T_1 and T_2 are imaged in the same motion-resolved scan, T_1 and T_2 maps are naturally coregistered and available in any combination of the cardiac phase and respiratory position.

concept with the XD flow method [9] by calculating multiple reconstructions with different combinations of cardiac, respiratory, flow, and DCE time dimensions from one scan.

This class of methods is characterized by a large predetermined basis $\mathcal{V}$ and sparse $\mathbf{U}$. The advantage of this model over its single-dynamic counterpart is that $\tilde{x}$ is potentially more compressible along an individual time dimension τ_ℓ, than along t. For example, we can reasonably expect the difference image along τ_ℓ,

$$\nabla_{\tau_\ell}\{\tilde{x}\} = \tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_\ell, \ldots, \tau_L) - \tilde{x}(\mathbf{r}, \tau_1, \tau_2, \ldots, \tau_\ell - \Delta\tau_\ell, \ldots, \tau_L), \quad (15)$$

to be sparser than the difference between successive real-time images,

$$\nabla_t\{x\} = x(\mathbf{r}, t) - x(\mathbf{r}, t - \Delta t) = \tilde{x}(\mathbf{r}, \tau_1(t), \tau_2(t), \ldots, \tau_L(t)) - \tilde{x}(\mathbf{r}, \tau_1(t - \Delta t), \tau_2(t - \Delta t), \ldots, \tau_L(t - \Delta t)); \quad (16)$$

this is because (15) describes the change from one dynamic effect, whereas (16) describes simultaneous changes from multiple dynamics. Image reconstruction is then possible according to

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathcal{X}) - \mathbf{b} \right\|_2^2 + \sum_{i=2}^{L+1} \lambda_i \left\| \operatorname{vec}(\mathbf{\Psi}_i \mathbf{X}_{(i)}) \right\|_1, \quad (17)$$

where $\mathbf{X}_{(i)}$ is the ith unfolding of $\mathcal{X}$ such that the rows of $\mathbf{X}_{(1)} \in \mathbb{C}^{M \times \Pi_{\ell=1}^{L} N_\ell}$ index $\mathbf{r}$, the rows of $\mathbf{X}_{(i)} \in \mathbb{C}^{N_{i-1} \times M \Pi_{\ell \neq i-1} N_\ell}$ index τ_{i-1} for $i > 1$, and the sparsifying transform $\mathbf{\Psi}_i$ operate along τ_{i-1}. This recovers an image tensor, $\mathcal{X}$, which is compressible, thereby reducing the scan time; however, algorithms to solve (17) generally do not explicitly store and operate upon $\mathcal{X}$ in compressed form, so the memory issues related to the curse of dimensionality remain.

An important consideration in multidynamic sampling is that $\{\tau_\ell(t)\}_{\ell=1}^{L}$ must be known to establish the mapping between t and $(\tau_1, \tau_2, \ldots, \tau_L)$. Any $\tau_\ell(t)$ that describes a sequence parameter is known a priori since it is a part of the pulse-sequence design. However, any $\tau_\ell(t)$ that is a physiological index, such as the cardiac phase (that is, the time point in the cardiac cycle) and respiratory position, cannot be known ahead of time. These physiological timings need to be monitored using ECG recordings and respiratory navigators or, alternatively, inferred from the acquired data (that is, "self-gated"). One of the primary benefits of multidynamic imaging is freedom from motion monitoring, so both the XD-GRASP and XD flow are self-gated methods, for which the physiological $\tau_\ell(t)$s are calculated from a subset of $\mathbf{b}$ that is referred to as the *self-gating signal*. Self-gating signals are collected throughout the scan at a high temporal-sampling rate from a limited subset of the $\mathbf{k}$-space, which is similar to the subject-specific training data to learn $\mathbf{V}$ that are used by some of the methods described in the previous section. For multidimensional CS, the self-gating signals serve only to define any unknown timings, $\{\tau_\ell(t)\}_{\ell=1}^{L}$, and have no bearing on $\mathcal{V}$.

The general partial separability model proposed by Liang can be used to interpret several fixed and learned spatio-temporal decomposition models in a unified manner.

LR tensor imaging

As in conventional single-dynamic imaging, a complementary alternative to CS that features learned representations is blind linear modeling, or in the context of multidynamic imaging, LR tensor (LRT) imaging. He et al. described the LRT with explicit subspace (LRTES) method [8] for multishell diffusion imaging, static T_1-T_2^* mapping, and 2D J-resolved spectroscopic imaging. The MR multitasking framework [10] extended the concept to handle motion-resolved imaging; performing motion-resolved quantitative imaging, such as non-ECG; free-breathing T_1–T_2 mapping in the heart, and cardiac- and DCE-resolved T_1 mapping for quantitative myocardial perfusion.

