



Published in final edited form as:

IEEE Trans Med Imaging. 2017 November ; 36(11): 2297–2307. doi:10.1109/TMI.2017.2723871.

A Kernel-based Low-rank (KLR) Model for Low-dimensional Manifold Recovery in Highly Accelerated Dynamic MRI

Ukash Nakarmi [Student Member, IEEE],

Department of Electrical Engineering, University at Buffalo, NY, 14260, USA

Yanhua Wang [Member, IEEE],

School of Information and Electronics, Beijing Institute of Technology, Beijing, China

Jingyuan Lyu [Student Member, IEEE],

Department of Electrical Engineering, University at Buffalo, NY, 14260, USA

Dong Liang [Senior Member, IEEE], and

Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Shenzhen, Guangdong 518055, China

Leslie Ying* [Senior Member, IEEE]

Department of Electrical Engineering, Department of Biomedical Engineering, University at Buffalo, NY, 14260, USA

Abstract

While many low rank and sparsity based approaches have been developed for accelerated dynamic magnetic resonance imaging (dMRI), they all use low rankness or sparsity in input space, overlooking the intrinsic nonlinear correlation in most dMRI data. In this paper, we propose a kernel-based framework to allow nonlinear manifold models in reconstruction from sub-Nyquist data. Within this framework, many existing algorithms can be extended to kernel framework with nonlinear models. In particular, we have developed a novel algorithm with a kernel-based low-rank (KLR) model generalizing the conventional low rank formulation. The algorithm consists of manifold learning using kernel, low rank enforcement in feature space, and preimaging with data consistency. Extensive simulation and experiment results show that the proposed method surpasses the conventional low-rank-modeled approaches for dMRI.

Index Terms

Low rank models; compressed sensing; kernel method; preimaging; manifold recovery

I. Introduction

DYNAMIC magnetic resonance imaging (dMRI) has been widely used in many clinical applications such as cardiac cine imaging and dynamic contrast enhanced imaging, due to its

Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

*Leslie Ying: (leiying@buffalo.edu).

ability to reveal spatial structures and kinetic information simultaneously. The data acquisition in MRI is modeled mathematically as sampling of the k -space, i.e., the Fourier transform of the desired image. The image is usually reconstructed from these samples using the well-known procedure of truncated Fourier series. In dMRI, to recover the image series using the conventional Fourier reconstruction, the Nyquist sampling criterion has to be satisfied in both k -space and the temporal direction. However, due to the slow data acquisition speed in MRI, this criterion is difficult to satisfy, which results in aliasing artifacts or motion artifacts. Research efforts have been made to reconstruct a high-quality image series from MRI data sampled below the Nyquist rate by exploiting spatial and/or temporal correlations in the image series [1]–[4]. A recent paradigm in sub-Nyquist sampling and signal recovery, Compressed Sensing (CS) has shown promising capability of accelerating data acquisition process without significant loss in signal recovery. Several types of constraints such as sparsity [5]–[17] and low-rankness [3], [18]–[21] have been explored at length in many literature. For example, the partial separable (PS) [18], [19] and sparsity with low rank (SLR) [20] methods suit dynamic data with slow variation, where the temporal basis are approximated nearly accurate from the navigator lines. This assumption is violated when the image series has higher temporal variation such as in perfusion imaging, dynamic imaging with fewer temporal frames and dMRI with motion.

Manifold models [22]–[27] have been widely used in machine learning for nonlinear dimensionality reduction, and it has been shown to be superior to linear approaches such as principal component analysis [28] and multidimensional scaling [29]. However, different from dimensionality reduction applications where the high-dimensional manifold-modeled signal is the input and is not of interest in the output, in image reconstruction, the signal is the output to recover. A few algorithms [30]–[35] have been developed for manifold-modeled signal recovery outside the MRI context. Recently, some papers have studied manifold for dMRI reconstruction [36]–[39]. Specifically, they use the idea of heat kernels and spectral graph theory [25], [40] to characterize the relationship between dMRI image series. For example, in [36], [38], a graph Laplacian is constructed to compute the image similarity, and the eigenfunction of the graph Laplacian provides a smooth embedding for dynamic image series. Such an approach is helpful when the image series consists of highly correlated and distributed image sequence along temporal direction, but might not be effective in the cases when very few temporal frames are available, and/or signal varies significantly along the temporal direction. In the meantime, kernel based methods have also been exploited to learn the intrinsic signal manifold and extensively used for signal embedding, classification, regression and denoising [41]–[43].

In this work, we establish a novel kernel-based framework to learn and incorporate a nonlinear manifold into constrained reconstruction from reduced acquisition. In particular, we use kernel principal component analysis (KPCA) [43], [44] to learn the manifold described by the principal components of the feature space. Low Rank in feature space is enforced by projecting undersampled signal into principal components of features space and preimaging technique is applied to project back the signal from feature space to input image space. We have previously shown that the kernel methods can be used to exploit the nonlinear correlation in the dMRI data [45]–[47]. Here, we put dMRI reconstruction in the context of manifold recovery from compressive measurements which is supported by

theoretical results [26], [43], [44]. We use kernel PCA to learn the manifold. Supported by numerical simulations, it is hypothesized that the model learned by kernel PCA can more efficiently (i.e., lower rank in feature space) represent the dynamic images when few temporal frames are acquired, and as a result, the reconstruction based on such a model is more faithful. We thereby propose an efficient manifold recovery algorithm in the kernel-based framework, extending results in our conference papers [46], [47]. Different from kernel compressed sensing [34] which is a nonlinear generalization of sparsity-based reconstruction using kernel dictionary learning, our proposed method is a nonlinear generalization of low rank matrix recovery using kernel principal component analysis. A number of simulation and experimental results are shown to support the hypothesis.

The rest of the paper is organized as follows. In section II-A, we provide a brief review to classical low rank approaches to dMRI. Section II-B presents some background on manifold learning and embeddings and section II-C provides the background of kernel methods, the foundation that our method resides on. Section III elaborates on our proposed method. In section IV, the proposed method is evaluated using extensive simulation and experimental results, followed by discussion on various aspects of the proposed method and results in section V. Finally, section VI concludes the paper.

II. Background Theory

A. Signal recovery under low-rank constraints

In dMRI, the k -space measurement at time t , denoted as $d(\mathbf{k}, t)$ can be represented as

$$d(\mathbf{k}, t) = \int \gamma(\mathbf{r}, t) e^{-j2\pi(\mathbf{k} \cdot \mathbf{r})} d\mathbf{r} \quad (1)$$

where $\gamma(\mathbf{r}, t)$ is the desired dynamic image at time t . The discrete version of dynamic image series can be written as an $M \times N$ Casorati matrix Γ , whose $(m, n)^{th}$ entry is defined as $\gamma(\mathbf{r}_m, t_n)$, representing the value at the corresponding spatial location and time point [3], [18]. Here, M is the number of voxels in the image and N is the number of frames in the data set. Assuming that data are collected in k -space at a sub-Nyquist sampling rate, the imaging equation can be written as:

$$\mathbf{y} = A(\Gamma) \quad (2)$$

where $A(\cdot)$ is a linear operator that performs spatial Fourier transform on the image series with sub-Nyquist sampling and stacks the resulting k -space measurements into a vector $\mathbf{y} = \mathbb{C}^D$, $D \ll MN$. Solving Eq. (2) for Γ is a highly ill-posed problem. To address this issue, the low-rank (or partial separability) model has been adopted as the constraint under which the dynamic image series is recovered from under-sampled data [3], [18]–[20]. Specifically, reconstructing $\gamma(\mathbf{r}, t)$ from the undersampled data $d(\mathbf{k}, t)$ can be regarded as recovering the low rank Casorati matrix Γ . One way is to use the convex nuclear norm or non-convex

Schatten p -norm ($p < 1$) [20] as surrogates of the rank of $\mathbf{\Gamma}$. Because the nuclear norm is a convex function, the problem can be solved via semi definite programming:

$$\text{minimize } \|\mathbf{\Gamma}\|_* \quad \text{s. t. } \mathbf{y} = \mathbf{A}(\mathbf{\Gamma}) \quad (3)$$

where $\|\cdot\|_*$ denotes the nuclear norm. An alternative solution is to enforce the low-rank constraint explicitly with a pre-defined rank l by

$$\mathbf{\Gamma}^{M \times N} = \mathbf{U}_s^{M \times l} \mathbf{V}_t^{l \times N} \quad (4)$$

where \mathbf{U}_s is a spatial coefficient matrix and \mathbf{V}_t a temporal basis. In partial-separability-based methods [3], [18], [19], [48], \mathbf{V}_t is usually estimated a priori by applying singular value decomposition (SVD) to the k -space data obtained with temporal Nyquist rate.