LRT imaging is characterized by learning a $\mathcal{V}$ for which $R < \min(M, \Pi_{\ell=1}^{L} N_\ell)$ and where $\mathbf{U}$ is not necessarily sparse. Compared to LR imaging with one time dimension, the LRT model places additional constraints on the structure of the temporal basis, specifically that the $\mathcal{V}$ itself can be factored, that is, that each $\tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L)$ can be further decomposed. Various tensor decompositions are available to model $\mathcal{V}$, for example, the canonical decomposition,

$$\tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L) = v_{1,i}(\tau_1) v_{2,i}(\tau_2) \ldots v_{N,i}(\tau_L), \quad (18)$$

where $\{v_{\ell,i}(\tau_\ell)\}_{i=1}^{R}$ spans the subspace for the ℓth time dimension, and the Tucker decomposition,

$$\tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L) = \sum_{i_1=1}^{R_1} \sum_{i_2=1}^{R_2} \cdots \sum_{i_L=1}^{R_L} c_{i i_1 i_2 \ldots i_L} v_{1,i_1}(\tau_1) v_{2,i_2}(\tau_2) \ldots v_{N,i_L}(\tau_L), \quad (19)$$

where $\{v_{\ell,i_\ell}(\tau_\ell)\}_{i_\ell=1}^{R_\ell}$ is the basis for the ℓth time dimension. The core tensor, $\mathcal{C} \in \mathbb{C}^{R \times R_1 \times \cdots \times R_L}$, has the elements $[\mathcal{C}]_{i i_1 i_2 \ldots i_L} = c_{i i_1 i_2 \ldots i_L}$, which determine the weight assigned to each basis function, which is similar to the role of the singular values in matrix decompositions. We also note the option of local [26] and patch-based tensor modeling [27], [28], which may have benefits related to nonlinear manifold modeling (which is discussed further in the "Manifold Models" section.). For a comprehensive description of the differences between various tensor decompositions, see [25]. The rest of this section assumes global modeling using (19), which permits a lower R than the canonical decomposition and is employed by LRTES and MR multitasking. In the global Tucker tensor form, the temporal model in (14) and (19) becomes

$$\mathcal{V} = \mathcal{C} \times_2 \mathbf{V}_1 \times_3 \mathbf{V}_2 \times_4 \cdots \times_{L+1} \mathbf{V}_L$$
$$\mathcal{X} = \mathcal{V} \times_1 \mathbf{U} = \mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{V}_1 \times_3 \mathbf{V}_2 \times_4 \cdots \times_{L+1} \mathbf{V}_L,$$

where the elements of $\mathbf{V}_\ell \in \mathbb{C}^{N_\ell \times R_\ell}$ are $[\mathbf{V}_\ell]_{ij} = v_{\ell,j}(\tau_{\ell,i})$.

There are various approaches to learning an LRT representation from sparsely sampled data. One approach is to implicitly impose the LRT model by penalizing the nuclear norm of each unfolding of $\mathcal{X}$, that is,

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathcal{X}) - \mathbf{b} \right\|_2^2 + \lambda \sum_{i=1}^{L+1} \left\| \mathbf{X}_{(i)} \right\|_*. \tag{20}$$

Similar to (17), this addresses the scan time by recovering an image tensor, $\mathcal{X}$, that is compressible (this time by tensor factorization) without explicitly storing $\mathcal{X}$ in a factorized form. To directly address the effect of curse of dimensionality on the memory and storage requirements, it is instead possible to recover the memory-efficient individual factors of the LRT model, $\mathbf{U}$, $\mathcal{C}$, and $\{\mathbf{V}_\ell\}_{\ell=1}^{L}$, without calculating their product, $\mathcal{X}$, in the sense of

$$\{\hat{\mathcal{C}}, \hat{\mathbf{U}}, \{\hat{\mathbf{V}}_\ell\}_{\ell=1}^{L}\} = \underset{\mathcal{C}, \mathbf{U}, \{\mathbf{V}_\ell\}_{\ell=1}^{L}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{V}_1 \times_3 \mathbf{V}_2 \times_4 \cdots \times_{L+1} \mathbf{V}_L) - \mathbf{b} \right\|_2^2. \tag{21}$$

Rather than jointly optimizing the cost function in (21) across all of the variables, it is often convenient to first learn the representations for the dynamic processes from subject-specific training data, as in the LRTES method. This method applies when $(\mathbf{k}, \tau_1, \tau_2, \ldots, \tau_L)$-space sampling can be prospectively controlled (for instance, when all of the $\{\tau_\ell(t)\}_{\ell=1}^{L}$ represents the sequence parameters). In this scenario, the sampling is designed to collect L subsets of training data, the ℓth set of which densely samples across τ_ℓ within a limited region of $(\mathbf{k}, \{\tau_q\}_{q \neq \ell})$-space. Each $\mathbf{V}_\ell$ is then learned by computing the SVD of each training-data subset before jointly recovering $\mathcal{C}$ and $\mathbf{U}$ from the remaining $(\mathbf{k}, \tau_1, \tau_2, \ldots, \tau_L)$-space information.

MR multitasking offers a solution to scenarios where the $(\mathbf{k}, \tau_1, \tau_2, \ldots, \tau_L)$-space sampling cannot be fully controlled (for example, when some of the $\{\tau_\ell(t)\}_{\ell=1}^{L}$ indexes physiological processes, such as motion). There, self-gating data are used to learn $\{\tau_\ell(t)\}_{\ell=1}^{L}$ before they are mapped into the $(\mathbf{k}, \tau_1, \tau_2, \ldots, \tau_L)$-space to serve as subject-specific training data. The training data typically only sparsely sample the $(\mathbf{k}, \tau_1, \tau_2, \ldots, \tau_L)$-space, so multitasking performs the temporal-feature extraction by learning $\mathcal{V}$ from all of the training data after they have been completed using nuclear-norm constraints, similar to (20). Because the training data are collected only in a limited region of the $\mathbf{k}$-space, their completion problem has far more modest memory and storage requirements than the entire-image completion problem in (20). After learning the feature-space $\operatorname{span}(\{\tilde{v}_i(\tau_1, \tau_2, \ldots, \tau_L)\}_{i=1}^{R})$, the image coordinates in that area (such as the $\mathbf{U}$) are recovered identically to the subspace-constrained LR matrix-image reconstruction problem, with $v_i(t) = \tilde{v}_i(\tau_1(t), \tau_2(t), \ldots, \tau_L(t))$.

Explicit motion estimation and compensation models

Several methods based on compensating the interframe object and/or subject motion have been proposed to improve the performance of accelerated dynamic MRI models. The general idea is to jointly estimate the deformation parameters and dynamic data set by imposing compactness priors on the motion-corrected time series rather than the original time series. An example optimization criterion for an explicit motion estimation and compensated dynamic MRI recovery scheme can be written as [11], [12]:

$$\{\hat{\mathbf{X}}, \hat{\theta}(\mathbf{r}, t)\} = \underset{\mathbf{X}, \theta(\mathbf{r}, t)}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{X}) - \mathbf{b} \right\|_2^2 + \lambda_1 \underbrace{\psi(\mathcal{T}_\theta(\mathbf{X}))}_{\substack{\text{regularization on motion-} \\ \text{compensated time series}}} + \lambda_2 \underbrace{\phi(\theta(\mathbf{r}, t))}_{\substack{\text{regularization} \\ \text{on motion maps}}}, \tag{22}$$

where, $\mathcal{T}_\theta$ denotes the deformation operator, described by the motion-field parameters, $\theta(\mathbf{r}, t)$. The motion field can be modeled as a rigid-body movement and nonrigid deformations, and ϕ is a regularizer on the motion field, such as a smooth spatial filter to penalize large, unrealistic deformations. Meanwhile, ψ specifies the arbitrary choices for the compactness priors in the motion-compensated time series, including the spatial-spectral sparsity prior, spatiotemporal finite-difference sparsity prior, and patch-based and/or global LR priors.

The optimization in (22) is challenging to solve due its nonlinearity and nonconvexity. The typical approaches include alternating between reconstruction and motion estimation by starting with reasonably good initial estimates of the reconstruction and the motion maps. In practice, strategies that correct for motion in a coarse-to-fine stage have shown to be robust against undesirable local minima. In a similar realm, other efficient methods based on variable splitting and continuation rules have been proposed to decouple the problem in (22) to simpler challenges and gradually update the complexity. Figure 4 presents an example of a motion-compensated reconstruction scheme that is applied to the recovery of free-breathing cardiac-perfusion data sets from fourfold undersampled data. While the explicit motion estimation and

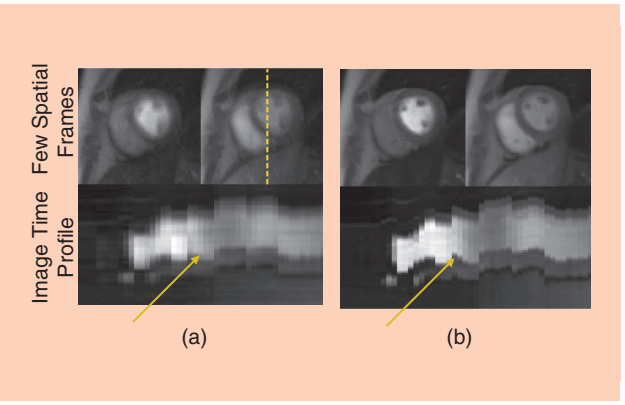


FIGURE 4. The demonstration of explicit motion estimation and the compensation models. This example displays fourfold-prospectively acquired myocardial perfusion data with heavy breathing. (a) Performing CS (via a patch-based low-rank regularizer) shows considerable motion artifacts. (b) With motion compensation, the reconstruction is robust against those artifacts and demonstrates an improved temporal fidelity. (Source: Xiao Chen, Siemens Healthineers; used with permission.)

compensation are extremely powerful in reducing motion artifacts, they are practically challenged by long computation times, nonconvexity, and the tuning of several parameters (for instance, the reconstruction and motion-estimation regularization parameters).

Manifold models

The manifold structure of data has been widely used to visualize the configuration of complex sets. Nonlinear dimensionality reduction and manifold-embedding methods assume that the data are points on a low-dimensional manifold (a smooth surface) in a higher-dimensional space. If the manifold is of low enough dimensionality, the data can be compactly visualized in low-dimensional space. For example, if the points are on a curve in 3D, each one can be associated on a straight line; nonlinear dimensionality reduction methods, including isometric mapping generate a nonlinear mapping between each point on the 3D curve to a point on a line [29]. Emerging research shows that the manifold structure can be used to regularize the recovery of dynamic MRI data from undersampled measurements.