B. Manifold

Loosely speaking, a manifold \mathcal{M} is a topological space which is locally similar to the Euclidean space. The broad class of manifold models arise both in parametric settings where a low-dimensional parameter θ controls the generation of the signal, as well as in non-parametric settings. In parametric settings, for example, manifold models arise when the signals of interest are generated continuously as a function of some d -dimensional parameter. Non-parametric models also arise in many applications where the parameter that controls the generation of manifolds are not known. Both parametric and nonparametric cases are guided by the fundamental assumption that an N dimensional signal $\mathbf{x}_j \in \mathbb{C}^N$ lie on or close to a smooth d dimensional manifold \mathcal{M} such that $d \ll N$, $\mathcal{M} \subset \mathbb{C}^N$. The goal of manifold learning is to find the low dimensional manifold \mathcal{M} through some nonlinear dimension reduction methods. Manifold learning has been extensively used in machine learning community for data classification and clustering [49], [50]. However, application of manifold learning in signal recovery is limited due to several challenges such as the lack of sufficient training data, the complexity in learning the manifold model, and the need for preimaging.

C. Kernel Method

Among many manifold learning algorithms, kernel principal component analysis (KPCA) is widely used and has the potential to be applied in signal recovery due to its capability to map testing data onto the embedding obtained from the training data, and the existence of explicit representation of preimages in some cases. Unlike many other approaches such as Isomap [22], Laplacian eigenmaps [25], locally linear embedding [23], and Hessian maps [24] which use local geometry to learn the manifold, KPCA [44] uses global correlation to learn such low-dimensional embedding. KPCA can be considered as a nonlinear generalization of principal component analysis [44].

The idea of kernel method is to nonlinearly transform the data from the original input space to a higher dimensional feature space such that linear operations in the feature space can

represent a class of nonlinear operations in the input space. The nonlinear mapping $\phi: \mathcal{X} \rightarrow \mathcal{H}$, $\mathbf{x} \rightarrow \phi(\mathbf{x})$ that maps the input \mathbf{x} from $\mathcal{X} \rightarrow \mathcal{H}$ establishes a bridge between the input space \mathcal{X} and feature space \mathcal{H} . In KPCA, performing the linear PCA operation in the feature space is equivalent to performing a nonlinear PCA in input space, thus achieving the goal of learning the nonlinear manifold.

Specifically, in KPCA [44], the eigenvectors \mathbf{v} of the covariance matrix

$\Sigma = \frac{1}{L} \sum_{i=1}^L \bar{\phi}(\mathbf{x}_i) \bar{\phi}(\mathbf{x}_i)^T$ is given by

$$\mathbf{v} = \sum_{i=1}^L \alpha_i \bar{\phi}(\mathbf{x}_i), \quad (5)$$

where $\bar{\phi}(\mathbf{x}_i) = \phi(\mathbf{x}_i) - \frac{1}{L} \sum_{i=1}^L \phi(\mathbf{x}_i)$ is the mean centered data and L is the numbers of signals.

These eigenvectors represent the principal components (PC) of the data in the feature space. The coefficients $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_L]^T$ can be found by solving the eigen system equation $\lambda \mathbf{v} = \Sigma \mathbf{v}$, or equivalently,

$$\mathbf{K}^c \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha}, \quad (6)$$

where $\mathbf{K}^c = \mathbf{K} - \mathbf{1}_T \mathbf{K} - \mathbf{K} \mathbf{1}_T + \mathbf{1}_T \mathbf{K} \mathbf{1}_T$ is the centered kernel matrix,

$$\mathbf{K} = \begin{pmatrix} \kappa(\mathbf{x}_1, \mathbf{x}_1) & \cdots & \kappa(\mathbf{x}_1, \mathbf{x}_L) \\ \vdots & \ddots & \vdots \\ \kappa(\mathbf{x}_T, \mathbf{x}_1) & \cdots & \kappa(\mathbf{x}_T, \mathbf{x}_L) \end{pmatrix}, \quad (7)$$

$\mathbf{1}_L$ is an $L \times L$ matrix with all its elements equals to $1/L$, λ and $\boldsymbol{\alpha}$ are the corresponding eigenvalue and eigenvector, and

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle_{\mathcal{H}}, \quad (8)$$

is the so-called kernel function with $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ denoting inner product in the feature space.

Different λ and $\boldsymbol{\alpha}$ contribute to different PCs in the feature space. To map a testing signal \mathbf{x} onto the low-dimensional embedding obtained by performing KPCA on the training data, we calculate inner product $\beta_q = \langle \mathbf{v}_q, \phi(\mathbf{x}) \rangle_{\mathcal{H}}$, where q is the index for the principal components.

It is seen that KPCA does not require explicit calculation of the nonlinear mapping $\phi(\mathbf{x})$, but only needs knowledge of the kernel functions $\kappa(\mathbf{x}_i, \mathbf{x}_j)$ [44]. The choice of the kernel function (as thus the nonlinear mapping) is critical for the capability of representing the

manifold effectively. For example, the Swiss roll manifold cannot be well represented using standard and widely used kernel functions such as Gaussian kernels $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/c)$ or polynomial kernels $\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle_{\mathcal{X}} + c)^d$, but can be well represented by constructing data-dependent kernel matrix [42], [44], [51].

In signal recovery, a corrupted testing signal is embedded on the learned low-dimensional manifold. Another fundamental challenge in using the kernel method for signal recovery is that the final embedded testing signal needs to be given in the input space rather than in the feature space. This necessitates the preimaging problem, that finds a so-called preimage $\mathbf{z} \in \mathcal{X}$ such that $\phi(\mathbf{z}) = \rho$, where ρ is the embedded signal in feature space. Since ϕ is usually highly nonlinear, the preimage does not necessarily exist nor is unique. The choice of kernel leads to different methods of preimaging. For example, for polynomial kernel of the form,

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + c)^d \quad (9)$$

with $c > 0$ and d being odd, there exists an invertible function f_{κ} such that $\kappa(\mathbf{x}_i, \mathbf{x}_j) = f_{\kappa}(\langle \mathbf{x}_i, \mathbf{x}_j \rangle)$. In this case, there exists an explicit expression for the preimage $\mathbf{z} \in \mathcal{X}$ of $\phi(\mathbf{x})$ and is given by [52, ch.18, pp 544–546]

$$\mathbf{z} = \sum_{n=1}^N f_{\kappa}^{-1} \left(\sum_{l=1}^L \gamma_l \kappa(\mathbf{x}_l, \xi_n) \right) \xi_n, \quad (10)$$

where ξ is any orthonormal basis of the original input space, and $\gamma_l = \sum_{q=1}^Q \beta_q \alpha_l^q$ where β_q is the projection coefficient of $\phi(\mathbf{x})$ onto the q^{th} PC \mathbf{v}_q of the feature space. For many other kernels, preimages can only be obtained through approximations [53], [54].

While KPCA has been widely used in pattern recognition [41], [44], data clustering and classification [49], these applications do not require the preimaging because the desired outputs are only the class that signals belong to, not the recovery of signals in the input space. Some image de-noising applications [34], [55], [56] employ KPCA and preimaging techniques, but require large numbers of training data.

III. Proposed Method

We assume the dynamic image series \mathbf{x} lies on a low-dimensional manifold \mathcal{M} . Let $\phi: \mathcal{X} \rightarrow \mathcal{H}$ be a nonlinear function that maps the signal \mathbf{x} from input space to $\phi(\mathbf{x})$ in feature space. Using kernel method, we formulate the problem to reconstruct the dynamic image series from undersampled (k, t) -space measurements \mathbf{y} as

$$\min_{\Gamma} \text{rank}(\phi(\Gamma)) \text{ s.t. } \mathbf{y} = \mathbf{A}(\Gamma), \quad (11)$$

where $\{\mathbf{x}_i\}_{i=1}^M$ are row vectors of $\mathbf{\Gamma}$ representing the temporal profile at each spatial location, $\phi(\mathbf{\Gamma}) = [\phi(\mathbf{x}_1^T), \phi(\mathbf{x}_2^T), \dots, \phi(\mathbf{x}_M^T)]^T$ is a representation of the row vectors of $\mathbf{\Gamma}$ in the feature space using kernel method, \mathbf{A} is combined partial Fourier and vectoring operator that gives a vectored representation \mathbf{y} of an undersampled k -space data. The undersampling pattern for \mathbf{y} is designed in such way that a few central k -space lines are fully sampled whereas higher frequencies are randomly undersampled at each frame and different random sampling patterns are adopted at different frames, as shown in the Fig. 1. Based on the properties that KPCA is able to reduce the dimension of data nonlinearly, we expect the Casorati matrix to have even lower rank when mapped to the feature space than in the input space with a proper choice of nonlinear mapping function ϕ . Since the manifold on which the dynamic MRI temporal profiles lie is unknown in general, some training data are acquired to learn the underlying manifold. Here we use KPCA to learn the manifold and enforce such manifold structure on the reconstructed image series while ensuring consistency with the k -space measurements. To solve Eq. (11) we use the following three distinct steps: (A) Manifold learning using KPCA, (B) Low rank enforcement in feature Space and (C) Preimaging with data consistency constraint. Step B and Step C are then iterated until convergence. The proposed method can be viewed as a generalization of the PS-based reconstruction method [3] in the feature space, where step (A) is to find the temporal basis in the feature space using training data, step (B) is to recover the spatial basis and thus the signal in the feature space, and finally step (C) is to map the signal from the feature space to obtain the final image series in the original space.