In applications such as free breathing and ungated cine MRI, the dynamic images can be associated with the nonlinear functions of two parameters: the cardiac and respiratory phases. Images with similar cardiac and respiratory phases are expected to be comparable; this property is used in gating methods. Hence, the images can be assumed to be points on a smooth low-dimensional manifold with a high ambient dimension, where the dimension is the total number of pixels in each image. Figure 5 demonstrates this concept on a simulated free-breathing cardiac cine phantom, where similar but distant time frames are mapped to lie in close proximity on the smooth manifold.

Manifold regularization exploits the similarity of image frames in terms of the proximity of the points on the smooth manifold. The reconstruction formulation can be written [13] as

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{X}) - \mathbf{b} \right\|_2^2 + \lambda \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\sqrt{W_{ij}} \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_p \right)^p, \tag{23}$$

where the matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ contains weighting factors that determine the degree of similarity between $\mathbf{x}_i$ and $\mathbf{x}_j$ and the ith and jth columns (time frames) of $\mathbf{X}$, $\| \cdot \|_p^p$ indicates the l_p norm, and $p <= 2$ indicates the image differences. These weights are inversely proportional to the distance between the corresponding points on the manifold. For pairs of points that are closer to each other, W_{ij} has a high value, while for pairs of points that are farther from each other, W_{ij} has a low value. One strategy to estimate these weights is use the navigator signals $y(\mathbf{k}, t)$ as

$$W_{ij} = \begin{cases} e^{-\frac{\left\| \mathbf{y}_i - \mathbf{y}_j \right\|_2^2}{\sigma^2}} & \text{if} \quad \left\| \mathbf{y}_i - \mathbf{y}_j \right\|_2^2 < 2\sigma, \\ 0 & \text{else} \end{cases} \tag{24}$$

where σ is a parameter that is dependent on the maximum curvature on the manifold. The acquisition scheme to collect the navigator data $y(\mathbf{k}, t)$ requires sampling of the $\mathbf{k}$-space at the same location for each time frame. Typically, the navigator data are simultaneously interspersed with the measurement data. For example, in free-breathing cardiac cine MRI [13], three to four radial spokes are acquired as navigator $\mathbf{k}$-space lines for each frame, while the remaining spokes are obtained by using a classical golden-angle radial scheme. This can be extended using other navigator-sampling methods (such as with spirals).

Other manifold regularizations have been proposed. As an example, in localized linear embedding [30], each image

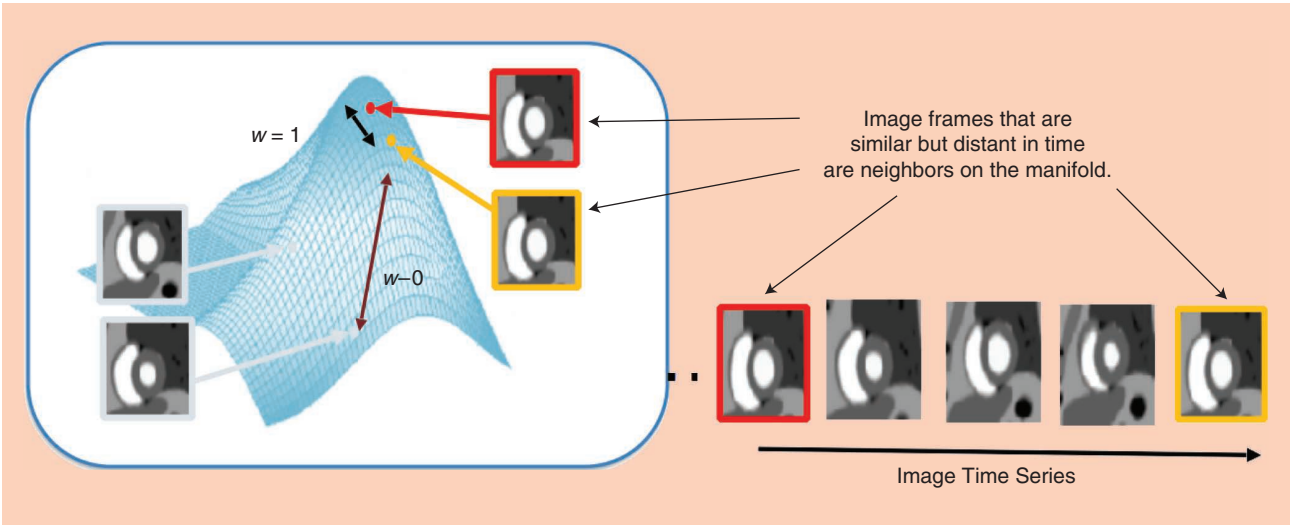


FIGURE 5. The manifold structure of dynamic images. In this example, a free-breathing, ungated cardiac cine phantom is considered for demonstration. The image frames can be considered to live as points on a smooth 2D manifold in a higher-dimensional space. Note that, in the sequence of images on the right, there are frames that are similar but distant in time (that is, of the similar cardiac and respiratory phase; see the red and yellow squares). On the manifold, the frames are mapped to points that are neighbors. Similarly, dissimilar frames correspond to points that are far away from each other on the manifold. The weighting factors that determine the degree of similarity between two frames are typically estimated via navigator signals, and the reconstruction is formulated as a penalized optimization [22].

(that is, a point on a manifold) is closely approximated by the weighted combination of the other images in the data set (that is, neighbors on the manifold). The weights are estimated from navigator data similar to smoothness regularization on manifolds (SToRM).

Based on the success of nonlocal means in denoising [31], patch-based regularization for implicit motion-compensated recovery has been proposed in dynamic MRI recovery [32]. Small spatial patches are defined in an image frame, and similar patches are located in a localized spatiotemporal neighborhood in subsequent time frames. Redundancies among the similar patches are then exploited. The recovery can be posed as a regularized reconstruction scheme, where the nonlocal regularization penalty is an unweighted sum of distances between image-patch pairs in the dynamic data set.

Representation learning has made important strides in accelerating dynamic MRI, and it has the potential to change the very structure of clinical MRI exams.

Kernel-based PCA approaches have been proposed that also fall under the umbrella of the manifold-regularization methods [14]. Kernel PCA maps the low-dimensional dynamic signal to a high-dimensional feature space via a nonlinear mapping. It exploits the LR structure in the high-dimensional feature space and maps the reconstruction from the feature space to the signal space. Recently, SToRM has been interpreted as a kernel PCA approach by exploiting a union-of-curves model [33].

Deep-learning models

Deep learning via multiple layers of hierarchical nonlinear representation modules (for example, neural networks) is a hugely active research area in many fields of science and engineering today. Deep-learning networks are capable of solving complex inverse problems noniteratively (or rather, with all of the iterations isolated to an initial training phase); as such, they offer an avenue for nonlinear image reconstruction at unprecedented calculation speeds. Neural networks are known to be highly effective at learning representations of data that lie on nonlinear manifolds [34], further making deep learning a promising fit for dynamic-image reconstruction. As covered elsewhere in this issue of *IEEE Signal Processing Magazine*, most deep-learning MR image-reconstruction networks have been designed to recover static images or individual frames from dynamic images (see [35]–[37]), thereby learning spatial representations. However, explicitly dynamic methods for learning spatiotemporal representations have started to appear in the literature as well, (as in [15]–[17]), presenting an exciting new avenue for accelerated dynamic MRI.

Rather than iteratively inverting the forward problem $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ to reconstruct $\hat{\mathbf{X}}$ for each individual subject, deep-learning reconstruction methods learn the parameters $\hat{\theta}$ of a nonlinear forward-reconstruction operator, $f(\cdot;\hat{\theta})$, from a large training set. This training set, $\mathcal{D}_{\text{train}}$, contains matched pairs of known ground-truth images and their undersampled **k**-space data, $(\mathbf{X}, \mathbf{b})$, from multiple subjects, enabling the network to be trained according to the objective function

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i|(\mathbf{X}_i, \mathbf{b}_i)\in\mathcal{D}_{\text{train}}} \mathcal{L}(\mathbf{X}_i, f(\mathbf{b}_i;\theta)), \tag{25}$$

where the loss function $\mathcal{L}(\mathbf{X}_i, f(\mathbf{b}_i;\theta))$ compares the ith reconstructed image $f(\mathbf{b}_i;\theta)$ to the corresponding labeled training image $\mathbf{X}_i$. In the context of dynamic MR image reconstruction, L2 has been the main loss function to be explored: $\hat{\theta} = \operatorname{argmin}_\theta \Sigma_{i|(\mathbf{X}_i, \mathbf{b}_i)\in\mathcal{D}_{\text{train}}} \| \mathbf{X}_i - f(\mathbf{b}_i;\theta) \|_2^2$. It should be noted that in practice, $\mathbf{b}$ is optionally transformed into image space by an initial preprocessing step, such as the adjoint operation $\mathcal{A}^*(\mathbf{b})$, before passing through the network. Here we absorb preprocessing steps into f as a means of simplifying the notation.

After they are trained by (25), future images can be reconstructed by a single pass through the deployed network, $\hat{\mathbf{X}} = f(\mathbf{b};\hat{\theta})$, or by performing data-consistent reconstruction that incorporates $f(\mathbf{b};\hat{\theta})$ as a prior on $\hat{\mathbf{X}}$ [15], [38], as in

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \| \mathcal{A}(\mathbf{X}) - \mathbf{b} \|_2^2 + \lambda \| \mathbf{X} - f(\mathbf{b};\hat{\theta}) \|_F^2. \tag{26}$$

A two-step process of 1) applying f to remove noise and artifacts and 2) enforcing the data consistency via (26) mirrors individual iterations of the optimization algorithms used for iterative nonlinear image reconstruction. As a result, deep-learning reconstruction pipelines that repeat this two-step process multiple times in series are analogous to "unrolling" iterative algorithms.