A. Manifold learning using KPCA

The objective of this step is to find the low-dimensional embedding of the dynamic MR images. Since we don't have knowledge of the desired images, we use the temporal profiles of the low resolution dynamic images as the training data to learn the low-dimensional embedding. Specifically, given a low resolution dynamic image series ($\mathbf{\Gamma}_{\text{low}}$) obtained from zero-filled reconstruction of some central k -space data (typically 15–25 lines), a set of T training signals, \mathbf{p}_t , $t = 1, 2, \dots, T$ are formed where each \mathbf{p}_t is an N dimensional temporal vector from a specific spatial location of the image series. These T (much smaller than the size of the image) training signals are randomly selected from all spatial locations. To learn the low dimensional manifold structure of the dynamic MRI series, we perform KPCA [44], [51] using the training data (\mathbf{x}_j in Eq. (5) becomes \mathbf{p}_j). The key idea here is that we represent each principal component in feature space as a weighted combination of nonlinear functions (mappings in feature space) of the training signals, that is

$$\mathbf{v}_q = \sum_{t=1}^T \alpha_t^q \bar{\phi}(\mathbf{p}_t). \quad (12)$$

Similar to the PS-based methods where the temporal basis are estimated initially, these principal components in feature space are equivalent to nonlinear temporal basis. If the desired dynamic image series has a low rank representation in feature space, then it can be represented efficiently by very few temporal basis.

B. Low Rank Enforcement in Feature Space

In this step, our objective is to recover the desired dynamic images in feature space. We start with the aliased dynamic images obtained by zero-filled reconstruction of all undersampled k -space data. Under the assumption that the true dynamic images should lie on the low-dimensional manifold learned from step A, we project the aliased dynamic images onto such a low-dimensional manifold. This is realized by enforcing the low rankness on the matrix formed by the aliased dynamic images in the feature space. Specifically, we construct test signals using the temporal profiles of the aliasing dynamic images. Each of the test signals \mathbf{x}_i , $i = 1, 2, \dots, M$ represent the corrupted version of the temporal signals of the desired dynamic images such that $\mathbf{x}_i = \Gamma(i, :)$. We then project the test signals \mathbf{x}_i onto the nonlinear temporal basis computed in step A, which is equivalent to mapping the test signals onto the learned low-dimensional manifold. Let \mathbf{v}_q represent the q^{th} PC in the feature space, the projection of the test signal \mathbf{x}_i on the q^{th} PC is obtained by

$$\beta_q^i = \langle \mathbf{v}_q, \phi(\mathbf{x}_i) \rangle_{\mathcal{H}} = (\mathbf{k}_{xp}^c)^T \boldsymbol{\alpha}^q, \quad (13)$$

where $\mathbf{k}_{xp}^c(i)$ is the mean centered version of $\mathbf{k}_{xp}(i)$ and measures the similarity between the test signal \mathbf{x}_i and all training signals \mathbf{p}_t such that,

$$\mathbf{k}_{xp}(i) = [\kappa(\mathbf{p}_1, \mathbf{x}_i) \kappa(\mathbf{p}_2, \mathbf{x}_i) \cdots \kappa(\mathbf{p}_T, \mathbf{x}_i)]^T, \quad (14)$$

$$\mathbf{k}_{xp}^c(i) = \mathbf{k}_{xp}(i) - \frac{1}{T} \sum_t \mathbf{K}_p(t, i) - \frac{1}{T} \sum_t \mathbf{k}_{xp}(t) + \frac{1}{T^2} \sum_{t_1=1}^T \sum_{t_2=1}^T \mathbf{K}_p(t_1, t_2). \quad (15)$$

It is seen from Eq. (13) that KPCA does not require explicit calculation of the nonlinear mapping $\phi(\mathbf{x})$, to compute the projection coefficients β_q^i but only requires the knowledge of the kernel functions $\kappa(\mathbf{x}_i, \mathbf{x}_j)$ and α which can be calculated using the eigen decomposition of kernel matrix as shown in Eq. (6). Similar to the PS-based reconstruction methods where a pre-determined number of temporal basis is chosen, we choose Q major principal components ($Q \ll T \ll M$) in the feature space corresponding to the Q largest eigenvalues as the nonlinear temporal basis to efficiently represent the temporal variations at all spatial locations. Hence, $\bar{\phi}(\mathbf{x}_i)$ is approximated as:

$$\bar{\phi}(\mathbf{x}_i) \approx \sum_{q=1}^Q \beta_q^i \mathbf{v}_q \approx \sum_{q=1}^Q \sum_{t=1}^T \beta_q^i \alpha_t^q \bar{\phi}(\mathbf{p}_t). \quad (16)$$

where $\bar{\gamma}_t^i = \sum_{q=1}^Q \beta_q^i \alpha_t^q$, and Eq. 16 takes the form of $\bar{\phi}(\mathbf{x}_i) \approx \sum_{t=1}^T \bar{\gamma}_t^i \bar{\phi}(\mathbf{p}_t)$. Then the feature space representation of the test signal takes the form of

$$\phi(\mathbf{x}_i) = \bar{\phi}(\mathbf{x}_i) + \frac{1}{T} \sum_{t=1}^T \phi(\mathbf{p}_t) \approx \sum_{t=1}^T \gamma_t^i \phi(\mathbf{p}_t), \quad (17)$$

where

$$\gamma_t^i = \bar{\gamma}_t^i + \left(1 - \sum_{t=1}^T \bar{\gamma}_t^i\right) / T. \quad (18)$$

The expression of the test signal in Eqs. 16 and (17) can be considered as the low rank representation of the dynamic image series in the feature space by truncating the numbers of principal components in feature space and further soft thresholding on β_q^i as

$$\beta_q^i = \begin{cases} \text{sign}(\beta_q^i) (|\beta_q^i| - S_{\text{th}}), & |\beta_q^i| - S_{\text{th}} > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

The soft threshold value is then updated for the next iteration as $S_{\text{th}} = S_{\text{th}} - (1 - (i_t - 1)C_{\text{step}})s_{\text{step}}$ where i_t is the number of iterations and s_{step} is the step size typically of order 10^{-5} [21], [57]. It should be noted that, although thresholding is typically not required in dimensionality reduction and classification problems because of the high quality test signals, in our case thresholding is desired because the test signals are corrupted temporal signals obtained from aliased image [21], [57]. Furthermore, our empirical results show that soft thresholding is slightly better than hard thresholding.

C. Preimaging with Data Consistency Constraint

After we calculate all the coefficients β_q^i , we have obtained $\bar{\phi}(\mathbf{x}_i)$, the desired image series in feature space. Different from the conventional PS-based method where the reconstruction is completed after Step B, our proposed method further needs to map the estimated image series $\bar{\phi}(\mathbf{x}_i)$ from the feature space back into the original input space. Such a mapping is obtained through the so-called preimaging, which relies on the specific kernel function. Here we use the polynomial kernel of $\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + c)^d$ such that the preimage \mathbf{z}_i of the temporal profile (test signal \mathbf{x}_i) at the i^{th} spatial location can be obtained by generalization of Eq. (10) as,

$$\mathbf{z}_i = \sum_{n=1}^N f_k^{-1} \left(\sum_{t=1}^T \gamma_t^i \kappa(\mathbf{p}_t, \xi_n) \right) \xi_n, \quad (20)$$

After computing preimages for all spatial locations, we have the image series $\mathbf{\Gamma}$, whose rows are the \mathbf{z}_i^T , $i = 1, \dots, M$. To ensure the reconstructed $\mathbf{\Gamma}$ is consistent with the measured undersampled data, the spatial Fourier transform of the dynamic images are replaced by the acquired k -space data at the specific sampled locations while keeping the values at unacquired locations unchanged. Specifically,

$$\mathbf{\Gamma}_k = \mathcal{S}' .* \mathbf{\Gamma}_k + \mathbf{\Gamma}_k^{\text{aq}}, \quad (21)$$

where, $\mathbf{\Gamma}_k$ is the k -space equivalent of $\mathbf{\Gamma}$, \mathcal{S}' is the binary inversion of the sampling mask \mathcal{S} , $\mathbf{\Gamma}_k^{\text{aq}}$, is the Casorati form of acquired k -space data (\mathbf{y}), and $(.*)$ represents element wise multiplication. Then the updated dynamic image series is obtained by inverse Fourier transform of the updated k -space data $\mathbf{\Gamma}_k$. Such an updating process has been widely used in MR image reconstruction [15], [16].

Algorithm 1

Kernel Low Rank (KLR) Model

Input:

Casorati matrix ($\mathbf{\Gamma}^{\text{aq}}$) for image series from undersampled data.

Casorati matrix ($\mathbf{\Gamma}_{\text{low}}$) from low resolution training image series.

Output:

Dynamic Image series ($\mathbf{\Gamma}$).

A. Manifold Learning Using Kernel

Construct: $\mathbf{p}_t = \mathbf{\Gamma}_{\text{low}}(t;)$, $t = 1, 2, \dots, M$.

Select T random training signals \mathbf{p}_t .

Compute kernel matrix \mathbf{K}_p and mean centered \mathbf{K}_p^C , using Eq. (7).

Eigen decomposition: $\mathbf{K}_p^C \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha}$.

$\mathbf{\Gamma} = \mathbf{\Gamma}^{\text{aq}}$.

B. Low Rank Enforcement in Feature Space

Construct test signals $\mathbf{x}_i = \mathbf{\Gamma}(i;)$, $i = 1, 2, \dots, M$.

for $i = 1 : M$ **do**

for $q = 1 : Q$ **do**

 Compute: Kernel vector \mathbf{k}_{xp}^C , using Eq. (15).

 Compute: Projection coefficients β_q^i , using Eq. (13).

 Soft threshold β_q^i , using Eq. (19).

 Compute: γ_t^i , using Eq. (18).

end for

```

end for
Update soft threshold  $S_{th}$ .
C. Preimaging with Data Consistency
for  $i = 1 : M$  do
    Compute preimage  $\mathbf{z}_i$ , using Eq. (20).

    Update  $\mathbf{\Gamma}$  using  $\{\mathbf{z}_i\}_{i=1}^M$ .

    end for
Update  $k$ -space data, using Eq. (21).
if Convergence then
    Output dynamic image series  $\mathbf{\Gamma}$ .
else
    Repeat steps B and C.
end if

```

The low rank enforcement (III-B) and the data consistency constraint (III-C) steps are then iterated until convergence. The iteration is continued until certain error tolerance criterion is

met, such as $\frac{\|\mathbf{\Gamma}_{i_t} - \mathbf{\Gamma}_{(i_t-1)}\|_F}{\|\mathbf{\Gamma}_{(i_t-1)}\|_F} < e_{tol}$ where $\|\cdot\|_F$ is the Frobenius norm. Typically, e_{tol} is set to

be in the order of 10^{-4} . Although convergence has not been proved theoretically due to many iterative variables, numerical results show that the proposed method converges empirically. The proposed method is summarized in the Algorithm 1.

IV. Results

We performed simulation and experiments to evaluate the proposed method. Because the temporal variations can be quite different from application to application, our focus was on dynamic perfusion imaging with fewer temporal frames. The pros and cons of the proposed method are also discussed for cardiac cine imaging when the number of temporal frames are higher.

We start with two simulations to validate the proposed model and framework independent of the reconstruction algorithm. The first one is to show kernel PCA can represent the dynamic perfusion phantom images more efficiently than linear PCA, and the second one is to show the training data from low resolution images and from full resolution images generate similar training results. We then use a dynamic perfusion phantom [58] and two sets of in vivo arterial spin labeled (ASL) [59], [60] perfusion dMRI data from calf muscle and myocardium. In all in vivo experiments, the fully sampled data in Cartesian (k, t) space was acquired from an MRI scanner and its conventional Fourier reconstruction was used as the “true” reference for comparison of the reconstruction results from different methods. Besides the proposed method, methods exploiting low rankness in the original spatial-temporal space were also used for reconstruction, including PS with sparsity constraint [18], SLR [20], and CS-PCA [6]. All methods reconstruct the desired dynamic MR images from

the retrospectively undersampled (k, t) space data obtained by using the undersampling pattern as shown in the Fig. 1. For the proposed method, all adjustable parameters including the number of principal components and the soft threshold value in feature space were all tuned heuristically. For all other methods, the results were obtained using the code from the corresponding research groups that developed the methods but were modified with the parameters tuned for the least root normalized mean square error (RNMSE) defined in Eq. (22),

$$\text{RNMSE} = \frac{\|\Gamma_{\text{REF}} - \Gamma_{\text{REC}}\|_F}{\|\Gamma_{\text{REF}}\|_F}, \quad (22)$$

where Γ_{REF} and Γ_{REC} are the matrix formed by the dynamic images from the reference and the reconstruction, respectively, and $\|\cdot\|_F$ denotes the Frobenius norm.

All computations were carried out on a DELL workstation with Intel(R) i7 3.40 GHz processor and 16 GB RAM, running MATLAB 2014.

A. Validation using simulations

We first use two simulations to validate the assumptions made in the proposed framework.

Simulation 1—The objective of this simulation is to demonstrate that kernel PCA is more efficient than linear PCA in representing the dynamic MR image series. We used a numeric perfusion phantom of size 256×256 with 50 time frames. The perfusion phantom represents liver, portal vein, inferior vena cava and pancreas based on enhancement kinetic modeling [58]. The injection rate, frequency and base vibration amplitude of 4 ml/sec, 500 and 30 respectively were chosen to mimic fast and abrupt temporal variation. From the original image series, 1000 training signals (temporal signals at 1000 spatial locations) were randomly selected and used to compute both linear PCs and kernel PCs. The 256×256 test signals (temporal signals at all spatial locations) were then projected onto 5% of the PCs in both linear and KPCA to generate approximations of the dynamic images. For KPCA, polynomial kernel was used and preimaging was performed to obtain the dynamic images for comparison. Figure 2 shows the comparison of spatial results and temporal curves of two regions of interest (ROI). It shows that with the same percentage of PCs, the RNMSE is much lower in KPCA than in linear PCA. The superiority of KPCA is also illustrated by the temporal curves of two ROIs, which suggests that KPCA represents the dynamic images more efficiently than PCA.

Simulation 2—The objective of this simulation is to demonstrate that the training data from low-resolution images is sufficient to learn the low-dimensional embedding of the true, full-resolution images. We replace the training data in Simulation 1 from the true images to the low-resolution images, and perform KPCA approximation again. In Fig. 3, we compare the approximations obtained by low-resolution and full-resolution training data. Both approximations are seen to be very close in the spatial images and temporal variation as

shown in Fig. 3. It is evident that the low-resolution images provide sufficient information for training.

B. Phantom Results

A dynamic perfusion phantom with aperiodic motion with 50 temporal frames (refer to section IV-A for data parameters) was used to simulate the reduced (k, t) data with a reduction factor (R) of 5. The spatial results, as shown in the Fig. 4 illustrate that the aliasing artifacts in the conventional state of art methods are significantly reduced by the proposed method. Figure 5 compares the temporal variation of the cross section along the yellow line indicated in the reference image. The improved performance of the proposed method can be attributed to the ability to reveal the low rankness of the dynamic data in a much higher dimensional feature space, whereas the rank in the original space is not as low when there are only a few temporal frames.

C. ASL Perfusion Results

To further evaluate the proposed method using in vivo experimental data, we tested it on two in-vivo Arterial Spin labeling (ASL) perfusion datasets [59], [60] where only very few frames were acquired and abrupt changes exist in temporal signals (temporally non-smooth). The first dataset was from calf muscle perfusion ($R=4$) and the second one was from myocardial perfusion ($R=3$). Both data were acquired from healthy volunteers on a 3T scanner with a single channel body coil. The calf muscle perfusion data was acquired using a dedicated ASL sequence based on the asymmetric spinecho sequence [59], and the myocardial perfusion data was acquired using an ASL sequence based on single-shot gradient echo techniques [60]. Acquisition parameters for muscle and cardiac perfusion data were $TR/TE = 2.8/1.2$ ms, $2.5/1.1$ ms, flip angle = 5° for both, data matrix size = $112 \times 100 \times 20$, $126 \times 120 \times 12$, $FOV = 160\text{mm} \times 112\text{mm}$, $220\text{mm} \times 200\text{mm}$, respectively. Figures 6 and 7 show images and temporal variations along the cross section for Muscle ASL perfusion data. The PS-sparse and SLR models are not good fit for these datasets because the rank is not low, with only very few time frames and those methods are primarily developed for dMRI modality with many temporal frames. Consequently, these two methods give images with more artifacts, and have comparatively higher RNMSEs than CS-PCA and the proposed method, as illustrated in the Table I. In contrast, the proposed method not only preserves the spatial information of images, but also effectively retains the kinetic information, both of which are essentials in ASL imaging. Similarly, the myocardial ASL perfusion results shown in Fig. 8 and Fig. 9 also suggest the superiority of the proposed method. The myocardium region in CS-PCA reconstruction is seen to have artifacts with blurred edges.