Deep-learning networks for imaging are generally structured so that early layers of f encode the input within the nonlinear feature space that was learned during training, and later layers decode the desired $\hat{\mathbf{X}}$. In the context of dynamic MRI, this can be understood as replacing handcrafted model choices (for example, whether $\mathbf{U}$ is sparse, what the size of $\mathbf{V}$ is, and whether to perform local or global modeling) with automatic modeling to some degree. These choices do not entirely disappear since they are reflected in the selection of the specific network architecture; as an example, the hierarchy of layers is usually designed to permit multiscale modeling, but the specific $\hat{\theta}$ learned during training would ultimately determine the balance between global and local modeling. Most of the deep-learning methods for dynamic MRI have been based on convolutional neural network (CNN) architectures, including those proposed by Schlemper et al. as a cascaded CNN with data sharing from previous time points [15], by Qin et al. in a recurrent CNN architecture [16], and by Biswas et al. [17] as an unfolded CNN with an additional spatiotemporal manifold constraint added to (26).

Outlook

Representation learning has made important strides in accelerating dynamic MRI, and it has the potential to change the very structure of clinical MRI exams. Today, MRI exams comprise a series of different scans that are acquired with a different

contrast weighting or interrogate a different dynamic process. All of the scans are reconstructed independently and typically stored for later analysis.

Multidynamic imaging has the potential to replace the series of independent scans with comprehensive, integrated single-scan examinations. Currently, the time between scans is used for technologists to make imaging decisions, such as the contrast-weighting selection and ECG trigger-window definition, as well as for patients to recover from holding their breath. Scanning multiple contrast weightings for a variety of motion states would shift these imaging decisions to after the reconstruction and eliminate breath holds. This would go beyond simply reducing the amount of data being acquired by also eliminating the delays between acquisitions. A motion-resolved multicontrast paradigm would be a major step toward a fully quantitative multiparameter exam for objective, comprehensive tissue characterization.

A motion-resolved multicontrast paradigm would be a major step toward a fully quantitative multiparameter exam for objective, comprehensive tissue characterization.

There are many signal processing opportunities for making this new paradigm a reality. Multidynamic CS has only explored the use of temporal, finite, different sparsifying transforms (for instance, temporal TV), and LRT imaging has primarily been limited to the Tucker decomposition. Major questions about the optimal sampling design remain, particularly how to prospectively plan the sampling when physiological timings are unpredictable, such as when a patient's physiology partially controls the sampling pattern. There are also open opportunities to extend other decompositions and models to their multidynamic counterparts, for example, LR-plus-sparse [39] tensor models, multidynamic BCS models, and multidynamic manifold models. The richness (and massive size) of the data will also require faster image reconstruction and necessitate new modes of viewing and analyzing high-dimensional scans, all of which are potential areas to incorporate more advanced representation-learning approaches, including deep learning.

As the newest of the methods described here, deep learning has perhaps the largest number of open technical problems and opportunities. Deep learning in dynamic MRI is currently dependent on supervised learning with huge sets of highly representative training data, which limits the ability to enable new modes of imaging where the labels are not yet available. This suggests the importance of developing unsupervised reconstruction networks and methods of generating realistic, representative training data (for example, self-supervised learning where the training images are generated by a secondary network).

There are opportunities to explore more advanced loss functions, such as with generative adversarial networks wherein $\mathcal{L}$ is output by a network, as well as opportunities to go beyond $(\mathbf{r}, t)$-space networks and explore $(\mathbf{k}, t)$-space and hybrid $(\mathbf{k}, t)$–$(\mathbf{r}, t)$ networks. The current reliance of deep learning on graphics-processing-unit training, where there are stricter memory limitations than there is CPU processing, presents challenges for reconstructing long image sequences and dynamic images with three spatial dimensions, and it creates an opportunity to explore low-resource architectures. The excellent performance of deep networks for image analysis tasks, although not covered in this article, also suggests the opportunity to build combined reconstruction and analysis networks that can make clinically useful inferences directly from input $(\mathbf{k}, t)$-space data, changing the imaging pipeline entirely.

Authors

Anthony G. Christodoulou (anthony.christodoulou@cshs.org) received his B.S. and M.S. degrees in electrical engineering from the University of Southern California, Los Angeles, in 2008 and 2009, respectively, and his Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign in 2015. He is an assistant professor with the Biomedical Imaging Research Institute at Cedars–Sinai Medical Center, Los Angeles. His research focuses on cardiovascular magnetic resonance (MR) imaging physics and image reconstruction. He has authored more than 20 peer-reviewed publications in outlets including *Nature Biomedical Engineering*, *Magnetic Resonance in Medicine*, and *IEEE Transactions on Medical Imaging*. He is a Member of the IEEE, Society for CMR (SCMR), International Society for Magnetic Resonance in Medicine, and American Heart Association. His work on 4D cardiovascular MR (CMR) received the best paper award at the 2011 IEEE Engineering in Medicine and Biology Conference; his research on multidimensional quantitative cardiovascular imaging, dubbed *MR multitasking*, received the 2016 Society for MR Angiography Best Oral Presentation Award, was a finalist for the 2017 SCMR Early Career Award, and won the 2018 SCMR Clinical Seed Grant Award.