Temporal intensity curves of Region of Interest (ROI) are typically of interest in ASL imaging. Figure 10 shows the temporal intensity curve of a particular ROI from the reference, CS-PCA and the proposed method. This shows the temporal curves from the proposed method follow the reference curve more consistently and precisely than the CS-PCA method.

D. Cardiac Cine Results

Besides perfusion imaging, we also studied the usefulness of the proposed method in cardiac cine imaging using a set of cardiac cine data. The complex-valued multi-cycle cardiac cine data was simulated from human cardiac MR data acquired using a balanced SSFP sequence and single channel body coil on a 3T scanner with retrospective ECG-gating during a single breathhold and acquisition parameters $TR = 3\text{ms}$, matrix size = 200×256 , field of view (FOV) = $273\text{mm} \times 350\text{mm}$, spatial resolution = $1.36\text{mm} \times 1.36\text{mm}$, slice thickness = 6mm . This data was then used to generate multiple time warps. Different heart rate variability and quasi-periodic spatial deformation along different cardiac cycle was obtained through thin plate spline mapping and nonlinear interpolation technique [61]. Each time warps were concatenated and a total of 40 frames were chosen representing about 3 cardiac cycle. It should be noted that the data in different cardiac cycles are not the same and such simulation has been used in [18]. Figure 11 shows the reconstruction results for the cine data with a reduction factor (R) of 5. The zoom-in views of the cardiac region show that the proposed method preserves the cardiac structures and intensity of myocardium better than the other methods. Figure 12 compares the temporal variation of the cross section along the yellow line indicated in the reference image. It demonstrates that the proposed method best preserves the temporal variations. Table I summarizes the RNMSEs and machine time for all test datasets. The qualitative reconstructions and quantitative measures signify that the proposed method outperforms the existing low-rank-based methods in these settings.

However, when the number of frames is increased, as is usually the case in real-time cardiac cine imaging, the proposed method might not bring much advantage. To illustrate such limitation of the proposed method, we increase the frame rate of cardiac cine data in IV-D by about 3 times mimicking smoother temporal variation and a larger number of temporal frames. Retrospectively undersampled data using 1-D Cartesian sampling with a high reduction factor of 7.5 was then used in all reconstructions. From Fig. 13 and Fig. 14 we can see that the conventional PS-sparse method performs better than the proposed method. This is because 1) when the number of temporal frames in dynamic imaging is substantially increased with smoother temporal variations, the conventional approaches are able to find the low rank approximations of the dynamic image series; 2) the achievable reduction factor in the proposed method is limited by the need for low resolution images to perform kernel PCA. The reduction factor cannot be increased as much as PS-sparse with a large number of frames because the proposed method needs quite several navigator k-space lines to perform kernel PCA, while PS-sparse only needs very few lines for PCA.

V. Discussions

A. Relationship with Existing Methods

While several dynamic image reconstruction methods based on the low rank model have demonstrated success in accelerating dMRI with very high reduction factors, a large number of temporal frames are usually required to reveal the low rankness. Motivated by the fact that the kernel method is able to reveal some non-obvious features of signals, here we presented a kernel-based framework to expand the temporal signal of the image series from a low dimensional input space to high dimensional feature space such that the low rankness model

still holds for the cases with few frames. Perfusion imaging is a dynamic MRI example with typically fewer frames, in which kernel-based framework shows improvement over existing methods. Such a framework opens up many possibilities to extend most existing reconstruction algorithms to new algorithms using manifold models. The proposed method is only one of possibilities, extending the low rank model to the nonlinear kernel PCA model.

B. Region of operation and Convergence

The performance of the proposed method depends on the choice of rank Q in feature space and the initial threshold value S_{th} . The choice of Q and threshold affect the tradeoff between the smoothness and details in the reconstructed images. Figure 15 shows the mean squared error $\left(\text{MSE} = \frac{\|\Gamma_{REF} - \Gamma_{REC}\|_F^2}{M \times N} \right)$ for muscle and myocardial ASL data as a function of Q and S_{th} while keeping other parameters fixed. We can clearly see that the error changes with the parameters, but is consistent over a wide range, suggesting the method is insensitive to small changes of the parameters. The values of the parameters giving the least MSE have been used in the results.

Figure 16 shows the performance of proposed method over different polynomial degrees d and constants c . Our experimental results show that the degree of order 3 gives the least MSE and the results are not sensitive significantly to constant c in the range of 100–1000. For better visualization, the MSE axis and the constant axis are shown in log scale in Fig. 16(a) and Fig. 16(b), respectively.

Although the convergence of the proposed algorithm has not been proved theoretically due to different iterative parameters, our results have shown empirically that the proposed algorithm consistently converges. Figure 17 illustrates the convergence curve of the proposed method for different Q values keeping other parameters fixed.

C. Computational Complexity and Machine Time

The proposed method requires longer machine time than the competing methods due to the inherent extra operations in decomposition of the kernel matrix which increases with increase in numbers of training signals, and preimaging process which is not required in competing methods. It is important to note that the computation complexity of eigen-decomposition of the kernel matrix in Eq. (7) is $\mathcal{O}(T^3)$ which only depends on the numbers of training data T (independent of the dimension of the feature space) [43], whereas the conventional low rank approximation costs $\mathcal{O}(MN^2)$ for computation of the covariance matrix and $\mathcal{O}(N^3)$ for singular value decomposition [62] which depends on the size of the image M and number of frames N . The increase in machine time is mainly attributed to the decomposition of kernel matrix and extra preimaging step which is not needed in other approaches.

D. Future Extensions

We present here a new kernel-based framework for dMRI reconstruction from sub-Nyquist acquisitions. The framework naturally motivates future studies on nonlinear sparsity models, and eventually combine nonlinear low rank and sparsity models using kernel method. The proposed method can also be integrated with existing parallel imaging techniques to further accelerate the data acquisition speed. Although our proposed method uses a single kernel for nonlinear low rank representation, multiple-kernel manifold models may be beneficial to application-specific dMRI reconstruction problems. Learning kernels from the undersampled data or data adaptive nonlinear dictionary learning instead of using generic kernels can be a challenging, yet interesting problem for future endeavors.

VI. Conclusion

In this paper we proposed a novel kernel-based framework for reconstructing dynamic MR images using manifold models. Within the framework, we extended the conventional low rank model to the feature space and developed a new method based on a kernel low-rank model. We have demonstrated that fewer nonlinear temporal basis than linear ones are needed to capture the temporal variations in dMRI. The method has shown its practical and successful application to the recovery of dynamic MR images from undersampled data, when existing low rank models fail, and its superior performance to the existing methods. It would be interesting to explore kernel extensions of other compressed-sensing based approaches in future studies.

Acknowledgments

This work is supported in part by the National Science foundation CBET-1265612, NSF CCF-1514403 and National Institute of Health R21EB020861.

We thank the anonymous reviewers for their comments and helping us to improve the quality of the paper significantly. We also thank Drs. Y. Wang (Cornell University), Z.-P. Liang (UIUC) and J. Zheng (Washington University) for helping with acquisition of different data sets.

References

1. Liang Z-P, Lauterbur PC. An efficient method for dynamic magnetic resonance imaging. *Medical Imaging, IEEE Transactions on*. 13(4):677–686.1994;
2. Madore B, Glover GH, Pelc NJ, et al. Unaliasing by fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. *Magnetic Resonance in Medicine*. 42(5):813–828.1999; [PubMed: 10542340]
3. Liang Z-P. Spatiotemporal imaging with partially separable functions. *Biomedical Imaging (ISBI), IEEE International Symposium on*. Apr.2007 :988–991.
4. Tsao J, Kozerke S. MRI temporal acceleration techniques. *Journal of Magnetic Resonance Imaging*. 36(3):543–560.2012; [PubMed: 22903655]
5. Tsao J, Boesiger P, Pruessmann KP. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magnetic Resonance in Medicine*. 50(5):1031–1042.2003; [PubMed: 14587014]
6. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. *Magnetic Resonance in Medicine*. 61(1):103–116.2009; [PubMed: 19097216]

7. Jung H, Ye JC, Kim EY. Improved k-t BLAST and k-t SENSE using FOCUSS. *Physics in medicine and biology*. 52(11):3201.2007; [PubMed: 17505098]
8. Doneva M, Börner P, Eggers H, Stehning C, Sénégas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magnetic Resonance in Medicine*. 64(4): 1114–1120.2010; [PubMed: 20564599]
9. Ravishankar S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *Medical Imaging, IEEE Transactions on*. 30(5):1028–1041.2011;
10. Wang Y, Ying L. Compressed sensing dynamic cardiac cine MRI using learned spatiotemporal dictionary. *Biomedical Engineering, IEEE Transactions on*. 61(4):1109–1120.2014;
11. Bilgic B, Goyal VK, Adalsteinsson E. Multi-contrast reconstruction with bayesian compressed sensing. *Magnetic Resonance in Medicine*. 66(6):1601–1615.2011; [PubMed: 21671267]
12. Chen L, Schabel MC, DiBella EV. Reconstruction of dynamic contrast enhanced magnetic resonance imaging of the breast with temporal constraints. *Magnetic Resonance Imaging*. 28(5): 637–645.2010; [PubMed: 20392585]
13. Ji J, Lang T. Dynamic MRI with compressed sensing imaging using temporal correlations. *Biomedical Imaging (ISBI), IEEE International Symposium on*. 2008:1613–1616.
14. Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI. *Magnetic Resonance in Medicine*. 59(2):365–373.2008; [PubMed: 18228595]
15. Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magnetic Resonance in Medicine*. 64(3): 767–776.2010; [PubMed: 20535813]
16. Liang D, DiBella EV, Chen R-R, Ying L. k-t ISD: Dynamic cardiac mr imaging using compressed sensing with iterative support detection. *Magnetic Resonance in Medicine*. 68(1):41–53.2012; [PubMed: 22113706]
17. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*. 58(6):1182–1195.2007; [PubMed: 17969013]
18. Zhao B, Haldar JP, Christodoulou AG, Liang Z-P. Image reconstruction from highly undersampled (k,t)-space data with joint partial separability and sparsity constraints. *Medical Imaging, IEEE Transactions on*. 31:1809–1820.2012;
19. Haldar J, Liang Z-P. Spatiotemporal imaging with partially separable functions: a matrix recovery approach. *Biomedical Imaging (ISBI), International Symposium on*. Apr.2010 :716–719.
20. Lingala S, Hu Y, Dibella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR. *Medical Imaging, IEEE Transactions on*. 30:1042–1054.2011;
21. Otazo R, Candès EJ, Sodickson DK. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine(Early view)*. 2014
22. Tenenbaum JB, De Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science*. 290(5500):2319–2323.2000; [PubMed: 11125149]
23. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*. 290(5500):2323–2326.2000; [PubMed: 11125150]
24. Donoho DL, Grimes C. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*. 100(10):5591–5596.2003;
25. Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*. 15(6):1373–1396.2003;
26. Muller K, Mika S, Ratsch G, Tsuda K, Scholkopf B. An introduction to kernel-based learning algorithms. *Neural Networks, IEEE Transactions on*. 12(2):181–201.2001;
27. Zhang Z, Zha H. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *Journal of Shanghai University (English Edition)*. 8(4):406–424.2004;
28. Cao L, Chua K, Chong W, Lee H, Gu Q. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing*. 55(12):321–336.2003;
29. Borg I, Groenen PJ, Mair P. *Applied Multidimensional Scaling* Springer Science & Business Media. 2012

30. Chen M, Silva J, Paisley J, Wang C, Dunson D, Carin L. Compressive sensing on manifolds using a nonparametric mixture of factor analyzers: Algorithm and performance bounds. *Signal Processing, IEEE Transactions on*. 58(12):6140–6155.2010;
31. Hegde C, Baraniuk RG. Signal recovery on incoherent manifolds. *Information Theory, IEEE Transactions on*. 58(12):7204–7214.2012;
32. Iwen MA, Maggioni M. Approximation of points on low-dimensional manifolds via random linear projections. *Information and Inference*. 2(1):1–31.2013;
33. Peyré G. Manifold models for signals and images. *Computer Vision and Image Understanding*. 113(2):249–260.2009;
34. Pourkamali Anaraki F, Hughes SM. Kernel compressive sensing. *Image Processing (ICIP), 2013 20th IEEE International Conference on IEEE*. 2013:494–498.
35. Qi H, Hughes S. Using the kernel trick in compressive sensing: Accurate signal recovery from fewer measurements. *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on IEEE*. 2011:3940–3943.
36. Poddar S, Jacob M. Dynamic MRI using smoothness regularization on manifolds (SToRM). *Medical Imaging, IEEE Transactions on*. 35(4):1106–1115.Apr; 2016
37. Schmidt JF, Santelli C, Kozerke S. MR image reconstruction using block matching and adaptive kernel methods. *PLoS one*. 11(4):1–14.2016;
38. Bhatia KK, Caballero J, Price AN, Hajnal JV, Rueckert D. Fast reconstruction of highly-undersampled dynamic MRI using random sampling and manifold interpolation. *ISMRM*. 2015
39. Usman M, Atkinson D, Kolbitsch C, Schaeffter T, Prieto C. Manifold learning based ECG-free free-breathing cardiac cine MRI. *Journal of Magnetic Resonance Imaging*. 41(6):1521–1527.2015; [PubMed: 25124545]
40. Chung FR. *Spectral Graph Theory Regional Conference Series in Mathematics*. 1992; (92)
41. Bishop, CM. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc; 2006.
42. Weinberger KQ, Sha F, Saul LK. Learning a kernel matrix for nonlinear dimensionality reduction. *Twenty First International Conference on Machine Learning (ICML)*. 2004
43. Mika, S, Schölkopf, B, Smola, A, Müller, KR, Scholz, M, Rätsch, G. Kernel PCA and de-noising in feature spaces. In: Kearns, MS, Solla, SA, Cohn, DA, editors. *Advances in Neural Information Processing Systems 11*. Morgan Kaufmann; 1998.
44. Schölkopf B, Smola A, Müller K-R. Kernel principal component analysis. *Artificial Neural Networks ICANN'97 Springer*. 1997:583–588.
45. Zhou Y, Wang Y, Ying L. A kernel-based compressed sensing approach to dynamic MRI from highly undersampled data. *Biomedical Imaging (ISBI), IEEE International Symposium on*. Apr. 2013 :716–719.
46. Wang Y, Ying L. Undersampled dynamic magnetic resonance imaging using kernel principal component analysis. *Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE*. 2014:1533–1536.
47. Nakarmi U, Wang Y, Lyu J, Ying L. Dynamic magnetic resonance imaging using compressed sensing with self-learned nonlinear dictionary (NL-D). *Biomedical Imaging (ISBI), IEEE International Symposium on*. 2015:331–334.
48. Pedersen H, Kozerke S, Ringgaard S, Nehrke K, Kim WY. k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis. *Magnetic Resonance in Medicine*. 62(3):706–716.2009; [PubMed: 19585603]
49. Wang R, Chen X. Manifold discriminant analysis. *Computer Vision and Pattern Recognition, IEEE Conference on*. 2009:429–436.
50. Belkin M, P N, Sindhwani V. *Manifold regularization: A geometric framework for learning from examples*. 2006
51. Ham J, Lee DD, Mika S, Schölkopf B. A kernel view of the dimensionality reduction of manifolds. *International Conference on Machine Learning (ICML)*. 2004
52. Schölkopf B, Smola AJ. *Learning with Kernels Cambridge: MIT Press*. 2001

53. Zheng WS, Lai J, Yuen PC. Penalized preimage learning in kernel principal component analysis. *Neural Networks, IEEE Transactions on*. 21(4):551–570.2010;
54. Honeine P, Richard C. Preimage problem in kernel-based machine learning. *IEEE Signal Processing Magazine*. 28(2):77–88.2011;
55. Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*. 54(11):4311–4322.2006;
56. Nguyen H, Patel VM, Nasrabadi NM, Chellappa R. Kernel dictionary learning. *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on IEEE*. 2012:2021–2024.
57. Donoho DL. De-noising by soft-thresholding. *Information Theory, IEEE Transactions on*. 41(3): 613–627.May; 1995
58. Xu B, Spincemaille P, Chen G, et al. Fast 3D contrast enhanced MRI of the liver using temporal resolution acceleration with constrained evolution reconstruction. *Magnetic Resonance in Medicine*. 69(2):370–381.2013; [PubMed: 22442108]
59. Zheng J, An H, Coggan AR, et al. Noncontrast skeletal muscle oximetry. *Magnetic Resonance in Medicine*. 71(1):318–325.2014; [PubMed: 23424006]
60. Northrup BE, McCommis KS, Zhang H, et al. Resting myocardial perfusion quantification with CMR arterial spin labeling at 1.5 T and 3.0 T. *Journal of Cardiovascular Magnetic Resonance*. 10(1)2008;
61. Tsai Y, Ling H, Hu Y, et al. Thin plate spline technique for medical image deformation. *Journal of medical and biological engineering*. 20:203–209.2000;
62. Du Q, Fowler JE. Low-complexity principal component analysis for hyperspectral image compression. *International Journal of High Performance Computing Applications*. 22(4):438–448.2008;

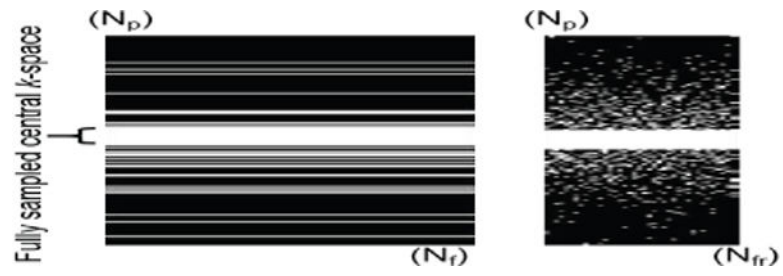


Fig. 1. Undersampling pattern used to simulate retrospective undersampling. Left: sampling pattern along phase (N_p) - frequency encoding (N_f). Right: phase encoding (N_p) - frame (N_{fr}) direction.