Sajan Goud Lingala (sajangoud-lingala@uiowa.edu) received his B.E. degree in biomedical engineering from the Osmania University, Hyderabad, India, in 2006, his M.Tech. degree in biomedical engineering from the Indian Institute of Technology in 2008, and his Ph.D. degree in biomedical engineering from the University of Iowa, Iowa City, in 2013. He is an assistant professor in the Department of Biomedical Engineering and the Department of Radiology at the University of Iowa. His research interests include the design of advanced acquisition and reconstruction methods for rapid and informative magnetic resonance imaging exams. He has contributed to more than 23 articles in technical and clinical journals and more than 70 peer-reviewed conference publications, including *IEEE Transactions on Medical Imaging*, *Magnetic Resonance in Medicine*, *Journal of Magnetic Resonance Imaging*, and *Medical Physics*, as well as two patents and one book chapter. He received awards including the Junior Fellow Distinction from the International Society of Magnetic Resonance in Medicine (ISMRM), Rex Montgomery Best Dissertation Prize from the University of Iowa, Provost's Postdoctoral Grant from the University of

Southern California, American Heart Association Pre-Doctoral Fellowship, and summa- and magna cum laude distinctions for several abstracts presented at ISMRM annual meetings. He is a Member of the IEEE and ISMRM.

References

[1] 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, Dec 2007. doi: 10.1002/mrm.21391.

[2] H. Jung, K. Sung, K. Nayak, E. Kim, and J. Ye, "*k-t* FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, Jan 2008. doi: 10.1002/mrm.21757.

[3] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in. *Proc. 2007 4th IEEE Int. Symp. Biomedical Imaging: From Nano to Macro*, pp. 988–991. doi: 10.1109/ISBI.2007.357020.

[4] H. Pedersen, S. Kozerke, S. Ringgaard, K. Nehrke, and W. Kim, "*k-t* PCA: Temporally constrained *k-t* BLAST reconstruction using principal component analysis," *Magn. Reson. Med.*, vol. 62, no. 3, pp. 706–716, Sept. 2009. doi: 10.1002/mrm.22052.

[5] S. 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. doi: 10.1109/TMI.2010.2100850.

[6] S. G. Lingala and M. Jacob, "Blind compressive sensing dynamic MRI," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1132–1145, June 2013. doi: 10.1109/TMI.2013.2255133.

[7] L. Feng, L. Axel, H. Chandarana, K. T. Block, D. K. Sodickson, and R. Otazo, "XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 775–788, Feb 2016. doi: 10.1002/mrm.25665.

[8] J. He, Q. Liu, A. G. Christodoulou, C. Ma, F. Lam, and Z.-P. Liang, "Accelerated high-dimensional MR imaging with sparse sampling using low-rank tensors," *IEEE Trans. Med. Imag.*, vol. 35, no. 9, pp. 2119–2129, Sept. 2016. doi: 10.1109/TMI.2016.2550204.

[9] J. Y. Cheng, T. Zhang, M. T. Alley, M. Uecker, M. Lustig, J. M. Pauly, and S. S. Vasanawala, "Comprehensive multi-dimensional MRI for the simultaneous assessment of cardiopulmonary anatomy and physiology," *Sci. Rep.*, vol. 7, no. 1, p. 5330, July 2017. doi: 10.1038/s41598-017-04676-8.

[10] A. G. Christodoulou, J. L. Shaw, C. Nguyen, Q. Yang, Y. Xie, N. Wang, and D. Li, "Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging," *Nat. Biomed. Eng.*, vol. 2, no. 4, pp. 215–226, Apr. 2018. doi: 10.1038/s41551-018-0217-y.

[11] X. Chen, M. Salerno, Y. Yang, and F. H. Epstein, "Motion-compensated compressed sensing for dynamic contrast-enhanced MRI using regional spatiotemporal sparsity and region tracking: Block low-rank sparsity with motion-guidance (BLOSM)," *Magn. Reson. Med.*, vol. 72, no. 4, pp. 1028–1038, Nov. 2013. doi: 10.1002/mrm.25018.

[12] S. Goud Lingala, E. DiBella, and M. Jacob, "Deformation corrected compressed sensing (DC-CS): A novel framework for accelerated dynamic MRI," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 72–85, Jan. 2015. doi: 10.1109/TMI.2014.2343953.

[13] S. Poddar and M. Jacob, "Dynamic MRI using smoothness regularization on manifolds (SToRM)," *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1106–1115, Apr. 2016. doi: 10.1109/TMI.2015.2509245.

[14] U. Nakarmi, Y. Wang, J. Lyu, D. Liang, and L. Ying, "A kernel-based low-rank (KLR) model for low-dimensional manifold recovery in highly accelerated dynamic MRI," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2297–2307, Nov. 2017. doi: 10.1109/TMI.2017.2723871.