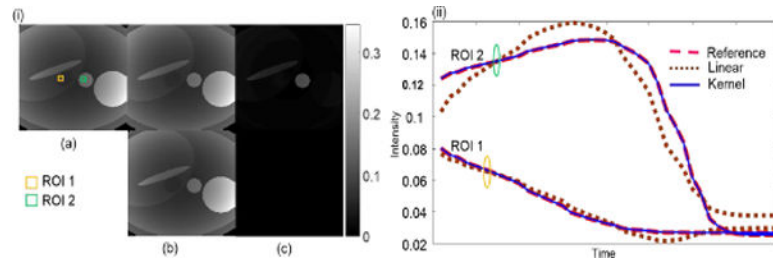


Fig. 2.

(i) Comparison of images and (ii) temporal curves of ROIs using linear and kernel PCA. (a) Reference. (b) Approximations using 5% PCs. Top: Linear PCA; Bottom: KPCA. (c) Corresponding error maps magnified $\times 20$. RNMSE: PCA 0.0182 and KPCA 2.9×10^{-8} , respectively.

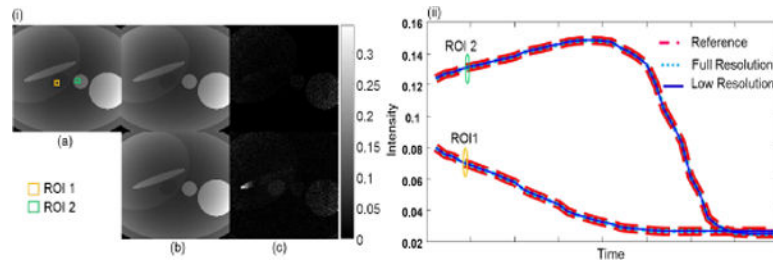


Fig. 3.

(i) Comparison of images and (ii) temporal curves of ROIs approximated using training signals from full resolution and low resolution images. (a) Reference. (b) Approximations using 5% of PCs. Top: using training signals from full resolution images; Bottom: using training signals from low resolution images. (c) Corresponding error maps magnified $\times 10^7$. RNMSE: Full resolution 1.12×10^{-8} , Low resolution 3.14×10^{-7} .

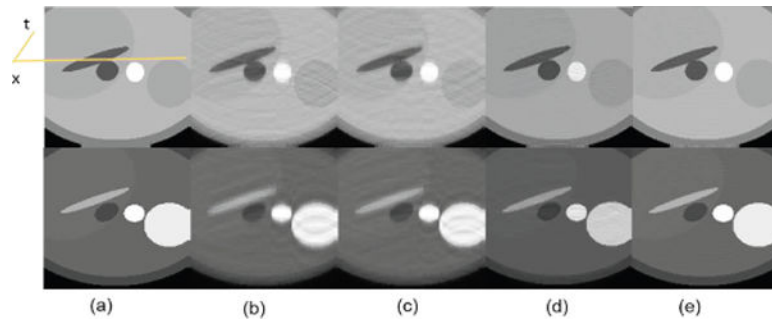


Fig. 4. Spatial results: reconstruction comparison for dynamic perfusion phantom with motion. Two representative frames (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

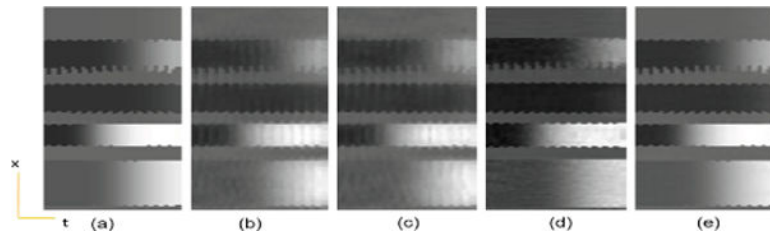


Fig. 5. Dynamic perfusion phantom with motion, temporal variation of a cross section. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

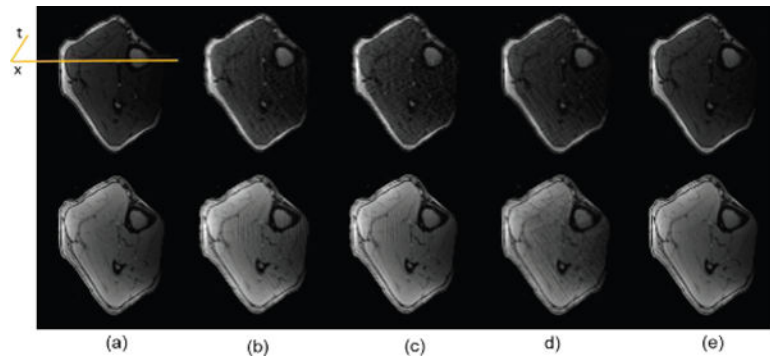


Fig. 6. Spatial results: reconstruction comparison for calf muscle ASL perfusion data. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR. Top: Frame 2; Bottom: Frame 7.

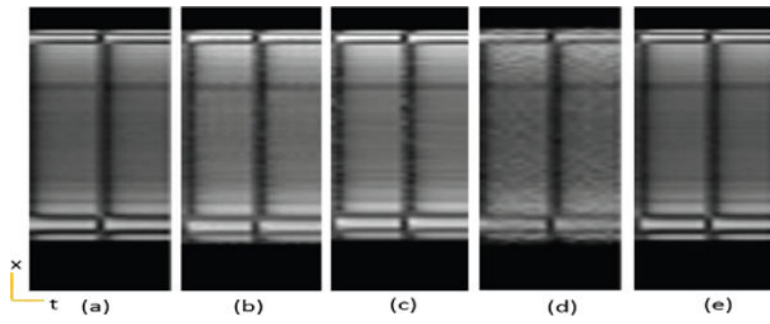


Fig. 7. Calf muscle ASL perfusion data temporal variation of a cross section. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

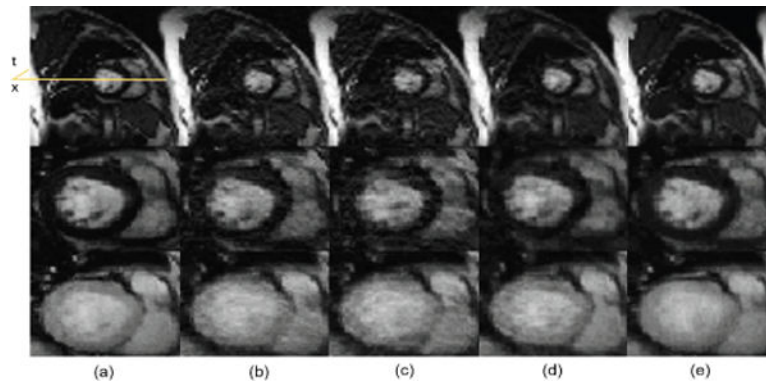


Fig. 8. Spatial results: reconstruction comparison for myocardial ASL perfusion data. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR. Top: Frame 2 Full FOV; Middle: Frame 2 ROI; Bottom: Frame 5 ROI.

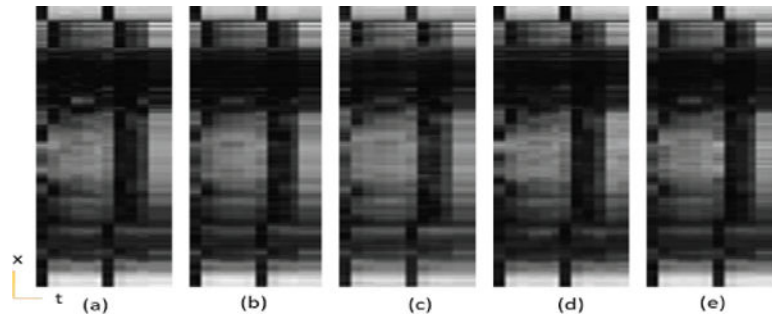


Fig. 9. Myocardial ASL perfusion data temporal variation of a cross section, (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

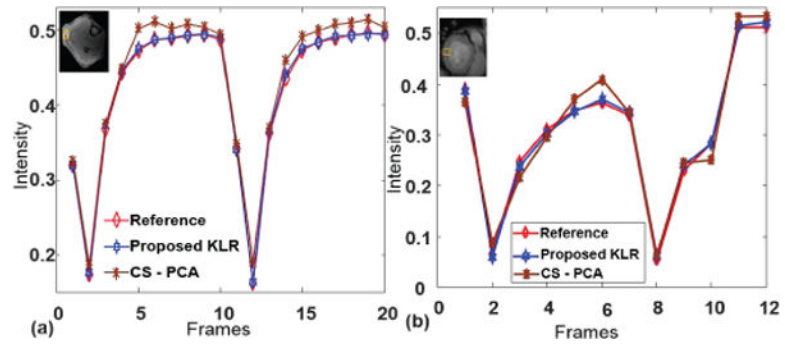


Fig. 10. Intensity curve of an ROI in ASL perfusion (a) Muscle ASL data (b) Cardiac ASL data.