[15] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, Feb. 2018. doi: 10.1109/TMI.2017.2760978.

[16] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional recurrent neural networks for dynamic MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 280–290, Jan. 2019. doi: 10.1109/TMI.2018.2863670.

[17] S. Biswas, H. K. Aggarwal, and M. Jacob, "Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM," *Magn. Reson. Med.*, vol. 82, no. 1, Mar. 2019. doi: 10.1002/mrm.27706.

[18] Z.-P. Liang, H. Jiang, C. Hess, and P. Lauterbur, "Dynamic imaging by model estimation," *Int. J. Imag. Syst. Technol.*, vol. 8, no. 6, pp. 551–557, 1997. doi: 10.1002/(SICI)1098-1098(1997)8:6<551::AID-IMA7>3.0.CO;2-9.

[19] N. Aggarwal and Y. Bresler, "Patient-adapted reconstruction and acquisition dynamic imaging method (PARADIGM) for MRI," *Inverse Probl.*, vol. 24, no. 4, p. 045015, July 2008. doi: 10.1088/0266-5611/24/4/045015.

[20] M. Doneva, P. Börnert, H. Eggers, C. Stehning, J. Sénégas, and A. Mertins, "Compressed sensing reconstruction for magnetic resonance parameter mapping," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1114–1120, June 2010. doi: 10.1002/mrm.22483.

[21] 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, Sept. 2012. doi: 10.1109/TMI.2012.2203921.

[22] 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.

[23] T. M. Quan and W.-K. Jeong, "Compressed sensing dynamic MRI reconstruction using GPU-accelerated 3D convolutional sparse coding," in *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2016)*, pp. 484–492. doi: 10.1007/978-3-319-46726-9_56.

[24] D. L. Donoho, "High-dimensional data analysis: The curses and blessings of dimensionality," in *Proc. AMS Conf. Math Challenges 21st Century*, 2000. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.329.3392

[25] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009. doi: 10.1137/07070111X.

[26] B. Yaman, S. Weingärtner, N. Kargas, N. D. Sidiropoulos, and M. Akçakaya, "Low-rank tensor models for improved multi-dimensional MRI: Application to dynamic cardiac T1 mapping," *IEEE Trans. Comput. Imaging*, to be published. doi: 10.1109/TCI.2019.2940916.

[27] M. Akçakaya, T. A. Basha, B. Goddu, L. A. Goepfert, K. V. Kissinger, V. Tarokh, W. J. Manning, and R. Nezafat, "Low-dimensional-structure self-learning and thresholding: Regularization beyond compressed sensing for MRI reconstruction," *Magn. Reson. Med.*, vol. 66, no. 3, pp. 756–767, Sept. 2011. doi: 10.1002/mrm.22841.

[28] A. Bustin, G. Lima da Cruz, O. Jaubert, K. Lopez, R. M. Botnar, and C. Prieto, "High-dimensionality undersampled patch-based reconstruction (HD-PROST) for accelerated multi-contrast MRI," *Magn. Reson. Med.*, vol. 81, no. 6, pp. 3705–3719, June 2019. doi: 10.1002/mrm.27694.

[29] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Sci.*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000. doi: 10.1126/science.290.5500.2319.

[30] U. Nakarmi, K. Slavakis, and L. Ying, "MLS: Joint manifold-learning and sparsity-aware framework for highly accelerated dynamic magnetic resonance imaging," in *Proc. 2018 IEEE 15th Int. Symp. Biomed Imaging (ISBI 2018)*, pp. 1213–1216. doi: 10.1109/ISBI.2018.8363789.

[31] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, Sept. 2012. doi: 10.1109/TIP.2012.2199324.

[32] Y. Q. Mohsin, S. G. Lingala, E. DiBella, and M. Jacob, "Accelerated dynamic MRI using patch regularization for implicit motion compensation," *Magn. Reson. Med.*, vol. 77, no. 3, pp. 1238–1248, Apr. 2016. doi: 10.1002/mrm.26215.

[33] S. Poddar, Y. Mohsin, D. Ansah, B. Thattaliyath, R. Ashwath, and M. Jacob, "Manifold recovery using kernel low-rank regularization: Application to dynamic imaging," *IEEE Trans. Comput. Imag.*, to be published.

[34] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug 2013. doi: 10.1109/TPAMI.2013.50.

[35] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *Proc. 2016 IEEE 13th Int. Symp. Biomed Imaging (ISBI)*, pp. 514–517. doi: 10.1109/ISBI.2016.7493320.

[36] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nat.*, vol. 555, no. 7697, pp. 487–492, Mar. 2018. doi: 10.1038/nature25988.

[37] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, June 2018. doi: 10.1002/mrm.26977.

[38] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019. doi: 10.1109/TMI.2018.2865356.

[39] R. Ramb, M. Zenge, L. Feng, M. Muckley, C. Forman, L. Axel, D. Sodickson, and R. Otazo, "Low-rank plus sparse tensor reconstruction for high-dimensional cardiac MRI," in *Proc. Int. Society Magnetic Resonance Medicine*, 2017, p. 1199.

SP