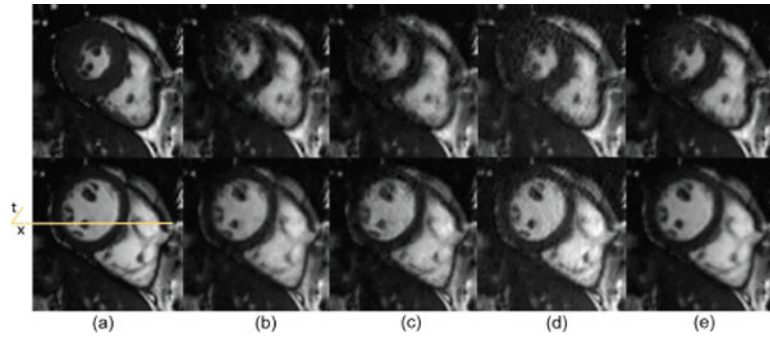


Fig. 11. Spatial results: representative systolic and diastolic frames comparison for dynamic cardiac cine. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

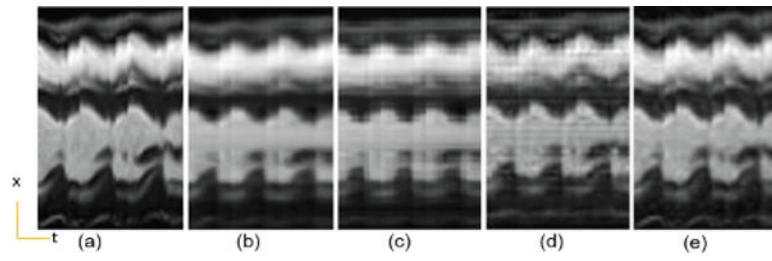


Fig. 12. Cardiac cine temporal variation of a cross section. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

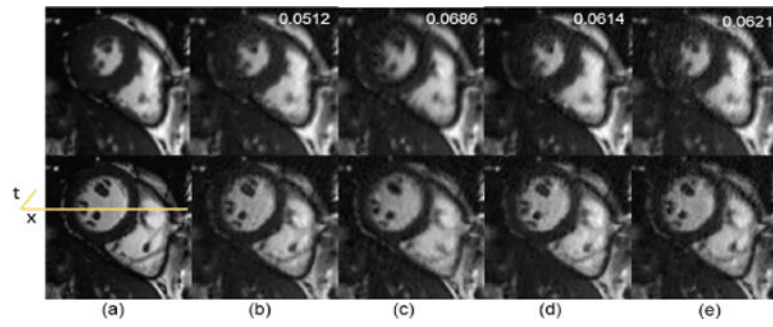


Fig. 13. Spatial results: representative systolic and diastolic frames to illustrate limitation of the proposed method in the case of high frame rate and large numbers of frames. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR. Numbers on top right of each column are respective RNMSE values.

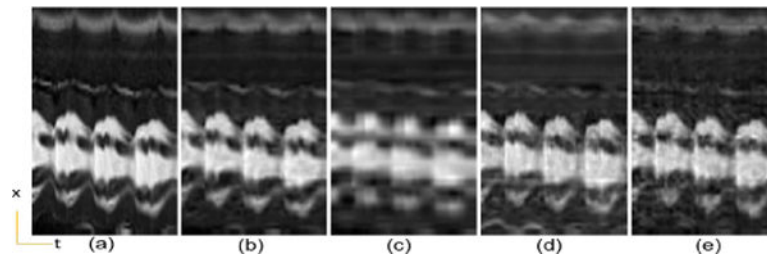


Fig. 14. Cardiac cine temporal variation of a cross section in the case of high frame rates and large number of frames. (a) Reference, (b) PS Sparse, (c) SLR, (d) CS-PCA, (e) Proposed-KLR.

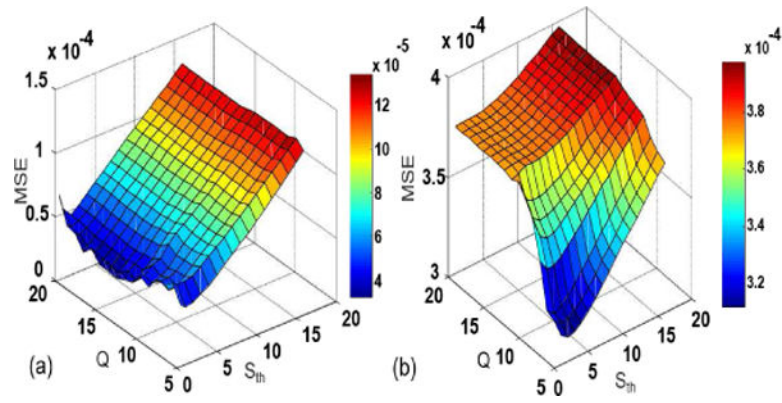


Fig. 15. Region of operation test: MSE for different Q and threshold values. (a) muscle ASL perfusion data, (b) myocardial ASL perfusion data.

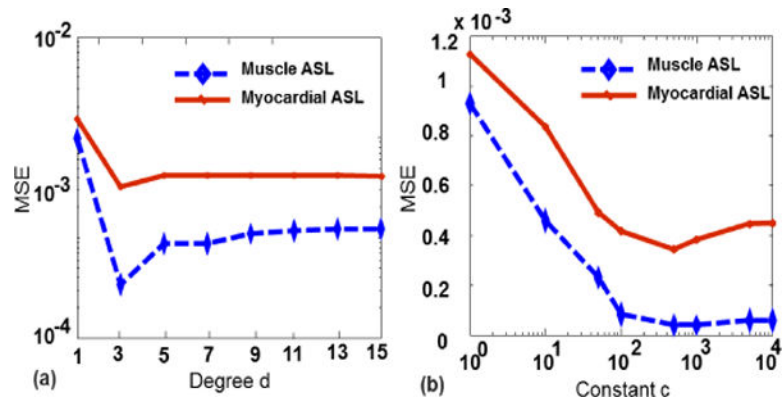


Fig. 16. Performance analysis for different polynomial degree and constant.(a) MSE vs degree d, (b) MSE vs constant c.

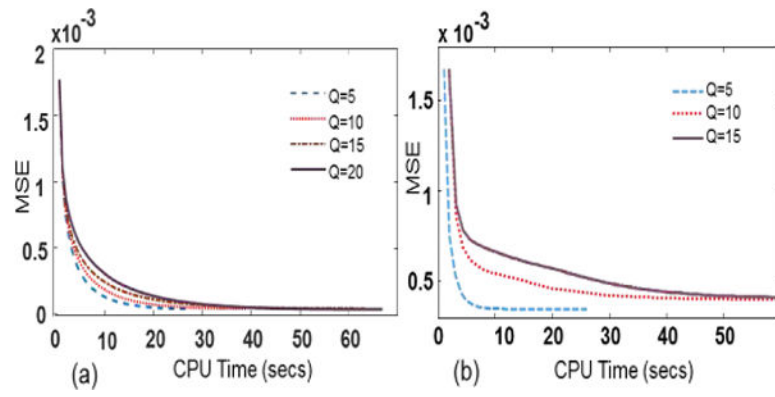


Fig. 17. Convergence test: convergence rate for different Q (a) muscle ASL perfusion data, (b) myocardial ASL perfusion data.

TABLE I

RNNMSE and Machine Time.

Data	RNNMSE						Machine Time (sec)			
	PS Sparse	SLR	CS-PCA	Proposed-KLR	PS Sparse	SLR	CS-PCA	SLR	CS-PCA	Proposed-KLR
Phantom	0.0620	0.0580	0.0315	0.0210	122	148	104	148	104	720
Cardiac cine	0.0314	0.0394	0.0416	0.0214	125	102	85	102	85	676
Muscle ASL	0.822	0.651	0.0711	0.0261	28	32	52	32	52	68
Myocardial ASL	0.868	0.883	0.0983	0.0623	26	25	43	25	43	54